

OPERATIONS MANAGEMENT OF CROWDSOURCING AND CROWD
BEHAVIOR

by

Lu Wang

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Rotman School of Management
University of Toronto

© Copyright 2017 by Lu Wang

ProQuest Number:10623537

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10623537

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Abstract

Operations Management of Crowdsourcing and Crowd Behavior

Lu Wang

Doctor of Philosophy

Graduate Department of Rotman School of Management

University of Toronto

2017

This thesis studies problems in operations management about how to manage crowd behavior. In Chapter 2, we investigate the crowdsourcing contest. In a crowdsourcing contest, innovation is outsourced by a firm to an open crowd that compete in generating innovative solutions. Since the projects typically consist of multiple attributes, we consider two alternative mechanisms. One is a simultaneous contest, where the best solution is chosen from the aggregate solutions simultaneously submitted by all contestants. The other is multiple sequential sub-contests, with each dedicated to one attribute and the contestants asked to build upon the best work in progress from previous sub-contests. The comparison of the expected *best* performances in the two contests depends on the project's characteristics.

In Chapter 3, we study the revenue sharing policy in the subscription platform. Subscription providers such as Spotify, Netflix and OneGo (an all-you-can-fly subscription service provider) crowdsource products/services from many vendors and bundle them for the price of one. The collected subscription fees for the bundle then are allocated according to the realized contributions by each crowdsourced product. However, this allocation scheme may create incentive incompatibility for vendors, given their options of not joining the bundle. We examine the incentive compatibility of different parties under various bundling strategies.

In Chapter 4, we study the customer behavior with an online reservation system. Online reservation system allows customers to join a queue and virtually wait for service before arriving on site. For example, some platforms have been designed to collect the

information of restaurants and show the real-time congestion levels (e.g., Nowait). We consider a model in which customers must travel from their location to the service area and incur a travelling cost. When customers intend to book service online, they are informed about their positions in the queue at the time of booking, so that they make their decision whether to join the queue taking into account both their travelling time and expected waiting time. With those customer behaviors, the optimal policy of a firm heavily relies on the conditions of the travelling and benefit of the service.

Acknowledgements

First and foremost, I would like to express my deepest gratitude and appreciation to my supervisors, Professor Ming Hu, Professor Oded Berman and Professor Opher Baron, for their support, enthusiasm and patience.

I am particularly grateful to Professor Ming Hu whose intellectual insights and constructive criticism help me complete this thesis smoothly. His talents as a scholar have left a profound influence on my thinking. In these five years, he has given me great support and guidance on my academic and career which are the most valuable treasure. I will never forget the time he encourages me to face every hardship and overcome difficulties. I am very thankful to Professor Oded Berman for his unwavering support and unparalleled mentorship. I appreciate his help in shaping the direction of my work and leading me to critically think about the problem. I have learned a lot from him how to be a rigorous researcher. My gratitude goes to Professor Opher Baron. His incisive comments and intellectual curiosity have improved this thesis immeasurably. His enthusiasm towards research continue to inspire me and have a lasting impact on my own career.

I am very fortunate to have received invaluable guidance from my supervisors. I would like to thank them for encouraging my research and for allowing me to grow as a researcher. Their advice on both research as well as on my career have been priceless. Without their commitment, knowledge and inspiring words, this thesis would not have been successful. With their encouragement, I feel confident in my future career.

I am deeply indebted to my committee members, Professor Dmitry Krass and Professor Gonzalo Romero, for offering me the opportunity to pursue my Ph.D. and for teaching me so much in my study. I am also grateful to my thesis external appraiser, Professor Yulan Wang for spending hours reading my thesis and providing valuable comments to improve it.

I would like to deeply thank other faculty members in the Operations Management group: Professor Joseph Milner, Professor Philipp Afèche, Professor Igor Averbakh, and Professor Azarakhsh Malekian, for their valuable encouragement along the way. I learned a lot from them for becoming a qualified academic.

I wish to acknowledge my fellow Ph.D. students: Yang, Mohammad, Vahid, Yun, Longyuan, Tianshu and Hojat. I feel very lucky to have entered the Ph.D. group in operations management. I enjoy the time we spent in this journey.

Finally, I owe much of my determination and sense of purpose to the love and support of my parents. I thank them for always believing in me.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 2 | Simultaneous vs. Sequential Crowdsourcing Contests | 4 |
| 2.1 | Introduction | 4 |
| 2.2 | Literature Review | 7 |
| 2.3 | Model Setup | 11 |
| 2.3.1 | Sequential Contest | 13 |
| 2.3.2 | Simultaneous Contest | 14 |
| 2.4 | Two-Person Model | 14 |
| 2.4.1 | Sub-Contests in the Sequential Contest | 14 |
| 2.4.2 | Simultaneous Contest | 15 |
| 2.4.3 | Performance Comparison | 16 |
| 2.4.4 | Curvature of Cost Functions | 21 |
| 2.4.5 | Optimal Allocation of Prizes in the Sequential Contest | 22 |
| 2.5 | Multiple-Person Model | 24 |
| 2.5.1 | Equilibrium Effort | 25 |
| 2.5.2 | Number of Contestants | 26 |
| 2.6 | Heterogeneous Contestants | 30 |
| 2.6.1 | Heterogeneous Cost Functions | 31 |
| 2.6.2 | Heterogeneous Starting Points | 32 |
| 2.7 | Conclusion | 34 |
| 2.8 | Proofs. | 35 |
| 2.9 | Appendix. | 44 |
| 2.10 | Numerical Experiments | 52 |
| 3 | Bundling with Crowdsourced Products | 55 |
| 3.1 | Introduction | 55 |
| 3.2 | Literature Review | 58 |

| | | |
|----------|--|------------|
| 3.3 | Model Setup | 60 |
| 3.4 | Revenue Comparison | 61 |
| 3.4.1 | Contingent Revenue Sharing Policy | 61 |
| 3.4.2 | Equilibrium | 64 |
| 3.5 | Pre-committed Revenue Sharing Policy | 65 |
| 3.6 | Extension: Mixed Bundling Sales | 66 |
| 3.6.1 | Partial Mixed Bundling Sales | 67 |
| 3.6.2 | Full Mixed Bundling Sales | 69 |
| 3.7 | Mean-Preserving Spread | 69 |
| 3.8 | Conclusion | 72 |
| 3.9 | Proofs. | 73 |
| 4 | Strategic Customer Behavior with A Reservation System | 93 |
| 4.1 | Introduction | 93 |
| 4.2 | Literature Review | 94 |
| 4.3 | Model Setup and Initial Results | 96 |
| 4.3.1 | Demand and Arrival Rate | 97 |
| 4.3.2 | Waiting Cost | 98 |
| 4.4 | No Travelling Cost ($C = 0$) | 100 |
| 4.4.1 | Case 1: $i_0 = i_1$ | 101 |
| 4.4.2 | Case 2: $i_0 < i_1$ | 103 |
| 4.5 | Sufficiently Large Travelling Cost ($C \geq 1$) | 108 |
| 4.5.1 | Arrival Rates | 109 |
| 4.6 | Intermediate Travelling Cost ($C \in (0, 1)$) | 111 |
| 4.7 | Discussion: Walk-in and Online Reservation Systems | 113 |
| 4.8 | Conclusion | 113 |
| 4.9 | Proofs. | 114 |
| | Bibliography | 119 |

Chapter 1

Introduction

Because of the convenience of the Internet, more and more companies are basing their businesses on large numbers of agents and consumers who are spread all over the world. Since business models are trending towards optimal control of every process, how to manage crowds is becoming one of the most important and challenging research topics. The feature of crowdsourced businesses is that they can gather a vast amount of information or products. Since the field of operations management is concerned with designing an efficient production process and optimally matching supply and demand, there are great many opportunities to research the management of crowdsourced products. Our research aims to characterize crowd behavior and to serve as a tool with which companies can manage their business processes better. We study problems in the following three fields:

- *Crowdsourcing Ideas*
- *Crowdsourcing Digital Products*
- *Crowd Behavior in Queueing System*

In the first chapter, we study the crowdsourcing innovation contest. Innovation contest is a powerful tool that companies can use to solve problems by outsourcing their tasks to the public. In tournament theory, the description of the contestant's behavior is well-developed whereas the optimality of various mechanisms is still under discussion. Instead of one big contest with a single prize, some firms design a number of contests or set up milestones with prizes at each of them. Some studies consider the task to be a one-dimensional project and investigate the motivational effect of those small prizes on contestants. However, they may ignore the fact that those projects often have many attributes, and the contest or milestone with a small prize focuses on just one of those

attributes. We compare the two alternative mechanisms for an innovative project involving multiple attributes. One crowdsourcing mechanism is to run a simultaneous contest where the best solution is chosen from the single solutions submitted simultaneously by all the contestants. The other mechanism is to run a number of sequential sub-contests, each of which deals with one element of the project.

There we find that, in contrast to the simultaneous contest, there are two opposing forces affecting performance in the sequential contest. One is the “*cooperation*” effect that enhances performance in the sequential contest since it combines the best solutions across different attributes. The other one is the “*effort reduction*” effect that weakens performance in the sequential contest. Unlike the sequential contest, the simultaneous contest has its advantage in pooling the risks of failure in multiple competitions, and thus it is better at motivating contestants to make an effort. With those two forces, I show that, depending on the nature of the problem, either of those mechanisms can be optimal.

In the second chapter, we study the crowdsourced digital products. In the information and entertainment industries, there is a trend towards bundling and subscription services. Previous studies have investigated the optimality of the bundling service and have found that it can achieve great economic efficiency. However, evidence shows that not all the product vendors are willing to join a subscription service. Since some firms (e.g., Spotify) allocate the subscription fees collected according to the contribution made by each crowdsourced product, the profit allocation scheme may be crucial in affecting incentive compatibility. In our work, we build a model to characterize the profit allocation scheme, and investigate the incentive compatibility for vendors of different products who are given the option of not joining the bundling platform.

The result shows that such allocation schemes will create incentive incompatibility. In particular, the vendors of high-quality products prefer the pure bundling subscription, whereas the vendors of low-quality products tend to favor the separate sales, in which their own product is sold separately in addition being part of the bundle. Moreover, if products are differentiated by their valuation dispersion, the vendor of the products with higher value dispersion (i.e., the popular vendor) tends to prefer the bundling sales, whereas the vendor of the products with lower value dispersion (i.e., the niche vendor) tends to favor the separate sales. We also compare the ex-post and ex-ante profit allocation schemes. The result will provide insight into whether product vendors should join a subscription platform (e.g., Netflix) that negotiate returns according to the expectation of a products’ popularity instead of the actual usage.

In the third chapter, we study the crowd behavior with an online reservation system. One of the most important features of an online reservation system is that it allows customers to line up without being on site. Some mobile phone applications (e.g., Nowait) allow customers to know the real-time congestion of the service area. Customers can then decide whether to join the queue, taking into account how long the queue is expected to be when they arrive at the service location. In other words, while a customer is travelling from his or her home to the service provider, the queue may become shorter. Since the amount by which the queue has become shorter is associated with the travelling distance, intuitively the more time customers spend in travelling, the less time they have to wait in line. Hence, there is a tradeoff between the cost of waiting in the queue and the cost of travelling.

We find that when the travelling cost is negligible, customers who live farther away from the service location are more likely to join the queue. However, when the travelling cost is high enough, customers who live farther away are less likely to join the queue. By considering online reservations, my study connects the queuing system with the Hotelling model. Contrary to the Hotelling model's assumption that the attraction of the service decreases in the distance between the service and the customer locations, We find that the attraction increases in the distance if the travelling cost is negligible. The result will provide managerial insights for choosing the optimal policy by considering the travelling costs and types of business.

Chapter 2

Simultaneous vs. Sequential Crowdsourcing Contests

2.1 Introduction

The crowdsourcing contest has been widely adopted by firms, non-profit organizations, and governments to solicit innovative solutions to complex problems. In a typical crowdsourcing contest, a contest organizer, by offering prizes, outsources its innovative project to the public, who compete to provide solutions.

Since a project often has multiple attributes or dimensions, one crowdsourcing scheme is to run a *sequential* contest where multiple sub-contests are launched, each dealing with one attribute or dimension of the project. For example, in 2013, the Pentagon launched a contest, through a web portal called Vehicleforge.mil, for the design of an amphibious vehicle for the U.S. Marines. The first sub-contest, with a one-million-dollar prize, involved mobility and drive-train subsystems for the vehicle. About six months later, came a sub-contest for the design of the chassis and other subsystems, a contest with another one-million-dollar prize (see Lohr 2012). Similar sequential schemes have taken place on many civilian crowdsourcing websites. For example, Quirky.com, a platform for crowdsourcing innovations, had the business model of distributing prizes in sub-contests at various stages of turning an idea into a final product. The stages start with idea generation, progress to product design, and may conclude with name and logo designs.

An alternative crowdsourcing mechanism is to run a *simultaneous* contest, in which every contestant is required to submit his or her solution for the whole project all at once, even though the project may require contestants to deal with problems in different attributes. For example, a class of so-called “reduction-to-practice” challenges on

InnoCentive.com, a leading innovation crowdsourcing platform, requires contestants to submit a prototype that shows an idea in actual practice, in other words, an aggregate solution that combines theoretical work of generating ideas and practical work of presenting physical evidence. For another example, right after the sequential contest for the military vehicle, the Pentagon launched a contest with a two-million-dollar prize in 2014. In contrast to the sequential contest held in 2013, this simultaneous contest requires contestants to submit a single solution for an entire vehicle (see Lohr 2012). One would assume that the sequential mechanism has the benefit of allowing the best for each aspect to build on the preceding best, thus leading to an overall best solution. Thus it is puzzling why the Pentagon switched to the simultaneous mechanism after using the sequential mechanism.

Motivated by the Pentagon example, we study the optimal crowdsourcing contest design for projects with multiple attributes. We consider two alternative mechanisms that can be potentially implemented by the contest organizer, referred to as the firm thereafter. One is to run a simultaneous contest where the best solution is chosen from the aggregate solutions submitted by contestants. The other is to run multiple sequential sub-contests, each dealing with one attribute of the project; the final design is made up of the best design for each dimension.

Our analysis shows that, other things being equal, there are two opposing forces affecting the comparison between the two contest mechanisms. On the one hand, since the sub-contests of the sequential contest each focus on a different attribute of the project, the best aggregate performance is made up of the best performance on those attributes. However, in the simultaneous contest, contestants submit a single solution for all the attributes; thus the best performance is the one submitted by a single contestant. Therefore, the combination of the best performance on different attributes in the sequential contest is more likely to have a *high* value than the best performance in the simultaneous contest. We call this, the *combination* effect. On the other hand, the contestants' effort depends on the incentive provided by the size of the prize and the risk of failure in the contest. In a winner-takes-all contest, all contestants incur a cost by making effort but gain nothing, except the winner. Thus, given the same amount of winner's reward, with a larger risk of failure in the contest, contestants tend to make less effort. In the simultaneous contest, the random factors that affect the performance or evaluation criterion across multiple attributes have been pooled together in the solutions provided by contestants. In view of a lower risk of failure relative to the same amount of reward, contestants have more incentive to exert an effort in each dimension in the simultaneous contest than in the sequential one. This is an intuitive explanation from an individual

contestant's perspective, and the exact reasoning is more involved as contestants compete with each other and need to take into account competitors' behavior. We call this the *pooling* effect driven by the pooling of multiple performances subject to random shocks across different attributes in the simultaneous contest. We show that "pooling," as a common theme in the operations literature (see, e.g., Bimpikis and Markakis 2015), has a notable application in the crowdsourcing contest design.

As a result, the comparison of the expected best performances in the two contest mechanisms boils down to a comparison of those two opposing effects.¹ The pooling effect occurs because of the difference in the incentives (rewards and risks) between the simultaneous and sequential contests, due to the different composition of the amount of the prize and the associated risk. We find that the magnitude of the pooling effect is influenced by the difficulty level of the project, which is captured by the curvature of the cost function. The latter reveals the relationship between effort making and its associated cost. If a project is relatively difficult, the contestants tend to make little effort since making a large effort will incur a significant cost. Due to little effort making for a difficult project, the difference in incentives is responsible for a small difference in effort between the two contest mechanisms. Hence, the pooling effect is relatively weak and can be easily dominated by the combination effect, which is not affected by the level of difficulty. As a result, the sequential contest tends to be optimal. On the other hand, if a project is relatively simple, contestants tend to make a considerable effort because it does not cost very much to do so; thus the difference in incentives between the two contest mechanisms can cause a large difference in the level of effort. Therefore, the pooling effect may be more likely to dominate the combination effect, and the simultaneous contest tends to be optimal.

We enhance our key insights by further exploring the comparison in the following directions. First, given the exogenously determined prize allocation in the base model, we examine the optimal allocation of prizes across sub-contests in the sequential contest, which can be viewed as the guided effort allocation by the firm, as opposed to the self-regulated effort allocation by contestants themselves in the simultaneous contest. The optimal allocation of prizes depends on the difficulty levels in different attributes. To induce effort making by contestants, the firm allocates a larger prize to the easier attribute. If the difficulty is sufficiently different across attributes, the sequential contest with the optimal prize allocation is more efficient in motivating contestants to make efforts than the simultaneous contest, and hence it performs better. Otherwise, the simultaneous contest may perform better.

¹In the absence of random factors, the sequential mechanism always weakly dominates.

Second, besides the two-person case that we examine in the base model for simplicity, we show that in the multiple-person model the interplay of the combination and pooling effects remains and all results carry over. Moreover, we investigate the magnitude of the combination effect for different numbers of contestants. We find that a greater number of contestants improves the combination effect but may not affect the pooling effect. Under some conditions, there exists a threshold for the number of contestants above which the combination effect dominates the pooling effect and the sequential contest is optimal, and below which the combination effect is dominated by the pooling effect and the simultaneous contest is optimal.

Finally, we examine the case in which contestants are heterogeneous in their expertise along different attributes but with the same aggregate expertise. We find that if the expertise across attributes is not too heterogeneous, for the project that is sufficiently difficult, the sequential contest is optimal, and for the project that is sufficiently easy, the simultaneous contest is optimal. These insights are consistent with our results from the base model.

2.2 Literature Review

The modeling of contestants' behavior has been an active research area in economics. It is also gaining traction in operations management, as part of managing the crowdsourcing of goods and services. There are many different models in contest theory, such as the random factor model, all-pay auction model, random trials model, and Tullock contest. Konrad (2007) conducts a comprehensive survey of those models. With the random factor model in which a contestant's performance is made up of his effort and a random factor, Lazear and Rosen (1981) show that a contestant's effort depends on the incentive provided by the prize and the cost incurred by exerting effort. In contrast to the traditional studies, we examine the design of contest with *multiple attributes*. In our sequential contest, each sub-contest focuses on one attribute of the project, and each can be viewed as a random factor model. However, in the simultaneous contest, we introduce the multi-dimensional, single-shot contest model in which contestants self-allocate their efforts to multiple dimensions and submit an aggregate solution simultaneously.

One of the main research questions in the contest design is how to design the optimal incentive scheme by allocating the total prize to contestants whose performance can be ranked. Rosen (1986) examines the elimination contest and finds that a large enough prize is needed for the best performer. Kalra and Shi (2001) show that the effort made by an individual contestant decreases in the number of contestants or the uncertainty

in the contest. If several contestants can be rewarded, the rank-order contest (i.e., a contestant with a better performance is awarded a larger prize) dominates the multiple-winner contest (i.e., several top contestants share the total prize *equally* even though their performance is different). The authors also consider different risk attitudes of contestants. They show that if contestants are risk-neutral, the “winner-takes-all” (WTA) (i.e., the best performer is awarded the total prize) becomes optimal. Moldovanu and Sela (2001) use the all-pay auction model to show that if the cost function is concave or not too convex, the WTA is optimal. Terwiesch and Xu (2008) show that the WTA is always optimal for the random factor model (referred to as a model of “ideation project” in their paper) but may or may not be optimal for the all-pay auction model (referred to as a model of “expertise-based project”). Ales et al. (2016b) combine the random trials and random factor models and find that the WTA is optimal if and only if the benefit of additional effort for increasing the probability of becoming the winner is greater than that for increasing the probability of attaining other ranks. Furthermore, they show that the WTA is optimal if the participation of contestants is guaranteed and the density function of the random factor is log-concave. Stouras et al. (2016) study service contest design in an on-demand service context where agents are ranked based on their service performance and higher performers receive priority over incoming service requests. The authors show that a coarse partition of priority classes, such as two priority classes, can be optimal. In contrast to the one-dimensional contest models, we consider a sequential contest as two parallel sub-contests with each focusing on one attribute of the project. We assume that the firm uses the WTA scheme in all sub-contests, since the WTA scheme has been proved to be optimal in most circumstances, especially with our assumption that the random factor is symmetric and log-concave. We also examine the optimal allocation of prizes by the firm in the sequential contest with the objective of achieving the best *aggregated* performance combined with the best performances from sub-contests.

Another research question is how the number of contestants affects the contestants’ behavior. Taylor (1995) examines a contest model in which each contestant conducts random trials to find his best shot that can meet a pre-determined level. The author shows that an open entry contest is not optimal since it reduces the effort of contestants in the equilibrium. Fullerton and McAfee (1999) suggest restricting the number of contestants and using the auctioning method to select the two best-qualified contestants to compete. Later, Che and Gale (2003) show that in designing the contest for procuring innovations, it is optimal to let the two most efficient innovators participate and compete, and to handicap the more efficient one if the contestants are asymmetric. All of those studies emphasize the role of random factors in the contest as we do, and they suggest that the

open entry contest may not be optimal so that the firm needs to restrict the number of contestants. With a different model setup, Terwiesch and Xu (2008) show that more contestants intensify the competition and thus lower the individual effort, but meanwhile the best performance can be *enhanced* by the diversity of contestants. In keeping with their result, Boudreau et al. (2011) show empirically that there is an effort-reducing effect by adding contestants. Ales et al. (2016a) use the random factor model in which the random factor follows a general distribution. They find that the effort-reducing effect may or may not exist, depending on the properties of the random factor's distribution. We examine both contest mechanisms with an *exogenous* number of contestants, to stay focused on the comparison between the two alternative mechanisms. We show that the interplay of the two opposing effects exists for any number of contestants and then investigate comparative statics on the number of contestants. Under some conditions, if the number of contestants is small enough, the simultaneous contest tends to be optimal, and if the number of contestants is large enough, the sequential contest tends to be optimal.

Some work examines the contestants' behavior when there are a series of contests. One stream of those studies considers *static* games in which contestants' behavior in different competitions is independent. Moldovanu and Sela (2006) study a tournament in which contestants are split among several competitions whose winners compete against one another in the final round. Konrad and Kovenock (2009) examine the equilibrium strategies in a series of competitions, in which in addition to the prize offered for each competition, there is a grand prize for overall performance. DiPalantino and Vojnovic (2009) consider a sequence of crowdsourcing contests where contestants can choose which contest(s) to enter. Though the three papers mentioned above are similar to ours in that the contest designer splits the contest into several sub-contests, their sub-contests are designed for selecting the two best contestants to compete in the final (e.g., NBA Playoffs), or for evaluating the total performance (e.g., English Premier League), or corresponding to different projects (e.g., Yahoo Answers). The sequential contest in our context refers to multiple sub-contests dealing with different attributes of a project (e.g., the Pentagon's example), and those sub-contests can have different cost functions and prizes.

The other stream of literature on multiple competitions focuses on the *dynamic* game in which the contestants' behavior in those competitions is correlated. Those papers mainly explore the strategic disclosure of information (on the contestants' progress) by the firm or by contestants themselves in the process of the multi-stage competition. Some papers study the effects of information disclosure among contestants in the R&D

competition (e.g., Harris and Vickers 1987, Choi 1991, Malueg and Tsutsui 1997, and Yildirim 2005). Perhaps the work most closely related to ours is on information disclosure by the contest organizer. The contest organizer can set some intermediate prizes as milestones throughout the process so that some contestants' performance will be revealed (e.g., Goltsman and Mukherjee 2011, Bimpikis et al. 2014 and Halac et al. 2016). Those studies characterize the strategic behavior of contestants in the intermediate stages of the contest and explore the optimal information disclosure strategies of the firm in anticipation of that strategic behavior. Jiang et al. (2016) run simulations based on empirical estimations and show that the disclosure of the evaluations of the performance of contestants throughout the contest may not be optimal but that the disclosure of those evaluations at a later time may lead to a better overall contest outcome. The sub-contests in our sequential contest can be different from "milestones" in their traditional sense. Our sub-contests focus on different attributes, and later sub-contests are built on the best outcome of previous sub-contests. Our contest has the feature that allows contestants to "cooperate" whereas the mechanisms in the aforementioned papers do not allow such "cooperation." The presence of multiple attributes in the contest is a unique feature of our model that has not been studied before. This feature makes possible a comparison between two alternative mechanisms: simultaneous and sequential contests. From a different angle, Acemoglu et al. (2014) emphasize that the exact difficulties of innovation tasks may not be known in advance. The authors take a mechanism design approach and show that the solution is a dynamic pricing mechanism that induces workers to self-select into different skill hierarchical layers.

Some work compares the simultaneous and sequential moves by agents in solving other related management problems. Hausch (1986) considers the situation where the seller can choose between two mechanisms: auctioning two identical objects at once or launching two auctions with each selling one object. The author shows that either mechanism can be optimal depending on the strategic behavior of bidders. Hu et al. (2013) compare the simultaneous and sequential group-buying mechanisms. They examine a two-period model in which consumers make sign-up decisions. The firm decides whether to disclose the number of sign-ups in the first period to the consumers arriving in the second period. They show that the sequential mechanism has a greater chance than the simultaneous mechanism of reaching the pre-determined threshold of the number of sign-ups.

Finally, there are two main criteria for evaluating the contest in the literature: expected *best* performance and expected *average* performance. Some studies investigate the average performance; see, e.g., Kalra and Shi (2001) and Moldovanu and Sela (2001). This criterion is appropriate for the project such as the sales contest. The benefit for the

firm that launches a sales contest is the total contributions made by all the contestants. Thus, the average performance measures the quality of a contest for a given number of contestants. Some other studies evaluate the contest on the basis of the expected best performance as well as the expected average performance; see, e.g., Terwiesch and Xu (2008) and Körpeoğlu and Cho (2017). The criterion of the expected best performance is better suited to a project like a research or brainstorming contest. There, a single outstanding solution can be more valuable than thousands of mediocre ones. Thus the expected best performance is the measure for evaluating those innovation contests, and how to enhance the best performance is of vital importance. We examine both the expected *average* performance and expected *best* performance in a comparison of those two contest mechanisms. However, the two opposing effects that we characterize exist when we identify the expected best performance, whereas only one effect, the pooling effect, exists when we identify the expected average performance. Our sequential contest resembles the *hybrid* structure of idea generation studied by Girotra et al. (2010) who focus on the best performance criterion. They show with a laboratory experiment that the hybrid structure, in which individuals first work independently and then work together, is able to generate more and better ideas than the team structure, in which the group works together in time and space. In our sequential contest, contestants work independently in each sub-contest, and the performances from sub-contests can be “assembled” to form a final solution. The main difference is that the crowdsourcing contest typically faces the general public who cannot work together and involves with monetary incentive that is likely absent in brainstorming within an organization.

2.3 Model Setup

In this section, we develop a base model for examining contestants’ behavior in crowdsourcing contests with multiple attributes. The firm outsources tasks to the public, and contestants make efforts to win a prize. Let A denote the total prize of the project. The number of contestants is $n \geq 2$, which is exogenously given. In other words, contestants do not endogenously make the entry decision, and they may incur a negative expected payoff. A contest with a fixed number of contestants is commonly seen in the literature (e.g., Lazear and Rosen 1981, Moldovanu and Sela 2001, 2006). It is also common in practice; for example, an employer may require all its employees to participate in a brainstorming contest. We discuss the impact of endogenized entry decisions on the comparison between the two alternative mechanisms in Section 2.5.2.

We consider a project with multiple attributes. Without loss of generality, we assume

that the project consists of two attributes, indexed by 1 and 2. In each attribute, the performance of a contestant is made up of two additive components. The first component is the *effort level*. The contestant decides his effort levels e^1 and e^2 , respectively in those two attributes, depending on the incentives. The second component is the *random factor*. The problem solving in innovation is often random. The random factors exist in the projects with unclear standards or projects in which contestants have random performances. For example, a logo-designing project on “99designs” may have unclear criteria because the judges have undisclosed artistic tastes. Therefore, it is unclear how a submission will be rated. Moreover, the design work could depend highly on a designer’s personal experience, random inspirations, or the designing environment; thus performance itself can be random.

There are random factors along the two attributes, denoted by ϵ^1 and ϵ^2 respectively. We assume that the two random factors have the same distribution with the cumulative density function (CDF) $\Psi(\cdot)$ and probability density function (PDF) $\psi(\cdot)$, which are common knowledge. (The qualitative insights would not change for the case in which random factors along the two attributes follow different distributions.) We assume that the random factors along the two attributes are independent, and they are identical and independent among all the contestants. Furthermore, we assume that $\psi(\cdot)$ is symmetric and log-concave with mean 0 and standard deviation $\sigma > 0$. The condition of symmetric log-concavity can be satisfied by commonly used distributions such as normal, logistic and uniform distributions.

The performance in each attribute is the sum of the corresponding effort and random factor. This additive form of individual performance in a contest is commonly seen in the literature; see, e.g., Lazear and Rosen (1981); Kalra and Shi (2001); Terwiesch and Xu (2008) and Ales et al. (2016a). We use subscript i to denote a specific contestant and superscript l to denote a specific attribute. If contestant i makes effort e_i^l , $l = 1, 2$, the performance of contestant i in those two attributes is given by

$$V_i^l = e_i^l + \epsilon_i^l.^2$$

The cost of exerting effort can be considered in the form of time consumption or

²Some studies, e.g., Kalra and Shi (2001), Terwiesch and Xu (2008), and Ales et al. (2016a), assume the performance in the form of $V = r(e) + \epsilon$, where $r(\cdot)$ is a concave function. Such a form of performance, together with the linear cost function, guarantees that the first-order condition characterizes the equilibrium strategy. With such a form, the effort is assumed to be non-negative, though $r(e)$ may be negative. For simplicity, we do not impose the non-negativity constraint on the effort. However, our model can be adapted to the performance form of $r(e) + \epsilon, e \geq 0$, with an appropriate defined performance function $r(\cdot)$.

monetary investment. Assume that cost functions along two attributes are, respectively, $C^1(\cdot)$ and $C^2(\cdot)$, which can be different. They are common knowledge. Moreover, we assume that $C^{l'}(\cdot) > 0$, $C^{l''}(\cdot) > 0$, $l = 1, 2$, the same as in Stouras et al. (2016). This assumption is also consistent with Terwiesch and Xu (2008) and Ales et al. (2016a). They assume a strictly increasing and strictly concave performance function that is equivalent to a strictly increasing and strictly convex cost function.

We assume that all the contests adopt a WTA scheme. Previous studies have found that WTA is optimal for a single contest in most circumstances. For example, Moldovanu and Sela (2001) demonstrate that WTA is optimal when the cost function is concave or not convex enough. Terwiesch and Xu (2008) show that WTA is always optimal provided that the performance is made up with the effort and random factor, and it may be optimal for other forms. Ales et al. (2016b) show that the WTA is optimal if the participation of contestants is guaranteed and the PDF of the random factor is log-concave.

2.3.1 Sequential Contest

In the sequential contest, the firm launches two sub-contests (indexed by 1 and 2), each of which focuses on one attribute of the project. It allocates the total prize A to two sub-contests with an exogenous weight $w \in (0, 1)$; hence prizes in those two sub-contests are $A^1 = wA$ and $A^2 = (1 - w)A$. Here we allow the weight w to be arbitrary. In Section 2.4.5, we will discuss the optimal allocation of prizes in the sequential contest, i.e., endogenizing the weight w .³

Furthermore, contestant i 's performance in sub-contests 1 and 2 are V_i^1 and V_i^2 respectively. The total performance of contestant i is in the additive form of the performance in each sub-contest, $V_i^{seq} = V_i^1 + V_i^2$. However, since the winners in those two sub-contests may be different, the realization of V_i^{seq} for any contestant i may not be the best total performance in the sequential contest. Moreover, the performance along the two attributes may have different levels of importance in the total performance. Nevertheless, since we do allow different awards and different cost functions in those two sub-contests, it is without loss of generality to normalize the relative importance of performances to 1, by changing the allocation of prizes and cost functions.

³If $A^1 > A^2$, A^2 in our context is not the so-called ‘‘second prize’’ in contest theory. In the previous studies, the second prize refers to a small prize awarded to the contestant whose total outcome ranks in the second place. Those studies are intended to solve the problem whether WTA or some other rewarding scheme is optimal for the firm (see, e.g., Kalra and Shi 2001, Moldovanu and Sela 2001, and Terwiesch and Xu 2008). In our context, a contestant may win the first sub-contest but lose the second one. While sub-contests focus on different aspects of the project, a contestant who ranks first in the second sub-contest wins the prize A^2 .

In sub-contest l , $l = 1, 2$, the payoff to contestant i is

$$u_i^l(e_i^l) = \begin{cases} A^l - C^l(e_i^l) & \text{if } i \text{ wins,} \\ -C^l(e_i^l) & \text{if } i \text{ loses.} \end{cases}$$

This mechanism has been widely adopted in the contest theory literature that assumes the WTA scheme a priori (e.g., Taylor 1995 and Fullerton and McAfee 1999). Note that in our model setting contestants are all a priori identical with the same random factors and cost functions. In Section 2.6, we discuss contestants with heterogeneous cost functions.

2.3.2 Simultaneous Contest

In the simultaneous contest, the firm launches a single contest to collect solutions, so the performance for contestant i is the aggregation of all performance along two attributes, denoted by $V_i^{sim} = V_i^1 + V_i^2 = (e_i^1 + \epsilon_i^1) + (e_i^2 + \epsilon_i^2)$. Unlike in the sequential contest, contestants make an aggregate submission instead of a solution for each sub-contest; therefore the best performance in the simultaneous contest is the realization of V_i^{sim} if contestant i is the winner. Since the contestant with the best performance wins the grand prize A , the payoff to contestant i is

$$u_i^{sim}(e_i^1, e_i^2) = \begin{cases} A - C^1(e_i^1) - C^2(e_i^2) & \text{if } i \text{ wins,} \\ -C^1(e_i^1) - C^2(e_i^2) & \text{if } i \text{ loses.} \end{cases}$$

2.4 Two-Person Model

In this section, we derive the contestants' equilibrium behavior in a two-person model. (In Section 2.5, we consider the n -person model.) Using the characterized behavior, we then compare the simultaneous and sequential contests.

2.4.1 Sub-Contests in the Sequential Contest

Each sub-contest in the sequential contest can be considered as a single-dimensional simultaneous contest. The existence of an equilibrium is guaranteed if the expected payoff function is unimodal (i.e., quasi-concave) in effort over the relevant range. However, the expected payoff function may not necessarily be unimodal. In Lazear and Rosen (1981) with the same model setup as ours, a pure equilibrium strategy solution exists provided that the standard deviation of the random factor is large enough. Moreover,

Dixit (1987) and Terwiesch and Xu (2008) make an even stronger assumption, namely, that the expected payoff function is concave in effort. Following the convention, we assume appropriate assumptions that can guarantee the existence of an equilibrium.

Index the two contestants by i and j . The difference of random factors between two contestants i and j is denoted by $\xi^l = \epsilon_i^l - \epsilon_j^l$, $l = 1, 2$ with PDF $g(\cdot)$ and CDF $G(\cdot)$. Because $\epsilon_i^l, \epsilon_j^l$ independently follow the distribution with mean 0 and standard deviation σ , the difference ξ^l has mean 0 and standard deviation $\sqrt{2}\sigma$. Since the PDF of random factors $\psi(\cdot)$ is symmetric at 0, so is the PDF $g(\cdot)$. The following lemma characterizes the contestants' behavior in the symmetric equilibrium. Since contestants are ex ante identical, the equilibrium efforts of two contestants are the same.

Lemma 2.1 *In the sub-contest l with two contestants, the equilibrium effort is $e^{l*} = C^{l'-1}(A^l g(0))$, $l = 1, 2$.⁴*

Lemma 2.1 has the same characterization of the contestants' behavior as Lazear and Rosen (1981). It shows that the effort in the equilibrium increases in the amount of prize. With a larger prize, contestants have more incentive to exert effort. The quantity $g(0)$ measures the marginal change in the probability of winning by exerting additional effort beyond the competitor, which can be interpreted as the risk taken by the contestant for making an extra effort. The higher the risk, the less effort contestants tend to make in the equilibrium. If ϵ^l , $l = 1, 2$, follows a normal distribution with standard deviation σ , then $e^{l*} = C^{l'-1}(A^l/(\sqrt{2\pi}\sigma))$, and the effort level is determined by the *return-risk ratio*, $A^l/(\sqrt{2\pi}\sigma)$. For this case, the equilibrium effort level is decreasing in σ , which measures the risk associated with exerting extra effort (see more discussion below).

2.4.2 Simultaneous Contest

Contestants in the simultaneous contest make efforts along two dimensions. The contestants' two-dimensional optimization problem can boil down to a single-dimensional optimization problem, with the help of the following lemma.

Lemma 2.2 (OPTIMAL EFFORT ALLOCATION BY CONTESTANT) *The cost function of the aggregate effort level, resulted from the optimal allocation of efforts on two attributes, $C^o(e^o) = \min_{e^1+e^2=e^o} \{C^1(e^1) + C^2(e^2)\}$, is a strictly increasing and strictly convex function. Given the aggregate effort level e^o , the optimal effort allocation $(\tilde{e}^1, \tilde{e}^2)$ satisfies $C^{o'}(e^o) = C^{1'}(\tilde{e}^1) = C^{2'}(\tilde{e}^2)$.*

⁴If $C^{l'}(\cdot)$, $l = 1, 2$ ranges over a bounded support, the symmetric equilibrium effort may be located at a corner; e.g., $e^{l*} = [C^{l'-1}(A^l g(0))]^+ = 0$, which boils down to a trivial case. As a result, we restrict our attention to the case in which the equilibrium effort is an interior point.

Lemma 2.2 characterizes the optimal allocation of efforts made by contestants in the simultaneous contest. It is analogous to the optimal allocation of a fixed budget across products to maximize the total profit in the economics literature. Given a fixed amount of aggregate effort, contestants optimally allocate the efforts to two dimensions. The marginal costs across the two dimensions are equal in the optimal allocation. Otherwise, say if $C^{1'}(e^1) > C^{2'}(e^2)$, the contestant can achieve a lower total cost by increasing his effort in the second attribute but reducing it in the first attribute. Moreover, we show that the total cost is strictly increasing and strictly convex in the aggregate effort.

Denote the difference of random factors between contestants i and j along the two dimensions by $\xi^\circ = \epsilon_i^1 + \epsilon_i^2 - \epsilon_j^1 - \epsilon_j^2 = (\epsilon_i^1 - \epsilon_j^1) + (\epsilon_i^2 - \epsilon_j^2) = \xi^1 + \xi^2$. By the symmetric property of ξ^1 and ξ^2 , the random variable ξ° has a symmetric PDF $g^\circ(\xi^\circ)$ and a CDF $G^\circ(\xi^\circ)$. Since ξ^1 and ξ^2 have mean 0 and standard deviation $\sqrt{2}\sigma$ and they are independent, their summation ξ° has mean 0 and standard deviation 2σ . By Lemma 2.2, the total cost $C^\circ(e^\circ)$ is strictly increasing and strictly convex. Then the derivation of the equilibrium effort in the simultaneous contest is analogous to what is in Lemma 2.1, provided the total prize A ,

$$e^{\circ*} = C^{\circ\prime-1}(Ag^\circ(0)). \tag{2.1}$$

In the simultaneous contest, the incentive for contestants to make an effort is the total prize. The $g^\circ(0)$ measures the risk, which depends on the two random factors along the two dimensions.

2.4.3 Performance Comparison

In this subsection, we compare the *expected best performances* and *expected average performances* between those two contest mechanisms with two contestants. For the two-person case, the highest-order statistics with sample size 2 is denoted by the subscript (2) . (For the general case, we denote the highest-order statistic with a sample size n by subscript (n) .) The expected best performances in the sequential and simultaneous contests are denoted by V^{seq} and V^{sim} , and

$$\begin{aligned} V^{seq} &= \mathbb{E}((e^{1*} + \epsilon^1)_{(2)}) + \mathbb{E}((e^{2*} + \epsilon^2)_{(2)}), \\ V^{sim} &= \mathbb{E}((e^{\circ*} + \epsilon^1 + \epsilon^2)_{(2)}). \end{aligned}$$

Denote the difference by

$$\Delta = V^{seq} - V^{sim}.$$

Because equilibrium efforts are deterministic (as in the base model, all contestants are homogeneous; we consider heterogeneous contestants in Section 2.6), we have $\mathbb{E}((e^{l*} + \epsilon^l)_{(2)}) = e^{l*} + \mathbb{E}(\epsilon^l_{(2)})$, $l = 1, 2$, and $\mathbb{E}((e^{\circ*} + \epsilon^1 + \epsilon^2)_{(2)}) = e^{\circ*} + \mathbb{E}((\epsilon^1 + \epsilon^2)_{(2)})$. The difference between the expected best performances can then be decomposed into two parts as $\Delta = \Delta^e + \Delta^\epsilon$, where

$$\begin{aligned}\Delta^e &= e^{1*} + e^{2*} - e^{\circ*}, \\ \Delta^\epsilon &= \mathbb{E}(\epsilon^1_{(2)}) + \mathbb{E}(\epsilon^2_{(2)}) - \mathbb{E}((\epsilon^1 + \epsilon^2)_{(2)}).\end{aligned}$$

The first part Δ^e is the difference in the equilibrium effort levels. The second part Δ^ϵ is the difference between the expected best random factors.

Lemma 2.3 *If $n \geq 2$ and ϵ^l , $l = 1, 2$, follows a symmetric and log-concave distribution,*

$$\mathbb{E}(\epsilon^1_{(n)}) + \mathbb{E}(\epsilon^2_{(n)}) > \mathbb{E}((\epsilon^1 + \epsilon^2)_{(n)}).$$

Lemma 2.3 can be simply shown as follows. Denote the realizations of ϵ^1 and ϵ^2 with sample size n by $\{\epsilon^1_1, \dots, \epsilon^1_n\}$ and $\{\epsilon^2_1, \dots, \epsilon^2_n\}$ respectively. Then $\max\{\epsilon^1_1, \dots, \epsilon^1_n\} + \max\{\epsilon^2_1, \dots, \epsilon^2_n\} \geq \max_{i', i'' \in \{1, \dots, n\}} \{\epsilon^1_{i'} + \epsilon^2_{i''}\}$. Since such inequality holds for any realization, $\mathbb{E}(\epsilon^1_{(n)}) + \mathbb{E}(\epsilon^2_{(n)}) \geq \mathbb{E}((\epsilon^1 + \epsilon^2)_{(n)})$. Moreover, Lemma 2.3 holds for more than two persons and it shows that if random factors follow a symmetric and log-concave distribution, the strict inequality holds.

Proposition 2.4 (EXPECTED BEST PERFORMANCE: TWO-PERSON) *Consider the simultaneous and sequential contests with two contestants.*

- (i) *The expected best random factor in the sequential contest is larger than that in the simultaneous contest, i.e., $\Delta^\epsilon > 0$;*
- (ii) *If $g^\circ(0) > \max\{wg(0), (1-w)g(0)\}$, the equilibrium effort in the simultaneous contest is higher than that in the sequential contest i.e., $\Delta^e < 0$;*
- (iii) *The condition in (ii) simplifies to $g^\circ(0) > g(0)/2$, if one of the following conditions holds:*
 - (a) $w = 1/2$; (b) $C^1(\cdot) = C^2(\cdot)$ and their derivatives are weakly convex.

The following example shows that the conditions in Proposition 2.4 parts (ii) and (iii) are satisfied if random factors follow a normal distribution.

Example 2.1 If $\epsilon^l \sim N(0, \sigma)$, $l = 1, 2$, by Lemma 2.1, the equilibrium effort for sub-contest l in the sequential contest is $e^{l*} = C^{l'-1}(A/(2\sqrt{\pi}\sigma))$. In the simultaneous contest, $\xi^\circ = \xi^1 + \xi^2$ follows $N(0, 2\sigma)$. Thus, $g^\circ(0) = 1/(2\sqrt{2\pi}\sigma)$. By (2.1), the equilibrium effort in the simultaneous contest is $e^{\circ*} = C^{\circ'-1}(A/(2\sqrt{2\pi}\sigma))$. For Proposition 2.4(ii), the difference between the equilibrium efforts in the two contest mechanisms is $\Delta^e = C^{1'-1}(wA/(2\sqrt{\pi}\sigma)) + C^{2'-1}((1-w)A/(2\sqrt{\pi}\sigma)) - C^{1'-1}(A/(2\sqrt{2\pi}\sigma)) - C^{2'-1}(A/(2\sqrt{2\pi}\sigma))$. One sufficient condition for $\Delta^e < 0$ is $g^\circ(0) > \max\{wg(0), (1-w)g(0)\}$, which is satisfied if $w \in (1 - \sqrt{2}/2, \sqrt{2}/2) \approx (0.29, 0.71)$. The inequality is due to $C^{l'}(\cdot) > 0$, $l = 1, 2$. For Proposition (iii), the condition $g^\circ(0) > g(0)/2$ can be naturally satisfied by the normal distribution because $2\sqrt{2}\sigma < 4\sigma$.

Proposition 2.4 characterizes two forces that affect the comparison of those two contest mechanisms. The sequential contest has an advantage in selecting the best performances mainly driven by the random factors. While the two sub-contests deal with different attributes of the project, each sub-contest selects the best performance in each attribute. Given the same effort by all the contestants, the sequential contest combines the performances with the best realizations of random factors in those two sub-contests. The aggregated best performance in the sequential mechanism is made up of the best realizations along the two attributes, and is more likely to have a high value. Moreover, the best solutions in the two sub-contests may be provided by different contestants. However, in the simultaneous contest, the contestants' performance depends on the sum of those random factors to the same individual; thus the sum of the random factors across the two attributes is less extreme. Then the sum of the most extreme random factors across the two sub-contests is larger than the extreme of the aggregated random factors in the simultaneous contest. As the old saying goes, "two heads are better than one." Since the sequential contest combines the best performance from each sub-contest, we call this the *combination* effect.

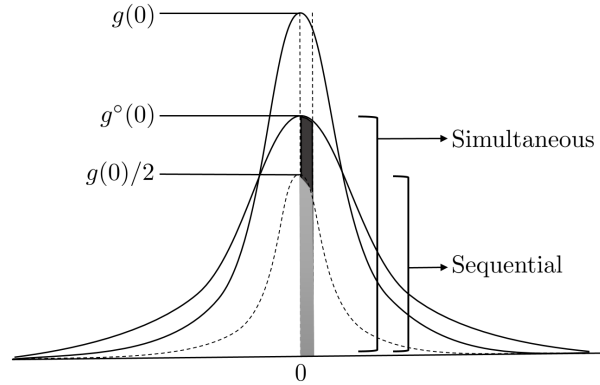
The simultaneous contest encourages a greater effort in the equilibrium under a mild sufficient condition in Proposition 2.4(ii), i.e., $g^\circ(0) > \max\{wg(0), (1-w)g(0)\}$. We begin our discussion by considering a special case as in Proposition 2.4(iii)(a), in which the firm evenly allocates prizes into two sub-contests, $w = 1/2$. If $\epsilon^l \sim N(0, \sigma)$, $l = 1, 2$, the condition $g^\circ(0) > g(0)/2$ can naturally hold since $2\sqrt{2}\sigma < 4\sigma$, which ensures $\Delta^e < 0$. In each sub-contest of the sequential contest, the return for the winner is one-half of the total prize. Since the prize encourages contestants to make an effort, the effort increases in the amount of prize. However, making efforts also has its risk. The risk can be measured by the marginal change in the probability of winning as a result of exerting additional effort, and it depends on the random factor. In the symmetric

equilibrium, both contestants follow the same strategy, thus the winning probability for each of them is $1/2$, i.e., $G(0) = 1/2$. Then $g(0)$ measures the marginal change of the winning probability by exerting additional effort beyond the competitor, in other words, the risk of making extra effort. For a normally distributed random factor, if σ is large, the extra effort can enhance the winning probability by only a little, which implies that the contestant takes a high-level risk of failure for making an extra effort. Hence, we can view the value of σ as the amount of risk that contestants bear when making extra efforts. Overall, a contestant's effort in each sub-contest depends on the return-risk ratio, i.e., $A/2 \cdot g(0) = (A/2)/(2\sqrt{\pi}\sigma)$. In fact, the larger the ratio, the greater incentive the contestant has to make an effort. Hence, seeing a larger return-risk ratio, the contestant makes more efforts in equilibrium.

In the simultaneous contest, the contestant submits an aggregate solution of two attributes in order to win the whole prize. Therefore, if the contestant wins, the return for his efforts in each attribute is the whole prize. As in the above analysis in each sub-contest of the sequential contest, in the simultaneous contest the contestant makes efforts in each attribute, taking into account not only the reward but also the risk. The risk in the simultaneous contest can also be measured by the marginal change of winning probability by exerting extra effort beyond the competitor, which now depends on the sum of random factors across the two attributes. If random factors are distributed normally with standard deviation σ , the summation of two random factors has the standard deviation $\sqrt{2}\sigma$. Thus, the effort in each attribute relies on the corresponding return-risk ratio; i.e., $A \cdot g^\circ(0) = A/(2\sqrt{2\pi}\sigma)$. Since the return-risk ratio is higher for each attribute in the simultaneous contest than in the sequential contest, i.e., $A/(2\sqrt{2\pi}\sigma) > (A/2)/(2\sqrt{\pi}\sigma)$, the effort for each attribute is higher in the simultaneous contest. Figure 2.1 illustrates the return-risk ratio in the equilibrium in each attribute by letting $A = 1$, $w = 1/2$ and $\epsilon^l \sim N(0, \sigma)$, $l = 1, 2$. In the equilibrium, the return-risk ratio is greater in the simultaneous contest than in the sequential contest. This implies that the contestant would make a greater effort in the simultaneous contest when $w = 1/2$ than in the sequential contest. $A = 1$, $w = 1/2$, $\epsilon^l \sim N(0, \sigma)$, $l = 1, 2$, $g(0) = 1/(2\sqrt{\pi}\sigma)$ and $g^\circ(0) = 1/(2\sqrt{2\pi}\sigma)$.

An explanation in plain words may be that if a contestant aims to win the whole prize in the sequential contest, he has to be the winner in both sub-contests, and that could be very difficult. However, in order to win the simultaneous contest, the contestant need not have the best performance in each attribute, but only the best total performance. That is to say, his not-so-great performance due to the random factor in one attribute can be compensated by his excellent performance in another attribute. Given the same return, the pooling of random factors has the result that the return-risk ratio for a contestant

Figure 2.1: Return-risk ratio in the equilibrium



in each attribute of the simultaneous contest is greater than that in each sub-contest of the sequential contest. Since the contestant makes a greater effort with a higher return-risk ratio, the effort for each attribute is greater in the simultaneous contest than in the sequential contest; we call this, the *pooling* effect.

For a general prize allocation w , if $w \neq 1/2$, the return-risk ratios in the two sub-contests of the sequential contest are different because the returns are different in those sub-contests; i.e., $wAg(0) \neq (1 - w)Ag(0)$. Proposition 2.4(ii) allows the cost functions along the two attributes to be different and the prize allocation in the sequential contest to be general; i.e., $w \in (0, 1)$. For a general prize allocation w , the sufficient condition $g^o(0) > \max\{wg(0), (1 - w)g(0)\}$ guarantees that the effort in each attribute of the simultaneous contest is greater than that in each sub-contest of the sequential contest.

The effort not only depends on the return-risk ratio, but also on the marginal cost function. Even though the return-risk ratio along the two attributes of the simultaneous contest is the same, the effort can be different due to the different cost functions along the two attributes. By allowing the cost functions to be identical and their derivative function to be convex, Proposition 2.4(iii)(b) shows that the sufficient condition for $\Delta^e < 0$ becomes as simple as $g^o(0) > g(0)/2$, and can be naturally satisfied by a normal distribution without any assumption on the prize allocation fraction w . The condition that $C^{l'}(\cdot)$, $l = 1, 2$, is convex can be satisfied by many cost functions, e.g., the exponential and quadratic form of the cost functions.

Many studies in the literature examine the *expected average performance*, e.g., Kalra and Shi (2001), Moldovanu and Sela (2001) and Terwiesch and Xu (2008). Since contestants are ex ante identical, the expected average performance is equivalent to the expected individual performance in our context. In some projects, e.g., a sales contest, every individual contestant's performance matters for the firm. The following corollary

compares the two contest mechanisms for each contestant's performance.

Corollary 2.5 (INDIVIDUAL PERFORMANCE) *For any contestant i , the performance in the simultaneous contest first-order stochastically dominates the performance in the sequential contest, i.e., $V_i^{sim} \geq_{st} V_i^{seq}$, if $g^\circ(0) \geq \max\{wg(0), (1-w)g(0)\}$.*

Corollary 2.5 shows that the performance of an individual contestant is more likely to be better in the simultaneous contest than in the sequential contest. By $V_i^{sim} \geq_{st} V_i^{seq}$, we immediately have $\mathbb{E}(V_i^{sim}) \geq \mathbb{E}(V_i^{seq})$. In contrast to the expected best performance, the combination effect does not play a role in the expected individual performance. For each given individual, the expected performance only relies on the equilibrium effort. With the pooling effect prevailing, the expected individual performance is higher in the simultaneous contest than in the sequential contest.

2.4.4 Curvature of Cost Functions

By the previous analysis, the comparison of the two contest mechanisms comes down to the comparison of the two opposing effects, combination and pooling effects. The magnitude of the combination effect has nothing to do with the cost functions, and depends only on the random factors. However, the magnitude of the pooling effect depends on the convexity of the cost functions along the two attributes. In this subsection, we examine the comparative statics on the degree of convexity (curvature) of the cost functions. The degree is measured by the Arrow-Pratt coefficient (see Mas-Colell et al. 1995). The degree of the curvature of the cost function has been used to discuss other problems in contest theory. For example, Moldovanu and Sela (2001) show that the Arrow-Pratt coefficient of the cost function determines whether WTA is optimal. For simplicity, we assume that the cost functions are in the exponential form, $C^l(e^l) = \exp(\rho^l e^l)$, $\rho^l > 0$, $l = 1, 2$, which allow us to use a single parameter to represent the Arrow-Pratt coefficient,

$$\frac{C^{l''}(e)}{C^{l'}(e)} = \frac{\rho^{l2} \exp(\rho^l e)}{\rho^l \exp(\rho^l e)} = \rho^l \text{ (degree of convexity),}$$

where $l = 1, 2$. Similar insights can be derived for the general form of cost functions. We compare those two contest mechanisms for different degrees of convexity.

Proposition 2.6 *Assume the exponential cost form. If $g^\circ(0) > \max\{wg(0), (1-w)g(0)\}$, there exist two thresholds $\bar{\rho} \geq \underline{\rho} > 0$ such that if $\rho^1, \rho^2 \geq \bar{\rho}$, the sequential contest dominates the simultaneous contest, i.e., $\Delta \geq 0$; and if $\rho^1, \rho^2 \leq \underline{\rho}$, the simultaneous contest dominates the sequential contest, i.e., $\Delta \leq 0$. Moreover, assume $\rho^1 = \rho^2 = \rho$.*

For ρ , there exists a threshold $\tilde{\rho} > 0$ above which the sequential contest is optimal, and below which the simultaneous contest is optimal.

Proposition 2.6 shows that the comparison of those two forces depends on the curvature of the cost functions. The pooling effect is due to the different incentives provided in the two contest mechanisms. Given the prizes and random factors, the difference in the incentives is fixed. However, the magnitude of such difference, which is reflected in the contestants' effort, depends on the convexity of the cost functions. The interpretation of the degree of convexity is that it measures the difficulty level of the project. If the degrees of convexity of the cost functions are large, the project for those contestants seems to be difficult since even a little extra effort will lead to a significant increase in cost. Thus the effort is small in both contest mechanisms. Then the difference in incentives makes little difference to the effort expended. As a result, the pooling effect will be weak and likely dominated by the combination effect, which has not been affected by the cost functions. Hence, the sequential contest tends to be optimal. However, if the degrees of convexity of those cost functions are small, contestants tend to make a large effort. Thus, the pooling effect will be significant and will likely outweigh the combination effect. Then the simultaneous contest tends to be optimal. If the cost functions along the two attributes are identical, i.e., $\rho^1 = \rho^2$, there exists a *single* threshold on the degree of convexity of the cost functions above which the sequential contest dominates and below which the simultaneous contest dominates.

Projects on different platforms are of varying difficulty. For example, the questions on Yahoo Answers or Amazon Mechanical Turk are mostly simple and do not require the problem solvers to demonstrate a strong ability in data analysis or logical thinking. However, for platforms such as InnoCentive or Kaggle, the projects are mostly accompanied with datasets. To solve those problems, contestants may need to search the academic literature, build mathematical models, and write programs. The heterogeneity of those platforms in the level of difficulty may lend some support to our results. Multi-stage questions are rare on platforms such as Yahoo Answers or Amazon Mechanical Turk, but multi-stage projects are commonly seen on InnoCentive or Kaggle.

2.4.5 Optimal Allocation of Prizes in the Sequential Contest

The difference in equilibrium efforts between two contest mechanisms, Δ^e , depends on the allocation of prizes in the sequential contest. In this subsection, we examine the value of Δ^e under the optimal allocation of prizes by the firm in the sequential contest. Again for simplicity, we adopt the exponential form of the cost functions, $C^l(e^l) = \exp(\rho^l e^l)$,

$\rho^l > 0$, $l = 1, 2$. Define the relative degree of convexity of cost functions along the two attributes by $t = \rho^1/\rho^2$.

Proposition 2.7 *Assume the exponential cost form.*

(i) *The optimal allocation of prizes to the sub-contest 1 in the sequential contest is*

$$w^* = \frac{\rho^2}{\rho^1 + \rho^2} = \frac{1}{t + 1}.$$

(ii) *Under the optimal allocation in the sequential contest, there exist $\underline{t} < 1$ and $\bar{t} > 1$ such that if $t \in (\underline{t}, \bar{t})$, $\Delta^e < 0$, otherwise $\Delta^e \geq 0$.*

Proposition 2.7(i) shows that in the sequential contest, the firm needs to allocate a larger prize to the easier attribute of the project. For example, if $\rho^1 < \rho^2$, it is more cost-effective for a contestant to make an effort in the sub-contest 1. Thus, sub-contest 1 achieves a higher efficiency in driving effort than sub-contest 2. Because the performance in those sub-contests are equally important, given a fixed total prize, it is more beneficial for the firm to allocate a larger prize to sub-contest 1. Moreover, since the cost functions $C^l(\cdot)$, $l = 1, 2$, are exponential, $C^{l'-1}(\cdot)$, $l = 1, 2$, is concave, and thus using rewards to encourage effort has diminishing returns. As a result, even though the efficiency is higher in the sub-contest 1, the firm does not want to allocate the whole prize to sub-contest 1. Therefore, the optimal allocation of prizes, w^* , exists as an interior point within $(0, 1)$.

Proposition 2.7(ii) compares the equilibrium efforts between two contest mechanisms under the optimal allocation of prizes in the sequential contest. In the simultaneous contest, the firm invests a grand prize in the simultaneous contest, where contestants optimally allocate their efforts to two dimensions. In the sequential contest, the firm optimally allocates prizes to two dimensions and contestants optimally react to the incentive provided by the size of the prize in each attribute. Our result shows that the equilibrium effort depends on the relative difficulty of the two attributes. If the difficulty is similar enough, the total equilibrium effort is larger in the simultaneous contest. Otherwise, the sequential contest has a larger total equilibrium effort. The intuitions are as follows.

For those scenarios in which the difficulty of the two attributes is far apart, we discuss, without loss of generality, the case that improving performance is much easier in the first attribute than in the second; i.e., t is sufficiently small (close to 0). On the one hand, under the optimal allocation of prizes in the sequential contest, $w^* = 1/(t + 1)$. Since t is sufficiently small, the firm allocates most of the prize to sub-contest 1 and

contestants make a significantly greater effort in sub-contest 1 than in sub-contest 2. Thus, the total effort in the sequential contest comes mostly from the effort in the sub-contest 1. Moreover, the contestant can win almost the whole prize by focusing on the first attribute of the project. On the other hand, in the simultaneous contest, under the optimal individual allocation of efforts, contestants tend to allocate almost all their efforts to the first attribute as well because it is much easier than the other one. Thus, the total effort in the simultaneous contest also relies mainly on the effort in the first attribute. However, in the simultaneous contest, the contestants' performance is influenced by the sum of random factors along the two attributes. Therefore, with almost the same return, a contestant in the simultaneous contest takes more risks than in sub-contest 1 of the sequential contest. That is, the return-risk ratio is lower in the simultaneous contest, and thus less effort is expended in the first attribute of the simultaneous contest than in sub-contest 1 of the sequential contest. Since the total effort relies largely on the effort in the first attribute in both contest mechanisms, the sequential contest has a higher total equilibrium effort than the simultaneous contest. That is, $\Delta^e \geq 0$, which is consistent with that the sufficient condition in Proposition 2.4(ii) no longer holds under the optimal prize allocation w^* for a sufficiently small t . We can apply a similar argument to the case that it is much easier to improve the performance in the second attribute than in the first; i.e., t is sufficiently large. In conclusion, if the difficulty is sufficiently different, $\Delta^e \geq 0$. Together with the combination effect, i.e., $\Delta^c > 0$, we have $\Delta = \Delta^e + \Delta^c > 0$. That is, the sequential contest dominates the simultaneous contest when the efficiency of the two attributes, in encouraging the contestants' efforts, is different enough.

If the difficulty along the two attributes is close to each other (i.e., t is close to 1), including the case of projects with symmetric attributes, then the total equilibrium effort is higher in the simultaneous contest than in the sequential contest. That is to say, the pooling effect sustains ($\Delta^c < 0$). Then the dominance of those two opposing effects, pooling and combination effects, depends on the convexity of the cost functions. (A similar analysis follows from Section 2.4.4).

2.5 Multiple-Person Model

In this section, we compare those two contest mechanisms with a general number of n contestants. Denote the equilibrium efforts in the sub-contests 1 and 2 of the sequential contest by $e^{1*}(n)$ and $e^{2*}(n)$ respectively. The equilibrium effort in the simultaneous contest is denoted by $e^{o*}(n)$. Thus, the expected best performance in the sequential and

simultaneous contests can be written as

$$\begin{aligned} V_n^{seq} &= \mathbb{E}((e^{1*}(n) + \epsilon^1)_{(n)}) + \mathbb{E}((e^{2*}(n) + \epsilon^2)_{(n)}), \\ V_n^{sim} &= \mathbb{E}((e^{\circ*}(n) + \epsilon^1 + \epsilon^2)_{(n)}). \end{aligned}$$

As in the two-person model, we decompose the difference between the expected best performance into the difference in effort and the difference contributed by different selections of random factors. That is, $\Delta_n = V_n^{seq} - V_n^{sim} = \Delta_n^e + \Delta_n^\epsilon$, where $\Delta_n^e = e^{1*}(n) + e^{2*}(n) - e^{\circ*}(n)$ and $\Delta_n^\epsilon = \mathbb{E}(\epsilon_{(n)}^1) + \mathbb{E}(\epsilon_{(n)}^2) - \mathbb{E}((\epsilon^1 + \epsilon^2)_{(n)})$. Note that by Lemma 2.3, the sum of the best realizations of random factors along the two attributes is more likely to be extreme than the best realization of the sum; i.e., $\Delta_n^\epsilon > 0$.

2.5.1 Equilibrium Effort

For the sub-contest l , $l = 1, 2$, in the sequential contest, suppose that all the contestants except contestant i make equilibrium effort $e^{l*}(n)$. The winning probability of contestant i , if he makes effort e_i^l , is

$$P(i \text{ wins with effort } e_i^l \mid \text{others make effort } e^{l*}(n)) = \int_{-\infty}^{+\infty} \Psi(e_i^l - e^{l*}(n) + \epsilon^l)^{n-1} \psi(\epsilon^l) d\epsilon^l.$$

For the simultaneous contest, suppose that all the contestants except contestant i make equilibrium effort $e^{\circ*}(n)$. Denote $\epsilon^\circ = \epsilon^1 + \epsilon^2$ with PDF $\psi^\circ(\epsilon^\circ)$ and CDF $\Psi^\circ(\epsilon^\circ)$. The winning probability of contestant i , if he makes effort e_i° , is

$$P(i \text{ wins with effort } e_i^\circ \mid \text{others make effort } e^{\circ*}(n)) = \int_{-\infty}^{+\infty} \Psi^\circ(e_i^\circ - e^{\circ*}(n) + \epsilon^\circ)^{n-1} \psi^\circ(\epsilon^\circ) d\epsilon^\circ.$$

When $e_i^l = e^{l*}(n)$ and $e_i^\circ = e^{\circ*}(n)$, one can easily verify that those winning probabilities are equal to $1/n$ since $\int_{-\infty}^{+\infty} \Psi^\circ(\epsilon^\circ)^{n-1} \psi^\circ(\epsilon^\circ) d\epsilon^\circ = 1/n$. It implies that contestants have equal chances of winning because they are ex ante identical. With those winning probabilities, we can characterize the equilibrium efforts for both mechanisms. Define functions $h(\cdot; n)$ and $h^\circ(\cdot; n)$ as:

$$h(\epsilon^l; n) = \int_{-\infty}^{+\infty} (n-1) \Psi(\epsilon^l)^{n-2} \psi(\epsilon^l)^2 d\epsilon^l \quad \text{and} \quad h^\circ(\epsilon^\circ; n) = \int_{-\infty}^{+\infty} (n-1) \Psi^\circ(\epsilon^\circ)^{n-2} \psi^\circ(\epsilon^\circ)^2 d\epsilon^\circ.$$

The functions $h(\cdot; n)$ and $h^\circ(\cdot; n)$ are well-defined functions that have been examined in the literature, e.g., the functions Ψ_1 in Kalra and Shi (2001), and I_N in Ales et al. (2016a). Since ϵ^1 and ϵ^2 are identical, the same form $h(\cdot; n)$ applies to both dimensions,

which we can denote by $h(\epsilon; n)$. Those functions measure the marginal change of the contestant's probability of winning by exerting additional effort. Note that $h(\epsilon; 2) = g(0)$ and $h^\circ(\epsilon^\circ; 2) = g^\circ(0)$. We have the following characterizations of equilibrium efforts.

Lemma 2.8 *Consider contests with n contestants. The equilibrium effort in sub-contest l , $l = 1, 2$, of the sequential contest is $e^{l*}(n) = C^{l'-1} (A^l h(\epsilon; n))$. The equilibrium effort in the simultaneous contest is $e^{\circ*}(n) = C^{\circ'-1} (A h^\circ(\epsilon^\circ; n))$.*

With Lemma 2.8, we characterize the combination and pooling effects in the n -person model.

Proposition 2.9 (EXPECTED BEST PERFORMANCE: n -PERSON)

- (i) (Combination effect) $\Delta_n^\epsilon > 0$.
- (ii) (Pooling effect) If $h^\circ(\epsilon^\circ; n) > \max\{wh(\epsilon; n), (1-w)h(\epsilon; n)\}$, $\Delta_n^\epsilon < 0$. In particular, if $\epsilon^l \sim N(0, \sigma)$, $l = 1, 2$, and $w \in (1 - \sqrt{2}/2, \sqrt{2}/2)$, $\Delta_n^\epsilon < 0$ for any n .

Proposition 2.9(i) is directly implied by Lemma 2.3. For (ii), if $h^\circ(\epsilon^\circ; n) > \max\{wh(\epsilon; n), (1-w)h(\epsilon; n)\}$, the equilibrium effort is higher in the simultaneous contest than in the sequential contest; i.e., $\Delta_n^\epsilon < 0$. Interestingly, if random factors follow a normal distribution (the arguably most commonly used distribution in the natural and social sciences), then $h^\circ(\epsilon^\circ; n)/h(\epsilon; n) = 1/\sqrt{2}$ for any number of contestants. Further, if $w \in (1 - \sqrt{2}/2, \sqrt{2}/2)$ then $\Delta_n^\epsilon < 0$. As a special case, if the firm allocates the prizes equally in the sequential contest, i.e., $w = 1/2$, then the simultaneous contest achieves a higher equilibrium effort level. In conclusion, the results in the two-person model can be carried over to the n -person model. The interplay of those two opposing forces, the combination effect and pooling effect, still applies in the n -person case.

2.5.2 Number of Contestants

In this subsection, we compare the two contest mechanisms for different numbers of contestants. The Pentagon's contest brought together specialized military contractors for designing a military vehicle. Thus, the number of contestants in the Pentagon's project is small compared to the platforms on which projects do not require sophisticated technique skills, such as InnoCentive or 99designs. On Kaggle, the problem solvers have been categorized into five tiers by the quality and quantity of their performances: novice, contributor, expert, master and grandmaster. Interestingly, some contests only allow the solvers who rank at the expert or higher tier to participate. It can be expected that such

contests must have fewer participants than the contests that are open to all solvers (we extend our base model to account for heterogeneous contestants and show that the main results carry over).

We provide managerial insights on contest design for projects with different numbers of contestants. By Lemma 2.3, the difference between the expected best random factors $\Delta_n^\epsilon > 0$. Now, we discuss how Δ_n^ϵ changes with an increasing number of contestants.

Proposition 2.10 *There exists $\bar{n} \geq 2$ such that Δ_n^ϵ is increasing in $n \in [2, \bar{n}]$.*

It is intuitive that the *best* performances in *both* contest mechanisms can be improved by having more contestants. However, it is not clear which contest mechanism benefits more from additional contestants. Proposition 2.10 shows that the marginal benefit of one additional contestant in boosting the expected best random factors is more significant for the sequential contest than for the simultaneous contest if the number of contestants is not too large. The intuition is as follows. For both contests, the best performance will be enhanced only if the additional solution is better than every single one of the solutions in the existing pool. For the simultaneous contest, contestants submit a single solution along the two dimensions, so the best performance will be improved if the additional aggregate solution is better. However, for the sequential contest, the best performance will be improved if the additional solution in either attribute is better. When the pool of contestants is small, it is more likely that the additional contestant is doing better than the existing pool of contestants in one of the attributes than that he is doing better in the whole project. Thus, if the contestant pool is not too large, the firm can benefit from obtaining a higher expected best random factors, by having more contestants in the sequential contest than in the simultaneous contest. When the contestant pool is large, having more contestants may lead to diminishing returns. Now we compare the two contest mechanisms in the expected best random factors, when the number of contestants is sufficiently large.

Proposition 2.11 (i) *If ϵ^1 and ϵ^2 have a bounded support $[-a, a]$, $\lim_{n \rightarrow \infty} \Delta_n^\epsilon = 0$.*

(ii) *If ϵ^1 and ϵ^2 are normally distributed, $\lim_{n \rightarrow \infty} \Delta_n^\epsilon = \infty$.*

Proposition 2.11 can be explained as follows. It is intuitive that when the number of contestants is large enough, the best performance in both contest mechanisms must be outstanding. Since contestants are ex ante identical, the effort is equal among all the contestants under each mechanism, so the firm selects the best random factor in each contest. Therefore, there must be a random factor approaching the upper limit provided

a sufficiently large number of contestants. It has been well known (see, e.g., David and Nagaraja 2003, pp. 80, (4.5.1)) that when n is sufficiently large, the expectation of the highest order statistics is approximately equal to the value of $\frac{n}{n+1}$ th quantile,

$$\mathbb{E}(\epsilon_{(n)}) \approx \Psi^{-1}\left(\frac{n}{n+1}\right), \tag{2.2}$$

where the $\Psi^{-1}(\cdot)$ is the quantile function of $\Psi(\cdot)$. When n is large enough, the term $\frac{n}{n+1}$ approaches 1 and $\mathbb{E}(\epsilon_{(n)})$ approaches the upper limit of the range of the random factor.

When the number of contestants is large enough and the random factors have a bounded support (e.g., two-sided truncated normal distribution), the expected best random factors in the two-contest mechanisms are approximately equal, since they are both close to the upper bound. However, with normally distributed random factors that have the unbounded support, the difference between the expected best random factors approaches infinity when the large pool of contestants grows even larger. Thus, the sequential contest can benefit more from an increasing number of contestants than the simultaneous contest, even when the pool of contestants is already very large. That is, the combination effect can be infinitely enhanced by more and more contestants.

In general, how the difference between the equilibrium efforts in those contest mechanisms, Δ_n^e , would change with one additional contestant can be ambiguous. However, we are able to obtain a clear-cut result for normally distributed random factors and exponential cost functions.

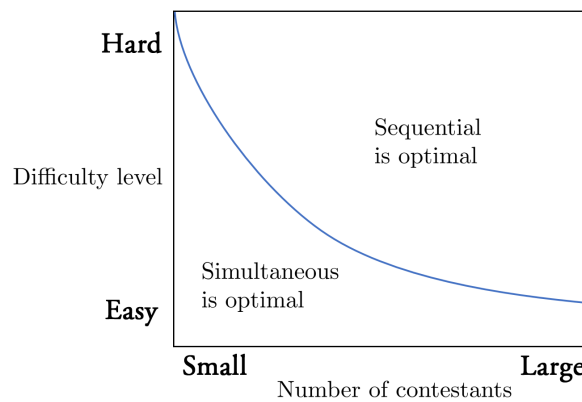
Proposition 2.12 (EXPECTED BEST PERFORMANCE: NUMBER OF CONTESTANTS)

Assume the exponential cost form. If $\epsilon^l \sim N(0, \sigma)$, $l = 1, 2$, and $w \in (1 - \sqrt{2}/2, \sqrt{2}/2)$, Δ_n^e is a constant for any n , and thus there exists a threshold $\tilde{n} \geq 2$ on the number of contestants, above which the sequential contest is optimal and under which the simultaneous contest is optimal.

Proposition 2.12 shows that if random factors follow a normal distribution (the arguably most commonly used distribution in the natural and social sciences), then the simultaneous contest is optimal when the number of contestants is relatively small, and the sequential contest is optimal when the number of contestants is relatively large. Ales et al. (2016a) study a one-dimensional contest and show that if the random factor follows a symmetric log-concave distribution (which the normal distribution satisfies), the effort is decreasing in the number of contestants (see their Proposition 1). That is because more contestants intensify the competition and reduce contestants' incentive to expend effort. With this result, the effort in both contest mechanisms is decreasing in

the number of contestants, but the monotonicity of the difference in the levels of effort may be ambiguous. However, if the random factors follow the normal distribution and the cost functions are in the exponential form, the difference in effort between the two contest mechanisms is a fixed value for any number of contestants. Nevertheless, for this specific case, we show that the difference in random factors is increasing in the number of contestants, a finding consistent with Propositions 2.10 and 2.11(ii) for the general case. The combination effect is reinforced by a larger number of contestants while the pooling effect exists but is not influenced by the number of contestants. Overall, everything else being equal, if the number of contestants is relatively small, the combination effect is weak and dominated by the pooling effect, and thus the simultaneous contest is optimal. Otherwise, if the number of contestants is relatively large, the combination effect becomes significant so that the sequential contest becomes optimal. This result may partially explain the puzzle of the Pentagon’s switching behavior. Since the number of contestants may not be large for a military project and that the combination effect is not significant for a small number of contestants, the simultaneous contest may perform better than the sequential contest. That may be one reason why the Pentagon switched to the simultaneous contest after experimenting with the sequential contest. Lastly, when the entry decisions by the contestants are endogenized, the number of entrants will be smaller and hence the simultaneous mechanism may tend to be favored. $\epsilon^l \sim N(0, \sigma)$, $C^l(e^l) = \exp(\rho^l e^l)$, $l = 1, 2$, and $w = \frac{1}{2}$.

Figure 2.2: Comparison between the simultaneous and sequential contests



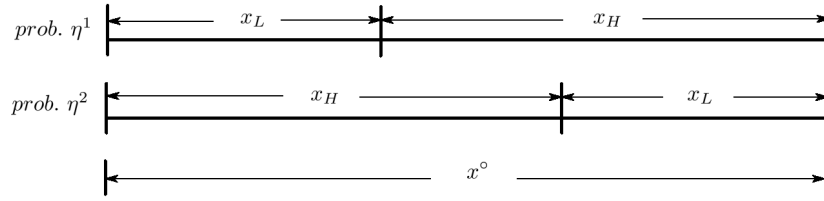
As a summary, Figure 2.2 illustrates the comparison of the two mechanisms depending on the difficulty level (see Proposition 2.6, which can be easily extended to the n -person case) and the number of contestants (see Proposition 2.12), for normally distributed random factors and exponential cost functions.

2.6 Heterogeneous Contestants

In this section, we consider a two-person model with two expertise types (high and low) in each attribute. In contrast to the base model in which all the contestants are assumed to be identical for each attribute, we assume here that contestants are endowed with expertise x_H for the first and x_L for the second attribute, or x_L for the first and x_H for the second attribute ($x_H \geq x_L > 0$). Here, the aggregate expertise for all the contestants is the same; i.e., a contestant has either (x_L, x_H) or (x_H, x_L) as the expertise along the two dimensions. This stylized assumption on the constant aggregate expertise being equal to $x^\circ = x_H + x_L$ is consistent with the common belief that the human beings are created equally but with different talents. In the simultaneous contest, all the contestants have the same aggregate expertise, consistent with the base model. The expertise follows a two-point distribution. Since we have studied the homogeneous expertise in the base model, the two-point expertise distribution is the *most* heterogeneous expertise distribution among the class of distributions that share the same mean and standard deviation. The probability that a contestant is endowed with (x_L, x_H) is η^1 , and the probability that a contestant is endowed with (x_H, x_L) is η^2 , where $\eta^1 + \eta^2 = 1$. Thus, we generalize the base model and allow contestants to have different expertise in different attributes. Allowing general expertise along the two dimensions would not qualitatively change our main results (see the footnote in the proof of Proposition 2.13).

In the sequential contest, we consider the situation in which the firm does *not* disclose the solutions of the first sub-contest during the second sub-contest, and will assemble the best performances from the two sub-contests at the end. This assumption requires the final solution to be modular, i.e., the solution can be divided into smaller modules that can be independently created in each dimension and then assembled. This assumption may be justified as follows. First, imposing this restriction would not change the results from our base model, because there all the contestants have an identical expertise, and nothing needs to be learned from one sub-contest to another. Second, in practice, this assumption may not be too restrictive, because information disclosure can lead to strategic behavior by contestants and as a result, the firm may have no incentive to do so. When the contestants' expertise in the first dimension is correlated with that in the second dimension, the revealing of solutions from the first sub-contest may lead to strategic behavior in both sub-contests. In anticipation of the solutions from the first sub-contest being revealed, contestants may distort their performance to hide their types in the first sub-contest; e.g., the contestant with high expertise may pretend to have less expertise. On learning that others are low-type in the first sub-contest, contestants may not exert

Figure 2.3: Two-point expertise distribution



their full effort in the second sub-contest. Lastly, for some projects, performance cannot be evaluated until the solutions along both attributes are collected. For example, in the Pentagon's contest, the military agency may only be able to provide feedback and evaluate performance after obtaining the complete vehicle design. $\eta^1 + \eta^2 = 1$ and $x^\circ = x_H + x_L$.

We assume that the cost functions along the two dimensions are identical in the exponential form, $C(\cdot) = C^1(\cdot) = C^2(\cdot) = \exp(\rho x)$, and moreover, $2g^\circ(0) > g(0)$, which is naturally satisfied by normal distributions. In each sub-contest, every contestant knows only his own expertise and that his opponents' expertise is drawn independently from the two-point distribution (see Figure 2.3). The game is a Bayesian game in the Harsanyi sense (see Harsanyi 1968) where "types" are defined by contestants' expertise. In the symmetric Bayesian equilibrium, contestants' behavior is determined by their types, regardless of their identities. Hence, we use type H or L to refer to a contestant's behavior in the equilibrium. We examine two settings where the expertise affects contestants' performances in different ways.

2.6.1 Heterogeneous Cost Functions

The first setting is that the expertise results in different efficiencies in making efforts. For exerting the same amount of effort, the high-type contestant incurs a lower cost than the low-type contestant. Such a model characterizes the heterogeneity of contestants in their innovation ability. Contestants with higher talents tend to spend less time in developing novel ideas. It is appropriate to use the heterogeneous cost model to characterize contestant behavior for projects that require innovative thinking, such as research and art designing contests. For each attribute, if a contestant is type $i = H, L$, his cost function is $C(\cdot)/x_i$. Similar characterizations have been adopted in Lazear and Rosen (1981), Moldovanu and Sela (2001) and Fey (2008) with slightly different model setups.

First, we characterize the contestants' behavior in equilibrium with a general $\eta^l \in (0, 1)$, $l = 1, 2$. Second, we examine a special case in which the fractions of both types of

contestants are equal, i.e., $\eta^l = 1/2$. One can easily show that if $\eta^l = 1/2$, the optimal allocation of prizes in the sequential contest is $A^1 = A^2 = A/2$ since the prior distributions of the expertise are the same and the cost functions along the two attributes are also the same. Thus, we compare the two contest mechanisms under the optimal allocation of prizes $w = 1/2$ for $\eta^l = 1/2$. Third, for $\eta^l \neq 1/2$, there is no closed form solution for the equilibrium effort and the comparison of the two mechanisms is intractable. We perform numerical tests in Online Appendix 2.10 and obtain consistent observations with Proposition 2.13(ii) for $\eta^l = 1/2$.

Proposition 2.13 *Assume the exponential cost form and consider two contestants.*

- (i) *(Equilibrium effort of different expertise in the sequential contest) In the sub-contest l , $l = 1, 2$, there exists an equilibrium such that $e_H^{l*} \geq e_L^{l*}$. If $\eta^l = 1/2$, such equilibrium is unique.*
- (ii) *(Comparison) If $\eta^l = 1/2$, and x_H and x_L are sufficiently close ($x_H/x_L \in [1, (2g^\circ(0)/g(0))^2]$), there exist $\underline{\rho}'$ and $\bar{\rho}'$ such that when $\rho \leq \underline{\rho}'$, the simultaneous contest is optimal, and when $\rho \geq \bar{\rho}'$, the sequential contest is optimal.*

Proposition 2.13(i) shows that at least in one equilibrium, the high-type contestants exert more effort than the low-type contestants. Since their equilibrium performances are $V_H^{l*} = e_H^{l*} + \epsilon^l$ and $V_L^{l*} = e_L^{l*} + \epsilon^l$, $l = 1, 2$, then $V_H^{l*} \geq_{st} V_L^{l*}$ because $e_H^{l*} \geq e_L^{l*}$. That is, the high-type contestants are more likely to have a higher performance than the low-type contestants. Proposition 2.13(ii) shows that when the high-type and low-type expertise levels are close enough to each other, if the convexity of the cost function is sufficiently large, the sequential contest dominates, and if the convexity of the cost functions is sufficiently small, the simultaneous contest dominates. This result is consistent with Proposition 2.6 for the case of homogeneous contestants.

2.6.2 Heterogeneous Starting Points

The second setting is that expertise provides different starting points. The performance of contestant i in sub-contest l , $l = 1, 2$, is $V_i^l = x_H + e_i^l + \epsilon^l$ if he has high expertise, and $V_i^l = x_L + e_i^l + \epsilon^l$ if he has low expertise. However, they have the same cost function. This model characterizes the heterogeneity of contestants in their skill levels or experience. A skilled programmer may possess several well-developed programming frameworks. An experienced salesperson may keep in contact with several clients so that

in the sales contest he can guarantee certain sales volume at the beginning of the competition. Thus, the heterogeneous starting point model can be implemented in characterizing the contestants' behavior for projects that require experience or technical skills, such as a technology competition and sales contest. Such a characterization has been adopted in Terwiesch and Xu (2008) and Körpeoğlu and Cho (2017).

As in Proposition 2.13, we characterize the equilibrium efforts made by contestants with a general $\eta^l \in (0, 1)$. We compare the two contest mechanisms under the optimal allocation of prizes $w = 1/2$ for $\eta^l = 1/2$. For $\eta^l \neq 1/2$, we perform numerical tests in Online Appendix 2.10 and obtain consistent observations with Proposition 2.14(ii) for $\eta^l = 1/2$.

Proposition 2.14 *Assume the exponential cost form and consider two contestants.*

- (i) *(Equilibrium effort of different expertise in the sequential contest) In sub-contest l , $l = 1, 2$, there exists an equilibrium such that (a) if $\eta^l < 1/2$, $e_H^* \geq e_L^*$; (b) if $\eta^l > 1/2$, $e_H^* \leq e_L^*$; (c) if $\eta^l = 1/2$, $e_H^* = e_L^*$. In all of the 3 cases, the expected equilibrium performances satisfy $\mathbb{E}(V_H^{l*}(e_H^{l*})) \geq \mathbb{E}(V_L^{l*}(e_L^{l*}))$.*
- (ii) *(Comparison) If $\eta^l = 1/2$, and x_H and x_L are sufficiently close ($x_H - x_L \in [0, 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^o)]$), there exist $\underline{\rho}''$ and $\bar{\rho}''$ such that when $\rho \leq \underline{\rho}''$, the simultaneous contest is optimal, and when $\rho \geq \bar{\rho}''$, the sequential contest is optimal.*

Proposition 2.14(i) shows that when the probability that high-type contestants will appear is high ($\eta^l < 1/2$), high-type contestants expend greater effort than low-type contestants. This is because high-type contestants have an inherently better starting point than low-type contestants. A higher chance of encountering a competitor with great expertise tends to intensify the competition and motivate the high-type contestant to make greater efforts. Meanwhile, a low-type contestant expects that there is little chance of winning because he is more likely to encounter a high-type contestant. As a result, low-type contestants tend to slack off. Combining the two sides, the performance of high-type contestant is better than that of low-type contestants.

When the probability that high-type contestants will appear is low ($\eta^l > 1/2$), surprisingly, low-type contestants make more effort than high-type contestants. High-type contestants tend to slack off because they expect there is little chance of encountering other high-type contestants. Meanwhile, low-type contestants find it more likely that they will win, and hence they exert more effort. In the literature, with the same model setup for a one-dimensional contest, Terwiesch and Xu (2008) demonstrate that for the general n -person case, when contestants are heterogeneous in expertise, contestants

with higher expertise may make a greater effort. Körpeoğlu and Cho (2017), by using a numerical test with a Gumbel distribution, show that the result can be reversed, i.e., that contestants with higher expertise may exert less effort. With a simple two-point distribution, we analytically show that the fraction of high or low contestants plays an important role in how expertise predicts effort, and we provide an intuitive explanation. Nevertheless, we show that the expected performance of the high-type contestants is still better than that of low-type contestants, at least in one equilibrium.

Proposition 2.14(ii) is analogous to Proposition 2.13(ii). Under similar conditions, it shows that if the convexity of the cost functions is large enough, the sequential contest is optimal, and that if the convexity of the cost functions is small enough, the simultaneous contest is optimal, which is again consistent with Proposition 2.6.

2.7 Conclusion

We compare the simultaneous and sequential contest mechanisms of crowdsourcing contests for projects with multiple attributes. With the characterization that a contestant's aggregate performance is made up of his effort levels and random factors across multiple dimensions, we find that the comparison comes down to a comparison of two opposing effects, the combination effect and the pooling effect. In addition, we obtain a set of managerial insights. First, the magnitude of the pooling effect depends on the difficulty of the project, i.e., the convexity of the cost functions. If the project is difficult, the sequential contest tends to be optimal, and if the project is simple, the simultaneous contest tends to be optimal. Second, we examine the optimal allocation of prizes across multiple sub-contests of the sequential contest. When the difficulty of attributes is different enough, the sequential contest in which the firm optimally allocates prizes to induce effort dominates the simultaneous contest in which contestants self-regulate their own effort in view of the big prize. Otherwise, the simultaneous contest may perform better. Third, we generalize our base model to consider the contest with more than two contestants, and investigate how the number of contestants affects the comparison. We find that the interplay of those two opposing effects exists with a general number of contestants. Under some conditions, if the number of contestants is large enough, the sequential contest tends to be optimal, and if the number of contestants is small enough, the simultaneous contest tends to be optimal. Lastly, in addition to the base model, which assumes all the contestants are homogeneous, we show that to a large extent, the results in considering the heterogeneous contestants are consistent with what we find in the base model.

Pooling is a theme widely seen in the operations literature. In the simultaneous contest, pooling of random factors reduces risk in effort making and incentivizes contestants to expend effort. Intuitively, the benefit of pooling, in favor of the simultaneous contest, increases with the number of attributes and the variability in the random factors. However, the combination effect, in favor of the sequential contest, is also expected to increase in those two factors. Still, the comparison of the two contest mechanisms comes down to the relative strength of the combination effect and pooling effect, which we leave for future research.

There are several limitations to our model. First, we assume that the random factors along different dimensions follow identical and independent distributions. Future research can consider correlated random factors across attributes. Second, in considering the heterogeneous contestants, we examine a special case in which contestants are endowed with heterogeneous expertise along different attributes, but with the same aggregate expertise. In other words, we assume that a contestant's expertise along different attributes is perfectly negatively correlated. Future research may consider more generally distributed joint expertise across attributes. Lastly, in the sequential contest with heterogeneous contestants, we assume that the firm does not disclose the performance of contestants in the earlier sub-contests. For those projects in which earlier performance can enhance the later performance, the firm may have an incentive to reveal earlier performance. However, such information disclosure may also induce strategic behavior among contestants. As a result, its overall effect on the comparison between the two mechanisms is not clear. Despite those limitations, our stylized model captures the core tradeoff in comparison of the two contest mechanisms for projects with multiple attributes and generates insights that seem consistent with many practical observations. Our results can be used to provide guidelines in designing crowdsourcing contests with multiple attributes.

2.8 Proofs.

Proof of Lemma 2.2. Consider the optimization problem below:

$$\min_{e^1, e^2} C^1(e^1) + C^2(e^2) \quad \text{s.t. } e^1 + e^2 = e^\circ.$$

The solution to this problem can be typically found by writing the Lagrangean, $L(e^1, e^2, e^\circ; \lambda) = C^1(e^1) + C^2(e^2) + \lambda(e^\circ - e^1 - e^2)$, and the FOCs are

$$\frac{\partial L}{\partial e^1} = C^1'(\tilde{e}^1) - \tilde{\lambda} = 0 \quad (3a), \quad \frac{\partial L}{\partial e^2} = C^2'(\tilde{e}^2) - \tilde{\lambda} = 0 \quad (3b),$$

$$\frac{\partial L}{\partial \lambda} = e^\circ - \tilde{e}^1 - \tilde{e}^2 = 0 \quad (2.3)$$

Solving the FOCs yields the Lagrange multiplier $\tilde{\lambda} = \lambda(e^\circ)$ and the optimal efforts $\tilde{e}^1(e^\circ)$, $\tilde{e}^2(e^\circ)$ along the two dimensions. Now plugging $\tilde{e}^1(e^\circ)$, $\tilde{e}^2(e^\circ)$ into the objective function and we can get a new function $C^\circ(e^\circ) = C^1(\tilde{e}^1(e^\circ)) + C^2(\tilde{e}^2(e^\circ))$ which yields the minimum value of C° for a given e° . Taking the derivative of C° with respect to e° , we obtain

$$\frac{dC^\circ(e^\circ)}{de^\circ} = C^{1'}(\tilde{e}^1(e^\circ)) \frac{d\tilde{e}^1(e^\circ)}{de^\circ} + C^{2'}(\tilde{e}^2(e^\circ)) \frac{d\tilde{e}^2(e^\circ)}{de^\circ}. \quad (2.4)$$

By (2.3a) and (2.3b), $C^{1'}(\tilde{e}^1) = \tilde{\lambda}$ and $C^{2'}(\tilde{e}^2) = \tilde{\lambda}$, we have (2.4) = $\tilde{\lambda} \left[\frac{d\tilde{e}^1(e^\circ)}{de^\circ} + \frac{d\tilde{e}^2(e^\circ)}{de^\circ} \right]$. By (2.3c), $e^\circ = \tilde{e}^1 + \tilde{e}^2$, and hence $\frac{d\tilde{e}^1(e^\circ)}{de^\circ} + \frac{d\tilde{e}^2(e^\circ)}{de^\circ} = 1$, thus $\frac{dC^\circ(e^\circ)}{de^\circ} = \tilde{\lambda}$. Because $C^1(\cdot)$ and $C^2(\cdot)$ are strictly increasing, again by (2.3a) and (2.3b), $\tilde{\lambda} > 0$, thus, $\frac{dC^\circ(e^\circ)}{de^\circ} > 0$, i.e., $C^\circ(e^\circ)$ is strictly increasing.

By $\frac{dC^\circ(e^\circ)}{de^\circ} = \tilde{\lambda}$, (2.3a) and (2.3b), we have

$$C^{\circ'}(e^\circ) = C^{1'}(\tilde{e}^1) = C^{2'}(\tilde{e}^2). \quad (2.5)$$

By the assumption $C^{2''}(\cdot) > 0$, $C^{2'^{-1}}(\cdot)$ is well-defined, hence $C^{2'^{-1}}(C^{1'}(\tilde{e}^1)) = \tilde{e}^2$. By (2.3c), $\tilde{e}^1 + \tilde{e}^2 = e^\circ$, we obtain $C^{2'^{-1}}(C^{1'}(\tilde{e}^1)) + \tilde{e}^1 = e^\circ$. Because $C^{1''}(\cdot) > 0$ and $C^{2''}(\cdot) > 0$, $C^{1'^{-1}}(\cdot)$ and $C^{2'^{-1}}(\cdot)$ are strictly increasing, thus \tilde{e}^1 is strictly increasing in e° by (2.5). Further by (2.3a), $C^{1'}(\tilde{e}^1) = \tilde{\lambda}$, $\tilde{\lambda}$ is strictly increasing in e° . Then because $\frac{dC^\circ(e^\circ)}{de^\circ} = \tilde{\lambda}$ that we have just proved, $\frac{dC^\circ(e^\circ)}{de^\circ}$ is strictly increasing in e° , i.e., $C^\circ(e^\circ)$ is strictly convex. ■

Proof of Lemma 2.3. Denote $e^\circ = \epsilon^1 + \epsilon^2$ and it has CDF $\Psi^\circ(\epsilon^\circ)$. Denote the quantile function of ϵ^l , $l = 1, 2$, by $\Psi^{-1}(u)$ and the quantile function of e° by $\Psi^{\circ^{-1}}(u)$. Write the formula of $\mathbb{E}(\epsilon_{(n)}^l)$, $l = 1, 2$,

$$\begin{aligned} \mathbb{E}(\epsilon_{(n)}^l) &= \int_{-\infty}^{+\infty} \epsilon^l n \Psi(\epsilon^l)^{n-1} \psi(\epsilon^l) d\epsilon^l = \int_{-\infty}^{+\infty} \epsilon^l n \Psi(\epsilon^l)^{n-1} d\Psi(\epsilon^l) \\ &= \int_0^1 \Psi^{-1}(u) n u^{n-1} du \end{aligned} \quad (2.6)$$

where the last equality is by substituting $\Psi^{-1}(u) = \epsilon^l$. Similarly, we have $\mathbb{E}(\epsilon_{(n)}^\circ) = \int_0^1 \Psi^{\circ^{-1}}(u) n u^{n-1} du$. Then

$$\begin{aligned} \mathbb{E}(\epsilon_{(n)}^1) + \mathbb{E}(\epsilon_{(n)}^2) - \mathbb{E}((\epsilon^1 + \epsilon^2)_{(n)}) &= \mathbb{E}(\epsilon_{(n)}^1) + \mathbb{E}(\epsilon_{(n)}^2) - \mathbb{E}(\epsilon_{(n)}^\circ) \\ &= \int_0^1 2\Psi^{-1}(u) n u^{n-1} du - \int_0^1 \Psi^{\circ^{-1}}(u) n u^{n-1} du \end{aligned}$$

$$= \int_0^1 (2\Psi^{-1}(u) - \Psi^{\circ-1}(u))nu^{n-1}du. \quad (2.7)$$

Recall the assumption that ϵ^l , $l = 1, 2$, follows a symmetric log-concave distribution, thus $\Psi(\epsilon^l) = 1 - \Psi(\epsilon^l)$, and then $\Psi^{-1}(u) = \Psi^{-1}(1 - u)$. Similarly, since ϵ° follows a symmetric distribution, then $\Psi^{\circ-1}(1 - u) = \Psi^{\circ-1}(u)$. As a result, by (2.7), we have the following result:

$$\begin{aligned} & \mathbb{E}(\epsilon_{(n)}^1) + \mathbb{E}(\epsilon_{(n)}^2) - \mathbb{E}((\epsilon^1 + \epsilon^2)_{(n)}) \\ &= \int_0^{\frac{1}{2}} (2\Psi^{-1}(u) - \Psi^{\circ-1}(u))nu^{n-1}du + \int_{\frac{1}{2}}^1 (2\Psi^{-1}(u) - \Psi^{\circ-1}(u))nu^{n-1}du \\ &= -\int_{\frac{1}{2}}^1 (2\Psi^{-1}(1-u) - \Psi^{\circ-1}(1-u))n(1-u)^{n-1}du + \int_{\frac{1}{2}}^1 (2\Psi^{-1}(u) - \Psi^{\circ-1}(u))nu^{n-1}du \\ &= -\int_{\frac{1}{2}}^1 (2\Psi^{-1}(u) - \Psi^{\circ-1}(u))n(1-u)^{n-1}du + \int_{\frac{1}{2}}^1 (2\Psi^{-1}(u) - \Psi^{\circ-1}(u))nu^{n-1}du \\ &= \int_{\frac{1}{2}}^1 (2\Psi^{-1}(u) - \Psi^{\circ-1}(u))n[u^{n-1} - (1-u)^{n-1}]du. \end{aligned} \quad (2.8)$$

By Bagnoli and Bergstrom (2005), Corollary 2, if the PDF is log-concave, then its hazard rate function is increasing over the support. By Watson and Gordon (1986), Theorem 1, if ϵ^1 and ϵ^2 are independent continuous random variables having symmetric distribution with non-decreasing hazard rate functions, then $2\Psi^{-1}(u) - \Psi^{\circ-1}(u) < 0$ if $u \in (0, 1/2)$, and $2\Psi^{-1}(u) - \Psi^{\circ-1}(u) > 0$ if $u \in (1/2, 1)$. Though $(2\Psi^{-1}(u) - \Psi^{\circ-1}(u))n[u^{n-1} - (1-u)^{n-1}] = 0$ if $u = 1/2$, it has a measure 0 for $u = 1/2$ in (2.8). Since $2\Psi^{-1}(u) - \Psi^{\circ-1}(u) > 0$ and $u^{n-1} - (1-u)^{n-1} > 0$ if $u \in (1/2, 1)$, when $n \geq 2$, by (2.8), $\mathbb{E}(\epsilon_{(n)}^1) + \mathbb{E}(\epsilon_{(n)}^2) - \mathbb{E}((\epsilon^1 + \epsilon^2)_{(n)}) > 0$. ■

Proof of Proposition 2.4. (i) is directly given by Lemma 2.3 by letting $n = 2$. To prove (ii), recall the equilibrium efforts in Lemma 2.1 and (2.1), $e^{l*} = C^{l' - 1}(A^l g(0))$, $l = 1, 2$, and $e^{\circ*} = C^{\circ' - 1}(Ag^\circ(0))$. The difference of the equilibrium efforts is

$$\begin{aligned} \Delta^e &= e^{1*} + e^{2*} - e^{\circ*} = C^{1' - 1}(A^1 g(0)) + C^{2' - 1}(A^2 g(0)) - C^{\circ' - 1}(Ag^\circ(0)) \\ &= C^{1' - 1}(wAg(0)) + C^{2' - 1}((1-w)Ag(0)) - C^{1' - 1}(Ag^\circ(0)) - C^{2' - 1}(Ag^\circ(0)) \\ &= [C^{1' - 1}(wAg(0)) - C^{1' - 1}(Ag^\circ(0))] \\ &\quad + [C^{2' - 1}((1-w)Ag(0)) - C^{2' - 1}(Ag^\circ(0))]. \end{aligned} \quad (2.9)$$

where the third equality is driven by Lemma 2.2 that $C^{\circ'}(e^\circ) = C^{1'}(\tilde{e}^1) = C^{2'}(\tilde{e}^2)$ and all the cost functions are strictly increasing and strictly convex. By (2.9), if $g^\circ(0) > wg(0)$ and $g^\circ(0) > (1-w)g(0)$, then $\Delta^e < 0$. Equivalently, the sufficient condition is $g^\circ(0) >$

$\max\{wg(0), (1-w)g(0)\}$.

(iii)(a) is directly given by letting $w = 1/2$. For (iii)(b), denote $C(\cdot) = C^1(\cdot) = C^2(\cdot)$. If $C'(\cdot)$ is convex, then $C'^{-1}(\cdot)$ is concave. By the concavity of $C'^{-1}(\cdot)$, we have $\frac{1}{2}(C'^{-1}(y_1) + C'^{-1}(y_2)) \leq C'^{-1}\left(\frac{y_1 + y_2}{2}\right)$ where y_1 and y_2 are in the domain of $C'^{-1}(\cdot)$. As a result, the inequality holds for $C^{1'-1}(wAg(0)) + C^{2'-1}((1-w)Ag(0)) = C'^{-1}(wAg(0)) + C'^{-1}((1-w)Ag(0)) \leq 2C'^{-1}(Ag(0)/2)$. Furthermore, by (2.9) and $C(\cdot) = C^1(\cdot) = C^2(\cdot)$, $\Delta^e = C'^{-1}(wAg(0)) + C'^{-1}((1-w)Ag(0)) - 2C'^{-1}(Ag(0)/2) \leq 2C'^{-1}(Ag(0)/2) - 2C'^{-1}(Ag(0)/2)$. Since $C(\cdot)$ is strictly convex, $C'(\cdot)$ is strictly increasing, and then $C'^{-1}(\cdot)$ is strictly increasing. Hence, if $g^\circ(0) > g(0)/2$, $\Delta^e < 0$. ■

Proof of Corollary 2.5. For contestant i , $V_i^{seq} = e^{1*} + \epsilon^1 + e^{2*} + \epsilon^2$ and $V_i^{sim} = e^{\circ*} + \epsilon^1 + \epsilon^2$. By (2.9), if $g^\circ(0) \geq \max\{wg(0), (1-w)g(0)\}$, $e^{1*} + e^{2*} \leq e^{\circ*}$. Thus, $P\{V_i^{sim} \geq z\} = P\{e^{\circ*} + \epsilon^1 + \epsilon^2 \geq z\} = P\{\epsilon^1 + \epsilon^2 \geq z - e^{\circ*}\} \geq P\{\epsilon^1 + \epsilon^2 \geq z - e^{1*} - e^{2*}\} = P\{V_i^{seq} \geq z\}$ for any z , where the inequality is due to $e^{1*} + e^{2*} \leq e^{\circ*}$. By the definition of usual stochastic order (see Shaked and Shanthikumar 2007, (1.A.1)), $V_i^{sim} \geq_{st} V_i^{seq}$. ■

Proof of Proposition 2.6. If the cost functions are in the exponential form $C^l(e^l) = \exp(\rho^l e^l)$, $l = 1, 2$, $C^{l'}(e^l) = \rho^l \exp(\rho^l e^l)$ and then $C^{l'-1}(y) = \ln(y\rho^l) / \rho^l$ where y is in the domain of $C^{l'-1}(\cdot)$. By (2.9), the difference of equilibrium efforts between two contests can be written as

$$\begin{aligned} \Delta^e &= \ln(wAg(0)/\rho^1) / \rho^1 + \ln((1-w)Ag(0)/\rho^2) / \rho^2 \\ &\quad - \ln(Ag^\circ(0)/\rho^1) / \rho^1 - \ln(Ag^\circ(0)/\rho^2) / \rho^2 \\ &= \ln(wg(0)/g^\circ(0)) / \rho^1 + \ln((1-w)g(0)/g^\circ(0)) / \rho^2. \end{aligned} \quad (2.10)$$

Without loss of generality, assume that $\rho^1 \geq \rho^2$ and denote $Q = \ln(wg(0)/g^\circ(0)) + \ln((1-w)g(0)/g^\circ(0))$. If $g^\circ(0) > \max\{wg(0), (1-w)g(0)\}$, then $Q < 0$. Thus we have the lower and upper bounds of Δ^e as $Q/\rho^1 \geq \Delta^e \geq Q/\rho^2$.

The upper bound Q/ρ^2 is increasing in ρ^2 and the lower bound Q/ρ^1 is increasing in ρ^1 . Since $\lim_{\rho^1 \rightarrow 0} Q/\rho^1 = -\infty$. By Proposition 2.4(i), $\Delta^\epsilon > 0$, there exists $\underline{\rho} > 0$ such that when $\rho^2 \leq \rho^1 \leq \underline{\rho}$, $\Delta = \Delta^e + \Delta^\epsilon \leq 0$. Moreover, $\lim_{\rho^2 \rightarrow \infty} Q/\rho^2 = 0$. There exists $\bar{\rho} > 0$ such that when $\rho^1 \geq \rho^2 \geq \bar{\rho}$, $\Delta = \Delta^e + \Delta^\epsilon \geq 0$.

If $\rho^1 = \rho^2$, denote $\rho = \rho^1 = \rho^2$. By (2.10), $\Delta^e = \frac{1}{\rho}(\ln(wg(0)/g^\circ(0)) + \ln((1-w)g(0)/g^\circ(0))) = Q/\rho$. If $g^\circ(0) > \max\{wg(0), (1-w)g(0)\}$, then $Q < 0$. The difference $\Delta^e = Q/\rho$ is increasing in $\rho > 0$. Since $\lim_{\rho \rightarrow 0} Q/\rho = -\infty$ and $\lim_{\rho \rightarrow \infty} Q/\rho = 0$, there exists $\tilde{\rho} > 0$ such that if $\rho \geq \tilde{\rho}$, $\Delta = \Delta^e + \Delta^\epsilon \geq 0$, and if $\rho \leq \tilde{\rho}$, $\Delta = \Delta^e + \Delta^\epsilon \leq 0$. ■

Proof of Proposition 2.7. For (i), by Lemma 2.1, the total effort in the sequential

contest is

$$e^{1^*}(w) + e^{2^*}(w) = \ln(wAg(0)/\rho^1)/\rho^1 + \ln((1-w)Ag(0)/\rho^2)/\rho^2.$$

Take derivative with respect to w , $\frac{d(e^{1^*}(w)+e^{2^*}(w))}{dw} = \frac{1}{\rho^1 w} - \frac{1}{\rho^2(1-w)}$. Hence, $\frac{d^2(e^{1^*}(w)+e^{2^*}(w))}{dw^2} = -\frac{1}{\rho^1 w^2} - \frac{1}{\rho^2(1-w)^2} \leq 0$, then $e^{1^*}(w) + e^{2^*}(w)$ is concave in $w \in (0, 1)$, and thus the optimal prize allocation is $w^* = \frac{\rho^2}{\rho^1 + \rho^2}$. With the optimal allocation, the total effort becomes

$$e^{1^*}(w^*) + e^{2^*}(w^*) = \ln\left(\frac{\rho^2 Ag(0)}{\rho^1(\rho^1 + \rho^2)}\right)/\rho^1 + \ln\left(\frac{\rho^1 Ag(0)}{\rho^2(\rho^1 + \rho^2)}\right)/\rho^2.$$

For (ii), by (2.1), the total effort in the simultaneous contest is $e^{o*} = \ln(Ag^\circ(0)/\rho^1)/\rho^1 + \ln(Ag^\circ(0)/\rho^2)/\rho^2$. Then, we have the difference of efforts under the optimal allocation of prizes in the sequential contest,

$$\begin{aligned} \Delta^e &= \frac{1}{\rho^1} \ln\left(\frac{\rho^2 g(0)}{(\rho^1 + \rho^2)g^\circ(0)}\right) + \frac{1}{\rho^2} \ln\left(\frac{\rho^1 g(0)}{(\rho^1 + \rho^2)g^\circ(0)}\right) \\ &= \left[\rho^2 \ln\left(\frac{\rho^2 g(0)}{(\rho^1 + \rho^2)g^\circ(0)}\right) + \rho^1 \ln\left(\frac{\rho^1 g(0)}{(\rho^1 + \rho^2)g^\circ(0)}\right) \right] / (\rho^1 \rho^2) \\ &= \ln\left[\left(\frac{\rho^2}{\rho^1 + \rho^2}\right)^{\rho^2} \left(\frac{\rho^1}{\rho^1 + \rho^2}\right)^{\rho^1} \left(\frac{g(0)}{g^\circ(0)}\right)^{\rho^1 + \rho^2} \right] / (\rho^1 \rho^2) \\ &= \left\{ \ln\left[\left(\frac{\rho^2}{\rho^1 + \rho^2}\right)^{\frac{\rho^2}{\rho^1 + \rho^2}} \left(\frac{\rho^1}{\rho^1 + \rho^2}\right)^{\frac{\rho^1}{\rho^1 + \rho^2}} \right] - \ln\left[\frac{g^\circ(0)}{g(0)}\right] \right\} / ((\rho^1 + \rho^2)\rho^1 \rho^2) \\ &= \left\{ \ln[s^s(1-s)^{1-s}] - \ln[g^\circ(0)/g(0)] \right\} / ((\rho^1 + \rho^2)\rho^1 \rho^2), \end{aligned}$$

where $s = \frac{\rho^2}{\rho^1 + \rho^2}$. The sign of Δ^e depends on the comparison of $s^s(1-s)^{1-s}$ ($0 < s < 1$) and $g^\circ(0)/g(0)$. Since $\frac{ds^s(1-s)^{1-s}}{ds} = (1-s)^{1-s}s^s(\ln(s) - \ln(1-s))$, $s^s(1-s)^{1-s}$ is decreasing in $s \in [0, 1/2]$ and increasing in $s \in [1/2, 1]$. If $s = 1/2$, then $s^s(1-s)^{1-s} = 1/2$, and if $s = 1$ or $s = 0$, then $s^s(1-s)^{1-s} = 1$. Therefore, $s^s(1-s)^{1-s} \in [1/2, 1)$. For a given value of $g^\circ(0)/g(0) \in (1/2, 1)$ (e.g., if $e^l \sim N(0, \sigma)$, $l = 1, 2$, $g^\circ(0)/g(0) = 1/\sqrt{2}$), there exist $\bar{s} \in (1/2, 1)$ and $\underline{s} \in (0, 1/2)$ such that when $s \in (\underline{s}, \bar{s})$, $s^s(1-s)^{1-s} < g^\circ(0)/g(0)$, i.e., $\Delta^e < 0$. Otherwise, if $s \in (0, \underline{s}] \cup [\bar{s}, 1)$, $\Delta^e \geq 0$.

Since $t = 1/s - 1$, $t \in (0, 1)$ if $s \in (1/2, 1)$, and $t \in (1, \infty)$ if $s \in (0, 1/2)$. Thus, there exist $\bar{t} \in (0, 1)$ and $\underline{t} \in (1, \infty)$, such that when $t \in (\underline{t}, \bar{t})$, $\Delta^e < 0$. Otherwise, if $t \in (0, \underline{t}] \cup [\bar{t}, \infty)$, $\Delta^e \geq 0$. ■

Proof of Proposition 2.9. (i) is directly given by Lemma 2.3. For (ii), by Lemma 2.8,

the difference of equilibrium efforts with n contestants is

$$\begin{aligned} \Delta_n^e &= e^{1*} + e^{2*} - e^{\circ*} = C^{1' - 1} (A^1 h(\epsilon; n)) + C^{2' - 1} (A^2 h(\epsilon; n)) - C^{\circ' - 1} (A h^\circ(\epsilon^\circ; n)) \\ &= C^{1' - 1} (A^1 h(\epsilon; n)) + C^{2' - 1} (A^2 h(\epsilon; n)) \\ &\quad - C^{1' - 1} (A h^\circ(\epsilon^\circ; n)) - C^{2' - 1} (A h^\circ(\epsilon^\circ; n)), \end{aligned} \quad (2.11)$$

where the third equality is driven by Lemma 2.3 that $C^{\circ'}(e^\circ) = C^{1'}(\tilde{e}^1) = C^{2'}(\tilde{e}^2)$ and all the cost functions are strictly increasing and strictly convex. Similar to Proposition 2.4(ii), the sufficient condition for $\Delta_n^e < 0$ is $h^\circ(\epsilon^\circ; n) > \max\{wh(\epsilon; n), (1 - w)h(\epsilon; n)\}$.

For (ii), denote $H(n) = \frac{h^\circ(\epsilon^\circ; n)}{h(\epsilon; n)}$. For any n , if $H(n) > \max\{w, (1 - w)\}$, then $\Delta_n^e < 0$. If $\epsilon^l \sim N(0, \sigma)$, $\epsilon^\circ = \epsilon^1 + \epsilon^2 \sim N(0, \sqrt{2}\sigma)$. We have

$$\begin{aligned} h(\epsilon; n) &= \int_{-\infty}^{+\infty} (n - 1)\Psi(\epsilon)^{n-2}\psi(\epsilon)^2 d\epsilon = \int_0^1 \psi(\epsilon)d\Psi(\epsilon)^{n-1} \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_0^1 \exp(-\epsilon^2/(2\sigma^2))d\Psi(\epsilon)^{n-1}. \end{aligned}$$

Substitute ϵ/σ with y , then $h(\epsilon; n) = \frac{1}{\sqrt{2\pi}\sigma} \int_0^1 \exp(-y^2/2)d\Phi(y)^{n-1}$ where $\Phi(y) \sim N(0, 1)$. Similarly, for $h^\circ(\epsilon^\circ; n)$, substitute $\epsilon/(\sqrt{2}\sigma)$ with \tilde{y} ,

$$\begin{aligned} h^\circ(\epsilon^\circ; n) &= \int_0^1 \psi^\circ(\epsilon^\circ)d\Psi^\circ(\epsilon^\circ)^{n-1} = \frac{1}{2\sqrt{\pi}\sigma} \int_0^1 \exp(-\epsilon^2/(4\sigma^2))d\Psi^\circ(\epsilon)^{n-1} \\ &= \frac{1}{2\sqrt{\pi}\sigma} \int_0^1 \exp(-\tilde{y}^2/2)d\Phi(\tilde{y})^{n-1}. \end{aligned}$$

Then $H(n) = \frac{h^\circ(\epsilon^\circ; n)}{h(\epsilon; n)} = 1/\sqrt{2}$. Thus if random factors follow normal distribution, $H(n) = 1/\sqrt{2}$. By $H(n) > \max\{w, (1 - w)\}$, if $w \in (1 - \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$, $\Delta_n^e < 0$ for any n . ■

Proof of Proposition 2.10. Denote the r th order statistic of a random variable with a sample size n by subscript $_{(r:n)}$ and denote $\Psi^{-1}(\cdot)$ as the quantile function of CDF $\Psi(\cdot)$. It is sufficient to show that there exists a \bar{n} such that

$$\mathbb{E}(\epsilon_{(n-1:n)}^1) + \mathbb{E}(\epsilon_{(n-1:n)}^2) - \mathbb{E}(\epsilon_{(n-1:n)}^\circ) \leq 0, \quad (2.12)$$

when $n \leq \bar{n}$. According to Chakraborty (1999), (2.12) holds for that the following *regularity condition* is satisfied: there exists $u_0 \in (0, 1)$ such that $\Psi^{-1}(u) + \Psi^{-1}(u) - \Psi^{\circ-1}(u) < 0$ if $u \in (0, u_0)$, and $\Psi^{-1}(u) + \Psi^{-1}(u) - \Psi^{\circ-1}(u) > 0$ if $u \in (u_0, 1)$.

Recall the assumption that ϵ^l , $l = 1, 2$, follows a symmetric log-concave distribution. By Bagnoli and Bergstrom (2005), Corollary 2, if the PDF is log-concave, then its hazard rate function is increasing over the support. By Watson and Gordon (1986), Theorem

1, one sufficient condition for the regularity condition to hold and $u_0 = 0.5$ is that $\Psi(\cdot)$ is a symmetric distribution with non-decreasing hazard rate function. Therefore, the existence of \bar{n} such that (2.12) holds is guaranteed under our assumption that the random factors follow a symmetric log-concave distribution.

Write the formulas of $\mathbb{E}(\epsilon_{(n)}^l)$ and $\mathbb{E}(\epsilon_{(n-1)}^l)$, $l = 1, 2$, $\mathbb{E}(\epsilon_{(n)}^l) = \int_{-\infty}^{+\infty} \epsilon^l n \Psi(\epsilon^l)^{n-1} \psi(\epsilon^l) d\epsilon^l$ and $\mathbb{E}(\epsilon_{(n-1)}^l) = \int_{-\infty}^{+\infty} \epsilon^l (n-1) \Psi(\epsilon^l)^{n-2} \psi(\epsilon^l) d\epsilon^l$. The following *recurrence relation* holds (David and Nagaraja, 2003, Chapter 3.4 Relation 1):

$$\begin{aligned} & n\mathbb{E}(\epsilon_{(n-1)}^l) - (n-1)\mathbb{E}(\epsilon_{(n)}^l) \\ &= \int_{-\infty}^{+\infty} \epsilon^l n(n-1) \Psi(\epsilon^l)^{n-2} \psi(\epsilon^l) d\epsilon^l - \int_{-\infty}^{+\infty} \epsilon^l n(n-1) \Psi(\epsilon^l)^{n-1} \psi(\epsilon^l) d\epsilon^l \\ &= \int_{-\infty}^{+\infty} \epsilon^l n(n-1) (\Psi(\epsilon^l)^{n-2} - \Psi(\epsilon^l)^{n-1}) \psi(\epsilon^l) d\epsilon^l = \mathbb{E}(\epsilon_{(n-1:n)}^l), \end{aligned} \quad (2.13)$$

where the last equality is because the PDF for the $(n-1)$ th order statistics of ϵ^l with sample size n is $n(n-1)\Psi(\epsilon^l)^{n-2}(1-\Psi(\epsilon^l))\psi(\epsilon^l)$. A similar relation can be applied to ϵ° , $n\mathbb{E}(\epsilon_{(n-1)}^\circ) - (n-1)\mathbb{E}(\epsilon_{(n)}^\circ) = \mathbb{E}(\epsilon_{(n-1:n)}^\circ)$. By the above relations, we have $\mathbb{E}(\epsilon_{(n-1)}^l) - \mathbb{E}(\epsilon_{(n)}^l) = \frac{1}{n}[\mathbb{E}(\epsilon_{(n-1:n)}^l) - \mathbb{E}(\epsilon_{(n)}^l)]$ ($l = 1, 2$) and $\mathbb{E}(\epsilon_{(n-1)}^\circ) - \mathbb{E}(\epsilon_{(n)}^\circ) = \frac{1}{n}[\mathbb{E}(\epsilon_{(n-1:n)}^\circ) - \mathbb{E}(\epsilon_{(n)}^\circ)]$. Then,

$$\begin{aligned} \Delta_{n-1}^\epsilon - \Delta_n^\epsilon &= [\mathbb{E}(\epsilon_{(n-1)}^1) + \mathbb{E}(\epsilon_{(n-1)}^2) - \mathbb{E}(\epsilon_{(n-1)}^\circ)] - [\mathbb{E}(\epsilon_{(n)}^1) + \mathbb{E}(\epsilon_{(n)}^2) - \mathbb{E}(\epsilon_{(n)}^\circ)] \\ &= [\mathbb{E}(\epsilon_{(n-1)}^1) - \mathbb{E}(\epsilon_{(n)}^1)] + [\mathbb{E}(\epsilon_{(n-1)}^2) - \mathbb{E}(\epsilon_{(n)}^2)] - [\mathbb{E}(\epsilon_{(n-1)}^\circ) - \mathbb{E}(\epsilon_{(n)}^\circ)] \\ &= \frac{1}{n} \left\{ [\mathbb{E}(\epsilon_{(n-1:n)}^1) - \mathbb{E}(\epsilon_{(n)}^1)] + [\mathbb{E}(\epsilon_{(n-1:n)}^2) - \mathbb{E}(\epsilon_{(n)}^2)] - [\mathbb{E}(\epsilon_{(n-1:n)}^\circ) - \mathbb{E}(\epsilon_{(n)}^\circ)] \right\} \\ &= \frac{1}{n} \left\{ [\mathbb{E}(\epsilon_{(n-1:n)}^1) + \mathbb{E}(\epsilon_{(n-1:n)}^2) - \mathbb{E}(\epsilon_{(n-1:n)}^\circ)] - [\mathbb{E}(\epsilon_{(n)}^1) + \mathbb{E}(\epsilon_{(n)}^2) - \mathbb{E}(\epsilon_{(n)}^\circ)] \right\} \\ &= \frac{1}{n} \left\{ [\mathbb{E}(\epsilon_{(n-1:n)}^1) + \mathbb{E}(\epsilon_{(n-1:n)}^2) - \mathbb{E}(\epsilon_{(n-1:n)}^\circ)] - \Delta_n^\epsilon \right\}. \end{aligned}$$

By Lemma 2.3, $\Delta_n^\epsilon \geq 0$. Then $\Delta_{n-1}^\epsilon - \Delta_n^\epsilon \leq 0$ if $\mathbb{E}(\epsilon_{(n-1:n)}^1) + \mathbb{E}(\epsilon_{(n-1:n)}^2) - \mathbb{E}(\epsilon_{(n-1:n)}^\circ) \leq 0$. By (2.12), there exists \bar{n} such that when $n \leq \bar{n}$, $\mathbb{E}(\epsilon_{(n-1:n)}^1) + \mathbb{E}(\epsilon_{(n-1:n)}^2) - \mathbb{E}(\epsilon_{(n-1:n)}^\circ) \leq 0$. Thus, there exists \bar{n} such that when $n \leq \bar{n}$, $\Delta_{n-1}^\epsilon - \Delta_n^\epsilon \leq 0$. The desired result holds. ■

Proof of Proposition 2.11. First, we prove the following lemma.

Lemma 2.15 *If PDF $\psi(\epsilon^1)$ is symmetric and log-concave, then*

$$\mathbb{E}(\epsilon_{(n)}^1) \geq \left(1 - \frac{1}{2^n}\right) \Psi^{-1}\left(\frac{n}{n+1} - \frac{1}{2^n}\right).$$

Proof of Lemma 2.15. The expectation of the highest order statistics of ϵ^1 can be written

as

$$\mathbb{E}(\epsilon_{(n)}^1) = \int_{-\infty}^{+\infty} \epsilon^1 n \Psi(\epsilon^1)^{n-1} d\Psi(\epsilon^1) = \int_0^1 \Psi^{-1}(u) n u^{n-1} du,$$

by letting $\epsilon^1 = \Psi^{-1}(u)$. Let $B(u) = nu^{n-1}$, then we have

$$\begin{aligned} \mathbb{E}(\epsilon_{(n)}^1) &= \int_0^1 \Psi^{-1}(u) B(u) du = \int_{1/2}^1 \Psi^{-1}(u) B(u) du + \int_0^{1/2} \Psi^{-1}(u) B(u) du \\ &= \int_{1/2}^1 \Psi^{-1}(u) B(u) du - \int_{1/2}^1 \Psi^{-1}(1-u) B(1-u) du \\ &= \int_{1/2}^1 \Psi^{-1}(u) B(u) du - \int_{1/2}^1 \Psi^{-1}(u) B(1-u) du \\ &= \int_{1/2}^1 \Psi^{-1}(u) [B(u) - B(1-u)] du \end{aligned}$$

where the fourth equality is by the symmetric property that $\psi(\epsilon^1) = \psi(-\epsilon^1)$, $\Psi(\epsilon^1) = 1 - \Psi(\epsilon^1)$ and $\Psi^{-1}(1-u) = \Psi^{-1}(u)$. By Lemma 2.16, PDF $\psi(\epsilon^1)$ is unimodal and symmetric at 0, thus $\psi(\epsilon^1)$ is decreasing in $\epsilon^1 \geq 0$. When $\epsilon^1 \geq 0$, the CDF $\Psi(\epsilon^1)$ is concave because $\Psi''(\epsilon^1) = \psi'(\epsilon^1) \leq 0$. As a result, $\Psi^{-1}(u)$ is convex in $u \in [1/2, 1]$. Let $K = \int_{1/2}^1 B(u) - B(1-u) du = \int_{1/2}^1 [nu^{n-1} - n(1-u)^{n-1}] du = 1 - 1/2^n$, thus $\int_{1/2}^1 (B(u) - B(1-u))/K du = 1$. Since $(B(u) - B(1-u))/K$ can be considered as a PDF, then $\int_{1/2}^1 \Psi^{-1}(u) \{[B(u) - B(1-u)]/K\} du$ is the expectation of $\Psi^{-1}(u)$ with such PDF. By the convexity of $\Psi^{-1}(u)$ and Jensen's inequality, we have

$$\mathbb{E}(\epsilon_{(n)}^1)/K = \int_{1/2}^1 \Psi^{-1}(u) \{[B(u) - B(1-u)]/K\} du \geq \Psi^{-1} \left(\int_{1/2}^1 u [B(u) - B(1-u)]/K du \right).$$

Integrating by parts, we have

$$\begin{aligned} \int_{1/2}^1 u [B(u) - B(1-u)] du &= \int_{1/2}^1 [nu^n - n(1-u)^{n-1}u] du \\ &= \left(\frac{n}{n+1} - \frac{n}{n+1} \frac{1}{2^{n+1}} \right) + \int_{1/2}^1 u d(1-u)^n \\ &= \left(\frac{n}{n+1} - \frac{n}{n+1} \frac{1}{2^{n+1}} \right) - \frac{1}{2^{n+1}} - \int_{1/2}^1 (1-u)^n du \\ &= \left(\frac{n}{n+1} - \frac{n}{n+1} \frac{1}{2^{n+1}} \right) - \frac{1}{2^{n+1}} - \frac{1}{n+1} \frac{1}{2^{n+1}} = \frac{n}{n+1} - \frac{1}{2^n}. \end{aligned}$$

Because $K = \int_{1/2}^1 B(u) - B(1-u) du = 1 - 1/2^n$,

$$\mathbb{E}(\epsilon_{(n)}^1) \geq K \Psi^{-1} \left(\int_{1/2}^1 u [B(u) - B(1-u)]/K du \right)$$

$$\geq \left(1 - \frac{1}{2^n}\right) \Psi^{-1}\left(\left(\frac{n}{n+1} - \frac{1}{2^n}\right) / \left(1 - \frac{1}{2^n}\right)\right) \geq \left(1 - \frac{1}{2^n}\right) \Psi^{-1}\left(\frac{n}{n+1} - \frac{1}{2^n}\right),$$

where the last inequality is because $\Psi^{-1}(\cdot)$ is increasing and $1 - \frac{1}{2^n} \leq 1$. The inequality holds. ■

Now we prove Proposition 2.11.

(i) If ϵ^1 and ϵ^2 have a bounded support $[-a, a]$, then ϵ° has the bounded support $[-2a, 2a]$. By Lemma 2.15, $\left(1 - \frac{1}{2^n}\right) \Psi^{-1}\left(\frac{n}{n+1} - \frac{1}{2^n}\right) \leq \mathbb{E}(\epsilon_{(n)}^1) \leq a$. We have $\lim_{n \rightarrow \infty} \left(1 - \frac{1}{2^n}\right) \Psi^{-1}\left(\frac{n}{n+1} - \frac{1}{2^n}\right) \leq \lim_{n \rightarrow \infty} \mathbb{E}(\epsilon_{(n)}^1) \leq \lim_{n \rightarrow \infty} a$. By the Squeeze Theorem,

$$\lim_{n \rightarrow \infty} \left(1 - \frac{1}{2^n}\right) \Psi^{-1}\left(\frac{n}{n+1} - \frac{1}{2^n}\right) = a = \lim_{n \rightarrow \infty} \mathbb{E}(\epsilon_{(n)}^1) = a.$$

Similar results can be extended to ϵ^2 and ϵ° , thus $\lim_{n \rightarrow \infty} \Delta_n^\epsilon = \lim_{n \rightarrow \infty} \mathbb{E}(\epsilon_{(n)}^1) + \lim_{n \rightarrow \infty} \mathbb{E}(\epsilon_{(n)}^2) - \lim_{n \rightarrow \infty} \mathbb{E}(\epsilon_{(n)}^\circ) = a + a - 2a = 0$. The result holds.

(ii) If $\epsilon^l \sim N(0, \sigma)$ ($l = 1, 2$), the quantile function can be written as

$$\Phi^{-1}(u) = \sigma \sqrt{2} \operatorname{erf}^{-1}(2u - 1), \quad (2.14)$$

where $u \in (0, 1)$ and $\operatorname{erf}(u)$ is the error function, $\operatorname{erf}(u) = \frac{2}{\sqrt{\pi}} \int_{-u}^u e^{-t^2} dt$. Then Δ_n^ϵ can be written as

$$\begin{aligned} \Delta_n^\epsilon &= \int_{-\infty}^{+\infty} \epsilon^1 n \Phi^{n-1}(\epsilon^1) d\Phi(\epsilon^1) + \int_{-\infty}^{+\infty} \epsilon^2 n \Phi^{n-1}(\epsilon^2) d\Phi(\epsilon^2) - \int_{-\infty}^{+\infty} \epsilon^\circ n \Phi^{\circ n-1}(\epsilon^\circ) d\Phi^\circ(\epsilon^\circ) \\ &= \int_0^1 n \Phi^{-1}(u) u^{n-1} du + \int_0^1 n \Phi^{-1}(u) u^{n-1} du - \int_0^1 n \Phi^{\circ-1}(u) u^{n-1} du \\ &= \int_0^1 n (2\Phi^{-1}(u) - \Phi^{\circ-1}(u)) u^{n-1} du. \end{aligned}$$

By (2.14), $2\Phi^{-1}(u) - \Phi^{\circ-1}(u) = (2\sqrt{2}\sigma - 2\sigma) \operatorname{erf}^{-1}(2u - 1)$ because $\epsilon^\circ \sim N(0, \sqrt{2}\sigma)$ and $\Phi^{\circ-1}(u) = 2\sigma \operatorname{erf}^{-1}(2u - 1)$. Hence, we can define a new random variable $\tilde{\epsilon}$ that is normally distributed with mean 0 and standard deviation $(2 - \sqrt{2})\sigma$. Denote its CDF by $\tilde{\Phi}(\tilde{\epsilon})$, then $2\Phi^{-1}(u) - \Phi^{\circ-1}(u) = \tilde{\Phi}^{-1}(u)$. Thus

$$\Delta_n^\epsilon = \int_0^1 n \tilde{\Phi}^{-1}(u) u^{n-1} du = \mathbb{E}(\tilde{\epsilon}_{(n)}). \quad (2.15)$$

By Lemma 2.15, $\lim_{n \rightarrow \infty} \left(1 - \frac{1}{2^n}\right) \tilde{\Phi}^{-1}\left(\frac{n}{n+1} - \frac{1}{2^n}\right) \leq \lim_{n \rightarrow \infty} \mathbb{E}(\tilde{\epsilon}_{(n)})$. Because the normal distribution is defined on the $(-\infty, \infty)$, then $\lim_{n \rightarrow \infty} \left(1 - \frac{1}{2^n}\right) \tilde{\Phi}^{-1}\left(\frac{n}{n+1} - \frac{1}{2^n}\right) = \infty$. Then if ϵ^1 and ϵ^2 follow normal distributin, $\lim_{n \rightarrow \infty} \Delta_n^\epsilon = \infty$. ■

Proof of Proposition 2.12. If the cost functions are in the exponential form, i.e.,

$C^l(e^l) = \exp(\rho^l e^l)$, $\rho^l > 0$, $l = 1, 2$, by (2.11), the difference of equilibrium efforts is

$$\Delta_n^e = \ln \left(\frac{wh(\epsilon; n)}{h^\circ(\epsilon^\circ; n)} \right) / \rho^1 + \ln \left(\frac{(1-w)h(\epsilon; n)}{h^\circ(\epsilon^\circ; n)} \right) / \rho^2. \quad (2.16)$$

By Proposition 2.9(ii), if $\epsilon^l \sim N(0, \sigma)$, $l = 1, 2$, then $1/H(n) = \frac{h(\epsilon; n)}{h^\circ(\epsilon^\circ; n)} = \sqrt{2}$ for any n . Since $w \in (1 - \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ and $\frac{h(\epsilon; n)}{h^\circ(\epsilon^\circ; n)} = \sqrt{2}$, $\Delta_n^e < 0$ by (2.16). Thus, Δ_n^e is a fixed non-positive value for any n . For Δ_n^e , by (2.15), $\Delta_n^e = \mathbb{E}(\tilde{\epsilon}_{(n)})$ where $\tilde{\epsilon} \sim N(0, (2 - \sqrt{2})\sigma)$. By (2.13), $\mathbb{E}(\tilde{\epsilon}_{(n-1)}) - \mathbb{E}(\tilde{\epsilon}_{(n)}) = \frac{1}{n}[\mathbb{E}(\tilde{\epsilon}_{(n-1:n)}) - \mathbb{E}(\tilde{\epsilon}_{(n)})] \leq 0$. Thus, $\mathbb{E}(\tilde{\epsilon}_{(n)})$ is increasing in n . By Proposition 2.11, $\lim_{n \rightarrow \infty} \Delta_n^e = \infty$, therefore the difference of the random factors Δ_n^e is a positive value increasing in n . There exists a threshold $\tilde{n} \in [2, \infty)$ such that when $n \leq \tilde{n}$, $\Delta_n = \Delta_n^e + \Delta_n^e \leq 0$, and when $n \geq \tilde{n}$, $\Delta_n = \Delta_n^e + \Delta_n^e \geq 0$. ■

2.9 Appendix.

Proof of Lemma 2.1. Consider a two-person model in which the contestants are denoted as i and j . For sub-contest l , $l = 1, 2$, because the random variables $\xi^l = \epsilon_i^l - \epsilon_j^l$ have CDF $G(\xi)$, the winning probability of a contestant i can be written by $P\{e_i^l + \epsilon_i^l > e_j^l + \epsilon_j^l\} = G(e_i^l - e_j^l)$. Assume that in the equilibrium, contestant j makes effort e^* , then the expected payoff to contestant i is $\mathbb{E}(u_i^l(e_i^l)) = A^l G(e_i^l - e_j^l) - C(e_i^l) = A^l G(e_i^l - e^{l*}) - C^l(e_i^l)$. Then the FOC is given by $A^l g(e_i^{l*} - e^{l*}) = C^{l'}(e_i^{l*})$. Since contestant i and j are homogeneous, in the symmetric equilibrium, contestant i makes the same effort e^{l*} as contestant j , thus $A^l g(0) = C^l(e^{l*})$. Since the cost function $C^{l''}(\cdot) > 0$, the equilibrium effort is $e^{l*} = C^{l''^{-1}}(A^l g(0))$. ■

Proof of Lemma 2.8. For the sub-contest l , $l = 1, 2$, the expected payoff to contestant i is $\mathbb{E}(u_i(e_i^l)) = A^l \int_{-\infty}^{+\infty} \Psi(e_i^l - e^{l*}(n) + \epsilon^l)^{n-1} \psi(\epsilon^l) d\epsilon^l - C^l(e_i^l)$. The FOC yields $A^l \int_{-\infty}^{+\infty} (n-1) \Psi(e_i^l - e^{l*}(n) + \epsilon^l)^{n-2} \psi(\epsilon^l - e^{l*}(n) + \epsilon^l) \psi(\epsilon^l) d\epsilon^l = C^{l'}(e_i^l)$. In the symmetric equilibrium, contestant i makes the same effort $e^{l*}(n)$ as other contestants. Since $C^{l''}(\cdot) > 0$, the equilibrium effort is given by $e^{l*}(n) = C^{l''^{-1}}(A^l \int_{-\infty}^{+\infty} (n-1) \Psi(\epsilon^l)^{n-2} \psi(\epsilon^l)^2 d\epsilon^l) = C^{l''^{-1}}(A^l h(\epsilon^l; n)) = C^{l''^{-1}}(A^l h(\epsilon; n))$.

Consider the simultaneous contest with n contestants. Given a fixed aggregate effort level e° , all the contestants follow the optimal effort allocation, $e^\circ = \tilde{e}^1 + \tilde{e}^2$. The strategy of a contestant is his aggregate effort level, e° . By Lemma 2.2, $C^\circ(e^\circ)$ is strictly increasing and strictly convex. Then the expected payoff to contestant i is $\mathbb{E}(u_i(e_i^\circ)) = A \int_{-\infty}^{+\infty} \Psi(e_i^\circ - e^{\circ*}(n) + \epsilon^\circ)^{n-1} \psi^\circ(\epsilon^\circ) d\epsilon^\circ - C^\circ(e_i^\circ)$. Similar to the derivation in the sub-contest 1, the equilibrium effort in the simultaneous contest is given by $e^{\circ*}(n) = C^{\circ''^{-1}}(A \int_{-\infty}^{+\infty} (n-1) \Psi(\epsilon^\circ)^{n-2} \psi^\circ(\epsilon^\circ)^2 d\epsilon^\circ) = C^{\circ''^{-1}}(A h^\circ(\epsilon^\circ; n))$. ■

Lemma 2.16 *If PDF $\psi(\epsilon)$ is twice continuously differentiable and log-concave, then it is unimodal.*

Proof of Lemma 2.16. By the definition of twice differentiable and log-concave function, we have $\frac{\partial^2 \ln[\psi(\epsilon)]}{\partial \epsilon^2} = \frac{\partial}{\partial \epsilon} \left[\frac{\psi'(\epsilon)}{\psi(\epsilon)} \right] \leq 0$. Thus, for any $\epsilon_1 \leq \epsilon_2$, $\frac{\psi'(\epsilon_1)}{\psi(\epsilon_1)} \geq \frac{\psi'(\epsilon_2)}{\psi(\epsilon_2)}$, and equivalently,

$$\frac{\psi'(\epsilon_1)\psi(\epsilon_2) - \psi(\epsilon_1)\psi'(\epsilon_2)}{\psi(\epsilon_1)\psi(\epsilon_2)} \geq 0. \quad (2.17)$$

First, consider the case that there exists an ϵ^* such that $\psi'(\epsilon^*) = 0$. In (2.17), let $\epsilon_2 = \epsilon^*$, then (2.17) implies that $\psi'(\epsilon_1)\psi(\epsilon^*) \geq 0$. Since $\psi(\epsilon^*) \geq 0$, $\psi'(\epsilon_1) \geq 0$ for $\epsilon_1 \leq \epsilon^*$. Similarly, in (2.17), let $\epsilon_1 = \epsilon^*$, then $-\psi'(\epsilon^*)\psi(\epsilon_2) \geq 0$. Thus, $\psi'(\epsilon_2) \leq 0$ for $\epsilon_2 \geq \epsilon^*$. Hence, if ϵ^* exists, PDF $\psi(\epsilon)$ is increasing for $\epsilon \leq \epsilon^*$ and decreasing for $\epsilon \geq \epsilon^*$. Second, if ϵ^* does not exist, because $\psi(\epsilon)$ is twice differentiable, $\psi(\epsilon)$ is either monotone increasing or decreasing. Thus, $\psi(\epsilon)$ is unimodal. ■

Proof of Proposition 2.13. Index two contestants by i and j .

(i) In the sub-contest l , $l = 1, 2$, of the sequential contest, if contestant j with type H or L makes effort e_H^{l*} or e_L^{l*} in the equilibrium, respectively, the winning probability of contestant i is $\eta^l G(e_i^l - e_L^{l*}) + (1 - \eta^l) G(e_i^l - e_H^{l*})$. The expected payoff to contestant i is $\mathbb{E}(u_i^l(e_i^l | x_i)) = A^l [\eta^l G(e_i^l - e_L^{l*}) + (1 - \eta^l) G(e_i^l - e_H^{l*})] - C(e_i^l) / x_i$. The FOC yields $A^l [\eta^l g(e_i^{l*} - e_L^{l*}) + (1 - \eta^l) g(e_i^{l*} - e_H^{l*})] = C'(e_i^{l*}) / x_i$. In the symmetric equilibrium, contestant i makes effort e_L^{l*} if he is low type and e_H^{l*} if he is high type, which lead to

$$A^l [\eta^l g(0) + (1 - \eta^l) g(e_L^{l*} - e_H^{l*})] = C'(e_L^{l*}) / x_L, \quad (2.18)$$

$$A^l [\eta^l g(e_H^{l*} - e_L^{l*}) + (1 - \eta^l) g(0)] = C'(e_H^{l*}) / x_H. \quad (2.19)$$

Now we prove that there exists an equilibrium such that $e_L^{l*} \leq e_H^{l*}$. For notation simplicity, we suppress the superscript l in the proof of $e_L^{l*} \leq e_H^{l*}$. Letting $e_H^* - e_L^* = \delta_{H-L}$, we want to show that there exists a $\delta_{H-L} \geq 0$. Divide (2.18) by (2.19),

$$\frac{\eta g(0) + (1 - \eta) g(-\delta_{H-L})}{\eta g(\delta_{H-L}) + (1 - \eta) g(0)} = \frac{x_H C'(e_L^*)}{x_L C'(e_H^*)}. \quad (2.20)$$

By the symmetric assumption of $g(\cdot)$, we have $g(\delta_{H-L}) = g(-\delta_{H-L})$. Then (2.20) becomes

$$\frac{\eta g(0) + (1 - \eta) g(\delta_{H-L})}{\eta g(\delta_{H-L}) + (1 - \eta) g(0)} - \frac{x_H C'(e_L^*)}{x_L C'(e_H^*)} = 0. \quad (2.21)$$

By $e_H^* = e_L^* + \delta_{H-L}$, the left hand side (LHS) of (2.21) is

$$\text{LHS of (2.21)} = \frac{\eta g(0) + (1 - \eta)g(\delta_{H-L})}{\eta g(\delta_{H-L}) + (1 - \eta)g(0)} - \frac{x_H C'(e_L^*)}{x_L C'(e_L^* + \delta_{H-L})}, \quad (2.22)$$

If $\delta_{H-L} = 0$, (2.22) = $1 - \frac{x_H}{x_L} \leq 0$. If $\delta_{H-L} \rightarrow \infty$, then (2.22) $\rightarrow \frac{\eta}{1-\eta} > 0$ because $\lim_{\delta_{H-L} \rightarrow \infty} g(\delta_{H-L}) = 0$. Since (2.22) is continuous in δ_{H-L} , there exists an intersection point $\delta_{H-L} \geq 0$ such that (2.21) is satisfied. As a result, there exists an equilibrium such that $e_H^* \geq e_L^*$.

(ii) We firstly derive the equilibrium effort in the simultaneous contest. Since contestants are endowed with expertise (x_H, x_L) or (x_L, x_H) . Both types of contestants have the cost function as $C(\cdot)/x_H + C(\cdot)/x_L$. Thus, the equilibrium effort for both contestants must be the same. If contestant i has expertise x_L in the first attribute and x_H in the second attribute, given the aggregate effort e_i° , there exists an optimal allocation of efforts $e_i^\circ = \tilde{e}_i^1 + \tilde{e}_i^2$. By Lemma 2.2, the optimal allocation of efforts satisfies $C'(\tilde{e}_i^1)/x_L = C'(\tilde{e}_i^2)/x_H$. Then,

$$C'(\tilde{e}_i^1)/C'(\tilde{e}_i^2) = x_L/x_H. \quad (2.23)$$

When $C'(e_i) = \rho \exp(\rho e_i)$, (2.23) becomes $\exp(\rho(\tilde{e}_i^1 - \tilde{e}_i^2)) = x_L/x_H$, equivalently $\tilde{e}_i^1 - \tilde{e}_i^2 = \ln(x_L/x_H)/\rho$. Since $e_i^\circ = \tilde{e}_i^1 + \tilde{e}_i^2$, we have $\tilde{e}_i^1 = [e_i^\circ + \ln(x_L/x_H)/\rho]/2$ and $\tilde{e}_i^2 = [e_i^\circ - \ln(x_L/x_H)/\rho]/2$. The total cost is $C^\circ(e_i^\circ) = C(\tilde{e}_i^1)/x_L + C(\tilde{e}_i^2)/x_H$. Then, the derivative of the total cost function is

$$\begin{aligned} C^{\circ'}(e_i^\circ) &= \frac{\rho}{2} \exp\left(\frac{\rho e_i^\circ + \ln(x_L/x_H)}{2}\right) / x_L + \frac{\rho}{2} \exp\left(\frac{\rho e_i^\circ - \ln(x_L/x_H)}{2}\right) / x_H \\ &= \rho \exp(\rho e_i^\circ / 2) / \sqrt{x_H x_L}. \end{aligned} \quad (2.24)$$

If contestant j makes equilibrium effort e^{o*} , then contestant i 's expected payoff function can be written as $\mathbb{E}(u_i^{sim}(e_i^\circ | x_L, x_H)) = Ag^\circ(e_i^\circ - e^{o*}) - C^\circ(e_i^\circ)$. In the equilibrium, contestant i makes the effort e^{o*} , so the FOC yields $Ag^\circ(0) = C^{\circ'}(e^{o*})$. By (2.24), we obtain $\rho \exp(\rho e_i^{o*} / 2) / \sqrt{x_H x_L} = Ag^\circ(0)$, thus

$$e^{o*} = C^{\circ'-1}(Ag^\circ(0)) = 2 \ln(Ag^\circ(0) \sqrt{x_H x_L} / \rho) / \rho. \quad (2.25)$$

Now we compare the expected best performances between the simultaneous and sequential contests. When $\eta^l = 1/2$, the LHS of (2.18) and (2.19) are equal. Since $C(e) = \exp(\rho e)$, $C'(e) = \rho \exp(\rho e)$. By (2.20) and $\eta^l = 1/2$, in sub-contest l ,

$l = 1, 2$, we have $\frac{x_H}{x_L} = \exp(\rho(e_H^{l*} - e_L^{l*}))$, equivalently, $e_H^{l*} - e_L^{l*} = \ln(x_H/x_L)/\rho$. Thus, if $w = 1/2$, and then the equilibrium efforts in sub-contest l of the sequential contest are $e_i^{l*} = \ln\left(\frac{Ax_i}{4\rho}(g(0) + g(\ln(x_H/x_L)/\rho))\right)/\rho$, where $i = H, L$.

Since $x_H \geq x_L$, $e_H^{l*} \geq e_L^{l*}$. Then the performances of the low-type and high-type contestants are $V_L^{l*} = e_L^{l*} + \epsilon^l$ and $V_H^{l*} = e_H^{l*} + \epsilon^l$. Recall the notation for the highest order statistic with sample size 2 is $(\cdot)_{(2)}$. Here, for example, if both contestants are low-type, then the expected best performance is $\mathbb{E}(\max\{V_L^{l*}, V_L^{l*}\}) = \mathbb{E}(V_{L(2)}^{l*})$. Since $\eta^l = 1/2$, the expected best performance in the sub-contest l is given by $V^l = \eta^{l2}\mathbb{E}(V_{L(2)}^{l*}) + 2\eta^l(1 - \eta^l)\mathbb{E}(\max\{V_L^{l*}, V_H^{l*}\}) + (1 - \eta^l)^2\mathbb{E}(V_{H(2)}^{l*}) = \mathbb{E}(V_{L(2)}^{l*})/4 + \mathbb{E}(\max\{V_L^{l*}, V_H^{l*}\})/2 + \mathbb{E}(V_{H(2)}^{l*})/4$. Since $e_H^{l*} \geq e_L^{l*}$, $V_H^{l*} \geq_{st} V_L^{l*}$, we have $\mathbb{E}(V_{H(2)}^{l*}) \geq \mathbb{E}(\max\{V_L^{l*}, V_H^{l*}\}) \geq \mathbb{E}(V_{L(2)}^{l*})$. Thus, we can characterize the lower bound and upper bound of V^l as:

$$\mathbb{E}(V_{L(2)}^{l*}) \leq V^l \leq \mathbb{E}(V_{H(2)}^{l*}). \quad (2.26)$$

Then the upper bound can be written as $\mathbb{E}(V_{H(2)}^{l*}) = \mathbb{E}((e_H^{l*} + \epsilon^l)_{(2)}) = e_H^{l*} + \mathbb{E}(\epsilon_{(2)}^l)$. Thus, we have the upper bound of the expected best performance in the sequential contest, $V^{seq} = V^1 + V^2 \leq e_H^{1*} + e_H^{2*} + \mathbb{E}(\epsilon_{(2)}^1) + \mathbb{E}(\epsilon_{(2)}^2)$. Since both sub-contests are identical, by denoting $\mathbb{E}(\epsilon_{(2)}) = \mathbb{E}(\epsilon_{(2)}^1) = \mathbb{E}(\epsilon_{(2)}^2)$ and $e_H^* = e_H^{1*} = e_H^{2*}$, we have $V^{seq} \leq 2e_H^* + 2\mathbb{E}(\epsilon_{(2)})$.⁵

The expected best performance in the simultaneous contest is $V^{sim} = \mathbb{E}((e^{\circ*} + \epsilon^{\circ})_{(2)}) = e^{\circ*} + \mathbb{E}(\epsilon_{(2)}^{\circ})$. The difference of the expected best performances between those two contest mechanisms is

$$\Delta = V^{seq} - V^{sim} \leq 2e_H^* + 2\mathbb{E}(\epsilon_{(2)}) - e^{\circ*} - \mathbb{E}(\epsilon_{(2)}^{\circ}). \quad (2.27)$$

Now we show that there exists a threshold $\underline{\rho}' > 0$ such that when $\rho \leq \underline{\rho}'$, $\Delta \leq 0$. If ϵ_i and ϵ_j follow a symmetric log-concave distribution, $\xi = \epsilon_i - \epsilon_j$ follows a symmetric log-concave distribution, $g(\xi)$, because the convolution of log-concave functions is log-concave, see Marshall et al. (2016), p.p.763. By Lemma 2.16, $g(\xi)$ is symmetric at 0 and unimodal, then it achieves the maximum value at 0. Denote $\ln(x_H/x_L) = r \geq 0$, then

$$2e_H^* - e^{\circ*} = \frac{2}{\rho} \ln\left(\frac{g(0) + g(\ln(x_H/x_L)/\rho)}{4g^{\circ}(0)} \sqrt{\frac{x_H}{x_L}}\right) = \frac{2}{\rho} \ln\left(\frac{g(0) + g(r/\rho)}{4g^{\circ}(0)}\right) + \frac{r}{\rho}$$

⁵We assume that the expertise of the high and low types in both attributes is the same. Alternatively, we could allow high or low types in both attributes to be different, with their sum fixed. Allowing the general expertise along the two dimensions only changes the upper bound or lower bound of V^{seq} in the proof, and thus the range of x_H/x_L in Proposition 2.13(ii) may change. It does not qualitatively alter our results that when the high and low expertise in each attribute is close enough, there exist $\bar{\rho}'$ and $\underline{\rho}'$ as thresholds on ρ such that if $\rho \geq \bar{\rho}'$, the sequential contest is optimal, and if $\rho \leq \underline{\rho}'$, the simultaneous contest is optimal. A similar remark applies to Proposition 2.14(ii).

$$\leq \frac{2}{\rho} \ln \left(\frac{g(0) + g(0)}{4g^\circ(0)} \right) + \frac{r}{\rho} = \frac{1}{\rho} \left[2 \ln \left(\frac{g(0)}{2g^\circ(0)} \right) + r \right], \quad (2.28)$$

where the inequality is because $g(\cdot)$ achieves its maximum value at 0. By $\ln(x_H/x_L) = r$, $x_H/x_L \in [1, (2g^\circ(0)/g(0))^2]$, is equivalent to $r \in [0, -2 \ln \left(\frac{g(0)}{2g^\circ(0)} \right)]$. Since $2g^\circ(0) > g(0)$, if $r \in [0, -2 \ln \left(\frac{g(0)}{2g^\circ(0)} \right)]$, $\frac{1}{\rho} \left[2 \ln \left(\frac{g(0)}{2g^\circ(0)} \right) + r \right] < 0$. Thus, $\frac{1}{\rho} \left[2 \ln \left(\frac{g(0)}{2g^\circ(0)} \right) + r \right]$ is increasing in ρ , and

$$\lim_{\rho \rightarrow 0^+} \frac{1}{\rho} \left[2 \ln \left(\frac{g(0)}{2g^\circ(0)} \right) + r \right] = -\infty. \quad (2.29)$$

By (2.27) and (2.28), $\Delta = V^{seq} - V^{sim} \leq 2e_H^* - e^{o*} + 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ) \leq \frac{1}{\rho} \left[2 \ln \left(\frac{g(0)}{2g^\circ(0)} \right) + r \right] + 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ)$. By Lemma 2.3, $2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ) > 0$. Thus, by (2.29), there exists $\underline{\rho}' > 0$ such that when $\rho \leq \underline{\rho}'$, $\Delta \leq 2e_H^* - e^{o*} + 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ) \leq 0$.

For the lower bound of V^{seq} , by (2.26), $V^{seq} = V^1 + V^2 \geq \mathbb{E}(V_{L(2)}^{1*}) + \mathbb{E}(V_{L(2)}^{2*})$. This lower bound can be written as $\mathbb{E}(V_{L(2)}^{1*}) + \mathbb{E}(V_{L(2)}^{2*}) = \mathbb{E}((e_L^{1*} + \epsilon^1)_{(2)}) + \mathbb{E}((e_L^{2*} + \epsilon^2)_{(2)}) = e_L^{1*} + \mathbb{E}(\epsilon_{(2)}^1) + e_L^{2*} + \mathbb{E}(\epsilon_{(2)}^2)$. Since two sub-contests are identical, denote $e_L^* = e_L^{1*} = e_L^{2*}$ and $\mathbb{E}(\epsilon_{(2)}) = \mathbb{E}(\epsilon_{(2)}^1) = \mathbb{E}(\epsilon_{(2)}^2)$. Then, $V^{seq} \geq 2e_L^* + \mathbb{E}(\epsilon_{(2)})$, and we have the difference between the two contest mechanisms

$$\Delta = V^{seq} - V^{sim} \geq 2e_L^* + 2\mathbb{E}(\epsilon_{(2)}) - e^{o*} - \mathbb{E}(\epsilon_{(2)}^\circ). \quad (2.30)$$

Now, we show that there exists $\bar{\rho}' > 0$ such that when $\rho \geq \bar{\rho}'$, $\Delta \geq 0$. Similar to the previous analysis, we have

$$\begin{aligned} 2e_L^* - e^{o*} &= \frac{2}{\rho} \ln \left(\frac{g(0) + g(\ln(x_H/x_L)/\rho)}{4g^\circ(0)} \sqrt{\frac{x_L}{x_H}} \right) - \frac{r}{\rho} \\ &\geq \frac{2}{\rho} \ln \left(\frac{g(0)}{4g^\circ(0)} \right) - \frac{r}{\rho} = \frac{1}{\rho} \left[2 \ln \left(\frac{g(0)}{4g^\circ(0)} \right) - r \right]. \end{aligned} \quad (2.31)$$

Because $x_H \geq x_L$, $r \geq 0$, and $2g^\circ(0) > g(0)$, $\frac{1}{\rho} \left[2 \ln \left(\frac{g(0)}{4g^\circ(0)} \right) - r \right] < 0$. Then, $\frac{1}{\rho} \left[2 \ln \left(\frac{g(0)}{4g^\circ(0)} \right) - r \right]$ is increasing in ρ , and

$$\lim_{\rho \rightarrow \infty} \frac{1}{\rho} \left[2 \ln \left(\frac{g(0)}{4g^\circ(0)} \right) - r \right] = 0. \quad (2.32)$$

By (2.30) and (2.31), $\Delta = V^{seq} - V^{sim} \geq 2e_L^* - e^{o*} + 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ) \geq \frac{1}{\rho} \left[2 \ln \left(\frac{g(0)}{4g^\circ(0)} \right) - r \right] + 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ)$. By Lemma 2.3, $2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ) > 0$. By (2.32), there exists $\bar{\rho}' > 0$ such that when $\rho \geq \bar{\rho}'$, $\Delta \geq 2e_L^* - e^{o*} + 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ) \geq 0$. ■

Proof of Proposition 2.14. Index the two contestants by i and j .

(i) In the sub-contest l , $l = 1, 2$, the winning probability of contestant i can be written as $P\{x_i + e_i^l + \epsilon_i^l \geq x_j + e_j^l + \epsilon_j^l\} = G(x_i + e_i^l - x_j - e_j^l)$. If contestant j with type H or L makes efforts e_L^* or e_H^* in the equilibrium, the expected payoff to the contestant i is $\mathbb{E}(u_i^l(e_i^l|x_i)) = A^l \eta^l G(x_i + e_i^l - x_L - e_L^*) + A^l (1 - \eta^l) G(x_i + e_i^l - x_H - e_H^*) - C(e_i^l)$. Because u_i is continuously differentiable, the FOC with respect to effort is given by

$$A^l [\eta^l g(x_i + e_i^{l*} - x_L - e_L^*) + (1 - \eta^l) g(x_i + e_i^{l*} - x_H - e_H^*)] = C'(e_i^{l*}). \quad (2.33)$$

For both types of contestants, in the symmetric equilibrium, the strategies e_H^* and e_L^* must satisfy

$$A^l [\eta^l g(0) + (1 - \eta^l) g(x_L + e_L^* - x_H - e_H^*)] = C'(e_L^*), \quad (2.34)$$

$$A^l [\eta^l g(x_H + e_H^* - x_L - e_L^*) + (1 - \eta^l) g(0)] = C'(e_H^*), \quad (2.35)$$

which can be obtained by setting $i = L$ or H in (2.33). To simplify the notation, we suppress the superscript l in the rest proof of (i).

Recall that $g(\cdot)$ is symmetric at 0, thus

$$g(x_L + e_L^* - x_H - e_H^*) = g(x_H + e_H^* - x_L - e_L^*). \quad (2.36)$$

And, by Lemma 2.16, $g(\cdot)$ is unimodal, then

$$g(x_L + e_L^* - x_H - e_H^*) \leq g(0), \quad (2.37)$$

because the symmetric unimodal function $g(\xi)$ is maximized at $\xi = 0$.

For (ii)(a), if $\eta < 1/2$, the LHSs of (2.34) and (2.35) satisfy

$$\eta g(0) + (1 - \eta) g(x_L + e_L^* - x_H - e_H^*) \leq \eta g(x_H + e_H^* - x_L - e_L^*) + (1 - \eta) g(0),$$

where the inequality is due to (2.37). This leads to $C'(e_H^*) \geq C'(e_L^*)$. By the assumption $C''(\cdot) > 0$, we must have $e_H^* \geq e_L^*$. Because of the assumption $x_H \geq x_L$, $\mathbb{E}(V_H^*) = x_H + e_H^* \geq x_L + e_L^* = \mathbb{E}(V_L^*)$.

For (ii)(b), if $\eta > 1/2$, we can show that the equilibrium efforts $e_H^* \leq e_L^*$ similar to (ii)(a). Now we show that there exists an equilibrium such that $\mathbb{E}(V_H^*) \geq \mathbb{E}(V_L^*)$ if

$\eta > 1/2$. Divide (2.34) by (2.35),

$$\frac{\eta g(0) + (1 - \eta)g(x_L + e_L^* - x_H - e_H^*)}{(1 - \eta)g(0) + \eta g(x_H + e_H^* - x_L - e_L^*)} - \frac{C'(e_L^*)}{C'(e_H^*)} = 0. \quad (2.38)$$

Denote $\delta_e = e_H^* - e_L^* \leq 0$ and $\delta = x_H - x_L \geq 0$. (2.38) can be written as

$$\frac{\eta g(0) + (1 - \eta)g(-\delta - \delta_e)}{(1 - \eta)g(0) + \eta g(\delta + \delta_e)} - \frac{C'(e_L^*)}{C'(e_L^* + \delta_e)} = 0. \quad (2.39)$$

Since $g(\cdot)$ is symmetric at 0, $g(-\delta - \delta_e) = g(\delta + \delta_e)$, and by (2.39) we have

$$\frac{\eta g(0) + (1 - \eta)g(\delta + \delta_e)}{(1 - \eta)g(0) + \eta g(\delta + \delta_e)} - \frac{C'(e_L^*)}{C'(e_L^* + \delta_e)} = 0. \quad (2.40)$$

If $\delta_e = -\delta$, LHS of (2.40) = $1 - \frac{C'(e_L^*)}{C'(e_L^* - \delta)} \leq 0$ because $C''(\cdot) > 0$. If $\delta_e = 0$, we have

$$\frac{\eta g(0) + (1 - \eta)g(-\delta)}{(1 - \eta)g(0) + \eta g(\delta)} - \frac{C'(e_L^*)}{C'(e_L^*)} = \frac{\eta g(0) + (1 - \eta)g(\delta)}{(1 - \eta)g(0) + \eta g(\delta)} - 1 \geq 0,$$

where the second equality is due to the symmetric property of $g(\cdot)$, and the inequality is due to $\eta g(0) + (1 - \eta)g(\delta) \geq (1 - \eta)g(0) + \eta g(\delta)$ because $\delta \geq 0$, $\eta > 1/2$ and $g(\xi)$ is decreasing in $\xi \in [0, \infty)$ since $g(\xi)$ is unimodal and symmetric at 0. Because the LHS of (2.39) is continuous, there exists a point δ_e^* in $[-\delta, 0]$ such that (2.38) is satisfied. It means that $e_H^* - e_L^* \geq -\delta$, i.e., $x_H + e_H^* \geq x_L + e_L^*$. Therefore, there exists an equilibrium such that $\mathbb{E}(V_H^*) = x_H + e_H^* \geq x_L + e_L^* = \mathbb{E}(V_L^*)$.

For (i)(c), if $\eta = 1/2$, again because $g(\cdot)$ is symmetric, by (2.36), the LHSs of (2.34) and (2.35) are equal. Hence, the RHSs of (2.34) and (2.35) satisfy $C'(e_H^*) = C'(e_L^*)$. By the assumptions $C''(\cdot) > 0$, we have $e_H^* = e_L^*$. Thus, by $C'(e_H^*) = C'(e_L^*) = \frac{A^l}{2}(g(0) + g(x_H - x_L))$, we obtain

$$e_H^{l*} = e_L^{l*} = C'^{-1} \left(\frac{A^l}{2}(g(0) + g(x_L - x_H)) \right). \quad (2.41)$$

To compare the expected performances, if $\eta = \frac{1}{2}$, $\mathbb{E}(V_H^*) = x_H + e_H^* \geq x_L + e_L^* = \mathbb{E}(V_L^*)$ because $e_H^* = e_L^*$ that was just proved and $x_H \geq x_L$ by the assumption.

(ii) We firstly derive the equilibrium effort in the simultaneous contest. For the simultaneous contest, both contestants are endowed with x_H and x_L , then the winning probability of contestant i is $P\{x_H + x_L + e_i^o + \epsilon_i^1 + \epsilon_i^2 \geq x_H + x_L + e_j^o + \epsilon_j^1 + \epsilon_j^2\} = G^o(e_i^o - e_j^o)$. If contestant j adopts the equilibrium strategy e^{o*} , then the expected payoff

to contestant i is $\mathbb{E}(u_i^{sim}(e_i^\circ)) = AG^\circ(e_i^\circ - e^{o*}) - C^\circ(e_i^\circ)$. By (2.1), the equilibrium effort is $e^{o*} = C^{\circ^{-1}}(Ag^\circ(0))$.

Now we compare the two contest mechanisms with $\eta^l = 1/2$, $l = 1, 2$. If $\eta^l = 1/2$, by (2.41), we have $V_i^{l*} = x_i + C'^{-1}\left(\frac{A}{2}(g(0) + g(x_L - x_H))/2\right) + \epsilon^l$, where $i = H$ or L . If $\eta^l = 1/2$, the expected best performance in the sub-contest l , $l = 1, 2$, is $V^l = \mathbb{E}(V_{L(2)}^{l*})/4 + \mathbb{E}(\max\{V_H^{l*}, V_L^{l*}\})/2 + \mathbb{E}(V_{H(2)}^{l*})/4$. Since $\mathbb{E}(V_H^*) \geq \mathbb{E}(V_L^*)$, $V_H^{l*} \geq_{st} V_L^{l*}$, we have $\mathbb{E}(V_{H(2)}^{l*}) \geq V^l \geq \mathbb{E}(V_{L(2)}^{l*})$. Because sub-contest 1 and 2 are identical, by denoting $V_{H(2)}^* = V_{H(2)}^{1*} = V_{H(2)}^{2*}$, $V_{L(2)}^* = V_{L(2)}^{1*} = V_{L(2)}^{2*}$, and $\mathbb{E}(\epsilon_{(2)}) = \mathbb{E}(\epsilon_{(2)}^1) = \mathbb{E}(\epsilon_{(2)}^2)$, we have $2\mathbb{E}(V_{H(2)}^*) \geq V^{seq} \geq 2\mathbb{E}(V_{L(2)}^*)$. Equivalently,

$$V^{seq} \leq 2x_H + 2C'^{-1}\left(\frac{A}{4}(g(0) + g(x_L - x_H))\right) + \mathbb{E}(\epsilon_{(2)}), \quad (2.42)$$

$$V^{seq} \geq 2x_L + 2C'^{-1}\left(\frac{A}{4}(g(0) + g(x_L - x_H))\right) + \mathbb{E}(\epsilon_{(2)}). \quad (2.43)$$

For the simultaneous contest, $e^{o*} = 2C'^{-1}(Ag^\circ(0))$, so the expected best performance in the simultaneous contest is $V^{sim} = x_H + x_L + 2C'^{-1}(Ag^\circ(0)) + \mathbb{E}(\epsilon_{(2)}^\circ)$ where $\epsilon^\circ = \epsilon^1 + \epsilon^2$. By (2.42),

$$\begin{aligned} V^{seq} - V^{sim} &\leq x_H - x_L + \frac{2}{\rho} \ln\left(\frac{g(0) + g(x_L - x_H)}{4g^\circ(0)}\right) + 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ) \\ &\leq x_H - x_L + \frac{2}{\rho} \ln\left(\frac{g(0)}{2g^\circ(0)}\right) + 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ), \end{aligned} \quad (2.44)$$

where the second inequality is because $g(\cdot)$ achieves the maximum value at 0. Since $2g^\circ(0) > g(0)$, $\ln\left(\frac{g(0)}{2g^\circ(0)}\right) < 0$, thus $\frac{2}{\rho} \ln\left(\frac{g(0)}{2g^\circ(0)}\right)$ is increasing in $\rho > 0$. Moreover, $\lim_{\rho \rightarrow 0} \frac{2}{\rho} \ln\left(\frac{g(0)}{2g^\circ(0)}\right) = -\infty$. Because $x_H - x_L \geq 0$, and by Lemma 2.3, $2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ) > 0$. By (2.44), there exists $\underline{\rho}'' > 0$ such that when $\rho \leq \underline{\rho}''$, $\Delta = V^{seq} - V^{sim} \leq 0$.

By (2.43), we have

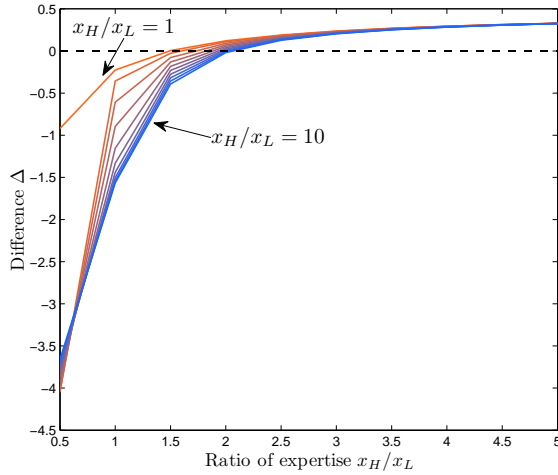
$$\begin{aligned} V^{seq} - V^{sim} &\geq x_L - x_H + \frac{2}{\rho} \ln\left(\frac{g(0) + g(x_L - x_H)}{4g^\circ(0)}\right) + 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ) \\ &\geq x_L - x_H + \frac{2}{\rho} \ln\left(\frac{g(0)}{4g^\circ(0)}\right) + 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ). \end{aligned} \quad (2.45)$$

Again, since $2g^\circ(0) > g(0)$, $\ln\left(\frac{g(0)}{4g^\circ(0)}\right) < 0$, thus $\frac{2}{\rho} \ln\left(\frac{g(0)}{4g^\circ(0)}\right)$ is increasing in ρ , and $\lim_{\rho \rightarrow \infty} \frac{2}{\rho} \ln\left(\frac{g(0)}{4g^\circ(0)}\right) = 0$. By Lemma 2.3, $2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ) > 0$. If $x_H - x_L < 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ)$, $x_L - x_H + 2\mathbb{E}(\epsilon_{(2)}) - \mathbb{E}(\epsilon_{(2)}^\circ) > 0$. By (2.45), there exists $\bar{\rho}'' > 0$, such that when $\rho \geq \bar{\rho}''$, $\Delta = V^{seq} - V^{sim} \geq 0$. ■

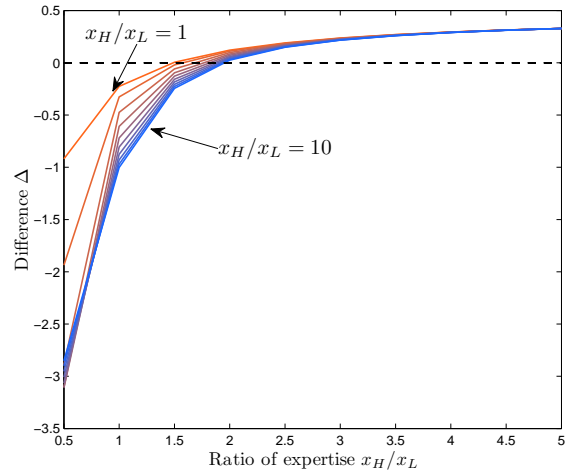
2.10 Numerical Experiments

For the heterogeneous expertise models studied in Section 2.6, if $\eta^l \neq 1/2$, $l = 1, 2$, there is no closed form solution for the equilibrium efforts in the sequential contest. We perform the following numerical tests. In these tests, we restrict our attention to the cases that x_H and x_L are close to each other. In particular, we let $A = 1$, $w = 1/2$, $\epsilon \sim N(0, 1)$, and $\eta^1 \in \{0.2, 0.4, 0.6, 0.8\}$ ($\eta^2 = 1 - \eta^1$). For the heterogeneous cost function model, in each of Figures B.1(a) – (d), different curves correspond to different values of Δ with a given $x_H/x_L \in \{1, 2, \dots, 10\}$. For the heterogeneous starting point model, in each of Figures B.2(a) – (d), different curves correspond to different values of Δ with $x_H - x_L \in \{0, 1, \dots, 10\}$. All the figures show that for $\eta^l \neq 1/2$, if x_H and x_L are sufficiently close, then there exists a threshold $\tilde{\rho}$ such that if $\rho \geq \tilde{\rho}$, the sequential contest dominates, i.e., $\Delta \geq 0$, and if $\rho \leq \tilde{\rho}$, the simultaneous contest dominates, i.e., $\Delta \leq 0$. These observations are consistent with our results in Section 2.4.4 with homogenous expertise.

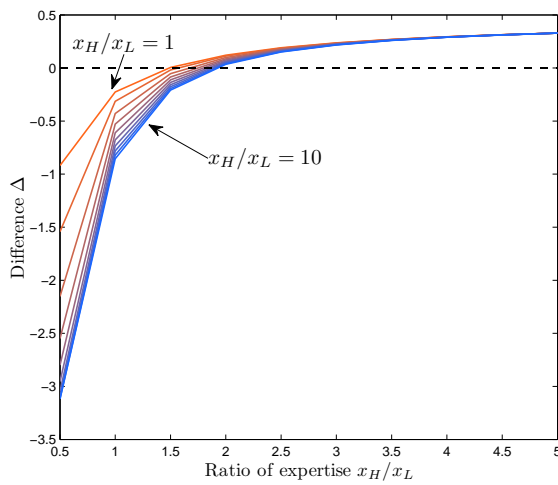
Figure 2.4: Heterogeneous Cost Functions



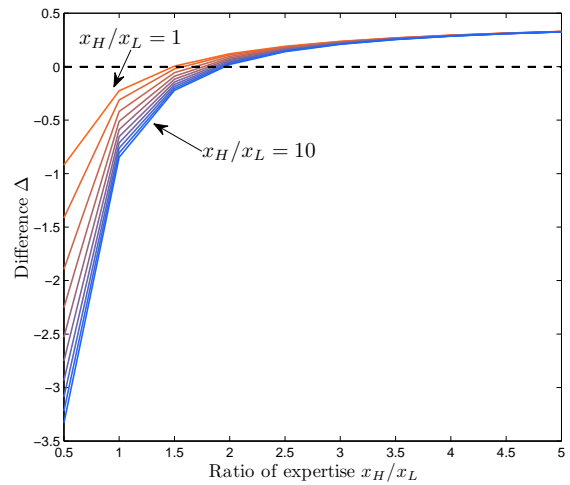
(a) $\eta^1 = 0.2, \eta^2 = 0.8, x_H/x_L \in \{1, 2, \dots, 10\}$



(b) $\eta^1 = 0.4, \eta^2 = 0.6, x_H/x_L \in \{1, 2, \dots, 10\}$

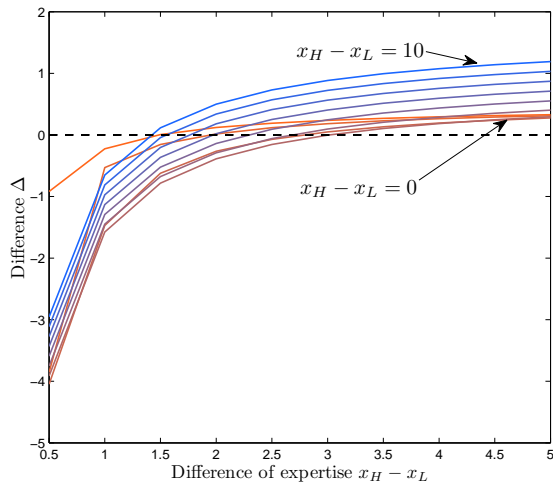


(c) $\eta^1 = 0.6, \eta^2 = 0.4, x_H/x_L \in \{1, 2, \dots, 10\}$

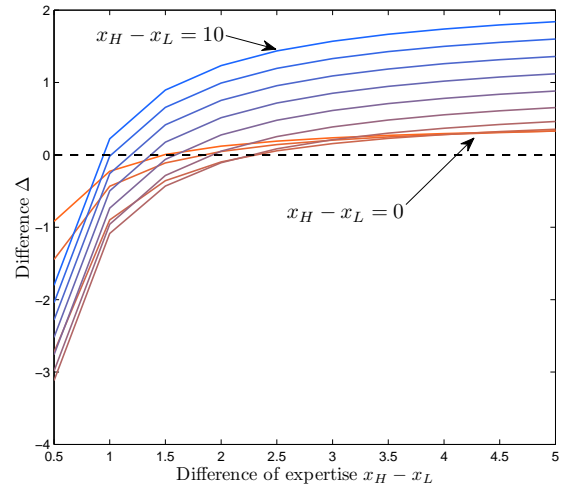


(d) $\eta^1 = 0.8, \eta^2 = 0.2, x_H/x_L \in \{1, 2, \dots, 10\}$

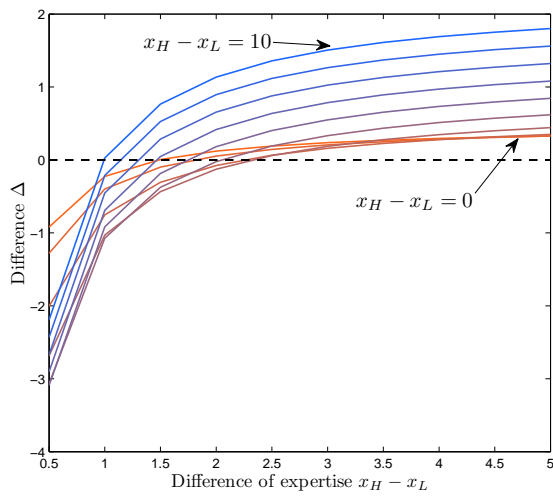
Figure 2.5: Heterogeneous Starting Points



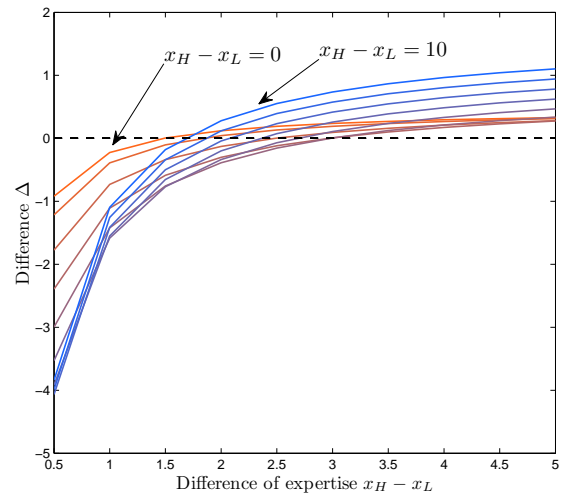
(a) $\eta^1 = 0.2, \eta^2 = 0.8, x_H - x_L \in \{0, 1, \dots, 10\}$



(b) $\eta^1 = 0.4, \eta^2 = 0.6, x_H - x_L \in \{0, 1, \dots, 10\}$



(c) $\eta^1 = 0.6, \eta^2 = 0.4, x_H - x_L \in \{0, 1, \dots, 10\}$



(d) $\eta^1 = 0.8, \eta^2 = 0.2, x_H - x_L \in \{0, 1, \dots, 10\}$

Chapter 3

Bundling with Crowdsourced Products

3.1 Introduction

As consumers buying habit are trending toward simpler and more convenient experience, more and more companies tend to provide subscription services especially in the e-commerce market. Some streaming media providers such as Spotify, Netflix and Apple music allow consumers to access to video or music collections by charging monthly fees. Other service providers such as OneGo offer the option of subscribing flight tickets, thus consumers whose work requires frequently travelling may benefit by the subscription service. As digital products are heterogeneous in their quality and popularity, media companies cannot evenly allocate the collected subscription fees. In some platforms, the total revenue is allocated according to the realized contributions by each crowdsourced product. For example, Spotify allocates its monthly revenue by the proportion of an artist's number of streams among the total number of streams (see Figure 3.1, Spotify's Royalty System¹). In other platforms, the subscription providers do not share the streaming information among the product suppliers, and they might allocate the total revenue according to the expected quality or popularity of different products (see Flint and Fritz (2015)).

In contrast to the traditional manufacturing industry, crowdsourced digital products or services have advantages of being easily spread and electronically stored. Without territory limit, consumers' heterogeneous preferences directly determine the royalties of artists with different popularity. Moreover, since the subscription platform often contains

¹<https://www.spotifyartists.com/spotify-explained/>

Figure 3.1: Spotify's Royalty System



a large number of digital products, the pricing strategy for the subscription fee is determined by the platform owner. Therefore, there is no “double marginalization” problem, but the product suppliers may lose the control of pricing their own products. Finally, the revenue of a supplier for streaming through the subscription service does not depend on the “pay-per-use” system in which the revenue of a supplier does not affect the revenue of another supplier. Under the revenue sharing policies mentioned above, the revenue of a supplier not only depends on the feature of his own product, but also depends on the feature of the products by other suppliers. Evidence shows that some music artists are not willing to join the subscription platform. Though such evidence may be caused by various reasons, one of the most salient ones may be the unequal revenue allocation. As a result, we intend to answer a question: how does the revenue sharing policy in the subscription platform influence the incentive compatibility of different product suppliers? When there is incentive incompatibility in the system, side payments need to be made to sustain the business.

Different subscription providers follow different ways of allocating the revenue among the product suppliers. Some subscription providers, such as Spotify, publish the monthly streams for all the product suppliers, and allocate the total revenue to those suppliers based on the contingent performance of their products. We call this “*contingent revenue sharing policy*.” Other subscription providers, such as Netflix, do not share the streaming information among the product suppliers, they may allocate the revenue according to the expected quality or popularity of the product. We call this “*pre-committed revenue sharing policy*.”

The subscription provider makes the pricing decision. However, the derivation of the optimal pricing strategy of a large number of heterogeneous products may be analytically intractable. We resort to a stylized model with two different products and examine two widely accepted specifications of products. In the first specification, the mean and variance of a consumer’s valuation of a product is higher than that of another product, but

both valuations have the same coefficient of variance (CV). Such specification has been widely used to characterize the high-valuation and low-valuation products, see, Bhargava (2013)². The second one is that those products are differentiated by the levels of the valuation dispersion. The popular product has a low valuation dispersion whereas the niche product has a high valuation dispersion. The valuations of the two products have the same mean value, but the valuation of the popular product has a lower dispersion than that of the niche product (see, e.g., Johnson and Myatt (2006) and Bar-Isaac et al. (2012)).

For the first specification, we find that the high-valuation product supplier may prefer to join the subscription platform while the low-valuation product supplier may prefer the separate sales under both revenue sharing policies. In contrast to the separate sales in which the revenue of each supplier only depends on consumer's valuation of his own product, the revenue sharing policies in the subscription platform introduces the interdependency of the firms' revenues on the consumer's valuations of both products. Thus, the high-valuation product can have an overwhelming proportion of streams if the two products are sufficiently different, i.e., the mean valuations are far apart. Our first result shows that the revenue allocation scheme for subscription service may create incentive incompatibility. Comparing between the contingent and pre-committed revenue sharing policies, we find that the high-valuation product supplier tends to prefer the pre-committed policy while the low-valuation product supplier tends to favor the contingent policy. The high-valuation product has a higher uncertainty of valuation in the market. By the contingent policy, the revenue of the high-valuation product supplier is largely influenced by the uncertainty of consumer's valuation. However, the pre-committed policy to some extent eliminates such uncertainty and the revenue allocation mainly depends on the mean valuation of those two products. Because the total revenue remains the same under both policies, if the high-valuation product supplier obtains a higher revenue under the pre-committed policy than under the contingent policy, then the low-valuation product supplier will earn a lower revenue.

For the second specification, the results are reversed. Under either revenue sharing policy, the popular product supplier tends to prefer the separate sales, while the niche product supplier may prefer to join the subscription platform. In the separate sales, since the valuation of the popular product is more clustered than the niche product, the popular product may cover the whole market but the niche product covers only a fraction of the market. However, in the subscription platform, even with the optimal

²In our model, the valuations of the two products follow uniform distributions $U(0, a_1)$ and $U(0, a_2)$, and $a_1 \leq a_2$.

pricing strategy, the bundling sales can only cover a fraction of the market. Under the contingent or pre-committed policies in the platform, the popular product supplier may earn a less revenue in the subscription service than in the separate sales. Since the total revenue is higher in the subscription service than in the separate sales, the niche product supplier can earn a higher revenue in the subscription platform. Comparing between the contingent and pre-committed revenue sharing policies, we find that the popular product supplier tends to prefer the contingent policy while the niche product supplier may prefer the pre-committed policy. Since the valuation of the niche product has a larger dispersion than the popular product, the small sales number of the niche product may be alleviated when the popular product's sales is also realized as low under the contingent policy. Under the pre-committed policy, the revenue allocation is almost determined by the mean valuation. Therefore, the niche product supplier may earn a higher revenue under the pre-committed policy than under the contingent policy. Meanwhile, the popular product supplier will earn less revenue under the pre-committed policy than under the contingent policy.

Furthermore, we examine an alternative mechanism, the *mixed bundling sales*, by some subscription providers to mitigate the unequal revenue allocation under the contingent revenue sharing policy. Mixed bundling is such that some products are sold separately in addition to the bundle in the subscription platform. For example, Apple Music also allows consumers to purchase single songs other than signing up for a subscription service. Amazon offers kindle unlimited reading service while readers have the option to buy single volume of e-books. In the mixed bundling, the product supplier earns the revenue in the separate sales of his product in addition to a fraction of the revenue in the bundling sales. As mixed bundling further changes the market segmentation, we find that both product suppliers may have incentive to join the subscription platform.

3.2 Literature Review

Product bundling is extensively studied in the marketing, economics, and information system. The earliest work on bundling by Stigler (1963) indicates the advantage of packaging two or more products to capture a larger return. Later Adams and Yellen (1976) introduced a two-dimensional graphical framework for analyzing bundling as a device for price discrimination. They suggest that bundling can increase profits if the valuations of the two goods are negatively correlated. The more formal analysis by Schmalensee (1984) and McAfee et al. (1989) also focused on bundles of two goods. Schmalensee (1984) assumed a bivariate Gaussian distribution of reservation prices, and

found that bundling dominates separate sales if the valuations are independent or slightly positively correlated. Later on, McAfee et al. (1989) derived a set of conditions under which mixed bundling of two goods dominates unbundled sales.

More recently, there are many streams of work over the bundling. One stream of studies focuses on the pricing strategy and comparison of mixed bundling and pure bundling. Eckalbar (2010) provides an analytical expression for optimal mixed bundling under zero marginal costs. While our research focuses on the crowdsourced digital products, the marginal cost is assumed to be zero. Our paper is based on the optimal pricing theory in Eckalbar (2010). Since the pricing strategy for bundling products with positive marginal cost may be intractable for a closed form result, Bhargava (2013) derives a closed-form approximation of the optimal bundling price. Hitt and Chen (2005) allowed consumers to choose a fixed quantity of goods for a fixed price and showed that bundling of low marginal cost goods outperforms separate sales. Bhargava (2012) examine the different distribution structures for the bundling sales. In the structure that a downstream retailer combines component goods sourced from separate manufacturers. The retailer may not prefer the bundling sales because of the channel conflicts.

Another stream of studies investigates asymptotic results when a large number of information goods are being bundled. Hanson and Martin (1990) derive the computational optimal bundling pricing strategy for a large number of products. Bakos and Brynjolfsson (1999) find that bundling very large numbers of unrelated information goods achieves greater profit and greater economic efficiency by assuming that the valuation of the products are uniformly distributed. Later on, Bakos and Brynjolfsson (2000) extend their model to consider the competition between bundling providers. Fang and Norman (2006) examine the optimality of the bundling strategy by considering a large number of products with symmetric log-concave distributed valuations. Geng et al. (2005) examines the optimality of bundling when consumer's values decline with the number of information goods consumed. Prasad et al. (2010) compared the pure bundling and a special case of partial mixed bundling where one product is set apart and other products are bundled.

For the subscription service, Wang et al. (2015) examine the pay-per-use and subscription of the main products and ancillary services. By the analytical characterization of those services, they derive the conditions on which the subscription is optimal. Allon et al. (2011b) study whether the firm should bundle the main service and ancillary service together, or unbundle them and set single price for each. Randhawa and Kumar (2008) employ a Markov chain model to characterize customer behavior and compare the benefit between the subscription option that limits the number of concurrent rentals in return

for a flat fee per-unit time and a pay-per-use option with no such restriction.

As far as we know, the existing literature did not look into the revenue sharing problem of bundling in the digital product market. Our work investigates the welfare of product suppliers when contracting with bundling providers. While we provide analytical expressions of profit allocation schemes, our research provides some explanations for why product suppliers may prefer or not prefer to join the subscription platform. With the result that the revenue sharing policy in the subscription platform creates incentive incompatibility, the platform owner can make side payments to some suppliers to sustain the operations. Thus, our work provides insights for the firm about what to offer to different product suppliers.

3.3 Model Setup

Assume that there are two product suppliers, supplier 1 and 2. The products produced by the two suppliers are denoted by A_1 and A_2 respectively. A potential customer holds the valuation of products A_1 and A_2 as X and Y which are random variables. Random variables X and Y are independently following uniform distributions over $[0, a_1]$ and $[0, a_2]$ respectively, $a_1, a_2 > 0$. The cumulative probability density functions (CDF) of X and Y are denoted by $F(\cdot)$ and $G(\cdot)$ respectively, and the probability density functions (PDF) of them are denoted by $f(\cdot)$ and $g(\cdot)$. Assume that $a_1 \leq a_2$, then X is first order stochastically dominated by Y , i.e., $X \leq_{FSD} Y$, see, Shaked and Shanthikumar (2007). Since the CV is $\sqrt{3}/3$ for X and Y , the valuations of both products have the same variability. In this specification, product A_1 is the low-valuation product and product A_2 is the high-valuation product. Without loss of generality, the market size is assumed to be 1. The assumptions about two products and uniform distributions have been widely accepted in the literature, see, e.g., Bhargava (2013) and Eckalbar (2010). Such specification can be used to characterize digital products that are differentiated by their quality.

In the separate sales, the two suppliers do not join the subscription platform, thus customers can only buy those two products separately. Denote the prices of A_1 and A_2 in the separate sales by p_1 and p_2 , respectively. Since the products are digit-type, we neglect production cost for those products. A potential customer will buy the product A_1 only if his realized value $x \geq p_1$, thus the expected revenue of product A_1 can be characterized by $R_1 = [1 - F(p_1)]p_1$. Similarly, the expected revenue of product A_2 can be characterized by $R_2 = [1 - G(p_2)]p_2$.

In the pure bundling sales, the two suppliers commit to join the subscription platform.

The platform sells those two products together as a bundle and sets a single price p_0 for it. Then, a potential customer will buy the bundle only if his realized valuation of those two products is greater than or equal to the bundle price. Denote the random variable $V = X + Y$ with CDF $H(\cdot)$ and PDF $h(\cdot)$ which is the convolution of the PDFs $f(\cdot)$ and $g(\cdot)$, i.e., $h = f \circ g$. The expected revenue for the pure bundling sales is then $R_u = [1 - H(p_0)]p_0$.

3.4 Revenue Comparison

Lemma 3.1 shows the optimal pricing strategies for the separate sales and pure bundling sales, p_1^* , p_2^* and p_0^* , and the revenue under the optimal pricing strategy, R_1^* , R_2^* and R_u^* .

Lemma 3.1 (PRICING STRATEGY I) (i) *The optimal pricing strategies for the separate sales is $p_1^* = a_1/2$ and $p_2^* = a_2/2$. The revenue under the optimal pricing strategy is $R_1^* = a_1/4$ and $R_2^* = a_2/4$.*

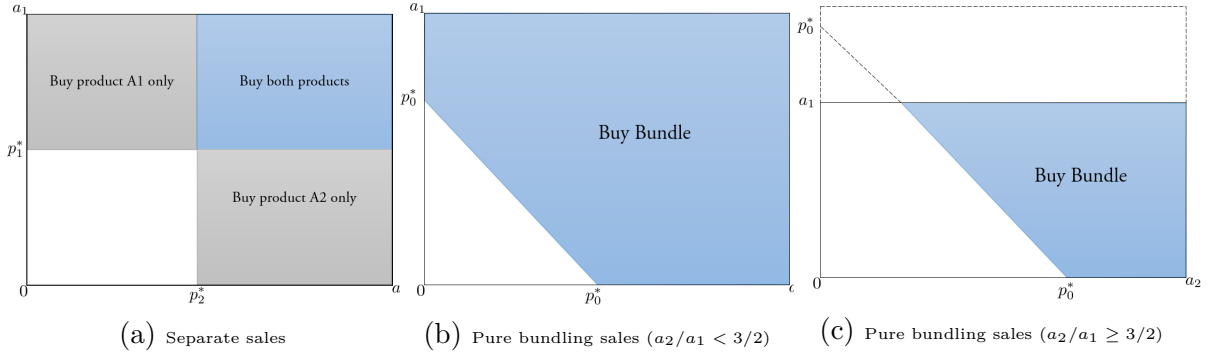
(ii) *The optimal pricing and revenue for the pure bundling sales is $p_0^* = \sqrt{\frac{2a_1a_2}{3}}$ and $R_u^* = \frac{2}{3}\sqrt{\frac{2a_1a_2}{3}}$ if $1 \leq a_2/a_1 < 3/2$, and $p_0^* = \frac{2a_2+a_1}{4}$ and $R_u^* = \frac{(2a_2+a_1)^2}{16a_2}$ if $a_2/a_1 \geq 3/2$.*

Figure 3.2 illustrates the optimal pricing strategies for the separate and pure bundling sales. In Figure 3.2(a), the customer whose valuation of the two products is in the grey area will buy a single product, and the customer whose valuation is in the dark grey area will buy both products. For the pure bundling sales, Figure 3.2(b) and (c) characterize two situations: if $a_2/a_1 < 3/2$, $p_0^* < a_1$, and if $a_2/a_1 \geq 3/2$, $a_1 \leq p_0^* \leq a_2$. For the separate sales, the revenue for each product is shared by its supplier and its distribution company. For the pure bundling sales, the revenue under the optimal pricing strategy is the total revenue that will be shared by the retailer and two suppliers if the retailer commits to provide the subscription service on the platform.

3.4.1 Contingent Revenue Sharing Policy

In this section, we compare the revenue of each supplier between the separate sales and pure bundling sales. If those suppliers commit to join the subscription platform, the revenue of the retailer is a fraction of the total revenue. Denote the fraction by $r \in [0, 1]$. Then the revenue for the retailer is rR_u^* in the pure bundling sales. If those suppliers do not join the platform, they have to pay a distribution fee as a fraction of the revenue

Figure 3.2: Optimal Pricing Strategy I



in the separate sales to their distribution companies. Here, we assume that the fraction of the revenue for the distribution fee is r , i.e., the royalty rates are the same in the separate and pure bundling sales. If the revenues of suppliers 1 and 2 in the separate sales under the optimal pricing strategy are denoted by R_{s1} and R_{s2} respectively, then $R_{s1} = (1 - r)R_1^*$ and $R_{s2} = (1 - r)R_2^*$. Now we model the contingent revenue sharing policy (Spotify’s strategy) in the pure bundling sales.

With the contingent revenue sharing policy, denote the revenue of supplier i , $i = 1, 2$, by R_{ui}^c . The revenue allocation depends on the amount of streams for each product over the total amount of streams. We assume that the amount of streams is a function of the valuation $\mathcal{V}(\cdot)$ which is increasing. That is, the higher the valuation, the more times the customer will enjoy the digital product. For simplicity, we examine the function $\mathcal{V}(x) = x$. Similar result in this section can be obtained with a general function. Denote the fraction of the revenue obtained by supplier 1 by α , thus the fraction of the revenue for supplier 2 is $1 - \alpha$. If the bundle price is p_0^* , the α under the contingent revenue sharing policy is

$$\begin{aligned} \alpha &= \mathbb{E}(\text{Proportion of streams} | \text{customer buys the bundle}) \\ &= \mathbb{E}\left(\frac{X}{X + Y} \mid X + Y \geq p_0^*\right). \end{aligned} \tag{3.1}$$

Then the revenue under the contingent revenue sharing policy for those two suppliers would be $R_{u1}^c = \alpha(1 - r)R_u^*$ and $R_{u2}^c = (1 - \alpha)(1 - r)R_u^*$. Since the optimal pricing strategy p_0^* is a function of the variables a_1 and a_2 , the allocation α only depends on the values of a_1 and a_2 . Proposition 3.2 compares R_{si} and R_{ui}^c , $i = 1, 2$, for different values of a_2/a_1 .

Proposition 3.2 (SEPARATE SALES VS. PURE BUNDLING SALES) *There exists a threshold $k^* > 1$, if $1 \leq a_2/a_1 \leq k^*$, both suppliers will prefer the pure bundling sales, i.e.,*

$R_{ui}^c \geq R_{si}$, $i = 1, 2$. However, if $a_2/a_1 > k^*$, supplier 1 will prefer the separate sales while supplier 2 will prefer the pure bundling sales, i.e., $R_{u1}^c < R_{s1}$ and $R_{u2}^c > R_{s2}$.

Proposition 3.2 shows that if the two products are similar, i.e., $1 \leq a_2/a_1 \leq k^*$, then both suppliers can earn a higher revenue in the pure bundling sales than in the separate sales. If those two products are sufficiently different, i.e., $a_2/a_1 > k^*$, then the high-valuation product supplier will earn more revenue in the pure bundling sales than in the separate sales, but the low-valuation product supplier will earn less revenue in the pure bundling sales than in the separate sales.

If those two products are similar, the revenue of those two suppliers are similar in the separate sales. In the pure bundling sales, since the two products are similar, by the contingent revenue sharing policy, the revenue shares of the two supplier must be similar, i.e., α is approximately equal to $1/2$. Because the pure bundling sales expands the market with a lower total price than the separate sales, the pure bundling sales captures a larger return. Therefore, the total revenue of the pure bundling sales is greater than or equal to that of the separate sales, i.e., $R_u^* \geq R_1^* + R_2^*$. If those suppliers evenly share the total revenue, both can benefit by joining the subscription platform.

If those two products are sufficiently different, the proportion of the revenue share for supplier 1 can be extremely small. However, the benefit of the pure bundling sales in capturing a larger return is bounded.³ In the separate sales, the revenue of each supplier depends only on the valuation of his or her product, however in the pure bundling sales the revenue of each product supplier depends on the conditional expected proportion of the valuations for both products. As a result, if the two products are sufficiently different, the revenue share of supplier 1 can be extremely small under the contingent revenue sharing policy. Therefore, supplier 1 will prefer the separate sales and supplier 2 will prefer the pure bundling sales. The intuition is that in the pure bundling sales, the allocation of revenue depends on the relative valuations of those two products, thus the market share of the niche product is cannibalized by the popular product.

Another explanation is that if the two products are sufficiently different, then the total revenues of the separate sales and pure bundling are similar, because the total revenue mainly depends on the sales of the popular product. In the separate sales, the proportion of the revenue for supplier 1 is $\frac{a_1}{a_1+a_2}$. However, in the pure bundling sales with contingent revenue sharing policy, the proportion of the revenue for supplier 1 is $\mathbb{E}(\frac{X}{X+Y} | X+Y \geq p_0^*)$. By Figure 3.2 (c), we find that if both products are sufficiently different, $p_0^* > a_1$. Since

³The result that the benefit of the pure bundling sales in capturing a larger return is bounded can be simply verified by considering the extreme case that the whole market has been covered with the optimal pricing strategy.

the expectation is conditional on $X + Y \geq p_0^*$, it is high probability that the reason for consumers to buy the bundle is that they favor the popular product. That is to say, the majority of consumers are essentially intended to buy the popular product. As a result, the proportion of revenue for supplier 1 is lower in the pure bundling sales than in the separate sales, i.e., $\mathbb{E}(\frac{X}{X+Y} | X + Y \geq p_0^*) < \frac{a_1}{a_1+a_2}$. Since the total revenues of both sales are similar, supplier 2 earns more revenue in the pure bundling sales than in the separate sales.

3.4.2 Equilibrium

Now we examine the behavior of those two suppliers in the equilibrium. The sequence of events is as follows. In stage 1, the retailer commits to use the pure bundling sales on the subscription platform, and then he selects the royalty rate $r \in [0, 1]$. In stage 2, the two suppliers decide whether to join the platform by considering the revenue in the separate and pure bundling sales. We restrict our attention on the case that the retailer picks the royalty rate r that is the same with the royalty rate of the distribution company. In the non-cooperative game, the action of those two suppliers are “join” or “not join” the platform. Table 3.1 summarizes this game given that the retailer picks the royalty rate r .

Table 3.1: Payoff Matrix

| | supplier 1 does not join | supplier 1 joins |
|--------------------------|--------------------------|------------------------|
| supplier 2 does not join | (R_{s1}, R_{s2}) | (R_{s1}, R_{s2}) |
| supplier 2 joins | (R_{s1}, R_{s2}) | (R_{u1}^c, R_{u2}^c) |

The equilibrium depends on the value of a_2/a_1 . If $a_2/a_1 < k^*$, A strong equilibrium exists such that both suppliers join the platform, and a weak equilibrium exists such that both suppliers do not join the platform. If $a_2/a_1 = k^*$, then a weak equilibrium exists such that both suppliers join or do not join the platform. If $a_2/a_1 > k^*$, then by Proposition 3.2, $R_{u1}^c < R_{s1}$, only the weak equilibrium exists such that supplier 1 does not join the platform, and supplier 2 may or may not join the platform.

In the equilibrium, if the two products are similar, then both suppliers may prefer to join the subscription platform. If the two products are sufficiently different, then the high-valuation product supplier may join or not join the platform but the low-valuation product supplier will not join the platform.

3.5 Pre-committed Revenue Sharing Policy

In this section, we examine the case in which the retailer commits to provide the pre-committed revenue sharing policy (Netflix's strategy) in the pure bundling sales. In contrast to the contingent policy that determines the revenue allocation according to the ex post expectation of the proportion of streams, in the pre-committed policy, the revenue of each supplier relies on the ex ante expectation of the proportion of streams.

Under the pre-committed revenue sharing policy, denote the revenue of supplier i , $i = 1, 2$ in the pure bundling sales by R_{ui}^p . Denote the proportion of the revenue obtained by supplier 1 by β , and then the proportion of the revenue for supplier 2 is $1 - \beta$. Under the optimal bundling pricing strategy p_0^* , the proportion β is

$$\begin{aligned}\beta &= \frac{\mathbb{E}(\text{Streams of product } A_1 | X + Y \geq p_0^*)}{\mathbb{E}(\text{Total streams} | X + Y \geq p_0^*)} \\ &= \frac{\mathbb{E}(X | X + Y \geq p_0^*)}{\mathbb{E}(X | X + Y \geq p_0^*) + \mathbb{E}(Y | X + Y \geq p_0^*)}.\end{aligned}\quad (3.2)$$

The revenue under the pre-committed revenue sharing policy for the two suppliers would be $R_{u1}^p = \beta(1 - r)R_u^*$ and $R_{u2}^p = (1 - \beta)(1 - r)R_u^*$. To compare the revenue of each suppliers in different revenue sharing policies, we have the following results.

- Proposition 3.3 (PRE-COMMITTED REVENUE SHARING)** (i) *There exists a threshold $\bar{k}^* > 1$, if $1 \leq a_2/a_1 \leq \bar{k}^*$, both suppliers will prefer the pure bundling sales, i.e., $R_{ui}^p \geq R_{si}$, $i = 1, 2$. However, if $a_2/a_1 > \bar{k}^*$, supplier 1 will prefer the separate sales while supplier 2 will prefer the pure bundling sales, i.e., $R_{u1}^p < R_{s1}$ and $R_{u2}^p > R_{s2}$.*
- (ii) *The pre-committed revenue sharing policy β is less than the contingent revenue sharing policy α . Supplier 1 prefers the contingent policy while supplier 2 prefers the pre-committed policy, i.e., $R_{u1}^p \leq R_{u1}^c$ and $R_{u2}^p \geq R_{u2}^c$.*

Proposition 3.3 (i) compares the separate sales and the pure bundling sales under the pre-committed revenue sharing policy. The result is similar to that of Proposition 3.2. The pure bundling sales expands the market size by a lower total price and thus captures a larger return than the separate sales. If those two products are similar, then they approximately evenly share the total revenue evenly, and thus both suppliers can benefit by joining the subscription platform. However, if those two products are sufficiently different, the low-valuation product supplier obtains less revenue in the pure bundling sales than in the separate sales, because the revenue allocation policy is based on the comparison of the amount of streams of those two products. Since the number of streams

is increasing in the valuation of the product, in the separate sales, the revenue of the low-valuation product depends on the conditional expectation of its own valuation, whereas in the pure bundling sales, the proportion of the total revenue can be extremely small due to the large difference of the two products. As a result, the low-valuation product supplier may not prefer to join the subscription platform.

Proposition 3.3 (ii) compares the revenue of each supplier between the contingent and pre-committed policies. It shows that the high-valuation product supplier tends to prefer the pre-committed policy while the low-valuation product supplier tends to prefer the contingent policy. By our assumption of the distributions, the valuation of the high-valuation product has the larger variation and mean than that of the low-valuation product. The difference of the variations has a great influence on the proportion of the revenue obtained by the high-valuation product supplier under the contingent policy. Even though the high-valuation product has a larger mean valuation than the low-valuation product, the higher variation of the valuation leads to a relatively low proportion of the total revenue. However, under the pre-committed policy, the proportion of the total revenue mainly depends on the mean valuation of both products, but is almost not influenced by the variation. As a result, the popular product can achieve a higher revenue under the pre-committed policy than under the contingent policy.

For the retailer, since those revenue sharing policies have no influence on the demand side, the optimal pricing strategy does not change. With the same royalty rate r , the retailer earn the same revenue in the pure bundling sales with both policies.

3.6 Extension: Mixed Bundling Sales

In this section, we relax the pure bundling sales “constraint”, and allow both suppliers to have their separate sales in addition to the bundling sales in the subscription platform. However, in practice, the mixed bundling sales may not be launched due to various reasons. According to Bhargava (2012) and Fang and Norman (2006), the retailer may prefer to stick to the pure bundling sales because of the historical practice. Moreover, the mixed bundling sales may not be implemented due to the technological challenges. Finally, the mixed bundling sales may not be implemented because of the antitrust concern by the retailer.

In the mixed bundling sales, the revenue for each supplier would be the summation of the separate sales and a fraction of the revenue in the bundling sales. We firstly show the optimal pricing strategy in the mixed bundling sales. Then we derive the revenue of both suppliers under the contingent revenue sharing policy.

Denote the prices of the products A_1 and A_2 in the mixed bundling sales by p_{m1} and p_{m2} respectively. The price of the bundle is denoted by p_{m0} . The revenue of the bundle is R_{m0} , and the revenue for supplier 1 and 2 in the bundling is denoted by R_{m1} and R_{m2} respectively, i.e., $R_{m0} = R_{m1} + R_{m2}$. Now we show the optimal pricing strategy p_{m1}^* , p_{m2}^* and p_{m0}^* for the mixed bundling sales. To achieve the feasible solution, several reasonable constraints will be added. First, the platform sets individual and bundle prices so that $p_{m0} \leq p_{m1} + p_{m2}$, otherwise no one would have reason to buy the bundle. Second, $p_{m0} \geq p_{m1}$ and $p_{m0} \geq p_{m2}$, because otherwise consumers could get both products for less than the price of one. The revenue under the optimal pricing strategy is denoted by R_{m0}^* , R_{m1}^* and R_{m2}^* . The verification follows the way proposed by Eckalbar (2010).

Given the option of separate sales in the subscription platform, providers endogenously decide which mechanism to implement (partial mixed bundling or full mixed bundling) so as to achieve the most total revenue.

Lemma 3.4 (PRICING STRATEGY II) *The optimal pricing strategy of the mixed bundling sales is*

- (i) $p_{m1}^* = \frac{2a_1}{3}$, $p_{m2}^* = \frac{2a_2}{3}$, $p_{m0}^* = \frac{1}{3}(2a_1 + 2a_2 - \sqrt{2a_1a_2})$ if $a_2/a_1 < 2$,
- (ii) $p_{m1}^* = \frac{2a_1}{3}$, $p_{m2}^* = \frac{1}{6}(3a_2 + 2a_1)$, $p_{m0}^* = \frac{1}{6}(3a_2 + 2a_1)$ if $a_2/a_1 \geq 2$.

Proposition 3.4(i) characterizes the full mixed bundling sales and (ii) characterizes the partial mixed bundling sales. Figure 3.3(a) and (b) illustrate the optimal pricing strategy in the mixed bundling sales. If the two products are sufficiently different, then the partial mixed bundling sales will be optimal, and if the two products are similar, then the full mixed bundling sales will be optimal.

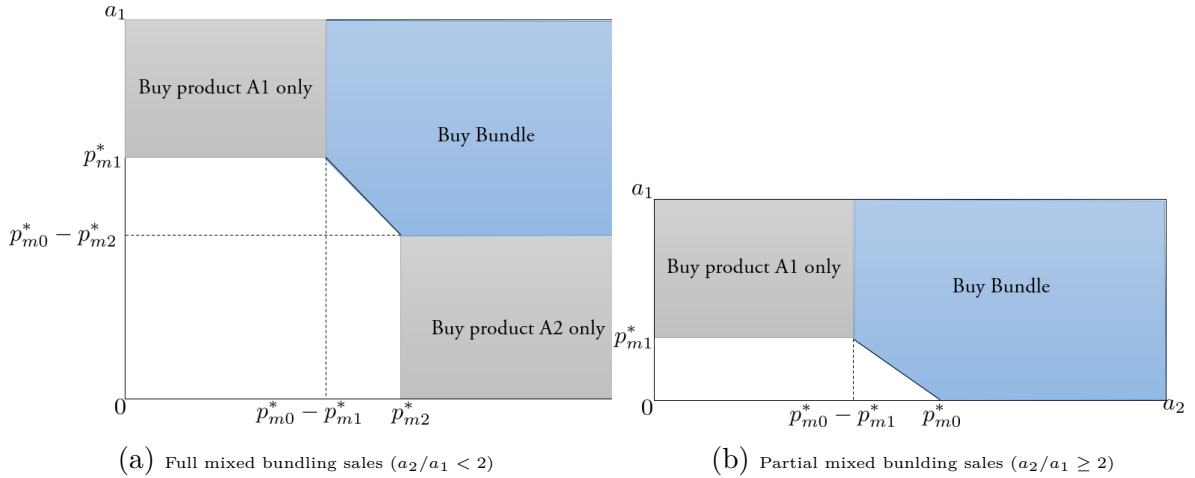
3.6.1 Partial Mixed Bundling Sales

In this section, we examine the revenue of the two suppliers in the partial mixed bundling sales. With the contingent revenue sharing policy, denote the fraction of the revenue in the bundling sales for product A_1 in the partial mixed bundling sales by γ_p^c , i.e.,

$$\gamma_p^c = \mathbb{E} \left(\frac{X}{X+Y} \middle| X+Y \geq p_{m0}^*, Y \geq p_{m0}^* - p_{m1}^* \right).$$

Then the fraction of the revenue for product A_2 is $1 - \gamma_p^c$. The total revenue of supplier 1 is $R_{p1}^c = (1 - r)(\gamma_p^c R_{m0}^* + R_{m1}^*)$ and the revenue of supplier 2 is $R_{p2}^c = (1 - r)(1 - \gamma_p^c)R_{m0}^*$. Proposition 3.5 compares the revenue in the different sales strategies for supplier 1 and 2.

Figure 3.3: Optimal Pricing Strategy II



Proposition 3.5 (PARTIAL MIXED BUNDLING SALES) (i) Compared with the separate sales, both suppliers prefer the partial mixed bundling sales, i.e., $R_{p1}^c \geq R_{s1}$ and $R_{p2}^c \geq R_{s2}$.

(ii) Compared with the pure bundling sales, supplier 1 prefers the partial mixed bundling sales, while supplier 2 prefers the pure bundling sales, i.e., $R_{p1}^c \geq R_{u1}^c$ and $R_{p2}^c \leq R_{u2}^c$.

Now it is straightforward to compare the revenue in different sales strategies. Since the partial mixed bundling will be implemented only if the two products are sufficiently different, Table 3.2 shows that the revenue ranking of the separate sales, pure bundling sales and partial mixed bundling sales, if the two products are sufficiently different $a_2/a_1 \geq 2$.

Table 3.2: Revenue Ranking ($a_2/a_1 \geq 2$)

| | Retailer | Supplier 1 | Supplier 2 |
|------------------------|----------|------------|------------|
| Separate | 3rd | 2nd | 3rd |
| Pure Bundling | 2nd | 3rd | 1st |
| Partial Mixed Bundling | 1st | 1st | 2nd |

By Proposition 3.5, we find that the partial mixed tends to be better than the separate sales. However, the popular supplier can earn the most revenue in the pure bundling sales.

3.6.2 Full Mixed Bundling Sales

Now we derive the revenue of those two suppliers in the full mixed bundling sales. Under the contingent revenue sharing policy, the fraction of the revenue for product A_1 in the full mixed bundling sales is denoted by γ_f^c , i.e.,

$$\gamma_f^c = \mathbb{E} \left(\frac{X}{X+Y} \middle| X+Y \geq p_{m0}^*, X \geq p_{m0}^* - p_{m2}^*, Y \geq p_{m0}^* - p_{m1}^* \right)$$

Then the fraction of the revenue for product A_2 is $1 - \gamma_f^c$. The total revenue for supplier 1 will be $R_{f1}^c = (1 - r)(\gamma_f^c R_{m0}^* + R_{m1}^*)$, and the total revenue for supplier 2 will be $R_{f2}^c = (1 - r)((1 - \gamma_f^c)R_{m0}^* + R_{m2}^*)$. Proposition 3.6 compares the revenue in different sales strategies for supplier 1 and 2.

Proposition 3.6 (FULL MIXED BUNDLING SALES) (i) Compared with the separate sales, both suppliers prefer the full mixed bundling sales, i.e., $R_{f1}^c \geq R_{s1}$ and $R_{f2}^c \geq R_{s2}$.

(ii) Compared with the pure bundling sales, supplier 1 prefers the full mixed bundling sales, i.e., $R_{f1}^c \geq R_{u1}^c$. There exists a threshold $\hat{k}^* \in [1, 2]$ such that if $a_2/a_1 \leq \hat{k}^*$, supplier 2 prefers the full mixed bundling sales, i.e., $R_{f2}^c \leq R_{u2}^c$. If $a_2/a_1 > \hat{k}^*$, the second supplier prefers the pure bundling sales, i.e., $R_{f2}^c > R_{u2}^c$.

Table 3.3 exhibits the revenue ranking for the separate sales, pure bundling sales and full mixed bundling sales. We find that the full mixed bundling sales tends to be optimal if the two products are similar. The number in the brackets shows the ranks if $a_2/a_1 \in (\hat{k}^*, 2]$.

Table 3.3: Revenue Ranking ($a_2/a_1 \leq 2$)

| | Retailer | Supplier 1 ($a_2/a_1 > \hat{k}^*$) | Supplier 2 ($a_2/a_1 > \hat{k}^*$) |
|---------------------|----------|--------------------------------------|--------------------------------------|
| Separate | 3rd | 3rd (2nd) | 3rd |
| Pure Bundling | 2nd | 2nd (3rd) | 2nd (1st) |
| Full Mixed Bundling | 1st | 1st | 1st (2nd) |

3.7 Mean-Preserving Spread

In this section, we follow a different way of specifying the products on the subscription platform, and examine the revenues of those suppliers in different sales strategies. We

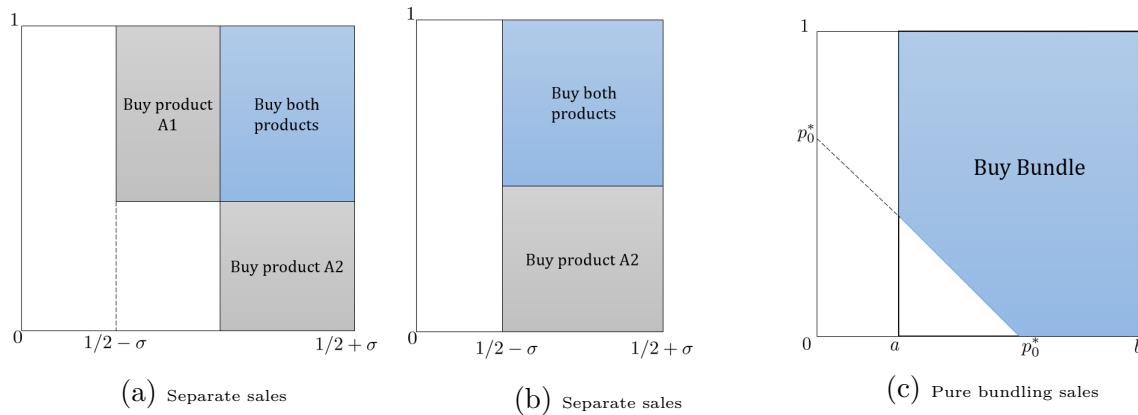
follow Johnson and Myatt (2006) in assuming that different products induce demand rotations. To make the model be consistent with the model in the previous sections, we assume that the valuations of those two products X and Y follow the uniform distributions over $[0, 1]$ and $[a, b]$ respectively. We impose that $a > 0$ and $b < 1$ (see Bar-Isaac et al. (2012), p.p. 1152.)⁴ To simplify the analysis, we focus on the case that $a + b = 1$, that is, X and Y have the same mean but different variances. As a result, X is second order stochastically dominated by Y , i.e., $X \leq_{SSD} Y$. According to Johnson and Myatt (2006), X is the niche product and Y is the popular product, and those two products are differentiated in the dispersion of the consumer valuation.

Lemma 3.7 shows that the optimal pricing strategy in the separate and pure bundling sales. Since $a + b = 1$, the mean value of the X is $\frac{1}{2}$, thus we can denote $X \sim U(\frac{1}{2} - \sigma, \frac{1}{2} + \sigma)$ in which σ is half of the length of the uniform distribution. Thus, the condition on the optimal pricing strategy depends on the single parameter σ . We determine the pricing strategy for different values of σ in the pure bundling sales.

Lemma 3.7 (PRICING STRATEGY III) (i) *The optimal pricing strategy for the separate sales is $p_1^* = \frac{1}{2}$ and $p_2^* = \max\{\frac{1}{4} + \sigma, \frac{1}{2} - \sigma\}$.*

(ii) *The optimal pricing strategy for the pure bundling sales is $p_0^* = \frac{3}{4}$ if $\sigma < \frac{1}{4}$, and $p_0^* = \frac{2a + \sqrt{a^2 - 6a + 6b}}{3}$ if $\sigma \geq \frac{1}{4}$.*

Figure 3.4: Optimal Pricing Strategy III



Figures 3.4 (a) and (b) illustrate the optimal pricing strategy for the separate sales with different values of σ . If σ is sufficiently small, the whole market will buy product 2

⁴Bar-Isaac et al. (2012) follow a more general way of assuming the two uniform distributions, i.e., $X \sim U(\tilde{a}, \tilde{b})$, $Y \sim U(a, b)$ with $a > \tilde{a}$ and $b < \tilde{b}$. Since we consider the two-product case, any assumption with the general form can be normalized to our model, i.e. $X \sim U(0, 1)$ and $Y \sim U(a, b)$.

while a fraction of the market will buy product 1. This is consistent with our intuition that the popular product may have a larger market coverage than the niche product. Figure 3.4 (c) illustrates the optimal pricing strategy for the pure bundling sales.

Based on the optimal pricing strategies in Lemma 3.7, we compare the revenues of the those two suppliers between the separate sales and pure bundling sales. Following the same way in studying the first specification, we discuss two ways of allocating the revenue in the pure bundling sales, the contingent revenue sharing policy and pre-committed revenue sharing policy.

The revenue share under the contingent revenue sharing policy is determined by (3.1) with different assumptions on the CDF of random variables X and Y . The revenue share under the pre-committed revenue sharing policy is determined by (3.2). Proposition 3.8 shows the comparison of the revenues under different sales strategies for different values of σ .

Proposition 3.8 (i) *Comparing the pure bundling and separate sales, supplier 1 prefers the pure bundling sales, i.e., $R_{u1}^c \geq R_{s1}$. There exists a threshold $\bar{\sigma}$ such that if $\sigma \geq \bar{\sigma}$, supplier 2 prefers the pure bundling sales, i.e., $R_{u2}^c \geq R_{s2}$, and if $\sigma < \bar{\sigma}$, supplier 2 prefers the separate sales, i.e., $R_{u2}^c < R_{s2}$.*

(ii) *Supplier 1 prefers the pre-committed policy, i.e., $R_{u1}^c \leq R_{u1}^p$, while supplier 2 prefers the contingent policy, i.e., $R_{u2}^p \leq R_{u2}^c$.*

For Proposition 3.8(i), the result is similar to the result in Proposition 3.2. In Proposition 3.2, the valuation of product 1 is first order stochastically dominated by the valuation of product 2, i.e., $X \leq_{FSD} Y$. In Proposition 3.8, the valuation of product 1 is second order stochastically dominated by the valuation of product 2, i.e., $X \leq_{SSD} Y$. Unlike before, Proposition 3.8(i) shows that the popular supplier will prefer the separate sales if those two products are sufficiently different. The result has nothing to do with the mean value of the valuation of the two products since we assume that X and Y have the same mean but different variations. If the two products are sufficiently different, the popular supplier tends to prefer the separate sales, whereas the niche supplier tends to favor the bundling sales.

In the separate sales, the popular product almost covers the whole market by the optimal pricing strategy, but the niche product only covers a fraction of the market. However, in the pure bundling sales, even with the optimal bundling pricing strategy, the bundle covers a fraction of the market. For the popular product supplier, the market coverage is less in the bundling sales than in the separate sales. Though the bundling

sales may earn a higher total revenue than the separate sales, the revenue of the popular supplier is less in the pure bundling sales than in the separate sales.

For Proposition 3.8(ii), since the valuation of the two products have different dispersion. The niche supplier prefers the pre-committed policy since the valuations of the niche product has a relatively high uncertainty, and pre-committed policy to some extent eliminates such uncertainty. For the popular supplier, the valuation of the popular product has a relatively low uncertainty, but the advantage of the low uncertainty valuation has not been taken under the pre-committed policy. Because the total revenue remains to be the same in both policies, if the niche supplier earns more revenue in the pre-committed policy than in the contingent policy, then the popular supplier earns less revenue in the pre-committed policy than in the contingent policy.

3.8 Conclusion

We investigate the influence of different revenue sharing policies on different suppliers on the subscription platform. Given that some online platforms share the revenue by calculating the proportion of streams for a digital product, we characterize two policies that are commonly implemented. One is the contingent revenue sharing policy, and the other is the pre-committed revenue sharing policy. By assuming that the subscription provider commits to provide the pure bundling sales, we have the following results. First, with the option of separate sales out of the subscription platform, we find that the high-valuation product supplier may have the incentive to join the platform while the low-valuation product supplier may not have the incentive to join the platform under both policies. Moreover, the high-valuation product supplier may prefer the pre-committed policy while the low-valuation product supplier may prefer the contingent policy. Second, if the products are differentiated by the dispersion of consumer's valuation, then the popular product supplier may prefer the separate sales while the niche product supplier may prefer to join the subscription platform. In the comparison between the two policies, the popular supplier tends to prefer the contingent policy while the niche supplier tends to prefer the pre-committed policy. Finally, we examine the mixed bundling sale in which the subscription provider may include the separate sales of a product in addition to the bundling sales. Results shows that under the contingent revenue sharing policy the mixed bundling sales may be optimal.

We employ a stylized two-product model to characterize the subscription service. Those results show how the heterogeneity of the products affects the revenue earned by the suppliers. There are several limitations to our model. First, we assume that

consumer's valuations about those two products are independent. By the classic result about the bundling model, if the valuations of the two products are positively correlated, the bundling sales may fail to achieve a higher revenue than the separate sales. In that case, both suppliers may not want to join the subscription platform since the bundling sales is no longer profitable. Thus, future research can take into account the correlated valuation across those products. Second, we assume that the two products are known to the whole market. In reality, perhaps the niche products may be only known by a fraction of market, and this can be a possible reason of the heterogeneous valuation. Moreover, the subscription platform can enhance the popularity of some products. Our model does not consider the causality of the heterogeneous valuations but directly assume it, thus future research may characterize the subscription service in a more detailed level and explore its influence over products' popularity. Despite those limitations, our model characterizes the revenue sharing policies in the subscription service, and captures the unequal allocation of revenues due to the heterogeneity of products. Our results may explain the practice that some product suppliers do not prefer to join the subscription, and help the subscription providers better design the revenue sharing contract.

3.9 Proofs.

Proof of Lemma 3.1. (i) A potential customer will buy the product A_1 only if his realized value $x \geq p_1$, thus the expected revenue of A_1 can be characterized by $R_1 = [1 - F(p_1)]p_1$. The FOC is given by $1 - F(p_1^*) - f(p_1^*)p_1^* = 0$. Thus, $p_1^* = (1 - F(p_1^*)) / f(p_1^*)$. If the valuation X follows the uniform distribution over $[0, a_1]$, $p_1^* = a_1/2$. Similarly, the optimal single price of product A_2 is $p_2^* = a_2/2$. The revenue under the optimal pricing strategy is $R_1^* = a_1/4$ and $R_2^* = a_2/4$.

(ii) The price for the pure bundling sales is denoted by p_0 . Each customer holds the realized value of the products in the bundle as x and y . The revenue for the bundle is $R_u = [1 - H(p_0)]p_0$ where $H(\cdot)$ is the CDF of $V = X + Y$ and $h(\cdot)$ is its PDF, i.e., $h = f \circ g$. The PDF and CDF of V are

$$h(v) = (f \circ g)(v) = \int_0^v f(v-y)g(y)dy = \begin{cases} \frac{v}{a_1 a_2}, & \text{for } 0 \leq v \leq a_1, \\ \frac{1}{a_2}, & \text{for } a_1 \leq v \leq a_2, \\ \frac{a_1 + a_2 - v}{a_1 a_2} & \text{for } a_2 \leq v \leq a_1 + a_2. \end{cases}$$

and

$$H(v) = \begin{cases} \frac{v^2}{2a_1a_2}, & \text{for } 0 \leq v \leq a_1, \\ \frac{v^2 - (v - a_1)^2}{2a_1a_2}, & \text{for } a_1 \leq v \leq a_2, \\ 1 - \frac{(a_1 + a_2 - v)^2}{2a_1a_2} & \text{for } a_2 \leq v \leq a_1 + a_2. \end{cases}$$

The FOC of R_u is $1 - H(p_0^*) - h(p_0^*)p_0^* = 0$. For $0 \leq p_0 \leq a_1$, $1 - H(p_0) - h(p_0)p_0 = 1 - \frac{3p_0^2}{2a_1a_2}$. Therefore, if $1 \leq a_2/a_1 < 3/2$, $p_0^* = \sqrt{\frac{2a_1a_2}{3}}$ and p_0^* is the maximum point over $[0, a_1]$. Since it can be simply verify that if $1 \leq a_2/a_1 < 3/2$, $[1 - H(p_0)]p_0$ is decreasing over $p_0 \in [a_1, a_1 + a_2]$, $p_0^* = \sqrt{\frac{2a_1a_2}{3}}$ is the global maximum point if $1 \leq a_2/a_1 < 3/2$. Similarly, for $a_1 \leq p_0 \leq a_2$, $1 - H(p_0) - h(p_0)p_0 = 1 - \frac{p_0^2 - (p_0 - a_1)^2}{2a_1a_2} - \frac{p_0}{a_2}$. If $3/2 \leq a_2/a_1 < \infty$, the maximum point is $p_0^* = \frac{2a_2 + a_1}{4}$ which solves the equation $1 - \frac{p_0^2 - (p_0 - a_1)^2}{2a_1a_2} - \frac{p_0}{a_2} = 0$. It can be simply verify that $p_0^* = \frac{2a_2 + a_1}{4}$ is the global maximum point if $3/2 \leq a_2/a_1 < \infty$.

Therefore, the revenue of bundling sales under the optimal pricing strategy is

$$R_u^* = \begin{cases} \frac{2}{3} \sqrt{\frac{2a_1a_2}{3}}, & \text{for } 1 \leq a_2 < \frac{3}{2}a_1, \\ \frac{(2a_2 + a_1)^2}{16a_2}, & \text{for } a_2 \geq \frac{3}{2}a_1. \end{cases}$$

which are the results in Lemma 3.1. ■

Proof of Proposition 3.2. First, we derive the joint distribution of $\frac{X}{X+Y}$ and $X + Y$. Denote that $U = \phi_1(X, Y) = \frac{X}{X+Y}$ and $V = \phi_2(X, Y) = X + Y$. Then the inverse transformation is given by $X = \varphi_1(U, V) = UV$ and $Y = \varphi_2(U, V) = V - UV$. The joint density function of U and V is $f_{U,V}(u, v) = f_{X,Y}(x, y)|J(x, y)|^{-1}$ where $J(x, y)$ denotes the Jacobian of the functions $\phi_1(X, Y)$ and $\phi_2(X, Y)$,

$$J(x, y) = \det \begin{pmatrix} \frac{\partial \phi_1}{\partial x} & \frac{\partial \phi_1}{\partial y} \\ \frac{\partial \phi_2}{\partial x} & \frac{\partial \phi_2}{\partial y} \end{pmatrix} = \frac{y}{(x+y)^2} - \frac{-x}{(x+y)^2} = \frac{1}{x+y}.$$

As a result, the joint PDF of U and V is $f_{U,V}(u, v) = f_{X,Y}(x, y)|J(x, y)|^{-1} = (x + y)f(x)g(y) = vf(uv)g(v - uv)$. Since the random variable $U \in [0, 1]$. If X and Y are uniformly distributed over $[0, a_1]$ and $[0, a_2]$ respectively, the joint PDF of U and V is

$$f_{U,V}(u, v) = \begin{cases} \frac{v}{a_1a_2}, & \text{for } 0 \leq uv \leq a_1 \text{ and } 0 \leq v(1-u) \leq a_2, \\ 0 & \text{otherwise.} \end{cases} \quad (3.3)$$

The ranges in (3.3) can be written as $0 \leq v \leq \frac{a_1}{u}$ and $0 \leq v \leq \frac{a_2}{1-u}$. Only one or other of these ranges needs to be retained, depending on whether u is over $\left[0, \frac{a_1}{a_1+a_2}\right]$ or $\left[\frac{a_1}{a_1+a_2}, 1\right]$. Thus, we have

$$f_{U,V}(u, v) = \begin{cases} \frac{v}{a_1 a_2}, & \text{for } 0 \leq v \leq \frac{a_2}{1-u} \text{ and } 0 \leq u \leq \frac{a_1}{a_1+a_2}, \\ \frac{v}{a_1 a_2}, & \text{for } 0 \leq v \leq \frac{a_1}{u} \text{ and } \frac{a_1}{a_1+a_2} \leq u \leq 1, \\ 0, & \text{otherwise.} \end{cases} \quad (3.4)$$

Second, we derive the revenue sharing policy α . Given the optimal bundling price p_0^* , the PDF conditional on $X + Y \geq p_0^*$ is $f_{U|V}(u|V \geq p_0^*) = \int_{p_0^*}^{\infty} f_{U,V}(u, v) dv / (1 - H(p_0^*))$. Given the optimal bundling price p_0^* , the revenue sharing policy $\alpha = \mathbb{E}\left(\frac{X}{X+Y} \mid X + Y \geq p_0^*\right) = \frac{1}{1-H(p_0^*)} \int_0^1 \int_{p_0^*}^{\infty} u f_{U,V}(u, v) dv du$. The revenue sharing policy α depends on the value of p_0^* . If $0 < p_0^* < a_1$, the ranges of v in (3.4) and (3.5) become $p_0^* \leq v \leq \frac{a_2}{1-u}$ and $p_0^* \leq v \leq \frac{a_1}{u}$. Therefore,

$$\begin{aligned} \alpha &= \frac{1}{1-H(p_0^*)} \int_0^1 \int_{p_0^*}^{\infty} u f_{U,V}(u, v) dv du \\ &= \frac{1}{1-H(p_0^*)} \left[\int_0^{\frac{a_1}{a_1+a_2}} \int_{p_0^*}^{\frac{a_2}{1-u}} \frac{uv}{a_1 a_2} dv du + \int_{\frac{a_1}{a_1+a_2}}^1 \int_{p_0^*}^{\frac{a_1}{u}} \frac{uv}{a_1 a_2} dv du \right] \\ &= \frac{1}{2(1-H(p_0^*))} \left[-\frac{a_2}{a_1} \ln\left(\frac{a_1+a_2}{a_2}\right) + \frac{a_1}{a_2} \ln\left(\frac{a_1+a_2}{a_1}\right) - \frac{p_0^{*2}}{2a_1 a_2} + 1 \right]. \end{aligned} \quad (3.6)$$

By (3.6), we can obtain the fraction of the revenue for product A_2 ,

$$1 - \alpha = \frac{1}{2(1-H(p_0^*))} \left[\frac{a_2}{a_1} \ln\left(\frac{a_1+a_2}{a_2}\right) - \frac{a_1}{a_2} \ln\left(\frac{a_1+a_2}{a_1}\right) - \frac{p_0^{*2}}{2a_1 a_2} + 1 \right]. \quad (3.7)$$

If $a_1 \leq p_0^* \leq a_2$, the ranges in (3.4) becomes $p_0^* \leq v \leq \frac{a_2}{1-u}$ and $0 \leq u \leq \frac{a_1}{a_1+a_2}$ and the ranges in (3.5) become $p_0^* \leq v \leq \frac{a_1}{u}$ and $\frac{a_1}{a_1+a_2} \leq u \leq \frac{a_1}{p_0^*}$. Then we have

$$\begin{aligned} \alpha &= \frac{1}{1-H(p_0^*)} \int_0^1 \int_{p_0^*}^{\infty} u f_{U,V}(u, v) dv du \\ &= \frac{1}{1-H(p_0^*)} \left[\int_0^{\frac{a_1}{a_1+a_2}} \int_{p_0^*}^{\frac{a_2}{1-u}} \frac{uv}{a_1 a_2} dv du + \int_{\frac{a_1}{a_1+a_2}}^{\frac{a_1}{p_0^*}} \int_{p_0^*}^{\frac{a_1}{u}} \frac{uv}{a_1 a_2} dv du \right] \\ &= \frac{1}{2(1-H(p_0^*))} \left[-\frac{a_2}{a_1} \ln\left(\frac{a_1+a_2}{a_2}\right) + \frac{a_1}{a_2} \ln\left(\frac{a_1+a_2}{a_1}\right) - \frac{a_1}{2a_2} \right. \\ &\quad \left. + \frac{a_1}{a_2} \ln\left(\frac{a_1}{p_0^*}\right) + 1 \right]. \end{aligned} \quad (3.8)$$

The fraction of the revenue for product A_2 is

$$1 - \alpha = \frac{1}{2(1 - H(p_0^*))} \left[\frac{a_2}{a_1} \ln \left(\frac{a_1 + a_2}{a_2} \right) - \frac{a_1}{a_2} \ln \left(\frac{a_1 + a_2}{a_1} \right) + \frac{a_1}{2a_2} - \frac{a_1}{a_2} \ln \left(\frac{a_1}{p_0^*} \right) + 1 + \frac{a_1 - 2p_0^*}{a_2} \right]. \quad (3.9)$$

Third, we incorporate the optimal bundling price p_0^* obtained in Lemma 3.1. If $a_1 \leq a_2 < \frac{3}{2}a_1$, $p_0^* = \sqrt{\frac{2a_1a_2}{3}} \in [0, a_1)$. The revenue of A_1 , $R_{u1}^c = (1 - r)\alpha R_u^*$. By (3.6), we have

$$\begin{aligned} \alpha R_u^* &= \alpha(1 - H(p_0^*))p_0^* = \frac{p_0^*}{2} \left[-\frac{a_2}{a_1} \ln \left(\frac{a_1 + a_2}{a_2} \right) + \frac{a_1}{a_2} \ln \left(\frac{a_1 + a_2}{a_1} \right) - \frac{p_0^{*2}}{2a_1a_2} + 1 \right] \\ &= \frac{1}{2} \sqrt{\frac{2a_1a_2}{3}} \cdot \left[-\frac{a_2}{a_1} \ln \left(\frac{a_1 + a_2}{a_2} \right) + \frac{a_1}{a_2} \ln \left(\frac{a_1 + a_2}{a_1} \right) + \frac{2}{3} \right]. \end{aligned}$$

Now we examine the value of R_{u1}^c/R_{s1} . Since $R_{u1}^c/R_{s1} = \alpha R_u^*/R_1^*$, we examine the value of $\alpha R_u^*/R_1^*$. Divide αR_u^* by R_1^* and substitute a_2/a_1 with k where $1 \leq k < \frac{3}{2}$, by Lemma 3.1, $R_1^* = a_1/4$, thus

$$\frac{\alpha R_u^*}{R_1^*} = 2\sqrt{\frac{2k}{3}} \left[\frac{1}{k} \ln(1 + k) - k \ln \left(1 + \frac{1}{k} \right) + \frac{2}{3} \right]. \quad (3.10)$$

Take derivative with respect to k , then $\frac{d(R_{u1}^c/R_{s1})}{dk} = \sqrt{\frac{2}{3k}} \left[-\frac{1}{k} \ln(1 + k) - 3k \ln \left(1 + \frac{1}{k} \right) + \frac{8}{3} \right] < 0$ if $k \in [1, \frac{3}{2})$. As a result, $\alpha R_u^*/R_1^*$ is decreasing in k . When $k = \frac{3}{2}$, $\alpha R_u^*/R_1^* = 1.0226 > 1$. Therefore, if $1 \leq k < \frac{3}{2}$, $\alpha R_u^* > R_1^*$. Because $R_{u1}^c/R_{s1} = \alpha R_u^*/R_1^*$, $R_{u1}^c > R_{s1}$ if $1 \leq a_2/a_1 < \frac{3}{2}$.

For product A_2 , by (3.7), we have

$$\begin{aligned} (1 - \alpha)R_u^* &= (1 - \alpha)(1 - H(p_0^*))p_0^* \\ &= \frac{p_0^*}{2} \left[\frac{a_2}{a_1} \ln \left(\frac{a_1 + a_2}{a_2} \right) - \frac{a_1}{a_2} \ln \left(\frac{a_1 + a_2}{a_1} \right) - \frac{p_0^{*2}}{2a_1a_2} + 1 \right] \\ &= \frac{1}{2} \sqrt{\frac{2a_1a_2}{3}} \left[\frac{a_2}{a_1} \ln \left(\frac{a_1 + a_2}{a_2} \right) - \frac{a_1}{a_2} \ln \left(\frac{a_1 + a_2}{a_1} \right) + \frac{2}{3} \right]. \end{aligned}$$

Substitute a_2/a_1 with k ,

$$\frac{(1 - \alpha)R_u^*}{R_2^*} = 2\sqrt{\frac{2}{3k}} \left[k \ln \left(1 + \frac{1}{k} \right) - \frac{1}{k} \ln(1 + k) + \frac{2}{3} \right]. \quad (3.11)$$

Take derivative with respect to k ,

$$\frac{d[(1-\alpha)R_u^*/R_2^*]}{dk} = \sqrt{\frac{2}{3k}} \left[\ln\left(1 + \frac{1}{k}\right) + \frac{3}{k^2} \ln(1+k) - \frac{8}{3k} \right]. \quad (3.12)$$

If $k = 1$, (3.12) is greater than 0, and if $k = \frac{3}{2}$, (3.12) is less than 0. Taking derivative with respect to k again, we find that $\frac{d^2[(1-\alpha)R_u^*/R_2^*]}{dk^2} < 0$ for $k \in [1, 3/2]$. Thus, there exists only one point \tilde{k} such that $\frac{d[(1-\alpha)R_u^*/R_2^*]}{dk}|_{k=\tilde{k}} = 0$. Since $\frac{(1-\alpha)R_u^*}{R_2^*}|_{k=1} = 1.0887 > 1$ and $\frac{(1-\alpha)R_u^*}{R_2^*}|_{k=3/2} = 1.0961 > 1$, we find that $(1-\alpha)R_u^*/R_2^* > 1$ for $k \in [1, \frac{3}{2})$, that is, $(1-\alpha)R_u^* > R_2^*$. Since $R_{u2}^c/R_{s2} = (1-\alpha)R_u^*/R_2^*$, $R_{u2}^c > R_{s2}$ if $1 \leq a_2/a_1 < \frac{3}{2}$.

By Lemma 3.1, if $a_2 \geq \frac{3}{2}a_1$, $p_0^* = \frac{2a_2+a_1}{4} \in [a_1, a_2]$. For product A_1 , by (3.8),

$$\begin{aligned} \alpha R_u^* &= \alpha(1 - H(p_0^*))p_0^* = \frac{p_0^*}{2} \left[-\frac{a_2}{a_1} \ln\left(\frac{a_1+a_2}{a_2}\right) + \frac{a_1}{a_2} \ln\left(\frac{a_1+a_2}{a_1}\right) - \frac{a_1}{2a_2} + \frac{a_1}{a_2} \ln\left(\frac{a_1}{p_0^*}\right) + 1 \right] \\ &= \frac{2a_2+a_1}{8} \left[-\frac{a_2}{a_1} \ln\left(\frac{a_1+a_2}{a_2}\right) + \frac{a_1}{a_2} \ln\left(\frac{a_1+a_2}{a_1}\right) - \frac{a_1}{2a_2} + \frac{a_1}{a_2} \ln\left(\frac{4a_1}{2a_2+a_1}\right) + 1 \right]. \end{aligned}$$

Substitute a_2/a_1 with k where $k \geq \frac{3}{2}$ in $\alpha R_u^*/R_1^*$,

$$\begin{aligned} \frac{\alpha R_u^*}{R_1^*} &= \left(\frac{a_2}{a_1} + \frac{1}{2}\right) \left[-\frac{a_2}{a_1} \ln\left(1 + \frac{a_1}{a_2}\right) + \frac{a_1}{a_2} \ln\left(1 + \frac{a_2}{a_1}\right) - \frac{a_1}{2a_2} - \frac{a_1}{a_2} \ln\left(\frac{a_2}{2a_1} + \frac{1}{4}\right) + 1 \right] \\ &= \left(k + \frac{1}{2}\right) \left[-k \ln\left(1 + \frac{1}{k}\right) + \frac{1}{k} \ln(1+k) - \frac{1}{2k} - \frac{1}{k} \ln\left(\frac{k}{2} + \frac{1}{4}\right) + 1 \right]. \quad (3.13) \end{aligned}$$

Taking derivative with respect to k , $\frac{d(R_u^*/R_1^*)}{dk} = \frac{1}{2} \left[-\ln\left(1 + \frac{1}{k}\right) - \frac{1}{k^2} \ln(1+k) + \frac{1}{k^2} \ln\left(\frac{k}{2} + \frac{1}{4}\right) \right] + \frac{1-2k}{4k^2}$. Because $\ln\left(\frac{k}{2} + \frac{1}{4}\right) < \ln(1+k)$ and $\frac{1-2k}{4k^2} < 0$ for $k \in [3/2, \infty)$, $\frac{d(\alpha R_u^*/R_1^*)}{dk} < 0$. Since $\alpha R_u^*/R_1^* > 1$ if $k = \frac{3}{2}$ and $\alpha R_u^*/R_1^* < 1$ if $s = 2$, and $\alpha R_u^*/R_1^*$ is continuous in k , there exists an k^* such that if $\frac{3}{2} \leq k \leq k^*$, $\alpha R_u^*/R_1^* \geq 1$, and if $k > k^*$, $\alpha R_u^*/R_1^* < 1$. Since $R_{u1}^c/R_{s1} = \alpha R_u^*/R_1^*$, the result about the threshold k^* holds for R_{u1}^c/R_{s1} .

For product A_2 , by (3.9), we have

$$\begin{aligned} (1-\alpha)R_u^* &= (1-\alpha)(1 - H(p_0^*))p_0^* \\ &= \frac{p_0^*}{2} \left[\frac{a_2}{a_1} \ln\left(\frac{a_1+a_2}{a_2}\right) - \frac{a_1}{a_2} \ln\left(\frac{a_1+a_2}{a_1}\right) + \frac{a_1}{2a_2} - \frac{a_1}{a_2} \ln\left(\frac{a_1}{p_0^*}\right) + 1 + \frac{a_1 - 2p_0^*}{a_2} \right] \\ &= \frac{2a_2+a_1}{8} \left[\frac{a_2}{a_1} \ln\left(\frac{a_1+a_2}{a_2}\right) - \frac{a_1}{a_2} \ln\left(\frac{a_1+a_2}{a_1}\right) + \frac{a_1}{2a_2} \right. \\ &\quad \left. - \frac{a_1}{a_2} \ln\left(\frac{4a_1}{2a_2+a_1}\right) + \frac{a_1}{2a_2} \right]. \quad (3.14) \end{aligned}$$

Again divide $(1 - \alpha)R_u^*$ by R_2^* and substitute a_2/a_1 with k ,

$$\frac{(1 - \alpha)R_u^*}{R_2^*} = \left(1 + \frac{1}{2k}\right) \left[k \ln\left(1 + \frac{1}{k}\right) - \frac{1}{k} \ln(1 + k) + \frac{1}{k} + \frac{1}{k} \ln\left(\frac{k}{2} + \frac{1}{4}\right) \right].$$

Take derivative with respect to k ,

$$\begin{aligned} & \frac{d[(1 - \alpha)R_u^*/R_2^*]}{dk} \\ &= \ln\left(1 + \frac{1}{k}\right) + \left(\frac{1}{k^2} + \frac{1}{k^3}\right) \ln(1 + k) - \left(\frac{1}{k^2} + \frac{1}{k^3}\right) \ln\left(\frac{2k + 1}{4}\right) - \left(\frac{1}{k^2} + \frac{1}{k^3}\right) + \frac{1 - 2k}{2k^2}. \end{aligned}$$

By the well-known inequalities that $\frac{x}{1+x} \leq \ln(1+x) \leq x$, $x > -1$, if $k \geq 3$, we have

$$\begin{aligned} & \left(\frac{1}{k^2} + \frac{1}{k^3}\right) \ln(1 + k) \leq \left(\frac{1}{k} + \frac{1}{k^2}\right) \leq \left(1 + \frac{1}{k}\right), \\ & \left(\frac{1}{k^2} + \frac{1}{k^3}\right) \ln\left(\frac{2k + 1}{4}\right) \geq \left(1 + \frac{1}{k}\right) \frac{4k - 6}{2k + 5} \geq \left(1 + \frac{1}{k}\right). \end{aligned}$$

Moreover, $\ln\left(1 + \frac{1}{k}\right) \leq \frac{1}{k}$. As a result, $\frac{d[(1-\alpha)R_u^*/R_2^*]}{dk} \leq \frac{1}{k} - \left(\frac{1}{k^2} + \frac{1}{k^3}\right) + \frac{1-2k}{2k^2} < 0$ for $k \in [3, \infty)$. If $k = 3$, $(1 - \alpha)R_u^*/R_2^* = 1.0743$, and if $k \rightarrow \infty$, $(1 - \alpha)R_u^*/R_2^* \rightarrow 1$. For $k \in [3/2, 3]$, by the numerical test over $[3/2, 3]$, we find that $(1 - \alpha)R_u^*/R_2^* > 1$. Therefore, $(1 - \alpha)R_u^* > R_2^*$ if $k \in [3/2, \infty)$. Since $R_{u2}^c/R_{s2} = (1 - \alpha)R_u^*/R_2^*$, $R_{u2}^c > R_{s2}$ if $k \in [3/2, \infty)$. ■

Proof of Lemma 3.4. Denote the demand of product A_1 and A_2 in the mixed bundling sales by D_1 and D_2 respectively. The demand of the bundle is denoted by D_0 . Then, $D_1 = (a_1 - p_1)(p_0 - p_1)$, $D_2 = (a_2 - p_2)(p_0 - p_2)$ and $D_0 = -\frac{1}{2}(p_1 + p_2 - p_0)^2 + (p_2 - p_0 + a_1)(p_1 - p_0 + a_2)$. The total revenue is $R_m = D_1p_1 + D_2p_2 + D_0p_0$. The optimal pricing strategies p_1^* , p_2^* and p_0^* are endogenously determined. The FOCs are given by

$$\begin{aligned} \frac{\partial R_m}{\partial p_1} &= (p_1 - p_0)(3p_1 - 2a_1) = 0, \\ \frac{\partial R_m}{\partial p_2} &= (p_2 - p_0)(3p_2 - 2a_2) = 0, \\ \frac{\partial R_m}{\partial p_0} &= \frac{1}{2} \left(-3p_1^2 - 3p_2^2 + 3p_0^2 + 4p_1a_1 - 4p_0a_1 + 4p_2a_2 - 4p_0a_2 + 2a_1a_2 \right) = 0. \end{aligned}$$

We focus on the points with positive value. The candidate points for achieving the maximum revenue are as follows: 1) $p_0 = p_1 = p_2 = \sqrt{\frac{2a_1a_2}{3}}$; 2) $p_1 = p_0 = \frac{a_1}{2} + \frac{a_2}{3}$, $p_2 = \frac{2a_2}{3}$; 3) $p_2 = p_0 = \frac{a_1}{3} + \frac{a_2}{2}$, $p_1 = \frac{2a_1}{3}$; 4) $p_1 = \frac{2a_1}{3}$, $p_2 = \frac{2a_2}{3}$, $p_0 = \frac{2a_1}{3} + \frac{2a_2}{3} - \frac{\sqrt{2a_1a_2}}{3}$; 5) $p_1 = \frac{2a_1}{3}$, $p_2 = \frac{2a_2}{3}$, $p_0 = \frac{2a_1}{3} + \frac{2a_2}{3} + \frac{\sqrt{2a_1a_2}}{3}$. To assess those points, we construct

the Hessian matrix \mathbf{H}_0 , and evaluate the second order conditions. \mathbf{H}_0 and its principal minors are

$$\begin{aligned}\mathbf{H}_0 &= \begin{bmatrix} 6p_1 - 3p_0 - 2a_1 & 0 & 2a_1 - 3p_1 \\ 0 & 6p_2 - 3p_0 - 2a_2 & 2a_2 - 3p_2 \\ 2a_1 - 3p_1 & 2a_2 - 3p_2 & -2a_1 - 2a_2 + 3p_0 \end{bmatrix}, \\ \mathbf{H}_1 &= 6p_1 - 3p_0 - 2a_1, \\ \mathbf{H}_2 &= \begin{bmatrix} 6p_1 - 3p_0 - 2a_1 & 0 \\ 0 & 6p_2 - 3p_0 - 2a_2 \end{bmatrix}.\end{aligned}$$

The determinants are

$$\begin{aligned}\det(\mathbf{H}_0) &= -3 \left[18p_1(p_2 - p_0)^2 + 18p_2p_0^2 - 9p_0^3 - 3p_2^2(3p_0 + 2a_1) \right. \\ &\quad \left. + 3p_1^2(6p_2 - 3p_0 - 2a_2) + 4p_0a_1a_2 \right], \\ \det(\mathbf{H}_1) &= 6p_1 - 3p_0 - 2a_1, \\ \det(\mathbf{H}_2) &= (6p_1 - 3p_0 - 2a_1)(6p_2 - 3p_0 - 2a_2).\end{aligned}$$

The point gives a global maximum if $\det(\mathbf{H}_0) \leq 0$, $\det(\mathbf{H}_1) \leq 0$ and $\det(\mathbf{H}_2) \geq 0$ at that point. For point 1, $\det(\mathbf{H}_1) = 6p_1 - 3p_0 - 2a_1 = 3\sqrt{\frac{2a_1a_2}{3}} - 2a_1 > 2\sqrt{a_1a_2} - 2a_1 \geq 0$. For point 2, $\det(\mathbf{H}_1) = a_2 - \frac{a_1}{2} \leq 0$ which contradicts the assumptions that $a_1 > 0$, $a_2 > 0$ and $a_2 \geq a_1$. Point 3 holds for $\det(\mathbf{H}_0) \leq 0$, $\det(\mathbf{H}_1) \leq 0$ and $\det(\mathbf{H}_2) \geq 0$ if $a_2/2 \geq a_1$. Point 4 holds if $a_2/2 < a_1$. For point 5, $\det(\mathbf{H}_0) = (6p_1 - 3p_0 - 2a_1)(6p_2 - 3p_0 - 2a_2)(-2a_1 - 2a_2 + 3p_0) > 0$. Thus, points 3 and 4 achieve the global maximum. ■

Proof of Proposition 3.5. For simplicity, we use p_0 and p_1 to denote p_{m0}^* and p_{m1}^* . Rewrite the expression of γ_p^c with p_0 and p_1 ,

$$\gamma_p^c = \mathbb{E} \left(\frac{X}{X+Y} \mid X+Y \geq p_0, Y \geq p_0 - p_1 \right). \quad (3.15)$$

First, we derive the joint PDF of $X+Y$ and $\frac{X}{X+Y}$ conditional on $Y \geq p_0 - p_1$ only. Denote $U = \frac{X}{X+Y}$, $V = X+Y$. Note that the following conditions must be satisfied: $0 \leq VU \leq a_1$, $p_0 - p_1 \leq V(1-U) \leq a_2$. Therefore, we have

$$f_{U,V}(u, v) = \begin{cases} \frac{v}{a_1a_2} & \text{for } \frac{p_0-p_1}{1-u} \leq v \leq \frac{a_1}{u}, \frac{a_1}{a_1+a_2} \leq u \leq \frac{a_1}{p_0-p_1+a_1}, \\ \frac{v}{a_1a_2} & \text{for } \frac{p_0-p_1}{1-u} \leq v \leq \frac{a_2}{1-u}, 0 \leq u \leq \frac{a_1}{a_1+a_2}, \\ 0 & \text{otherwise.} \end{cases} \quad (3.16)$$

$$(3.17)$$

Second, we derive the expectation of $\frac{X}{X+Y}$ conditional on both $X+Y \geq p_0$ and $Y \geq p_0-p_1$. For the range in (3.16), $\frac{p_0-p_1}{1-u} \leq v \leq \frac{a_1}{u}$, $\frac{a_1}{a_1+a_2} \leq u \leq \frac{a_1}{a_1+p_0-p_1}$, the lower bound of v is $\frac{(p_0-p_1)(a_1+a_2)}{a_2}$ if $u = \frac{a_1}{a_1+a_2}$, and the lower bound of v is $p_0 - p_1 + a_1$ if $u = \frac{a_1}{a_1+p_0-p_1}$. By Lemma 3.4(i), we find that $\frac{(p_0-p_1)(a_1+a_2)}{a_2} < p_0$ and $p_0 - p_1 + a_1 > p_0$. Thus, there exists \hat{u} such that for $\frac{a_1}{a_1+a_2} \leq u \leq \hat{u}$, $p_0 \leq v \leq \frac{a_1}{u}$, and for $\hat{u} \leq u \leq \frac{a_1}{a_1+p_0-p_1}$, $\frac{p_0-p_1}{1-u} \leq v \leq \frac{a_1}{u}$. We can obtain the value of \hat{u} by the relationship $\frac{p_0-p_1}{1-\hat{u}} = p_0$, i.e., $\hat{u} = p_1/p_0$. Denote t_1 as the integration with the first range and t_2 as the second range, that is,

$$\begin{aligned} t_1 &= \int_{\frac{a_1}{a_1+a_2}}^{\hat{u}} \int_{p_0}^{\frac{a_1}{u}} \frac{uv}{a_1 a_2} dv du = \int_{\frac{a_1}{a_1+a_2}}^{\frac{p_1}{p_0}} \int_{p_0}^{\frac{a_1}{u}} \frac{uv}{a_1 a_2} dv du \\ t_2 &= \int_{\hat{u}}^{\frac{a_1}{a_1+p_0-p_1}} \int_{\frac{p_0-p_1}{1-u}}^{\frac{a_1}{u}} \frac{uv}{a_1 a_2} dv du = \int_{\frac{p_1}{p_0}}^{\frac{a_1}{a_1+p_0-p_1}} \int_{\frac{p_0-p_1}{1-u}}^{\frac{a_1}{u}} \frac{uv}{a_1 a_2} dv du \end{aligned}$$

For the range in (3.17), $\frac{p_0-p_1}{1-u} \leq v \leq \frac{a_2}{1-u}$ and $0 \leq u \leq \frac{a_1}{a_1+a_2}$, since the lower bound of v is less than p_0 for $0 \leq u \leq \frac{a_1}{a_1+a_2}$, we can simply substitute the lower bound with p_0 . Denote t_3 as the integration with this range,

$$t_3 = \int_0^{\frac{a_1}{a_1+a_2}} \int_{p_0}^{\frac{a_2}{1-u}} \frac{uv}{a_1 a_2} dv du = \frac{1}{2a_1 a_2} \left[a_2^2 \ln \left(\frac{a_2}{a_1 + a_2} \right) + a_1 a_2 - \frac{p_0^2}{2} \left(\frac{a_1}{a_1 + a_2} \right)^2 \right].$$

Third, we derive the probability that $P(X + Y \geq p_0, Y \geq p_0 - p_1)$. Note that several conditions must be satisfied. They are $p_0 \leq v \leq a_1 + a_2$, $p_0 - p_1 \leq y \leq a_2$, $0 \leq v - y \leq a_1$. Denote the probability as t_4 , we have

$$t_4 = \int_{p_0-p_1}^{p_0} \int_{p_0}^{a_1+y} \frac{1}{a_1 a_2} dv dy + \int_{p_0}^{a_2} \int_y^{a_1+y} \frac{1}{a_1 a_2} dv dy = \frac{a_1(a_2 - p_0 + p_1) - p_1^2/2}{a_1 a_2} = \frac{D_0^*}{a_1 a_2}.$$

where D_0^* is the demand for the bundling sales under the optimal pricing strategy. As a result, we can obtain the expectation of $\frac{X}{X+Y}$ with those conditions. By (3.15), $\gamma_p^c = (\sum_{i=1}^3 t_i)/t_4$. Since $R_{m0}^* = D_0^* p_0 / (a_1 a_2)$, $\gamma_p^c R_{m0}^* = (\sum_{i=1}^3 t_i) p_0$, and thus

$$\begin{aligned} \gamma_p^c R_{m0}^* &= \frac{p_0}{2} \left[-\frac{a_2}{a_1} \ln \left(\frac{a_1 + a_2}{a_2} \right) + \frac{a_1}{a_2} \ln \left(\frac{a_1 + a_2}{a_1} \right) - \frac{a_1}{a_2} \ln \left(\frac{p_0 - p_1 + a_1}{a_1} \right) \right. \\ &\quad \left. + \frac{(p_0 - p_1)^2}{a_1 a_2} \ln \left(\frac{p_0 - p_1 + a_1}{p_0} \right) + 1 - \frac{p_1^2}{2a_1 a_2} - \frac{(a_1 - p_1)(p_0 - p_1)}{a_1 a_2} \right]. \end{aligned}$$

Then the total revenue of supplier 1 would be the fraction of the revenue in the bundling sales and the revenue in the separate sales, $R_{p1}^c = (1 - r)(\gamma_p^c R_{m0}^* + R_{m1}^*)$. We have

$$\gamma_p^c R_{m0}^* + R_{m1}^* = \frac{p_0}{2} \left[-\frac{a_2}{a_1} \ln \left(\frac{a_1 + a_2}{a_2} \right) + \frac{a_1}{a_2} \ln \left(\frac{a_1 + a_2}{a_1} \right) - \frac{a_1}{a_2} \ln \left(\frac{p_0 - p_1 + a_1}{a_1} \right) \right]$$

$$+ \frac{(p_0 - p_1)^2}{a_1 a_2} \ln \left(\frac{p_0 - p_1 + a_1}{p_0} \right) + 1 - \frac{p_1^2}{2a_1 a_2} - \frac{(a_1 - p_1)(p_0 - p_1)}{a_1 a_2} \Big] + \frac{p_1(a_1 - p_1)(p_0 - p_1)}{a_1 a_2}.$$

Divide $\gamma_p^c R_m^* + R_{m1}^*$ by R_1^* , and substitute a_2/a_1 with k ,

$$\begin{aligned} \frac{\gamma_p^c R_{m0}^* + R_{m1}^*}{R_1^*} &= \left(k + \frac{2}{3} \right) \left[-k \ln \left(1 + \frac{1}{k} \right) + \frac{1}{k} \ln \left(\frac{6 + 6k}{4 + 3k} \right) + \left(\frac{k}{4} + \frac{1}{9k} - \frac{1}{3} \right) \ln \left(\frac{3k + 4}{3k + 2} \right) \right. \\ &\quad \left. + 1 - \frac{2}{9k} - \left(\frac{1}{6} - \frac{1}{9k} \right) \right] + \frac{8}{3} \left(\frac{1}{6} - \frac{1}{9k} \right). \end{aligned} \quad (3.18)$$

We can obtain the revenue share $\mathbb{E} \left(\frac{Y}{X+Y} \mid X + Y \geq p_0, Y \geq p_0 - p_1 \right)$ by a similar manner or simply by calculating $1 - \gamma_p^c$. For supplier 2, the revenue would be a fraction of the revenue in the bundling sales, $R_{p2}^c = (1 - r)(1 - \gamma_p^c)R_{m0}^*$. We have

$$\begin{aligned} (1 - \gamma_p^c)R_{m0}^* &= \frac{p_0}{2} \left[-\frac{a_1}{a_2} \ln \left(\frac{a_1 + a_2}{a_1} \right) + \frac{a_2}{a_1} \ln \left(\frac{a_1 + a_2}{a_2} \right) + \frac{a_1}{a_2} \ln \left(\frac{p_0 - p_1 + a_1}{a_1} \right) \right. \\ &\quad \left. - \frac{(p_0 - p_1)^2}{a_1 a_2} \ln \left(\frac{a_1 + p_0 - p_1}{p_0} \right) + 1 + \frac{p_1^2}{2a_1 a_2} - \frac{(p_0 - p_1)}{a_2} - \frac{p_0 p_1}{a_1 a_2} \right]. \end{aligned}$$

Divide $(1 - \gamma_p^c)R_{m0}^*$ by R_2^* , and substitute a_2/a_1 with k ,

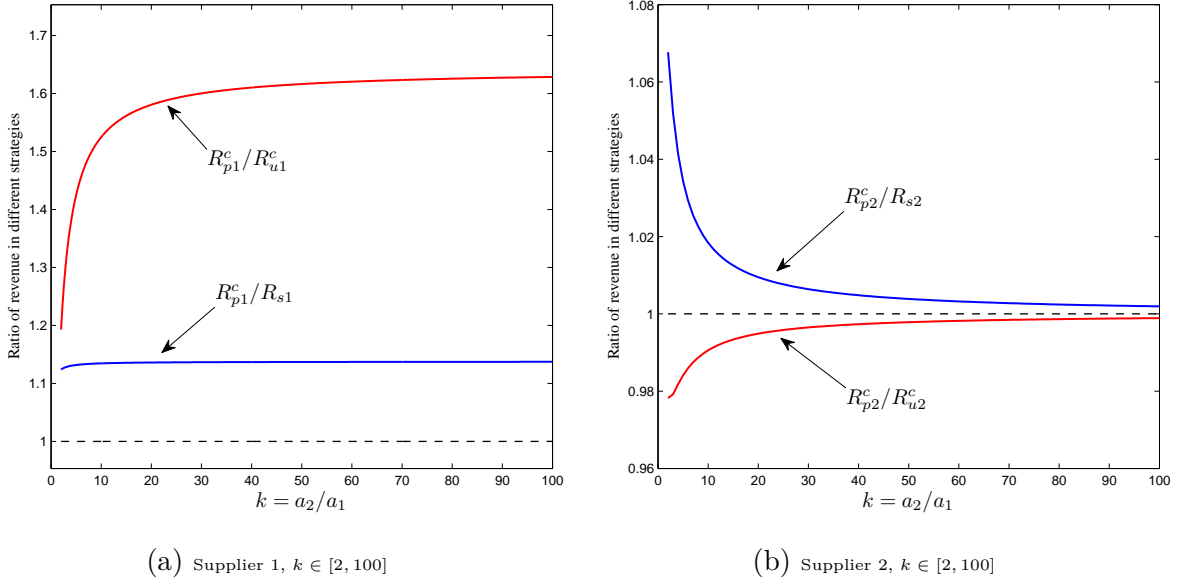
$$\begin{aligned} \frac{(1 - \gamma_p^c)R_{m0}^*}{R_2^*} &= \left(1 + \frac{2}{3k} \right) \left[k \ln \left(1 + \frac{1}{k} \right) - \frac{1}{k} \ln \left(\frac{6 + 6k}{4 + 3k} \right) \right. \\ &\quad \left. - \left(\frac{k}{4} + \frac{1}{9k} - \frac{1}{3} \right) \ln \left(\frac{3k + 4}{3k + 2} \right) + \frac{1}{6} + \frac{1}{3k} \right]. \end{aligned} \quad (3.19)$$

Since $R_{p1}^c/R_{s1} = (\gamma_p^c R_m^* + R_{m1}^*)/R_1^*$, $R_{p2}^c/R_{s2} = (1 - \gamma_p^c)R_m^*/R_2^*$, by (3.18) and (3.19), we perform the numerical test for $k = a_2/a_1 \in [2, 10^5]$. Similarly, since $R_{p1}^c/R_{u1} = ((\gamma_p^c R_{m0}^* + R_{m1}^*)/R_1^*)/(\alpha R_u^*/R_1^*)$ and $R_{p2}^c/R_{u2} = (((1 - \gamma_p^c)R_{m0}^*)/R_2^*)/((1 - \alpha)R_u^*/R_2^*)$, by (3.18), (3.19), (3.13) and (3.14), we perform the numerical test for $k = a_2/a_1 \in [2, 10^5]$. To exhibit clean results, Figure 3.5(a) and (b) only show the results for $k \in [2, 100]$. One can simply obtain the same results for $k \in [2, 10^5]$. The results in Proposition 3.5 hold. ■

Proof of Proposition 3.6. For simplicity, we use p_0, p_1 and p_2 to denote p_{m0}^*, p_{m1}^* and p_{m2}^* respectively in this proof. Rewrite the expression of γ_f^c with p_0, p_1 and p_2 ,

$$\gamma_f^c = \mathbb{E} \left(\frac{X}{X+Y} \mid X + Y \geq p_0, Y \geq p_0 - p_1, X \geq p_0 - p_2 \right). \quad (3.20)$$

Figure 3.5: Parital Mixed Bundling Sales



Denote $U = \frac{X}{X+Y}$ and $V = X+Y$. First, we derive the joint PDF of U and V conditional on $X+Y \geq p_0$, $Y \geq p_0 - p_1$ and $X \geq p_0 - p_2$. It is equivalent to derive the joint PDF of U and V if $p_0 - p_2 \leq VU \leq a_1$, $p_0 - p_1 \leq V(1-U) \leq a_2$ and $p_0 \leq V \leq a_1 + a_2$. Denote $f_{U,V}^{full}(u, v)$ as the joint PDF of events $U = u, V = v, p_0 - p_2 \leq VU \leq a_1, p_0 - p_1 \leq V(1-U) \leq a_2, p_0 \leq V \leq a_1 + a_2$. The value of the random variable U ranges from $\frac{a_1}{a_1 + p_0 - p_1}$ to $\frac{p_0 - p_2}{p_0 - p_2 + a_2}$. The upper bound and lower bound of V is determined by $\min\{\frac{a_1}{u}, \frac{a_2}{1-u}, a_1 + a_2\}$ and $\max\{\frac{p_0 - p_2}{u}, \frac{p_0 - p_1}{1-u}, p_0\}$ respectively for different realized values of U . Thus, we can obtain that

$$f_{U,V}^{full}(u, v) = \begin{cases} \frac{v}{a_1 a_2} & \text{for } \frac{p_0 - p_1}{1-u} \leq v \leq \frac{a_1}{u}, \frac{p_1}{p_0} \leq u \leq \frac{a_1}{a_1 + p_0 - p_1}, & (3.21) \\ \frac{v}{a_1 a_2} & \text{for } p_0 \leq v \leq \frac{a_1}{u}, \frac{a_1}{a_1 + a_2} \leq u \leq \frac{p_1}{p_0}, & (3.22) \\ \frac{v}{a_1 a_2} & \text{for } p_0 \leq v \leq \frac{a_2}{1-u}, \frac{p_0 - p_2}{p_0} \leq u \leq \frac{a_1}{a_1 + a_2}, & (3.23) \\ \frac{v}{a_1 a_2} & \text{for } \frac{p_0 - p_2}{u} \leq v \leq \frac{a_2}{1-u}, \frac{p_0 - p_2}{p_0 - p_2 + a_2} \leq u \leq \frac{p_0 - p_2}{p_0}, & (3.24) \\ 0 & \text{otherwise.} \end{cases}$$

With the ranges in (3.21)-(3.24), denote

$$t_1 = \int_{\frac{p_1}{p_0}}^{\frac{a_1}{a_1 + p_0 - p_1}} \int_{\frac{p_0 - p_1}{1-u}}^{\frac{a_1}{u}} \frac{uv}{a_1 a_2} dv du, \quad t_2 = \int_{\frac{a_1}{a_1 + a_2}}^{\frac{p_1}{p_0}} \int_{p_0}^{\frac{a_1}{u}} \frac{uv}{a_1 a_2} dv du,$$

$$t_3 = \int_{\frac{p_0-p_2}{p_0}}^{\frac{a_1}{a_1+a_2}} \int_{p_0}^{\frac{a_2}{1-u}} \frac{uv}{a_1 a_2} dv du, \quad t_4 = \int_{\frac{p_0-p_2}{p_0-p_2+a_2}}^{\frac{p_0-p_2}{p_0}} \int_{\frac{p_0-p_2}{u}}^{\frac{a_2}{1-u}} \frac{uv}{a_1 a_2} dv du.$$

Moreover, denote $t_5 = P(p_0 - p_2 \leq VU \leq a_1, p_0 - p_1 \leq V(1-U) \leq a_2, p_0 \leq V \leq a_1 + a_2)$. Then by (3.20), we have $\gamma_f^c = (\sum_{i=1}^4 t_i)/t_5$. It is straightforward to show that $t_5 = D_0^*/(a_1 a_2)$. Since $R_{m0}^* = p_0 D_0^*/(a_1 a_2)$, the revenue share of supplier 1 in the bundling sales is $\gamma_f^c R_{m0}^* = (\sum_{i=1}^4 t_i)/t_5 \cdot D_0^*/(a_1 a_2) \cdot p_0 = (\sum_{i=1}^4 t_i)p_0$, i.e.,

$$\begin{aligned} \gamma_f^c R_{m0}^* &= \frac{p_0}{2a_1 a_2} \left[a_2^2 \ln \left(\frac{a_2}{a_1 + a_2} \right) \right. \\ &\quad - a_1^2 \ln \left(\frac{a_1}{a_1 + a_2} \right) + a_2^2 \ln \left(\frac{p_0 - p_2 + a_2}{a_2} \right) - a_1^2 \ln \left(\frac{p_0 - p_1 + a_1}{a_1} \right) \\ &\quad - (p_0 - p_2)^2 \ln \left(\frac{p_0 - p_2 + a_2}{p_0} \right) + (p_0 - p_1)^2 \ln \left(\frac{p_0 - p_1 + a_1}{p_0} \right) \\ &\quad \left. + a_2(a_1 - p_0 + p_2) + \frac{(p_0 - p_2)^2}{2} - \frac{p_1^2}{2} + (p_0 - p_1)(p_1 - a_1) \right]. \end{aligned}$$

Then, we can obtain $(1 - \gamma_f^c)R_{m0}^*$ by the similar process,

$$\begin{aligned} (1 - \gamma_f^c)R_{m0}^* &= \frac{p_0}{2a_1 a_2} \left[-a_2^2 \ln \left(\frac{a_2}{a_1 + a_2} \right) \right. \\ &\quad + a_1^2 \ln \left(\frac{a_1}{a_1 + a_2} \right) - a_2^2 \ln \left(\frac{p_0 - p_2 + a_2}{a_2} \right) + a_1^2 \ln \left(\frac{p_0 - p_1 + a_1}{a_1} \right) \\ &\quad + (p_0 - p_2)^2 \ln \left(\frac{p_0 - p_2 + a_2}{p_0} \right) - (p_0 - p_1)^2 \ln \left(\frac{p_0 - p_1 + a_1}{p_0} \right) \\ &\quad \left. + a_1(a_2 - p_0 + p_1) + \frac{(p_0 - p_1)^2}{2} - \frac{p_2^2}{2} + (p_0 - p_2)(p_2 - a_2) \right]. \end{aligned}$$

The total revenue for supplier 1 is the fraction of the revenue in the bundling sales and the separate sales, i.e., $R_{f1}^c = (1-r)(\gamma_f^c R_{m0}^* + R_{m1}^*)$. By Lemma 3.4, $R_{m1}^* = \frac{p_1(a_1-p_1)(p_0-p_1)}{a_1 a_2}$ for the full mixed bundling sales, thus

$$\begin{aligned} \gamma_f^c R_{m0}^* + R_{m1}^* &= \frac{p_0}{2} \left[\frac{a_2}{a_1} \ln \left(\frac{a_2}{a_1 + a_2} \right) - \frac{a_1}{a_2} \ln \left(\frac{a_1}{a_1 + a_2} \right) + \frac{a_2}{a_1} \ln \left(\frac{p_0 - p_2 + a_2}{a_2} \right) \right. \\ &\quad - \frac{a_1}{a_2} \ln \left(\frac{p_0 - p_1 + a_1}{a_1} \right) - \frac{(p_0 - p_2)^2}{a_1 a_2} \ln \left(\frac{p_0 - p_2 + a_2}{p_0} \right) \\ &\quad + \frac{(p_0 - p_1)^2}{a_1 a_2} \ln \left(\frac{p_0 - p_1 + a_1}{p_0} \right) + \frac{(a_1 - p_0 + p_2)}{a_1} + \frac{(p_0 - p_2)^2}{2a_1 a_2} \\ &\quad \left. - \frac{p_1^2}{2a_1 a_2} + \frac{(p_0 - p_1)(p_1 - a_1)}{a_1 a_2} \right] + \frac{p_1(a_1 - p_1)(p_0 - p_1)}{a_1 a_2}. \end{aligned}$$

Now we compare the revenue of supplier 1 in the full mixed bundling and separate sales. Divide $\gamma_f^c R_{m0}^* + R_{m1}^*$ by $R_1^* = a_1/4$ and substitute a_2/a_1 with k ,

$$\begin{aligned} \frac{\gamma_f^c R_{m0}^* + R_{m1}^*}{R_1^*} &= \left[-k \ln \left(1 + \frac{1}{k} \right) + \frac{1}{k} \ln(1+k) + k \ln \left(1 + \frac{2}{3k} - \frac{1}{3} \sqrt{\frac{2}{k}} \right) \right. \\ &\quad - \frac{1}{k} \ln \left(1 + \frac{2k}{3} - \frac{\sqrt{2k}}{3} \right) - \left(\frac{4}{9k} - \frac{4}{9} \sqrt{\frac{2}{k}} + \frac{2}{9} \right) \ln \left(\frac{2 - \sqrt{2k} + 3k}{2 - \sqrt{2k} + 2k} \right) \\ &\quad + \left(\frac{4k}{9} - \frac{4\sqrt{2k}}{9} + \frac{2}{9} \right) \ln \left(\frac{2k - \sqrt{2k} + 3}{2k - \sqrt{2k} + 2} \right) + \frac{2}{9} - \frac{1}{9} \sqrt{\frac{2}{k}} \\ &\quad \left. + \frac{\sqrt{2k}}{3} \right] \left(\frac{4}{3} + \frac{4k}{3} - \frac{2\sqrt{2k}}{3} \right) + \frac{8}{3} \left(\frac{2}{9} - \frac{1}{9} \sqrt{\frac{2}{k}} \right). \end{aligned} \quad (3.25)$$

Similarly, for supplier 2, the total revenue is $R_{f2}^c = (1-r)((1-\gamma_f^c)R_{m0}^* + R_{m2}^*)$. By Lemma 3.4, $R_{m2}^* = \frac{p_2(a_2-p_2)(p_0-p_2)}{a_1 a_2}$ for the full mixed bundling sales, thus

$$\begin{aligned} (1-\gamma_f^c)R_{m0}^* + R_{m2}^* &= \frac{p_0}{2} \left[-\frac{a_2}{a_1} \ln \left(\frac{a_2}{a_1 + a_2} \right) + \frac{a_1}{a_2} \ln \left(\frac{a_1}{a_1 + a_2} \right) - \frac{a_2}{a_1} \ln \left(\frac{p_0 - p_2 + a_2}{a_2} \right) \right. \\ &\quad + \frac{a_1}{a_2} \ln \left(\frac{p_0 - p_1 + a_1}{a_1} \right) + \frac{(p_0 - p_2)^2}{a_1 a_2} \ln \left(\frac{p_0 - p_2 + a_2}{p_0} \right) \\ &\quad - \frac{(p_0 - p_1)^2}{a_1 a_2} \ln \left(\frac{p_0 - p_1 + a_1}{p_0} \right) + \frac{(a_2 - p_0 + p_1)}{a_2} + \frac{(p_0 - p_1)^2}{2a_1 a_2} \\ &\quad \left. - \frac{p_2^2}{2a_1 a_2} + \frac{(p_0 - p_2)(p_2 - a_2)}{a_1 a_2} \right] + \frac{p_2(a_2 - p_2)(p_0 - p_2)}{a_1 a_2}. \end{aligned}$$

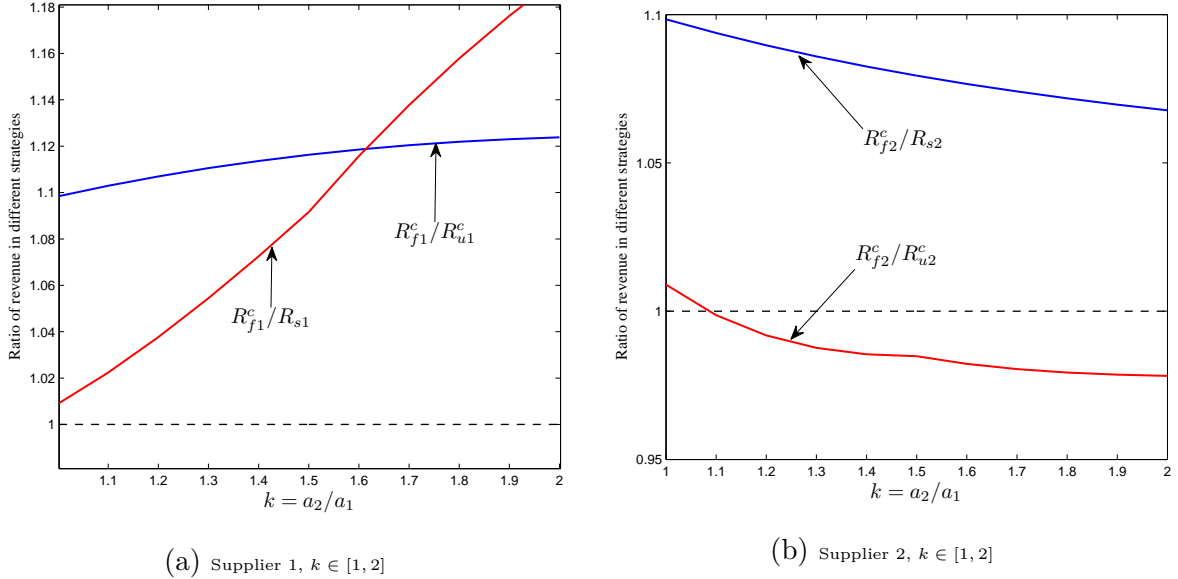
Divide $(1-\gamma_f^c)R_{m0}^* + R_{m2}^*$ by $R_2^* = a_2/4$ and substitute a_2/a_1 with k ,

$$\begin{aligned} &\frac{(1-\gamma_f^c)R_{m0}^* + R_{m2}^*}{R_2^*} \\ &= \left[k \ln \left(1 + \frac{1}{k} \right) - \frac{1}{k} \ln(1+k) - k \ln \left(1 + \frac{2}{3k} - \frac{1}{3} \sqrt{\frac{2}{k}} \right) + \frac{1}{k} \ln \left(1 + \frac{2k}{3} - \frac{\sqrt{2k}}{3} \right) \right. \\ &\quad \left(\frac{4}{9k} - \frac{4}{9} \sqrt{\frac{2}{k}} + \frac{2}{9} \right) \ln \left(\frac{2 - \sqrt{2k} + 3k}{2 - \sqrt{2k} + 2k} \right) - \left(\frac{4k}{9} - \frac{4\sqrt{2k}}{9} + \frac{2}{9} \right) \ln \left(\frac{2k - \sqrt{2k} + 3}{2k - \sqrt{2k} + 2} \right) \\ &\quad \left. + \frac{2}{9} - \frac{\sqrt{2k}}{9} + \frac{1}{3} \sqrt{\frac{2}{k}} \right] \left(\frac{4}{3} + \frac{4}{3k} - \frac{2}{3} \sqrt{\frac{2}{k}} \right) + \frac{8}{3} \left(\frac{2}{9} - \frac{\sqrt{2k}}{9} \right). \end{aligned} \quad (3.26)$$

Since $R_{f1}^c/R_{s1} = (\gamma_f^c R_{m0}^* + R_{m1}^*)/R_1^*$, $R_{f2}^c/R_{s2} = (1-\gamma_f^c)R_{m0}^*/R_2^*$, by (3.25) and (3.26), we perform the numerical test for $k = a_2/a_1 \in [1, 2]$. Similarly, since $R_{f1}^c/R_{u1} = ((\gamma_f^c R_{m0}^* + R_{m1}^*)/R_1^*)/(\alpha R_u^*/R_1^*)$ and $R_{f2}^c/R_{u2} = (((1-\gamma_f^c)R_{m0}^*)/R_2^*)/((1-\alpha)R_u^*/R_2^*)$, by (3.25),

(3.26), (3.10), (3.11), (3.13) and (3.14), we perform the numerical test for $k = a_2/a_1 \in [1, 2]$. The results in Proposition 3.6 hold. ■ *Proof of Proposition 3.3.* First, we derive

Figure 3.6: Full Mixed Bundling Sales



the joint distribution of X and $X + Y$. Recall that $V = X + Y$. Then the inverse transformation is given by $Y = V - X$. The joint density function of X and V is $f_{X,V}(x, v) = f_{X,Y}(x, y)|J_a(x, y)|^{-1}$ where $J_a(x, y)$ denotes the Jacobian of the functions $X = X$ and $V = X + Y$. As a result, the joint distribution of X and V is

$$f_{X,V}(x, v) = \begin{cases} \frac{1}{a_1 a_2}, & \text{for } 0 \leq x \leq a_1 \text{ and } 0 \leq v - x \leq a_2, \\ 0, & \text{otherwise.} \end{cases} \quad (3.27)$$

The ranges can be written as $0 \leq x \leq a_1$ and $x \leq v \leq a_2 + x$. If $0 < p_0^* \leq a_1$, $\mathbb{E}(X|X + Y \geq p_0^*) = \frac{1}{1-H(p_0^*)} \left[\int_{p_0^*}^{a_1} \int_x^{a_2+x} x f_{X,V}(x, v) dv dx + \int_0^{p_0^*} \int_{p_0^*-x}^{a_2+x} x f_{X,V}(x, v) dv dx \right] = \frac{1}{1-H(p_0^*)} \left(\frac{a_1}{2} - \frac{p_0^{*3}}{6a_1 a_2} \right)$. By the similar method, if $0 < p_0^* < a_1$, we have $\mathbb{E}(Y|X + Y \geq p_0^*) = \frac{1}{1-H(p_0^*)} \left(\frac{a_2}{2} - \frac{p_0^{*3}}{6a_1 a_2} \right)$.

As a result, if $0 < p_0^* < a_1$, the pre-committed revenue sharing policy β is

$$\beta = \frac{\mathbb{E}(X|X + Y \geq p_0^*)}{\mathbb{E}(X|X + Y \geq p_0^*) + \mathbb{E}(Y|X + Y \geq p_0^*)} = \frac{3a_1^2 a_2 - p_0^{*3}}{3a_1^2 a_2 + 3a_1 a_2^2 - 2p_0^{*3}}. \quad (3.28)$$

By Lemma 3.1, if $a_1 \leq a_2 < \frac{3}{2}a_1$, $p_0^* = \sqrt{\frac{2a_1 a_2}{3}} \in [0, a_1)$. For product A_1 , the revenue is $R_{u_1}^p = (1-r)\beta(1-H(p_0^*))p_0^* = \frac{3a_1^2 a_2 - p_0^{*3}}{3a_1^2 a_2 + 3a_1 a_2^2 - 2p_0^{*3}} \cdot \frac{2p_0^*}{3}$. By substituting a_2/a_1 with k ,

we have

$$\frac{\beta R_u^*}{R_1^*} = \left(\frac{3 - \frac{2}{3}\sqrt{\frac{2k}{3}}}{3 + 3k - \frac{4}{3}\sqrt{\frac{2k}{3}}} \right) \frac{8}{3} \sqrt{\frac{2k}{3}}, \quad \frac{(1-\beta)R_u^*}{R_2^*} = \left(\frac{3k - \frac{2}{3}\sqrt{\frac{2k}{3}}}{3 + 3k - \frac{4}{3}\sqrt{\frac{2k}{3}}} \right) \frac{8}{3} \sqrt{\frac{2}{3k}}. \quad (3.29)$$

If $a_1 \leq p_0^* \leq a_2$, we have

$$\begin{aligned} & \mathbb{E}(X|X+Y \geq p_0^*) \\ &= \frac{1}{1-H(p_0^*)} \int_0^{a_1} \int_{p_0^*}^{a_2+x} x f_{X,V}(x,v) dv dx = \frac{1}{1-H(p_0^*)} \left(\frac{a_1}{2} - \frac{a_1 p_0^*}{2a_2} + \frac{a_1^2}{3a_2} \right), \\ & \mathbb{E}(Y|X+Y \geq p_0^*) \\ &= \frac{1}{1-H(p_0^*)} \left[\int_{p_0^*}^{a_2} \int_y^{a_1+y} y f_{Y,V}(y,v) dv dy + \int_{p_0^*-a_1}^{p_0^*} \int_{p_0^*}^{a_1+y} y f_{Y,V}(y,v) dv dy \right] \\ &= \frac{1}{1-H(p_0^*)} \left(\frac{a_2}{2} - \frac{p_0^{*2}}{2a_2} + \frac{a_1 p_0^*}{2a_2} - \frac{a_1^2}{6a_2} \right). \end{aligned}$$

Under the pre-committed revenue sharing policy, the fraction β is given by

$$\beta = \frac{\mathbb{E}(X|X+Y \geq p_0^*)}{\mathbb{E}(X|X+Y \geq p_0^*) + \mathbb{E}(Y|X+Y \geq p_0^*)} = \frac{3a_1 a_2 - 3a_1 p_0^* + 2a_1^2}{3a_1 a_2 + 3a_2^2 + a_1^2 - 3p_0^{*2}}. \quad (3.30)$$

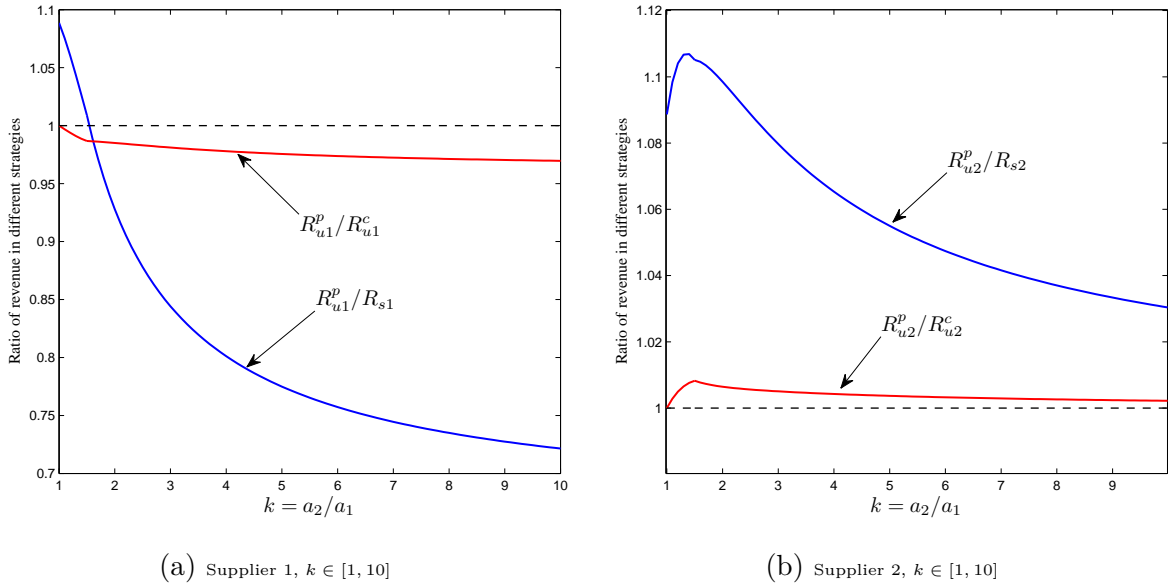
By Lemma 3.1, if $a_2 \geq \frac{3}{2}a_1$, $p_0^* = \frac{2a_2+a_1}{4} \in [a_1, a_2]$,

$$\begin{aligned} \frac{\beta R_u^*}{R_1^*} &= \left(\frac{24k+20}{36k^2+36k+13} \right) \frac{4k^2+4k+1}{4k}, \\ \frac{(1-\beta)R_u^*}{R_2^*} &= \left(\frac{36k^2+12k-7}{36k^2+36k+13} \right) \frac{4k^2+4k+1}{4k^2}. \end{aligned} \quad (3.31)$$

Since $R_{u1}^p/R_{s1} = \beta R_u^*/R_1^*$ and $R_{u2}^p/R_{s2} = (1-\beta)R_u^*/R_2^*$, by (3.29) and (3.31), we perform the numerical test for $a_2/a_1 \in [1, 10^5]$. Since $R_{u1}^p/R_{u1}^c = \beta/\alpha$ and $R_{u2}^p/R_{u2}^c = (1-\beta)/(1-\alpha)$, by (3.28), (3.30), (3.6), (3.7), (3.8) and (3.9), we perform the numerical test for $a_2/a_1 \in [1, 10^5]$. To exhibit clean results, Figure 3.7a and 3.7b only show the results for $k = a_2/a_1 \in [1, 10]$. One can simply show the same results for $k \in [1, 10^5]$. ■

Proof of Lemma 3.7. (i) Recall that X follows the uniform distribution over range $[0, 1]$ with CDF $F(\cdot)$ and PDF $f(\cdot)$, and Y follows the uniform distribution over $[a, b]$ where $a \geq 0$ and $b \leq 1$ with CDF $G(\cdot)$ and PDF $g(\cdot)$. Note that we examine the case in which $a+b=1$. The optimal pricing strategy for supplier 1 is by solving the FOC $[1-F(p_1^*)]/f(p_1^*) = p_1^*$ in which $F(p_1^*) = p_1^*$ and $f(p_1^*) = 1$. Thus, we have $p_1^* = 1/2$. The optimal pricing strategy for supplier 2 is by solving the FOC $[1-G(p_2^*)]/g(p_2^*) = p_2^*$ in

Figure 3.7: Pre-committed vs. Contingent



which $G(p_2^*) = \frac{p_2^* - a}{b - a}$ and $g(p_2^*) = \frac{1}{b - a}$. Then we have $p_2^* = b/2$. The price should be greater than or equal to the bottom value of the product 2, $p_2^* \geq a$, thus $p_2^* = \max\{b/2, a\}$.

(ii) The PDF of the random variable $Z = X + Y$ is given by

$$h(z) = (f \circ g)(z) = \int_0^z f(z - y)g(y)dy = \begin{cases} \frac{z - a}{b - a}, & \text{for } a \leq z \leq b, \\ 1, & \text{for } b < z \leq a + 1, \\ \frac{b + 1 - z}{b - a}, & \text{for } a + 1 < z \leq b + 1. \end{cases}$$

Its CDF is given by

$$H(z) = \begin{cases} \frac{(z - a)^2}{2(b - a)}, & \text{for } a \leq z \leq b, \\ \frac{(z - a)^2 - (z - b)^2}{2(b - a)}, & \text{for } b < z \leq a + 1, \\ 1 - \frac{(b + 1 - z)^2}{2(b - a)}, & \text{for } a + 1 < z \leq b + 1. \end{cases}$$

If the price for the bundle is p_0 , the demand is $1 - H(p_0)$ and the expected revenue for the bundling sales is $R = (1 - H(p_0))p_0$. By the FOC, the optimal pricing strategy is satisfied by $p_0^* = (1 - H(p_0^*))/h(p_0^*)$. For $p_0^* > b$, $p_0^* = (1 - H(p_0^*))/h(p_0^*)$ is equivalent to $p_0^* = 1 - \frac{(p_0^* - a)^2 - (p_0^* - b)^2}{2(b - a)}$, which yields $p_0^* = 1/2 + (a + b)/4$. Since $a + b = 1$, we have

$p_0^* = 3/4$. It is simple to verify that $p_0^* = 3/4$ is the global maximum point if $b < 3/4$.

For $p_0^* \leq b$, $p_0^* = (1 - H(p_0^*))/h(p_0^*)$ is equivalent to $p_0^* = [1 - \frac{(p_0^*-a)^2}{2(b-a)}] \frac{p_0^*-a}{b-a}$, which yields $p_0^* = \frac{2a + \sqrt{a^2 - 6a + 6b}}{3}$. Now we verify that p_0^* is indeed less than or equal to b : $b \geq 3/4, a + b = 1 \Leftrightarrow (3b - a) \geq 2 \Leftrightarrow (3b - a)(b - a) \geq 2(b - a) \Leftrightarrow a^2 - 4ab + 3b^2 + 2a - 2b \geq 0 \Leftrightarrow 3b - 2a \geq \sqrt{a^2 - 6a + 6b} \Leftrightarrow b \geq \frac{2a + \sqrt{a^2 - 6a + 6b}}{3}$. It is simple to verify that $p_0^* = \frac{2a + \sqrt{a^2 - 6a + 6b}}{3}$ is indeed the global maximum point if $b \geq 3/4$. ■

Proof of Proposition 3.8. First, we derive the joint distribution of $\frac{X}{X+Y}$ and $X+Y$ by following the same way in the proof of Proposition 3.2. With exactly the same method in the proof of Proposition 3.2 from the beginning to (3.3), we directly give the joint PDF of $U = \frac{X}{X+Y}$ and $V = X+Y$

$$f_{U,V}(u, v) = (x + y)f(x)g(y) = \begin{cases} \frac{v}{b-a}, & \text{for } 0 \leq vu \leq 1, a \leq v(1-u) \leq b, \\ 0, & \text{otherwise.} \end{cases}$$

Note that random variable $U \in [0, 1]$. If there is a condition that $V \geq p_0$, the conditional PDF is

$$f_{U|V}(u|V \geq p_0) = \frac{\Pr(U = u, V \geq p_0)}{\Pr(V \geq p_0)} = \frac{\int_{p_0}^{\infty} f_{U,V}(u, v)dv}{1 - H(p_0)}.$$

And thus, the conditional expectation is

$$\alpha_{pn} = \mathbb{E}\left(\frac{X}{X+Y} | X+Y \geq p_0\right) = \frac{1}{1 - H(p_0)} \int_0^1 \int_{p_0}^{\infty} f_{U,V}(u, v)dvdu.$$

Since the optimal total revenue is $R_u^* = (1 - H(p_0^*))p_0^*$, the revenue for supplier 1 in the pure bundling sales is $R_{u1}^c = \alpha_{pn}R_u^*(1 - r) = (1 - r)p_0^* \int_0^1 \int_{p_0}^{\infty} f_{U,V}(u, v)dvdu$. And the revenue for supplier 2 in the pure bundling sales is $R_{u2}^c = (1 - \alpha_{pn})R_u^*(1 - r)$.

Second, we derive the value of $R_{u1}^c = (1 - r)p_0^* \int_0^1 \int_{p_0}^{\infty} f_{U,V}(u, v)dvdu$ in which p_0^* is the optimal pricing strategy in Lemma 3.7. The ranges $0 \leq vu \leq 1$ and $a \leq v(1-u) \leq b$ are equivalent to $0 \leq v \leq \frac{1}{u}$ and $\frac{a}{1-u} \leq v \leq \frac{b}{1-u}$. Since $\frac{a}{1-u} \geq 0$, we have $\frac{a}{1-u} \leq v \leq \frac{1}{u}$ and $\frac{a}{1-u} \leq v \leq \frac{b}{1-u}$. Only one or other of these ranges needs to be retained, depending on whether u is over $[\frac{1}{b+1}, \frac{1}{a+1}]$ or $[0, \frac{1}{b+1}]$. Thus,

$$f_{U,V}(u, v) = \begin{cases} \frac{v}{b-a}, & \text{for } \frac{a}{1-u} \leq v \leq \frac{b}{1-u}, 0 \leq u \leq \frac{1}{b+1}, \\ \frac{v}{b-a}, & \text{for } \frac{a}{1-u} \leq v \leq \frac{1}{u}, \frac{1}{b+1} \leq u \leq \frac{1}{a+1}, \\ 0 & \text{otherwise.} \end{cases} \quad (3.32)$$

$$f_{U,V}(u, v) = \begin{cases} \frac{v}{b-a}, & \text{for } \frac{a}{1-u} \leq v \leq \frac{1}{u}, \frac{1}{b+1} \leq u \leq \frac{1}{a+1}, \\ 0 & \text{otherwise.} \end{cases} \quad (3.33)$$

The customer whose value $v \geq p_0^*$ will purchase the bundle. Now we discuss what the

lower bound of v should be. By Lemma 3.7, if $b > 3/4$, $p_0^* = \frac{2a+\sqrt{a^2-6a+6b}}{3}$. We find that if $b > 3/4$, then $\frac{a(b+1)}{b} < 3/4 \leq p_0^* = \frac{2a+\sqrt{a^2-6a+6b}}{3}$, so the ranges in (3.32) become $p_0^* \leq v \leq \frac{b}{1-u}$, $0 \leq u \leq \frac{1}{b+1}$. The ranges in (3.33) become $p_0^* \leq v \leq \frac{1}{u}$, $\frac{1}{b+1} \leq u \leq 1 - \frac{a}{p_0^*}$ and $\frac{a}{1-u} \leq v \leq \frac{1}{u}$, $1 - \frac{a}{p_0^*} \leq u \leq \frac{1}{a+1}$. Therefore, we have

$$\begin{aligned}
R_{u1}^c &= (1-r)p_0^* \int_0^1 \int_{p_0^*}^{\infty} uvf(uv)g(v-uv)dvdu \\
&= (1-r)p_0^* \int_0^{\frac{1}{b+1}} \int_{p_0^*}^{\frac{b}{1-u}} \frac{uv}{b-a} dvdu + \int_{\frac{1}{b+1}}^{1-\frac{a}{p_0^*}} \int_{p_0^*}^{\frac{1}{u}} \frac{uv}{b-a} dvdu + \int_{1-\frac{a}{p_0^*}}^{\frac{1}{a+1}} \int_{\frac{a}{1-u}}^{\frac{1}{u}} \frac{uv}{b-a} dvdu \\
&= \frac{(1-r)p_0^*}{2(b-a)} \left[b^2 \ln\left(\frac{b}{b+1}\right) - a^2 \ln\left(\frac{a}{a+1}\right) - \ln(a+1) + \ln(b+1) + a^2 \ln\left(\frac{a}{p_0^*}\right) \right. \\
&\quad \left. + b - a - \frac{p_0^{*2}}{2} - \frac{3a^2}{2} + 2ap_0^* \right], \tag{3.34}
\end{aligned}$$

where $p_0^* = \frac{2a+\sqrt{a^2-6a+6b}}{3}$.

If $b \leq 3/4$, then by Lemma 3.7, $p_0^* = 3/4$. The lower bound of v in (3.32) depends on whether $\frac{a(b+1)}{b}$ is greater than $3/4$. By $a+b=1$ and $\frac{a(b+1)}{b} = 3/4$, we can obtain an $\hat{b} \leq 3/4$ such that if $b \leq \hat{b}$, $\frac{a(b+1)}{b} \geq 3/4$, and if $b \geq \hat{b}$, $\frac{a(b+1)}{b} \leq 3/4$.

If $3/4 \geq b \geq \hat{b}$, $p_0^* = 3/4 \geq b$, and the ranges in (3.32) become $p_0^* \leq v \leq \frac{b}{1-u}$ and $1 - \frac{b}{p_0^*} \leq u \leq \frac{1}{b+1}$. The ranges in (3.33) become $p_0^* \leq v \leq \frac{1}{u}$ and $\frac{1}{b+1} \leq u \leq 1 - \frac{a}{p_0^*}$, and $\frac{a}{1-u} \leq v \leq \frac{1}{u}$ and $1 - \frac{a}{p_0^*} \leq u \leq \frac{1}{a+1}$. By substituting p_0^* with $3/4$, we can obtain

$$\begin{aligned}
R_{u1}^c &= (1-r)p_0^* \int_0^1 \int_{p_0^*}^{\infty} uvf(uv)g(v-uv)dvdu \\
&= (1-r)p_0^* \int_{1-\frac{4b}{3}}^{\frac{1}{b+1}} \int_{3/4}^{\frac{b}{1-u}} \frac{uv}{b-a} dvdu + \int_{\frac{1}{b+1}}^{1-\frac{4a}{3}} \int_{3/4}^{\frac{1}{u}} \frac{uv}{b-a} dvdu + \int_{1-\frac{4a}{3}}^{\frac{1}{a+1}} \int_{\frac{a}{1-u}}^{\frac{1}{u}} \frac{uv}{b-a} dvdu \\
&= \frac{3(1-r)}{8(b-a)} \left[b^2 \ln\left(\frac{b}{b+1}\right) - a^2 \ln\left(\frac{a}{a+1}\right) - \ln(a+1) + \ln(b+1) \right. \\
&\quad \left. + a^2 \ln\left(\frac{4a}{3}\right) - b^2 \ln\left(\frac{4b}{3}\right) + \frac{3b^2}{2} - \frac{b}{2} - \frac{3a^2}{2} + \frac{a}{2} \right]. \tag{3.35}
\end{aligned}$$

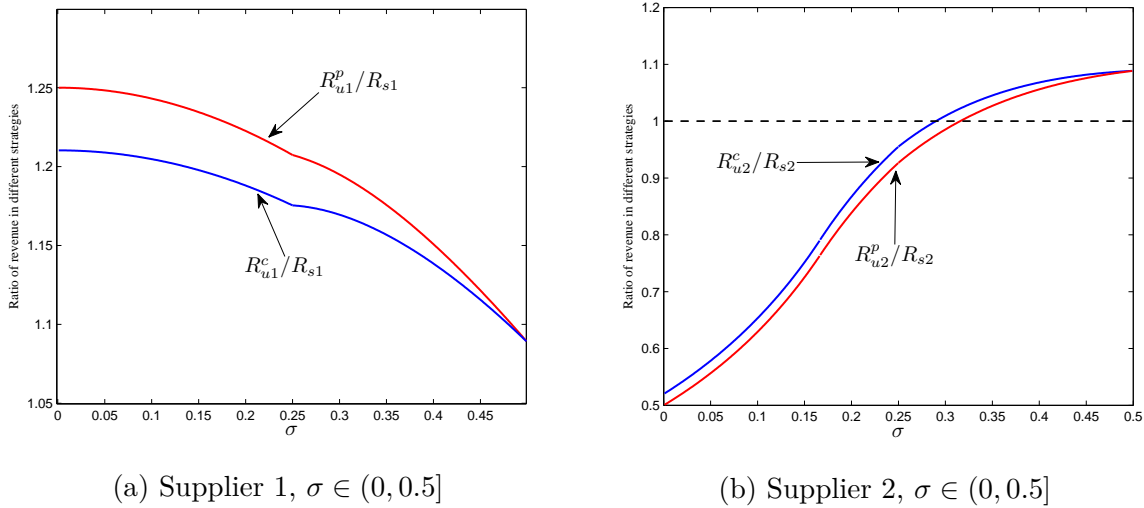
If $b \leq \hat{b}$, $\frac{a(b+1)}{b} \leq 3/4$, and then the ranges in (3.32) become $p_0^* \leq v \leq \frac{b}{1-u}$, $1 - \frac{b}{p_0^*} \leq u \leq 1 - \frac{a}{p_0^*}$, and $\frac{a}{1-u} \leq v \leq \frac{b}{1-u}$, $1 - \frac{b}{p_0^*} \leq u \leq \frac{1}{b+1}$. The ranges in (3.33) become $\frac{a}{1-u} \leq v \leq \frac{1}{u}$, $\frac{1}{b+1} \leq u \leq \frac{1}{a+1}$. By substituting p_0^* with $3/4$, we can obtain

$$R_{u1}^c = (1-r)p_0^* \int_0^1 \int_{p_0^*}^{\infty} uvf(uv)g(v-uv)dvdu$$

$$\begin{aligned}
 &= (1-r)p_0^* \int_{1-\frac{4b}{3}}^{1-\frac{4a}{3}} \int_{p_0^*}^{\frac{b}{1-u}} \frac{uv}{b-a} dvdu + \int_{1-\frac{4a}{3}}^{\frac{1}{b+1}} \int_{\frac{a}{1-u}}^{\frac{b}{1-u}} \frac{uv}{b-a} dvdu + \int_{\frac{1}{b+1}}^{\frac{1}{a+1}} \int_{\frac{a}{1-u}}^{\frac{1}{1-u}} \frac{uv}{b-a} dvdu \\
 &= \frac{3(1-r)}{8(b-a)} \left[b^2 \ln\left(\frac{b}{b+1}\right) - a^2 \ln\left(\frac{a}{a+1}\right) - \ln(a+1) + \ln(b+1) + a^2 \ln\left(\frac{4a}{3}\right) \right. \\
 &\quad \left. - b^2 \ln\left(\frac{4b}{3}\right) + \frac{3b^2}{2} - \frac{b}{2} - \frac{3a^2}{2} + \frac{a}{2} \right]. \tag{3.36}
 \end{aligned}$$

The revenue of supplier 2 R_{u2}^c can be obtained by the similar method or simply $R_{u2}^c = (1-r)R_u^* - R_{u1}^c$. Now we can compare the revenue of suppliers 1 and 2 between the separate and pure bundling sales. By Lemma 3.7(i), for the separate sales, supplier 1's revenue under the optimal pricing strategy is $R_{s1} = (1-r)/4$. Supplier 2's revenue is $R_{s2} = (1-r)a$ if $b/2 \leq a$ (i.e., $\sigma \leq 1/3$), and $R_{s2} = \frac{(1-r)b^2}{4(b-a)}$ if $b/2 \geq a$ (i.e., $\sigma \geq 1/3$). Since r does not play a role in R_{u1}^c/R_{s1} and R_{u2}^c/R_{s2} , and by the assumption $a+b=1$, those functions about a and b can be presented by a single variable a or b or σ . We show the numerical results for R_{u1}^c/R_{s1} and R_{u2}^c/R_{s2} in Figure 3.8. We find that $R_{u1}^c \geq R_{s1}$. There exists a threshold $\bar{\sigma}$ such that if $\sigma \geq \bar{\sigma}$, $R_{u2}^c \geq R_{s2}$, and if $\sigma < \bar{\sigma}$, $R_{u2}^c < R_{s2}$.

Figure 3.8: Comparison



We derive the joint distribution of X and $X + Y$. Recall that $V = X + Y$. Then the inverse transformation is given by $Y = V - X$. The joint density function of X and V is $f_{X,V}(x, v) = f_{X,Y}(x, y)|J_a(x, y)|^{-1}$ where $J_a(x, y)$ denotes the Jacobian of the functions $X = X$ and $V = X + Y$. As a result, the joint distribution of X and V is

$$f_{X,V}(x, v) = \begin{cases} \frac{1}{b-a}, & \text{for } 0 \leq x \leq 1 \text{ and } a \leq v - x \leq b, \\ 0, & \text{otherwise.} \end{cases} \tag{3.37}$$

The ranges can be written as $0 \leq x \leq 1$ and $a + x \leq v \leq b + x$. The conditional expectation is $\mathbb{E}(X|X + Y \geq p_0^*) = \frac{1}{1-H(p_0^*)} \int_0^1 \int_{p_0^*}^{\infty} x f_{X,V}(x, v) dv dx$. Now we derive the $\int_0^1 \int_{p_0^*}^{\infty} x f_{X,V}(x, v) dv dx$ with different values of b .

If $b < 3/4$, by Lemma 3.7, $b < p_0^* = 3/4$. Therefore,

$$\begin{aligned} \int_0^1 \int_{p_0^*}^{\infty} x f_{X,V}(x, v) dv dx &= \int_{p_0^*-b}^{p_0^*-a} \int_{p_0^*}^{b+x} \frac{x}{b-a} dv dx + \int_{p_0^*-a}^1 \int_{a+x}^{b+x} \frac{x}{b-a} dv dx \\ &= \frac{(p_0^* - b)^3}{6(b-a)} - \frac{(p_0^* - a)^3}{6(b-a)} + \frac{1}{2}, \end{aligned} \quad (3.38)$$

where $p_0^* = 3/4$. Now we derive $\mathbb{E}(Y|X + Y \geq p_0^*) = \frac{1}{1-H(p_0^*)} \int_0^1 \int_{p_0^*}^{\infty} y f_{X,V}(x, v) dv dy$. If $b < p_0^* = 3/4$, by the similar method above, we have

$$\begin{aligned} \int_0^1 \int_{p_0^*}^{\infty} y f_{X,V}(x, v) dv dy &= \int_a^b \int_{p_0^*}^{1+y} \frac{y}{b-a} dv dy \\ &= \frac{(1 - p_0^*)(b^2 - a^2)}{2(b-a)} + \frac{b^3 - a^3}{3(b-a)}. \end{aligned} \quad (3.39)$$

As a result, by (3.38) and (3.39), the fraction of the revenue for supplier 1 if $b < 3/4$ is

$$\begin{aligned} \beta_{pn} &= \frac{\mathbb{E}(X|X + Y \geq p_0^*)}{\mathbb{E}(X + Y|X + Y \geq p_0^*)} \\ &= \frac{(p_0^* - b)^3 - (p_0^* - a)^3 + 3(b-a)}{(p_0^* - b)^3 - (p_0^* - a)^3 + 3(b-a) + 3(1 - p_0^*)(b^2 - a^2) + 2(b^3 - a^3)}, \end{aligned}$$

where $p_0^* = 3/4$.

If $b \geq 3/4$, by lemma 3.7, $b \geq p_0^*$. Therefore,

$$\begin{aligned} \int_0^1 \int_{p_0^*}^{\infty} x f_{X,V}(x, v) dv dx &= \int_0^{p_0^*-a} \int_{p_0^*}^{b+x} \frac{x}{b-a} dv dx + \int_{p_0^*-a}^1 \int_{a+x}^{b+x} \frac{x}{b-a} dv dx \\ &= -\frac{(p_0^* - a)^3}{6(b-a)} + \frac{1}{2}. \end{aligned} \quad (3.40)$$

If $b \geq 3/4$, by the similar method above, we have

$$\begin{aligned} \int_0^1 \int_{p_0^*}^{\infty} y f_{X,V}(x, v) dv dy &= \int_a^{p_0^*} \int_{p_0^*}^{1+y} \frac{y}{b-a} dv dy + \int_{p_0^*}^b \int_y^{1+y} \frac{y}{b-a} dv dy \\ &= \frac{(1 - p_0^*)(p_0^{*2} - a^2)}{2(b-a)} + \frac{p_0^{*3} - a^3}{3(b-a)} + \frac{b^2 - p_0^{*2}}{2(b-a)}. \end{aligned} \quad (3.41)$$

As a result, by (3.40) and (3.41), the fraction of the revenue for supplier 1 if $b \geq 3/4$ is

$$\begin{aligned}\beta_{pn} &= \frac{\mathbb{E}(X|X+Y \geq p_0^*)}{\mathbb{E}(X+Y|X+Y \geq p_0^*)} \\ &= \frac{-(p_0^* - a)^3 + 3(b - a)}{-(p_0^* - a)^3 + 3(b - a) + 3(1 - p_0^*)(p_0^{*2} - a^2) + 2(p_0^{*3} - a^3) + 3(b^2 - p_0^{*2})},\end{aligned}$$

where $p_0^* = \frac{2a + \sqrt{a^2 - 6a + 6b}}{3}$. Now we can compare the revenue for both suppliers under the pre-committed and contingent revenue sharing policies. ■

Chapter 4

Strategic Customer Behavior with A Reservation System

4.1 Introduction

An online reservation system allows customers to join the queue and virtually wait at the service site. Some online reservation systems, e.g., Nowait, inform customers of the queue length when they are about to make the reservation. Such a system reduces customers' uncertainty about the queue length, and helps them in planning their time schedule. As a result, firms that offer the online reservation systems may attract more customers. However, such system may lead to the strategic behavior of customers. If customers make their decision whether to join the queue taking into account both their waiting and travelling costs, the market coverage may not be as large as the firm expects it to be.

In this setting, we consider the following research questions: 1. What is the influence of the online reservation system on customers' behavior? 2. Given certain conditions, (i.e., the benefit of the service, the travelling distance, and etc.) what is the optimal service operation policy for the firm?

We consider a model in which customers must travel from their location to the service site and incur a linear travelling cost. When customers intend to book service, they are informed about the current queue length. Because customers reserve a position in the queue before travelling, the more time they spend in travelling the less time they would wait at the service site. Therefore, there exists a tradeoff between the travelling and waiting costs. The travelling cost is exogeneously determined, and it depends on the environmental factors, such as weather, road condition, and traffic. For simplicity of the analysis, we assume that other arriving customers cannot overtake a customer during her

traveling time to the service site and during this traveling time, the server is only idle if the facility is empty.

Intuitively, if the travelling cost per unit distance is sufficiently low, customers who reside far away from the service location may prefer to join the queue even if the queue is long. Customers who reside near the service location may prefer not to join the queue if the queue length is long since their travelling cost is negligible and they can join the queue whenever the queue length is short. With such customers' behavior, the attraction of the service is increasing in the distance between the customer and service locations. Indeed, our analysis in Section 4.4 supports this intuition.

If the travelling cost per unit distance is sufficiently high, customers' behavior may differ. Customers who reside near the service site may choose to join the queue, while the customers who reside far away from the service site may prefer not to join the queue because the travelling cost becomes too high. Then the discussion for the high travelling cost is similar to the traditional analysis that the attraction of the service is decreasing in the distance between the customer and service locations. Indeed, this intuition is supported by our analysis in Section 4.5.

The interesting and more applicable case is that the travelling cost is intermediate. In this case, the attraction of the service is first increasing and then decreasing in the distance. In other words, the customers who reside in an intermediate distance from the service site are more likely to join the queue, but the customers who reside near or far away from the service site are less likely to join the queue because the queue waiting cost is too high for the customers who are near the service site and the travelling cost is too high for the customers who are far away from the service site. The market coverage for different queue lengths tends to be a reversed "U-shape" with respect to the travelling distance. Our analysis in Section 4.6 is dedicated to this case and provides the exact characterization of the customers' choice and its dependency of the model's parameters.

This paper builds a connection between the queueing games and Hotelling models. By characterizing customers' behavior, we can examine the influence of different types of the delay information, capacity and location policies over the firm's profit. We leave the investigation of how firms should optimally take these decision for future research.

4.2 Literature Review

Customer behavior in queueing model has been extensively studied since Naor (1969) who showed that if customers self-decide whether to join the queue taking into account the queue length, then in equilibrium, they will join if and only if the queue is shorter than a

specific threshold. He further showed that the threshold determined by maximizing the social welfare is greater than or equal to the queue length threshold.

Recent work has extended Noar's observable queueing model to more general settings. There are many streams. Some papers focus on the customers with different classes. For example, Larsen (1998) examines a model with continuously randomly distributed service value. Adiri and Yechiali (1974) focus on the two priority classes and there are prices for becoming one of those classes. Mendelson and Whang (1990) study an M/M/1 queueing system in which customers are classified into different groups, and those groups are differentiated in their delay cost, expected service time and demand function. Afèche et al. (2013) study the optimal lead time pricing strategy for serving multiple time-sensitive customer groups. In their model, the type of customers are classified by their arrival rate, delay cost and utility function. For other related topics, Hassin and Haviv (2003) provide a comprehensive survey of this literature. In our model, customers are heterogeneous in the distance from their home to the service site. It is equivalent to that customers are differentiated by the valuation of the service. In contrast to the discrete classification of customers, the valuation of the service is uniformly distributed over a line.

For the advance reservation, Simhon (2016) examines the M/D/1 queue model with advance reservation. Oh and Su (2012) studied the optimal pricing strategy for punishing the no-shows in advance reservation. Alexandrov and Lariviere (2012) examine the role of reservation for a restaurant with a capacity constraint. The papers mentioned above examine a process with two periods. The first period is for the advance reservation, and after that the second period is for queueing. Therefore, the reservation and the queue waiting are separated in their model. The tradeoff of the travelling cost and queue waiting cost does not exist. In our model, the reservation and waiting form a coherent process and customers make their decision whether to join the queue based on both the travelling and queueing processes. Thus, our model captures a more complex and practical behavior of customers in presence of reservation systems.

For the delay information sharing, Guo and Zipkin (2007) examine different levels of delay information announcements and show the sufficient condition to ensure that more information makes the queue system better. Armony et al. (2009) compare the impact of making different delay announcements in customer contact centers in which customers cannot observe the queue. Allon et al. (2011a) study the strategic behavior of customers about when to join the queue by interpreting the announcements made by the firm. In their setting, the announcements could be vague and unverifiable, they investigate the strategic behavior of customers in response to different announcements. Our work share

the similarity that customers' behavior is influenced by the delay information. However, customers in our model make their decision based on the expectation of the queue length which is not only influenced by the queue information provided by the firm but also the distance from their home to the service site.

Hassin (2016) provides a survey about customer's rational decision in a queueing system and firm's optimal strategy in response to the customer behavior. Perhaps the most relevant paper to our work is by Hassin et al. (2016) who consider a queueing model in which customers incur a traveling cost. They assume that the customers' home locations are distributed over a line, and they can only join the queue after travelling the distance between their home and service locations. Though their work may be the first to build up the relationship between the Hotelling model and queueing model, they include no interaction between the traveling and waiting costs. Different from their study, we examine the online reservation system that allows customers to make advance reservations before traveling to the service site, thus the tradeoff between the traveling and waiting costs exists when customers make the decision whether to join the queue.

The Hotelling model has been extensively implemented to solve problems in marketing and operations management areas. The earliest work are by Hotelling (1929) and D'aspremont et al. (1979). The model characterizes the market by a line, and customers are uniformly distributed over it. The fundamental assumption is that the attraction of the service is decreasing in the distance between customers and the service site. By incorporating the queue waiting cost, our result shows that the attraction of the service is not necessarily decreasing in the distance. Intuitively, as we demonstrate in Section 4.4, if the travelling cost is sufficiently low, customers may be willing to spend much time on travelling, thus the attraction of the service can be increasing in the distance.

4.3 Model Setup and Initial Results

Consider a market on the $[0, 1]$ line with customers distributed on it. There is a facility located at 0^{th} end point of the line. Customers must travel from their location to the facility to be served. Assume that the facility has an infinite buffer size with a fixed service rate μ . In our model, customers call the server in the facility to make a reservation (or book service online) and are informed about their positions in the queue at the time of booking. More specifically, customers are told how many people are at the queue before them. They then make a decision whether to join the queue taking into account their travelling and queue waiting costs. If customers decide to join the queue they will be served and will obtain a reward R when the service is completed. For simplicity, we

ignore the price of the service. Without loss of generality if customers are risk neutral, then the *Reward* can be written as $Reward - Price$. Thus, the utility of a customer is

$$Utility = Reward - Service\ time\ cost - Waiting\ cost - Travelling\ cost.$$

Suppose that there is a customer whose home location is at a distance of x from the facility, $x \in [0, 1]$, and x is in units of distance. If the customer agrees to join the queue and is served, he will incur a travelling cost Cx where C is the cost per unit of time. Without loss of generality, we assume that the cost of waiting per unit of time is 1. The utility for the customer travelling from x and observing j customers ahead of her in the queue upon arrival is

$$U_i(x) = R - \frac{1}{\mu} - \frac{j}{\mu} - Cx.$$

4.3.1 Demand and Arrival Rate

The demand at any point $x \in [0, 1]$ is assumed to follow a renewal process. We assume that at each arrival epoch there is a single arrival with probability 1, and the total demand per period over the entire line is λ_T with $0 < \lambda_T < \infty$. We denote the number of calls for service at x up to time t by $Y_t(x)$ and the average demand arrival rate at x by $\lim_{t \rightarrow \infty} Y_t(x)/t = \lambda_T dF_{\lambda_T}(x)$. Thus, the proportion of demand at any point $x \in [0, 1]$ is given by $dF_{\lambda_T}(x)$ such that the Lebesgue integral is well defined and $\int_0^1 dF_{\lambda_T}(x) = 1$. Finally, if $\lim_{\epsilon \rightarrow 0} \int_x^{x+\epsilon} dF_{\lambda_T}(x) > 0$, we require that the demand at x follows a Poisson process, but we allow demand to follow a general renewal process whenever $\lim_{\epsilon \rightarrow 0} \int_x^{x+\epsilon} dF_{\lambda_T}(x) = 0$. Observe that $\lim_{\epsilon \rightarrow 0} \int_x^{x+\epsilon} dF_{\lambda_T}(x) = 0$ implies that the general renewal process at x has an infinitesimal rate. The latter requirement is that the inter-renewal time at x with $\lim_{\epsilon \rightarrow 0} \int_x^{x+\epsilon} dF_{\lambda_T}(x) = 0$ is distributed with a cumulative distribution function (CDF) $F_x(t)$ such that for any $\epsilon > 0$ and $t > 0$ we have $F_x(t) \leq \epsilon$.

With the above assumptions, the aggregated demand generation process over the entire line follows a Poisson process with arrival rate λ_T . Since $\lim_{t \rightarrow \infty} Y_t(x)/t = \lambda_T dF_{\lambda_T}(x)$, the actual arrival rate depends on the proportion of customers over the line who decide to join the queue. For example, if demand is symmetric over the line and half of it is covered by the firm, then the arrival rate will be $\lambda_T/2$. Those assumptions have been widely used in the literature, see, e.g., Baron et al. (2008) and Hassin et al. (2016).

4.3.2 Waiting Cost

Let $N_t(i)$ be the event that $i \geq 0$ customers are at the queue at time t (possibly some of them are still travelling to the queue), $i \geq 0$. Thus, when the customer decides to join the queue, the event $N_{t+x}(j)$ denotes that there are j people who are still waiting ahead of the customer when the customer arrives at the facility. If we allow no overtaking of customers during their traveling time, we have $0 \leq j \leq i$ where i is the number of customers in the queue when the customer is checking on line. Let $W_i(x)$ denote the expected waiting time in the queue upon arrival to the facility of a customer from x that is informed that i customers are waiting before him when checking online. Given i , the expected utility is

$$\mathbb{E}[U_i(x)] = R - \frac{1}{\mu} - W_i(x) - Cx.$$

We assume that customers make their decision whether to join the queue by considering the expected waiting cost. However, expressing $W_i(x)$ may be very hard because (i) overtaking, i.e., that customers at $y < x$ may join the on-line queue later than the customer from x , but arrive at the facility and join the physical queue before the customer from x , and (ii) starvation, i.e., the server may have no customers in the physical system even if they are in the on-line queue.

To address (i) and (ii), we consider as follows:

- (a) The server may keep its service order as is dictated by the on-line queue, so even if overtaking occurs physically, it does not affect the physical service order.
- (b) Because the rate of arrivals over $[0, x)$ may be small, overtaking should not be very common.
- (c) Customers have *bounded rationality* of customers (see Hassin (2016), p.p. 17). That is, in reality customers may not be sophisticated to capture this complexity.
- (d) If the online reservation system directly informs customers about the expected waiting time at the service site instead of the queue length, then it is possible that customers would arrive at the service site before the service starts, alleviating the starvation problem.
- (e) In make-to-order systems, such as the take-away service in the fast food restaurant, once customers order online, the products or service “join” the queue right away. In such applications, there is no starvation.

With all of these, we assume that the facility never idles if its virtual queue > 0 (i.e., all customers that are travelling to the facility get to the facility on or before their service time could start). Then, $W_i(x)$ is given, as discussed below, by

$$W_i(x) = \frac{i}{\mu} - \sum_{j=1}^i \frac{i-j}{\mu} \Pr\{N_{t+x}(j)|N_t(i), j > 0\} - \frac{i}{\mu} \Pr\{N_{t+x}(j)|N_t(i), j = 0\}. \quad (4.1)$$

Note that $W_i(0) = \frac{i}{\mu}$ for all $i \geq 0$. The customer will agree to join the queue if $\mathbb{E}[U_i(x)] \geq 0$, or make no reservation if $\mathbb{E}[U_i(x)] < 0$. If the number of customers at the facility is greater than 0, $j > 0$, then the departure process from the queue is Poisson with intensity μ during the travelling time x , therefore we have

$$\Pr\{N_{t+x}(j)|N_t(i), j > 0\} = \frac{(\mu x)^{(i-j)}}{(i-j)!} e^{-\mu x}.$$

The expected reduction of queue size during the travelling time is then

$$\begin{aligned} & \sum_{j=1}^i \frac{i-j}{\mu} \Pr\{N_{t+x}(j)|N_t(i), j > 0\} + \frac{i}{\mu} \Pr\{N_{t+x}(j)|N_t(i), j = 0\} \\ &= \sum_{k=0}^{i-1} \frac{k}{\mu} \frac{(\mu x)^k}{k!} e^{-\mu x} + \sum_{k=i}^{\infty} \frac{i}{\mu} \frac{(\mu x)^k}{k!} e^{-\mu x} \geq 0. \end{aligned}$$

where k is the number of departures of a Poisson process. Note that $W_0(x) = 0$ for all $x \in [0, 1]$.

To analyze $W_i(x)$ as given in (4.1), the following notation is required:

$$\begin{aligned} p_i(\mu) &= \frac{(\mu)^i}{i!} e^{-\mu} \text{ (Probability of } i \text{ events of Poisson distribution with intensity } \mu), \\ P_i(\mu) &= \sum_{n=0}^i \frac{(\mu)^n}{n!} e^{-\mu} \text{ (Probability of up to } i \text{ events of Poisson distribution with intensity } \mu), \\ Q_i(\mu) &= \sum_{n=i+1}^{\infty} \frac{(\mu)^n}{n!} e^{-\mu} = 1 - P_i(\mu). \end{aligned}$$

The Poisson distribution can be expressed by the incomplete gamma function,

$$P_i(\mu) = \frac{\Gamma(i+1, \mu)}{\Gamma(i+1)}, \quad Q_i(\mu) = \frac{\gamma(i+1, \mu)}{\Gamma(i+1)},$$

in which $\Gamma(i, \mu) = \int_{\mu}^{\infty} t^{i-1} e^{-t} dt$, $\gamma(i, \mu) = \int_0^{\mu} t^{i-1} e^{-t} dt$, and $\Gamma(i) = \Gamma(i, 0)$.

Some relevant properties of $W_i(x)$ are given by Lemma 4.1.

Lemma 4.1 For $x \in [0, 1]$, the expected queue waiting time is

$$W_i(x) = \begin{cases} \frac{1}{\mu} \sum_{n=0}^{i-1} P_n(\mu x) = \frac{1}{\mu} \left[-\frac{\mu x \Gamma(i-1, \mu x)}{\Gamma(i-1)} + \frac{i \Gamma(i, \mu x)}{\Gamma(i)} \right], & \text{if } x > 0, \\ \frac{i}{\mu} & \text{if } x = 0. \end{cases}$$

which is strictly increasing in i and decreasing in x .

As the number in queue for an M/M/1 is independent of future arrivals and by Lemma 4.1, for any $x \in [0, 1]$ there exists a threshold i_x such that if $i \leq i_x$, $\mathbb{E}[U(x)|i] \geq 0$ and if $i > i_x$, $\mathbb{E}[U(x)|i] < 0$. In other words, in equilibrium the fraction of customers from x who join the queue is $p_x = \sum_{i=0}^{i_x} p_i$ where p_i is the steady state probability of the number of customers in the queue. The fraction of customers who will make no reservation would be $p_{-x} = \sum_{i=i_x}^{\infty} p_i$. By definition of i_x , the following inequalities hold

$$W_{i_x+1}(x) + Cx \geq R - \frac{1}{\mu} \geq W_{i_x}(x) + Cx.$$

Now, we can characterize the behavior of i_x with the variation of x for different travelling costs. To avoid the trivial case, we assume that $R - \frac{1}{\mu} > 0$. Otherwise, the service has no attraction to any customer over the line for any queue length.

4.4 No Travelling Cost ($C = 0$)

If there is no travelling cost, only the queue waiting cost matters. Lemma 4.1 guarantees that given i , for any $\hat{x} > x$, if customers at x joins the queue, so will customers from \hat{x} . Given R , there exists i_0 such that for the customers who reside in the 0th end point will join the queue if $i \leq i_0$ and make no reservation if $i > i_0$, thus

$$i_0 = \arg \max_i \left\{ R - \frac{1}{\mu} - \frac{i}{\mu} \geq 0 \right\} = \lfloor R\mu - 1 \rfloor, \quad (4.2)$$

where $\lfloor x \rfloor$ denotes that the largest integer that is smaller than or equal to x .

Consider the customers who reside at the 1st end point, $x = 1$. By Lemma 4.1, $W_i(1) = \frac{1}{\mu} \sum_{n=0}^{i-1} P_n(\mu)$ is increasing in i and $\lim_{i \rightarrow +\infty} W_i(1) = +\infty$. Thus, there always exists i_1 such that for the customers who reside at $x = 1$ will join the queue if $i \leq i_1$ and

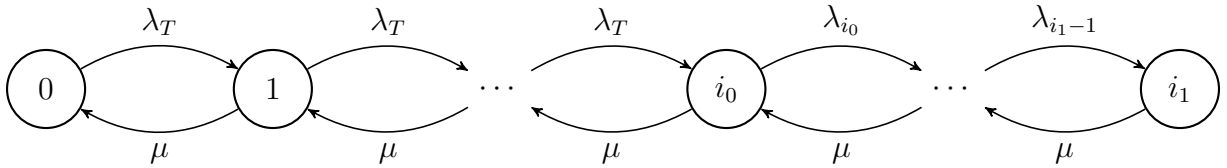
make no reservation if $i > i_1$, thus

$$i_1 = \arg \max_i \left\{ R - \frac{1}{\mu} - W_i(1) \geq 0 \right\}. \quad (4.3)$$

Because $W_i(1) = \frac{1}{\mu} \sum_{n=0}^{i-1} P_n(\mu) \leq \frac{i}{\mu} = W_i(0)$, we have $i_0 \leq i_1$.

To summarize, i_0 and i_1 defines the highest queue length that customers residing in 0th and 1st end points agree to join the queue, respectively. Thus, the maximum queue length for customers who reside in $(0, 1)$ to join the queue must be in $[i_0, i_1]$. If the queueing system is in state $i \in (i_0, i_1]$, a proportion of the customers in the market will join the queue. Thus, the arrival rate is not λ_T , but a fraction of it. Since the arrival rates are state-dependent, we denote the arrival rate for state i by λ_i . Figure 4.1 illustrates the arrival rates and the service rate for different states of the queueing model when there are no traveling cost.

Figure 4.1: Arrival rates when there is no travelling cost



4.4.1 Case 1: $i_0 = i_1$

If $i_0 = i_1$, all customers in $[0, 1]$ will not join the queue if $i > i_0$, but will join the queue if $i \leq i_0$. Thus, $i_0 = i_1$ is the case that the arrival rate $\lambda_i = \lambda_T$ if $i \leq i_0$, and $\lambda_i = 0$ if $i > i_0$. Now, we can characterize the sufficient and necessary conditions for $i_0 = i_1$. By (4.2) and (4.3), $i_0 = i_1$ implies

$$\frac{i_0+1}{\mu} > R - \frac{1}{\mu} \geq \frac{i_0}{\mu}, \quad (4.4)$$

$$W_{i_0+1}(1) > R - \frac{1}{\mu} \geq W_{i_0}(1). \quad (4.5)$$

From (4.4) and (4.5), we obtain

$$W_{i_0+1}(1) > R - \frac{1}{\mu} \geq \frac{i_0}{\mu}. \quad (4.6)$$

Inequalities (4.6) imply $W_{i_0+1}(1) > \frac{i_0}{\mu}$, equivalently, $\sum_{n=0}^{i_0} P_n(\mu) > i_0$. Therefore, the sufficient and necessary conditions for $i_0 = i_1$ are

$$\sum_{n=0}^{i_0} P_n(\mu) > i_0 \quad (a), \quad W_{i_0+1}(1) > R - \frac{1}{\mu} \geq \frac{i_0}{\mu} \quad (b). \quad (4.7)$$

We find that condition (a) depends on μ . And, condition (b) depends on both μ and R . The following lemma gives the values of μ that satisfy condition (a).

Lemma 4.2 (i) If $\mu \in (0, 1)$, we have $\sum_{n=0}^i P_n(\mu) > i$ for any $i \geq 0$.

(ii) If $\mu \in [1, +\infty)$, there exists a unique $\hat{i} \geq 0$ such that $\sum_{n=0}^i P_n(\mu) > i$ for all $i \leq \hat{i}$, and $\sum_{n=0}^i P_n(\mu) \leq i$ for all $i > \hat{i}$. Specifically, $\hat{i} = 0$ iff $\mu \in [\mu^*, +\infty)$ where $\mu^* = -2 - W(-1/e^2)$ and $W(\cdot)$ is the Lambert W function in the lower branch¹

The term $\mu = 1$ plays a special role in Lemma 4.2. In our discussion, we derive the conditions for $i_0 = i_1$ by comparing the waiting cost of customers at the 0th and 1st end points. Since we normalize the time to travel the entire length of the market to 1, $\mu = 1$ is an important threshold. If this time is not 1, the results would be different. Now, it is straightforward to classify R into three categories for different behaviors of customers in response to μ 's value.

By Lemma 4.2(i), if $\mu \in (0, 1)$, $\sum_{n=0}^i P_n(\mu) > i$ for all $i \geq 0$. Since $\sum_{n=0}^i P_n(\mu) = \mu W_{i_0+1}(1)$, $W_{i_0+1}(1) > \frac{i}{\mu}$. Moreover, $W_{i_0+1}(1) \geq W_{i_0}(1)$ because the expected waiting cost is increasing in i_0 . Thus, we can order the waiting cost of customers who reside in the 0th end point with different i values as follows:

$$0 < W_1(1) \leq \frac{1}{\mu} < W_2(1) \leq \frac{2}{\mu} < W_3(1) \leq \dots < W_i(1) \leq \frac{i}{\mu} < W_{i+1}(1) \leq \frac{i+1}{\mu} < \dots$$

Since $i_0 = i_1$, the inequalities hold for the waiting cost of customers who reside in the 1st end point. As a result, for all i , if $R \in [\frac{i+1}{\mu}, W_{i+1}(1) + \frac{1}{\mu})$, then conditions (a) and (b) hold, thus $i_0 = i_1 = i$. If $R \in [W_{i+1}(1) + \frac{1}{\mu}, \frac{i+2}{\mu})$, $i_0 < i_1$. Obviously, we have $i_0 = i - 1$ and $i_1 = i$.

Corollary 4.3 If $\mu \in (0, 1)$, $i_0 + 1 \geq i_1 \geq i_0$.

¹The Lambert W function $W(x)$ is double-valued if $x \in (-1/e, 0)$, thus it has two branches, i.e., the principle and lower branches. In the lower branch of the Lambert W function, the value of it monotonically decreases from $W(-1/e) = -1$ to $W(0^-) = -\infty$. We restrict our attention on the lower branch so that μ is non-negative.

If $\mu \in [1, +\infty)$, there exists $\hat{i} \geq 0$ such that for all $i \leq \hat{i}$, $\sum_{n=0}^i P_n(\mu) > i$. Thus, for any $i \leq \hat{i}$, we have similar results to these in $\mu \in (0, 1)$. As a result, for $i \leq \hat{i}$, if $R \in [\frac{i+1}{\mu}, W_{i+1}(1) + \frac{1}{\mu})$, (4.7)(a) and (4.7)(b) are satisfied, thus $i_0 = i_1 = i$. If $R \in [W_{i+1}(1) + \frac{1}{\mu}, \frac{i+2}{\mu})$, $i_0 = i - 1 < i = i_1$. If $R \in [W_{i+1}(1) + \frac{1}{\mu}, \infty)$, $i_0 < i_1$. If $\mu \in [\mu^*, \infty)$, by Lemma 4.2(ii), $\hat{i} = 0$. Table 4.1 summarizes the three categories of R resulting in $i_0 = i_1$ or $i_0 < i_1$ in response to different values of μ . Table 4.2 summarizes the arrival rates for different system states if $i_0 = i_1$. The term ‘‘QL’’ is short for queue length.

 Table 4.1: Three categories of μ and R 's value

| | $\mu \in (0, 1)$ | $\mu \in [1, \mu^*)$ | $\mu \in [\mu^*, \infty)$ |
|-------------|---|---|---|
| $i_0 = i_1$ | $R \in [\frac{i+1}{\mu}, W_{i+1}(1) + \frac{1}{\mu})$ for all $i \geq 0$ | $R \in [\frac{i+1}{\mu}, W_{i+1}(1) + \frac{1}{\mu})$ for $i \leq \hat{i}$ ($\hat{i} > 0$) | $R \in (\frac{1}{\mu}, \frac{e^{-\mu}}{\mu} + \frac{1}{\mu})$ |
| $i_0 < i_1$ | $R \in [W_{i+1}(1) + \frac{1}{\mu}, \frac{i+2}{\mu})$ for all $i \geq 0$ | $R \in [W_{i+1}(1) + \frac{1}{\mu}, \frac{i+2}{\mu})$ for $i \leq \hat{i}$ ($\hat{i} > 0$), or $R \in [W_{i+1}(1) + \frac{1}{\mu}, \infty)$ | $R \in [\frac{e^{-\mu}}{\mu} + \frac{1}{\mu}, \infty)$ |

 Table 4.2: Arrival rates ($i_0 = i_1$)

| $\mu \in (0, 1)$ and $R \in [\frac{i+1}{\mu}, W_{i+1}(1) + \frac{1}{\mu})$ | $\mu \in [1, \mu^*)$ and $R \in [\frac{i+1}{\mu}, W_{i+1}(1) + \frac{1}{\mu})$ for $i \leq \hat{i}$ ($\hat{i} > 0$) | $\mu \in [\mu^*, \infty)$ and $R \in (\frac{1}{\mu}, \frac{e^{-\mu}}{\mu} + \frac{1}{\mu})$ |
|---|---|--|
| λ_T , if QL $\leq i$; 0, if QL $> i$ | λ_T , if QL $\leq i$; 0, if $i < \text{QL} \leq \hat{i}$ | λ_T , if QL = 0; 0, if QL > 0 |

In conclusion, for $i_0 = i_1$, we have sufficient conditions that μ and R are in those three categories, and necessary conditions that μ and R are in at least one of those three categories.

4.4.2 Case 2: $i_0 < i_1$

Let $x_i = \{x \geq 0 : \mathbb{E}(U_i(x)) = 0\}$. Proposition 4.4 characterizes the ‘‘steps’’ property of the demand rates when $i_0 < i_1$.

Proposition 4.4 *If $i_0 < i_1$, for any integer $i \in (i_0, i_1]$, there exists $x_i \in [0, 1]$ such that $\mathbb{E}(U_i(x_i)) = 0$. Moreover, for any integers $a, b \in [i_0, i_1]$, if $a > b$, then $x_a > x_b$.*

Proposition 4.4 shows that if the queue length is large, the customers who reside far away from the service site is more likely to join the queue if the travelling cost is negligible. Because in the online reservation system, customers make the advance reservation and virtually wait in the line, the longer the distance they are travelling the less time they have to wait in the service site. Meanwhile, the travelling cost is zero, thus the waiting cost tends to be small if customers spend much time in travelling. Therefore, the expected utility of the customer who resides far away tends to be high. For the customers who reside near the service site, they can join the queue whenever the queue length is sufficiently short, thus they may not make the reservation if the queue length is high. However, it is possible that some customers make the reservation and then wait at home until the queue length becomes zero. It is then equivalent to that they do not have any cost for waiting at home, and thus the “travelling cost” for them is zero. Generally, by Proposition 4.4, the attraction of the service may be increasing in the distance between the customer and service locations.

Table 4.3 summarize the arrival rates for different system states if $i_0 < i_1$. In the 3rd and 4th rows of Table 4.3, we show the arrival rates for $R \in [\frac{i_0+1}{\mu}, W_{i_1}(1) + \frac{1}{\mu})$. However, if denoting $i_0+k = i_1, k \in N^+$, we did not specify the values of R for each k . The discussion of R for different k values boils down to the derivation of $[\sum_{i=0}^k P_i(\mu)]$. One can expect that $[\sum_{i=0}^k P_i(\mu)]$ highly relies on μ . The discussion of $[\sum_{i=0}^k P_i(\mu)]$ is well-defined and complicated, but we can show i_0 and i_1 and the arrival rates numerically.

Example 4.1 (a) Assume that customers are uniformly distributed over $[0, 1]$, then for customers who reside in $[x, 1]$ the arrival rate will be $\int_x^1 \lambda_T dF_{\lambda_T}(x) = (1-x)\lambda_T$. Consider $i_0 < i_1$, by Proposition 4.4, for any integer $i \in (i_0, i_1]$ there exists $x_i \in [0, 1]$ solving $R - \frac{1}{\mu} = W_i(x)$. And, the arrival rate from state i to $i + 1$ is $\lambda_i = \int_{x_i}^1 \lambda_T dF_{\lambda_T}(x) = (1 - x_i)\lambda_T$.

Consider $\mu \in [\mu^*, +\infty)$. When $R \in (\frac{1}{\mu}, \frac{e^{-\mu}}{\mu} + \frac{1}{\mu})$, for customers at $x = 0$, the expected utilities for $i = 0, 1$ are

$$\begin{aligned} (i = 0) \quad R - \frac{1}{\mu} &> 0 \\ (i = 1) \quad R - \frac{2}{\mu} &< \frac{e^{-\mu}}{\mu} - \frac{1}{\mu} < 0, \end{aligned}$$

Thus, $i_0 = \lfloor R\mu - 1 \rfloor = 0$. For customers at $x = 1$, the expected utilities for $i = 0, 1$ are

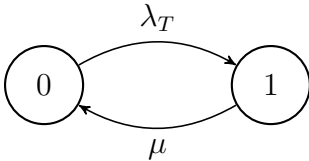
$$(i = 0) \quad R - \frac{1}{\mu} - W_0(1) = R - \frac{1}{\mu} > 0$$

Table 4.3: Arrival rates ($i_0 < i_1$)

| | |
|--|--|
| $\mu \in (0, 1)$ and $R \in [W_{i+1}(1) + \frac{1}{\mu}, \frac{i+2}{\mu})$ | $\mu \in [1, \mu^*)$ and $R \in [W_{i+1}(1) + \frac{1}{\mu}, \frac{i+2}{\mu})$ for $i \leq \hat{i}$ ($\hat{i} > 0$) |
| λ_T , if $QL \leq i$; $\lambda_{i+1} = (1 - x_{i+1}) \lambda_T$, if $QL = i + 1$; 0, if $QL > i + 1$ | λ_T , if $QL \leq i$; $\lambda_{i+1} = (1 - x_{i+1}) \lambda_T$, if $QL = i + 1$; 0, if $QL > i + 1$ |
| $\mu \in [1, \mu^*)$ and $R \in [\frac{i_0+1}{\mu}, W_{i_1}(1) + \frac{1}{\mu})$ ($\hat{i} < i_0 < i_1$) | $\mu \in [\mu^*, \infty)$ and $R \in [\frac{i_0+1}{\mu}, W_{i_1}(1) + \frac{1}{\mu})$ ($0 < i_0 < i_1$) |
| λ_T , if $QL \leq i_0$; $\lambda_{i_0+1} = (1 - x_{i_0}) \lambda_T$, if $QL = i_0 + 1$; \vdots $\lambda_{i_1} = (1 - x_{i_1-1}) \lambda_T$, if $QL = i_1$; 0, if $QL > i_1$ | λ_T , if $QL \leq i_0$; $\lambda_{i_0+1} = (1 - x_{i_0}) \lambda_T$, if $QL = i_0 + 1$; \vdots $\lambda_{i_1} = (1 - x_{i_1-1}) \lambda_T$, if $QL = i_1$; 0, if $QL > i_1$ |

$$(i = 1) R - \frac{1}{\mu} - W_1(1) = R - \frac{1}{\mu} - \frac{e^{-\mu}}{\mu} < 0,$$

thus $i_1 = 0$. When the system is in state $i = 0$, all the customers on the line $[0, 1]$ will join the queue; and when the system is in state $i = 1$, all the customers on the line will not join the queue. The system is illustrated by Figure 4.2.

 Figure 4.2: Example 1(a): $\mu \in [\mu^*, +\infty)$ and $R \in (\frac{1}{\mu}, \frac{e^{-\mu}}{\mu} + \frac{1}{\mu})$


By the classic result of the state-dependent queue, the queue can be modeled as a standard B&D process and thus the steady state probability has the relation $\lambda_T p_0 = \mu p_1$. By $p_0 + p_1 = 1$, we obtain $p_0 = \frac{1}{1 + \lambda_T/\mu}$ and $p_1 = \frac{\lambda_T/\mu}{1 + \lambda_T/\mu}$. The average arrival rate is $\frac{\lambda_T}{1 + \lambda_T/\mu}$, the average queue length is $\frac{\lambda_T/\mu}{1 + \lambda_T/\mu}$, and the average utility for customers is $\frac{R - 1/\mu}{1 + \lambda_T/\mu}$.

(b) If $R \in \left[\frac{e^{-\mu}+1}{\mu}, \frac{(\mu+2)e^{-\mu}+1}{\mu}\right)$, for $x = 0$, the expected utility for $i = 0, 1$ are

$$\begin{aligned} (i = 0) \quad R - \frac{1}{\mu} &> 0 \\ (i = 1) \quad R - \frac{2}{\mu} &< \frac{(\mu + 2)e^{-\mu} - 1}{\mu} \leq 0. \end{aligned}$$

The second inequality holds for $i = 1$ because $\mu \in [\mu^*, +\infty)$, $2 + \mu - e^\mu \leq 0$, so $i_0 = 0$. For customers at $x = 1$, the expected utilities for $i = 0, 1, 2$ are

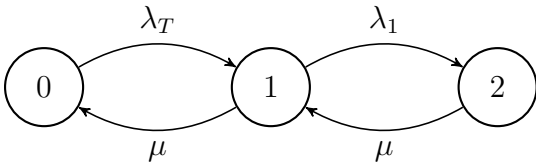
$$\begin{aligned} (i = 0) \quad R - \frac{1}{\mu} - W_0(1) &= R - \frac{1}{\mu} > 0 \\ (i = 1) \quad R - \frac{1}{\mu} - W_1(1) &= R - \frac{1}{\mu} - \frac{e^{-\mu}}{\mu} \geq 0 \\ (i = 2) \quad R - \frac{1}{\mu} - W_2(1) &= R - \frac{1}{\mu} - \frac{e^{-\mu}}{\mu} - \frac{e^{-\mu} + \mu e^{-\mu}}{\mu} < 0 \end{aligned}$$

Thus, we have $i_1 = 1$. When the system is in state $i = 1$, customers at 0th end point will not join the queue but customers at 1th end point will join the queue. There exists $x \in (0, 1)$ such that $\mathbb{E}(U_1(x)) = 0$. That is

$$R - \frac{1}{\mu} - W_1(x) = R - \frac{1}{\mu} - \frac{e^{-\mu x}}{\mu} = 0.$$

$x = -\frac{\ln(R\mu-1)}{\mu}$. Therefore, $\lambda_1 = \left(1 + \frac{\ln(R\mu-1)}{\mu}\right) \lambda_T$. The system is illustrated by Figure 4.3.

Figure 4.3: Example 1(b): $\mu \in [\mu^*, +\infty)$ and $R \in \left[\frac{e^{-\mu}+1}{\mu}, \frac{(\mu+2)e^{-\mu}+1}{\mu}\right)$



By the relation that $\lambda_T p_0 = \mu p_1$, $\lambda_1 p_1 = \mu p_2$ and $p_0 + p_1 + p_2 = 1$, we can obtain the steady state probability as $p_0 = \frac{1}{\lambda_1 \lambda_T / \mu^2 + \lambda_T / \mu + 1}$, $p_1 = \frac{\lambda_T / \mu}{\lambda_1 \lambda_T / \mu^2 + \lambda_T / \mu + 1}$ and $p_2 = \frac{\lambda_1 \lambda_T / \mu^2}{\lambda_1 \lambda_T / \mu^2 + \lambda_T / \mu + 1}$ in which $\lambda_1 = \left(1 + \frac{\ln(R\mu-1)}{\mu}\right) \lambda_T$. The average arrival rate is $\frac{\lambda_T + \lambda_1 \lambda_T / \mu}{\lambda_1 \lambda_T / \mu^2 + \lambda_T / \mu + 1}$, the average queue length is $\frac{\lambda_T / \mu + 2 \lambda_1 \lambda_T / \mu^2}{\lambda_1 \lambda_T / \mu^2 + \lambda_T / \mu + 1}$ and the average utility of customers is

$$\frac{R - 1/\mu + (R - 1/\mu - e^{-\mu}/\mu) \lambda_T / \mu}{\lambda_1 \lambda_T / \mu^2 + \lambda_T / \mu + 1}.$$

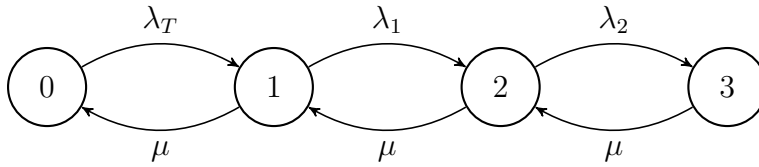
(c) For $R \in \left[\frac{(\mu+2)e^{-\mu}+1}{\mu}, \infty\right)$, we do the numerical experiments for $\mu = 2$. If $\mu = 2$, we obtain the queue waiting costs in Table 4.4 as follow:

Table 4.4: Queue waiting cost

| | $i = 0$ | 1 | 2 | 3 | 4 | 5 |
|----------|---------|--------|--------|--------|--------|--------|
| $W_i(0)$ | 0 | 0.5 | 1 | 1.5 | 2 | 2.5 |
| $W_i(1)$ | 0 | 0.0677 | 0.2707 | 0.6090 | 1.0376 | 1.5112 |

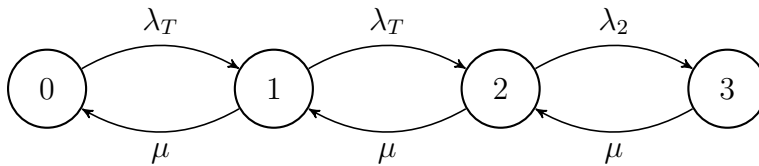
If $R \in [0.5, 0.5677)$, then $R - \frac{1}{\mu} \in [0, 0.0677)$, and the queue model is illustrated by Figure 4.2. If $R \in [0.5677, 0.7707)$, then $R - \frac{1}{\mu} \in [0.0677, 0.2707]$, and the queue model is illustrated by Figure 4.3. If $R - \frac{1}{\mu} \in [0.2707, 0.6090)$, then the queue model is illustrated by Figure 4.4. Since x_i defines the location of customers who have the 0 expected utility if the queue length is i , x_i solves the equation $R - \frac{1}{\mu} - W_i(x) = 0$, and then we obtain $\lambda_i = (1 - x_i)\lambda_T$. If $R = 0.8$, then we have $\lambda_1 = 0.7446\lambda_T$, and $\lambda_2 = 0.0690\lambda_T$.

Figure 4.4: Example 1(c): $\mu = 2$ and $R - \frac{1}{\mu} \in [0.2707, 0.5)$



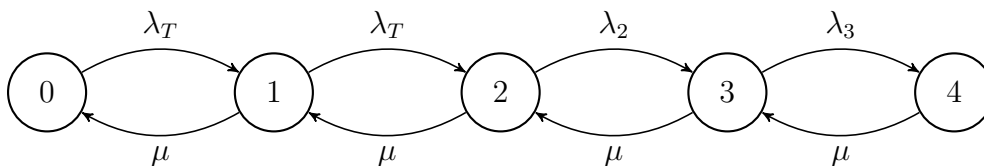
If $R - \frac{1}{\mu} \in [0.5, 0.6090)$, then the queue model is illustrated by Figure 4.5. By a similar manner, if $R = 1.1$, then $\lambda_2 = 0.5635\lambda_T$.

Figure 4.5: Example 1(c): $\mu = 2$ and $R - \frac{1}{\mu} \in [0.5, 0.6090)$



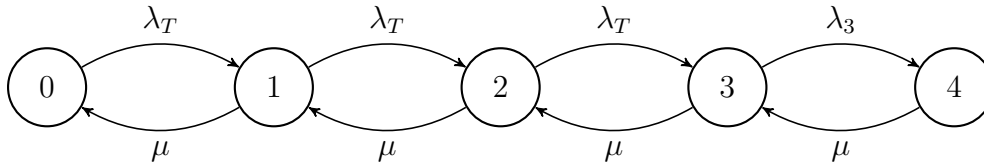
If $R - \frac{1}{\mu} \in [0.6090, 1)$, then the queue model is illustrated by Figure 4.6. If $R = 1.3$, then $\lambda_2 = 0.7953\lambda_T$, and $\lambda_3 = 0.2563\lambda_T$.

Figure 4.6: Example 1(c): $\mu = 2$ and $R - \frac{1}{\mu} \in [0.6090, 1)$



If $R - \frac{1}{\mu} \in [1, 1.0376)$, then the queue model is illustrated by Figure 4.7. If $R = 1.51$, then $\lambda_3 = 0.4982\lambda_T$.

Figure 4.7: Example 1(c): $\mu = 2$ and $R - \frac{1}{\mu} \in [1, 1.0376)$



For other values of μ and R , one can calculate the arrival rates for all the states by a similar method.

4.5 Sufficiently Large Travelling Cost ($C \geq 1$)

In the previous section, if $C = 0$, for any i , the expected total cost of a customer who resides in the 1st end point tends to be less than that of a customer who resides in the 0th end point. In this section, we examine the case that the travelling cost is sufficiently large. Thus, customers will consider also the travelling cost when deciding whether to join the queue. Denote that the expected total cost of customers who reside in x by $\eta_i(x)$,

$$\eta_i(x) = \frac{1}{\mu} + W_i(x) + Cx$$

Now, we show a lower bound of C so that for any i , the expected total cost is increasing in x . It guarantees that for any $\hat{x} < x$, if customers at x join the queue, then customers at \hat{x} will join. Thus, the expected total cost for customers who reside in the 1st end point is greater than that for customers who reside in the 0th end point.

Lemma 4.5 *If $C \geq 1$, $\eta_i(x)$ is increasing in x for any i .*

Lemma 4.5 shows the lower bound of C for the total cost to be increasing in distance for all i . The intuition is that the reduction of the expected queue waiting cost per unit distance is less than 1. If $C \geq 1$, then the increment of the travelling cost per unit distance is greater than or equal to 1. Thus, the total cost is increasing in the distance if $C \geq 1$.

There exists i_1^c such that customers who reside in 1st end point will join the queue if $i \leq i_1^c$, and make no reservation if $i > i_1^c$, thus

$$i_1^c = \arg \max_i \left\{ R - \frac{1}{\mu} - W_i(1) - C \geq 0 \right\}. \quad (4.8)$$

If $C \geq 1$, $W_i(1) + C \geq \frac{i}{\mu} = W_i(0)$ for any i , thus we have $i_0 \geq i_1^c$. For an extreme case that $C > R - \frac{1}{\mu}$, because $W_0(x) = 0$, customers who reside in the 1st end point will not join the queue, namely, $i_1^c < 0$. It implies that if $C > R - \frac{1}{\mu}$ even if there are no customers in system, the travelling cost for customers at 1st end point is too high. Since $i_0 \geq i_1^c$, i_0 defines the maximum queue size in the system. In contrast to the analysis in the case of $C = 0$, we find that if $C \geq 1$, $i_0 \neq i_1^c$ for any R and μ .

Lemma 4.6 *If $C \geq 1$, $i_0 \neq i_1^c$ for any R and μ .*

By Lemma 4.6, $i_0 \neq i_1^c$, thus the only relation for i_0 and i_1^c is $i_0 > i_1^c$. Next, we show the “steps” property for $C \geq 1$ and $i_0 > i_1^c$

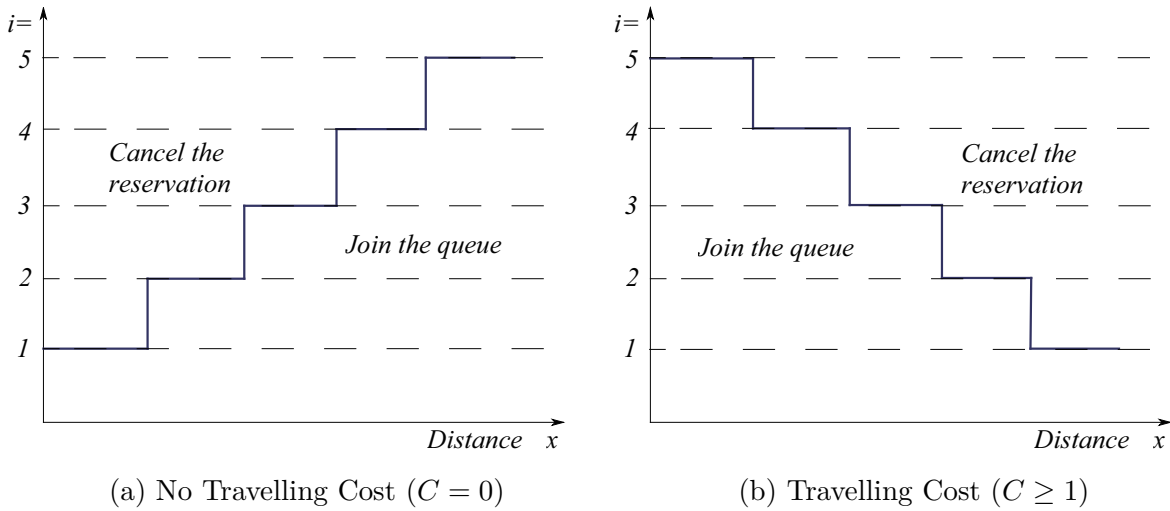
Proposition 4.7 *If $i_0 > i_1^c$, for any $i \in (i_1^c, i_0]$, there exists $x_i \in [0, 1]$ such that $\mathbb{E}(U_i(x_i)) = 0$. Moreover, for any integers $a, b \in [i_0, i_i]$, if $a > b$, then $x_a < x_b$.*

Proposition 4.7 shows an opposite result compared to the one in Proposition 4.4. If the travelling cost is sufficiently large, the customers who reside near the service site are more likely to make a reservation. As a result, the attraction of the service is decreasing in the distance between the customer and service locations. Such result is consistent with the traditional result in the literature that the market coverage is decreasing in the distance. Figure 4.8 shows the basic idea of the “steps” property for different travelling costs. Figure 4.8(a) illustrates the behavior of customers if the travelling cost is negligible, and Figure 4.8(b) illustrates the behavior of customers if the travelling cost is sufficiently large.

4.5.1 Arrival Rates

By (4.8), i_1^c defines the maximum queue length for customers who reside in 1st end point to join the queue. If $i \leq i_1^c$, then all the customers will join the queue, and the arrival rate will be λ_T . By (4.2) and $i_0 > i_1^c$, if $i > i_0$, all the customers will make no reservation, and the arrival rate will be 0. For the queue length $i_0 \geq i > i_1^c$, a fraction of the customers over the line will join the queue. Recall that $x_i = \mathbb{E}(U_i(x_i)) = 0$ which defines the location of customers who have the 0 expected utility. Thus, the arrival rate for $i_0 \geq i > i_1^c$ will

Figure 4.8: The “Steps” Property



be $\lambda_i = x_i \lambda_T$, namely, when the travelling cost is sufficiently large, customers who reside near the service location will join the queue.

Example 4.2 Let $\mu = 2$, $C = 2$, and $R = 3$, the utilities of customers with different locations are showed in 4.9. Thus, the queueing model can be illustrated by Figure 4.10.

Figure 4.9: Example 2: $\mu = 2$, $C = 2$, $R = 3$

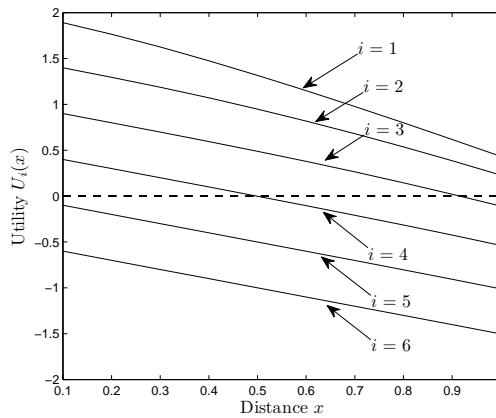
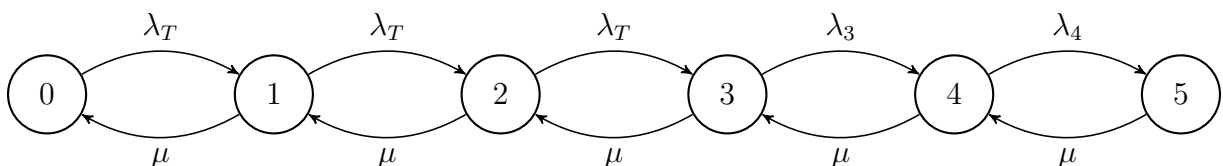


Figure 4.10: Example 2: $\mu = 2$, $C = 2$, $R = 3$



For the arrival rates, x_3 solves the equation $R - \frac{1}{\mu} - W_3(x) - Cx = 0$ and x_4 solves the equation $R - \frac{1}{\mu} - W_4(x) - Cx = 0$. Thus, we can obtain the arrival rates $\lambda_3 = x_3\lambda_T$ and $\lambda_4 = x_4\lambda_T$. In this example, $x_3 = 0.9162$ and $x_4 = 0.4979$, therefore, $\lambda_3 = 0.9162\lambda_T$ and $\lambda_4 = 0.4979\lambda_T$.

4.6 Intermediate Travelling Cost ($C \in (0, 1)$)

The analysis tends to be more complicated for the intermediate travelling cost. Intuitively, if $C \in (0, 1)$, the total cost may have an “U-shape” in response to the distance between the customer and service locations. However, in contrast to the result in the previous sections that the total cost is either decreasing or increasing for every queue length i , the total costs for this case do not have the the same shape. Lemma 4.8 shows that there exists a threshold \bar{i} such that the total cost has the “U-shape” only if $i \leq \bar{i}$.

Lemma 4.8 *If $C \in (0, 1)$, there exists $\bar{i} \geq 0$ such that if $i > \bar{i}$ the total cost $\eta_i(x)$ is increasing in $x \in [0, 1]$. If $i \leq \bar{i}$, there exists $\bar{x}(i)$ such that if $x \leq \bar{x}(i)$, $\eta_i(x)$ is decreasing in x , and if $x > \bar{x}(i)$, $\eta_i(x)$ is increasing in x .*

In Lemma 4.8, the threshold $\bar{x}(i)$ denotes the location with the lowest total cost. However, $\bar{x}(i)$ can be different for different i values. Therefore, the functions of the total cost with respect to the distance are different for different i values. By Lemma 4.8, it is straightforward to obtain the following result.

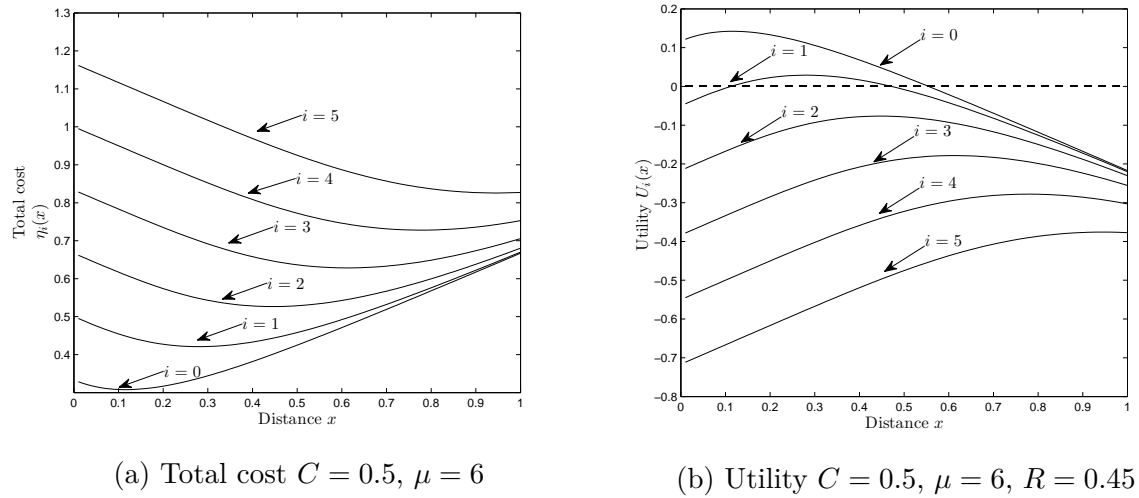
Proposition 4.9 *If $C \in (0, 1)$, there exists there exists $\bar{i} \geq 0$ such that if $i > \bar{i}$ the utility of customers $U_i(x)$ is decreasing in $x \in [0, 1]$. If $i \leq \bar{i}$, there exists $\bar{x}(i)$ such that if $x \leq \bar{x}(i)$, $U_i(x)$ is increasing in x , and if $x > \bar{x}(i)$, $U_i(x)$ is increasing in x .*

Proposition 4.9 shows that the utilities of customers can be a reversed “U-shape” for different distances. Similarly, the shape of $U_i(x)$ are different for different i values. If $C \in (0, 1)$, the customers who reside in an intermediate distance from the service site may join the queue because the travelling cost is too high for the customers who reside far away from the service site, and the queue waiting cost is too high for the customers who reside near the service site. Figure 4.11 illustrates the results in Lemma 4.8 and Proposition 4.9.

Example 4.3 *Let $C = 0.5$, $\mu = 6$ and $R = 0.45$. Figure 4.11(a) illustrates the curves of the total cost for different i values, and Figure 4.11(b) illustrates the curves of the utility for different i s. The dash line in Figure 4.11(b) is the lower boundary for that customers*

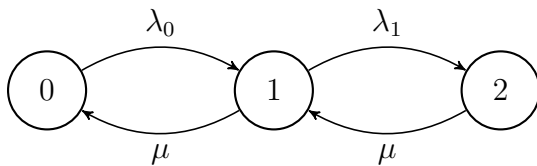
will join the queue. In this example, if $i = 0$, customers who reside near the service site will join the queue, and if $i = 1$, customers who reside in an intermediate distance will join the queue. For $i \geq 2$, no customer makes reservation. The queue model can be

Figure 4.11: Example 3



illustrated by Figure 4.12.

Figure 4.12: Example 3: $C = 0.5, \mu = 6$ and $R = 0.45$



If $i = 0$, a fraction of customers over the line will join the queue. Similar to the methods in previous examples, x_0 solves the equation $R - \frac{1}{\mu} - W_0(x) - Cx = 0$ in which $W_0(x) = 0$ because if there is no customers in the queue then there will be no queue waiting cost for the next customer. Thus, $x_0 = 0.5667$ and $\lambda_0 = 0.5667$. If $i = 1$, there are two solutions to the equation $R - \frac{1}{\mu} - W_1(x) - Cx = 0$. They are $x_{11} = 0.1123$ and $x_{12} = 0.4716$, and the arrival rate is $\lambda_1 = (x_{12} - x_{11})\lambda_T = 0.3593\lambda_T$.

4.7 Discussion: Walk-in and Online Reservation Systems

In this section, we consider the mixture of the walk-in and online reservation systems. We denote the fraction of customers who are going to make reservation online by α . Thus, the fraction of customers who do not make reservation online but walk in the service site is $1 - \alpha$. For simplicity, we call those customers “walk-in customers”. Assume that the arrival process of walk-in customers follows a Poisson distribution. Then, if the total arrival rate is λ_T , then the arrival rate for the walk-in customers is $(1 - \alpha)\lambda_T$, and the aggregated arrival process is also Poisson.

The simplest way to incorporate the walk-in customers in our model is by considering that all of them live in the 0th end point. Then, that the expected utility of those customers is

$$\mathbb{E}[U_i(0)] = R - \frac{1}{\mu} - W_i(0).$$

By the definition of i_0 , the customers who reside in the 0th end point will join the queue if the queue length $i \leq i_0$ and make no reservation if $i > i_0$. In this case, the 0th end point becomes a “mass point” where a positive fraction of customers reside. Since only the distribution of the market changes, the main results in sections 4.4-4.6 carry over.

The intuition for the difference between the mixture and pure reservation system is as follows. When the travelling cost is sufficiently large, i.e., $C \geq 1$, the arrival rate of the mixed system will be relatively higher than that of the pure reservation system for any queue length i , because there are more customers close to the facility. When the travelling cost is negligible, i.e., $C = 0$. The arrival rate of the mixed system will be higher than that of the pure reservation system if the queue length is short, and will be lower if the queue length is long. For the intermediate travelling cost, the result is similar to the negligible travelling cost case.

4.8 Conclusion

In this paper, we characterize the behavior of customers with an online reservation system. The online reservation system informs customers the real-time queue length allowing them to join the queue without physically being at the service site. Thus, customers can make advance reservations and travel to the service site. During their travel, the queue length tends to get shorter. Meanwhile, customers take into account both the traveling

and expected waiting costs when making their decision whether to join the queue. Intuitively, the longer the distance they must travel the less time they may wait in the line. By considering different levels of the traveling cost, we find that if the travelling cost is negligible, then the attraction of the service is increasing in the distance, and if the travelling cost is sufficiently large, then the attraction of the service is decreasing in the distance. For the intermediate travelling cost, the attraction of the service can be a reversed “U-shape” in response to the distance.

There are several limitations in our work. First, we assume that customers arrive at the service site when they are called to be served. Though the model with such assumption has its application, e.g., the take-away service in fast food restaurants or Starbucks, it does not characterize the online reservation system in many other facilities. Thus, the complete analysis for the online reservation system is a future research direction. Second, one can consider the non-linear travelling and waiting costs. One can expect that if the travelling cost is concave, then the attraction of the service may be increasing in the distance, and if the travelling cost is convex, the attraction of the service may be decreasing in the distance.

Further work could be dedicated to find the optimal location, capacity, information sharing policies of a firm with an online reservation system. First, since the customers with different distances from the service site may have different behavior, the firm can optimally select the service location to maximize its revenue. Second, some firms may not prefer to reveal the queue length. Based on our current result, we expect that the optimal strategy may be revealing the queue length if the queue length is small, but not revealing the queue length if the queue length is large. However, for different travelling costs, the threshold queue length for information revelation may be different. Third, the complete analysis of the mixture of the walk-in and reservation system will be made. Fourth, the main controlling parameters in our model is the service rate μ and the reward of the service R . Thus, to earn the optimal revenue, the firm can dynamically select those two parameters by increasing or decreasing the staff number or using dynamic pricing.

4.9 Proofs.

Proof of Lemma 4.1. For the waiting cost of customers who is located at $x \in [0, 1]$, $W_i(x)$, we have the following relation

$$W_{i+1}(x) - W_i(x) = \frac{1}{\mu} \left[1 - \sum_{n=i+1}^{\infty} \frac{(\mu x)^n}{n!} e^{-\mu x} \right] = \frac{1}{\mu} \sum_{n=0}^i \frac{(\mu x)^n}{n!} e^{-\mu x} = \frac{1}{\mu} P_i(\mu x) > 0.$$

By $W_0(x) = 0$ for $x \in [0, 1]$, we have

$$W_i(x) = \begin{cases} \frac{1}{\mu} \sum_{n=0}^{i-1} P_n(\mu x), & \text{if } x > 0, \\ \frac{i}{\mu} & \text{if } x = 0. \end{cases}$$

Rewrite the formula of $W_i(x)$, if $i = 1$, $W_1(x) = \frac{1}{\mu} e^{-\mu x}$, $W_1(x)$ is decreasing in x . If $i \geq 2$,

$$\begin{aligned} W_i(x) &= \frac{1}{\mu} \left(i - \sum_{k=0}^{i-1} k \frac{(\mu x)^k}{k!} e^{-\mu x} - \sum_{k=i}^{\infty} i \frac{(\mu x)^k}{k!} e^{-\mu x} \right) \\ &= \frac{1}{\mu} \left(i - \sum_{k=0}^{\infty} k \frac{(\mu x)^k}{k!} e^{-\mu x} + \sum_{k=i}^{\infty} (k - i) \frac{(\mu x)^k}{k!} e^{-\mu x} \right) \\ &= \frac{1}{\mu} [i - \mu x + \mu x Q_{i-2}(\mu x) - i Q_{i-1}(\mu x)] \\ &= \frac{1}{\mu} \left[-\frac{\mu x \Gamma(i-1, \mu x)}{\Gamma(i-1)} + \frac{i \Gamma(i, \mu x)}{\Gamma(i)} \right], \end{aligned}$$

where $Q_i(\mu x)$ is the normalized incomplete gamma function $Q_i(\mu x) = \frac{\gamma(i+1, \mu x)}{\Gamma(i+1)} = 1 - \frac{\Gamma(i+1, \mu x)}{\Gamma(i+1)}$. Take derivative with respect to x , by definition of the incomplete gamma function, $\frac{\partial \Gamma(i+1, \mu x)}{\partial x} = -\mu (\mu x)^i e^{-\mu x}$, we have

$$\begin{aligned} \frac{dW_i(x)}{dx} &= -\frac{1}{\mu} \left[\frac{\mu \Gamma(i-1, \mu x)}{\Gamma(i-1)} + \frac{\mu^2 x}{\Gamma(i-1)} (\mu x)^{i-2} e^{-\mu x} - \frac{i \mu}{\Gamma(i)} (\mu x)^{i-1} e^{-\mu x} \right] \\ &= -\frac{\Gamma(i-1, \mu x)}{\Gamma(i-1)} \leq 0. \end{aligned}$$

As a result, $W_i(x)$ is strictly increasing in i and decreasing in x . ■

Proof of Lemma 4.2. For any $\mu > 0$, if $i = 0$, $P_0(\mu) = e^{-\mu} > 0$. Thus, there exists at least one $\hat{i} = 0$ such that $\sum_{n=0}^{\hat{i}} P_n(\mu) > i$ if $0 \leq i \leq \hat{i}$. Now, we compare $\sum_{n=0}^i P_n(\mu)$ and i with μ in different ranges.

(i) Note that $\sum_{n=0}^i P_n(\mu) > i$ is equivalent to $P_0(\mu) > \sum_{n=1}^i Q_n(\mu)$. By the 9th property in Hadley and Whitin (1961), we have

$$\sum_{n=1}^{\infty} Q_n(\mu) = \mu Q_0(\mu) - Q_1(\mu) = \mu(1 - p_0(\mu)) - (1 - p_0(\mu) - p_1(\mu)) = \mu - 1 + e^{-\mu}.$$

If $\mu \in (0, 1)$, $P_0(\mu) = e^{-\mu} > \sum_{n=1}^{\infty} Q_n(\mu) > \sum_{n=1}^i Q_n(\mu)$. Therefore, $\sum_{n=0}^i P_n(\mu) > i$ for any $i \geq 0$.

(ii) If $\mu \geq 1$, let μ^* is the solution of $2 + \mu - e^\mu = 0$. Note that the function $2 + \mu - e^\mu$ is strictly decreasing if $\mu \geq 1$ because its derivative $1 - e^\mu < 0$. As a result, for any $\mu \in [1, \mu^*)$, $2 + \mu - e^\mu > 0$, and for any $\mu \in [\mu^*, +\infty)$, $2 + \mu - e^\mu \leq 0$. Because $P_0(\mu) - Q_1(\mu) = e^{-\mu}(2 + \mu - e^\mu)$, we have that for any $\mu \in [1, \mu^*)$, $P_0(\mu) - Q_1(\mu) > 0$, and for any $\mu \in [\mu^*, \infty)$, $P_0(\mu) - Q_1(\mu) \leq 0$.

If $\mu \in [\mu^*, +\infty)$, then for any $i > 0$, $P_0(\mu) \leq Q_1(\mu) \leq \sum_{n=1}^i Q_n(\mu)$, so $\hat{i} = 0$.

If $\mu \in [1, \mu^*)$, because $P_0(\mu) > Q_1(\mu)$, $P_0(\mu) \leq \sum_{n=1}^{\infty} Q_n(\mu)$ and $\sum_{n=1}^i Q_n(\mu)$ is monotonically increasing in i , there exists a unique \hat{i} such that $\sum_{n=0}^i P_n(\mu) > i$ when $i \leq \hat{i}$, and $\sum_{n=0}^i P_n(\mu) \leq i$ when $i > \hat{i}$.

Let μ^* be the solution of the equation $2 + \mu - e^\mu = 0$. Thus, $\mu^* = -2 - W(-1/e^2)$ in which $W(\cdot)$ is the Lambert W function in the lower branch. ■

Proof of Proposition 4.4. For $i_0 < i_1$, when the system is in state $i \in (i_0, i_1]$, customers residing in 0th end point will not join the queue because $R - \frac{1}{\mu} - W_i(0) = R - \frac{1}{\mu} - \frac{i}{\mu} < 0$. However, for the customers who reside in 1st end point, we have $R - \frac{1}{\mu} - W_i(1) \geq 0$ because $i \leq i_1$. While $W_i(x)$ is continuously decreasing in $x \in [0, 1]$, there exists $x \in [0, 1]$ such that $R - \frac{1}{\mu} - W_i(x) = 0$. Therefore, when there are i customers in the system, customers who live in $[0, x)$ will not join the queue, but customers who reside in $[x, 1]$ will join the queue.

Consider an integer b and $i_0 < b < i_1$. By above verification, for b there exists $x_b \in [0, 1]$ such that $R - \frac{1}{\mu} - W_b(x_b) = 0$. Now consider $b+1$, we have $R - \frac{1}{\mu} - W_{b+1}(x_b) < 0$ because $W_i(x)$ is strictly increasing in i . However, $R - \frac{1}{\mu} - W_{b+1}(1) \geq 0$ because $b+1 \leq i_1$, so there exists $x_{b+1} > x_b$ such that $R - \frac{1}{\mu} - W_{b+1}(x_{b+1}) = 0$. Therefore, when there are $b+1$ in the system, customers who reside in $[0, x_{b+1}]$ will not join the queue, but customers who live in $[x_{b+1}, 1]$ will join the queue, and $x_{b+1} > x_b$. It can be easily deduced that for any integers $a, b \in [i_0, i_1]$, if $a > b$, then $x_a > x_b$. ■

Proof of Lemma 4.5. Take derivative to $\eta_i(x)$ with respect to x , we obtain $\frac{d\eta_i(x)}{dx} = \frac{dW_i(x)}{dx} + C$. By lemma 1, when $i = 1$, $\frac{dW_1(x)}{dx} = -e^{-\mu x} \geq -1$ and $\frac{dW_i(x)}{dx} = -\frac{\Gamma(i-1, \mu x)}{\Gamma(i-1)} = Q_{i-2}(\mu x) - 1 \geq -1$, thus when $C \geq 1$, $\eta_i(x)$ is increasing in x . ■

Proof of Lemma 4.6. We prove this result by contradiction. By (4.4) and (4.8), $i_0 = i_1^c$ implies

$$\begin{aligned} \frac{i_0+1}{\mu} &> R - \frac{1}{\mu} \geq \frac{i_0}{\mu}, \\ W_{i_0+1}(1) + C &> R - \frac{1}{\mu} \geq W_{i_0}(1) + C. \end{aligned}$$

If $C \geq 1$, by Lemma 4.5, $W_{i_0+1}(1) + C \geq \frac{i_0+1}{\mu}$ and $W_{i_0}(1) + C \geq \frac{i_0}{\mu}$, thus equivalently we

have

$$\frac{i_0 + 1}{\mu} > R - \frac{1}{\mu} \geq W_{i_0}(1) + C. \quad (4.9)$$

By (4.9), $\frac{i_0+1}{\mu} > W_{i_0}(1) + C$, therefore, $i_0 + 1 > \sum_{n=0}^{i_0} P_n(\mu) + C\mu$. It is obvious that $i_0 + 1 > \sum_{n=0}^{i_0} P_n(\mu)$, therefore the upper bound of C is

$$\frac{1}{\mu} \left[i_0 + 1 - \sum_{n=0}^{i_0} P_n(\mu) \right] > C.$$

While $C \geq 1$, $\frac{1}{\mu} \left[i_0 + 1 - \sum_{n=0}^{i_0} P_n(\mu) \right] > 1$, then $i_0 + 1 - \sum_{n=0}^{i_0} P_n(\mu) > \mu$. However, $i_0 + 1 - \sum_{n=0}^{i_0} P_n(\mu) = \sum_{n=0}^{i_0} Q_n(\mu) < \sum_{n=0}^{\infty} Q_n(\mu) = \mu$. Therefore, it is impossible for $i_0 = i_1^c$. ■

Proof of Proposition 4.7. For $i_0 > i_1^c$, when the system is in state $i \in [i_1^c, i_0]$, customers residing in 1th end point will not join the queue because $R - \frac{1}{\mu} - W_i(1) - C < 0$. However, for the customers who reside in 0th end point, $R - \frac{1}{\mu} - W_i(0) = R - \frac{1}{\mu} - \frac{i}{\mu} \geq 0$. While $\eta_i(x) = W_i(x) + Cx$ is strictly increasing in x , there exists $x \in [0, 1]$ such that $R - \frac{1}{\mu} - W_i(x) - Cx = 0$. Therefore, when there are i customers in the system, customers who live in $[0, x]$ will join the queue, but customers who reside in $(x, 1]$ will not join the queue.

Consider $a \in (i_1^c, i_0)$, By above verification, for a there exists $x_a \in [0, 1]$ such that $R - \frac{1}{\mu} - W_a(x_a) - Cx_a = 0$. Now consider $a - 1$, we have $R - \frac{1}{\mu} - W_{a-1}(x_a) - Cx_a > 0$ because $W_i(x)$ is strictly increasing in i . However, $R - \frac{1}{\mu} - W_{a-1}(1) - C \leq 0$ because $a - 1 \geq i_1^c$, so there exists $x_{a-1} \in (x_a, 1]$ such that $R - \frac{1}{\mu} - W_{a-1}(x_{a-1}) - Cx_{a-1} = 0$. Therefore $x_{a-1} > x_a$. It can be easily deduced that for any integers $a, b \in [i_0, i_1^c]$, if $a > b$, then $x_a < x_b$. ■

Proof of Lemma 4.8. First, we prove that there exists an \bar{i} such that $\eta_i(x)$ is decreasing in x . Taking derivative of $\eta_i(x)$ with respect to x , we obtain that $\frac{d\eta_i(x)}{dx} = \frac{dW_i(x)}{dx} + C$. By Lemma 4.1, if $i = 1$, then $\frac{dW_1(x)}{dx} = -e^{-\mu x} \in [-1, -e^{-\mu}]$. Since $C \in (0, 1)$, if $C < e^{-\mu}$, then the result holds for $\bar{i} = 0$, and if $C \geq e^{-\mu}$, then $i > 0$. For $i > 1$, by Lemma 4.1, $\frac{dW_i(x)}{dx} = Q_{i-2}(\mu x) - 1 \in [-1, Q_{i-2}(\mu) - 1]$. By the property of $Q_{i-2}(\mu)$, $Q_{i-2}(\mu)$ is strictly decreasing in i , and meanwhile, $\lim_{i \rightarrow \infty} \frac{dW_i(x)}{dx} = \lim_{i \rightarrow \infty} Q_{i-2}(\mu x) = 0$.

Denote that $\bar{i} = \arg \max_i \{-Q_{i-2}(\mu) + C \geq 0\}$. The threshold \bar{i} depends on the value of C and μ . As a result, if $i \leq \bar{i}$, then there exists $\bar{x}(i)$ such that if $x \leq \bar{x}(i)$ $\eta_i(x)$ is decreasing in x , and if $x > \bar{x}(i)$ $\eta_i(x)$ is increasing in x . ■

Proof of Proposition 4.9. If $i > \bar{i}$, then the total cost $\eta_i(x)$ is decreasing in x . As result, if the customers who reside in 0th point join the queue, the whole market will join

the queue.

If $i \leq \bar{i}$, then the total cost $\eta_i(x)$ has an “U-shape” in x . Denote the threshold point $\bar{x}(i) = \{x : Q_{i-2}(\mu x) = C\}$. In other words, $\bar{x}(i)$ solves the equation $Q_{i-2}(\mu x) = C$ and defines the position with the lowest total cost. If $x > \bar{x}(i)$, the total cost $\eta_i(x)$ is increasing in x , and if $x \leq \bar{x}(i)$, the total cost $\eta_i(x)$ is decreasing in x . Note that $\bar{x}(i)$ can be different for different i . It would be straightforward to show that for each i , there exists a threshold $\bar{x}(i)$ such that utility of customers is increasing if $x \leq \bar{x}(i)$ but decreasing if $x > \bar{x}(i)$. ■

Bibliography

- Acemoglu, Daron, Mohamed Mostagir, Asuman Ozdaglar. 2014. Managing innovation in a crowd. Tech. rep., NBER Working Paper No. 19852. <https://ssrn.com/abstract=2382917>.
- Adams, William James, Janet L. Yellen. 1976. Commodity bundling and the burden of monopoly. *Quarterly J. Econom.* **90**(3) 475–498.
- Adiri, Igal, Ury Yechiali. 1974. Optimal priority-purchasing and pricing decisions in nonmonopoly and monopoly queues. *Oper. Res.* **22**(5) 1051–1066.
- Afèche, Philipp, Opher Baron, Yoav Kerner. 2013. Pricing time-sensitive services based on realized performance. *Manufacturing Service Oper. Management* **15**(3) 492–506.
- Ales, Laurence, Soo-Haeng Cho, Ersin Körpeoğlu. 2016a. Innovation tournaments with multiple contributors. <https://ssrn.com/abstract=2277215>.
- Ales, Laurence, Soo-Haeng Cho, Ersin Körpeoğlu. 2016b. Optimal award scheme in innovation tournaments. <https://ssrn.com/abstract=2748895>, forthcoming at *Oper. Res.*
- Alexandrov, Alexei, Martin A. Lariviere. 2012. Are reservations recommended? *Manufacturing Service Oper. Management* **14**(2) 218–230.
- Allon, Gad, Achal Bassamboo, Itai Gurvich. 2011a. “We will be right with you”: Managing customer expectations with vague promises and cheap talk. *Oper. Res.* **59**(6) 1382–1394.
- Allon, Gad, Achal Bassamboo, Martin Lariviere. 2011b. Would the social planner let bags fly free? https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1919807.
- Armony, Mor, Nahum Shimkin, Ward Whitt. 2009. The impact of delay announcements in many-server queues with abandonment. *Oper. Res.* **57**(1) 66–81.

- Bagnoli, Mark, Ted Bergstrom. 2005. Log-concave probability and its applications. *Econom. Theory* **26**(2) 445–469.
- Bakos, Yannis, Erik Brynjolfsson. 1999. Bundling information goods: Pricing, profits, and efficiency. *Management Sci.* **45**(12) 1613–1630.
- Bakos, Yannis, Erik Brynjolfsson. 2000. Bundling and competition on the internet. *Marketing Sci.* **19**(1) 63–82.
- Bar-Isaac, Heski, Guillermo Caruana, Vicente Cuñat. 2012. Search, design, and market structure. *Amer. Econom. Rev.* **102**(2) 1140–1160.
- Baron, Opher, Oded Berman, Dmitry Krass. 2008. Facility location with stochastic demand and constraints on waiting time. *Manufacturing Service Oper. Management* **10**(3) 484–505.
- Bhargava, Hemant K. 2012. Retailer-driven product bundling in a distribution channel. *Marketing Sci.* **31**(6) 1014–1021.
- Bhargava, Hemant K. 2013. Mixed bundling of two independently valued goods. *Management Sci.* **59**(9) 2170–2185.
- Bimpikis, Kostas, Shayan Ehsani, Mohamed Mostagir. 2014. Designing dynamic contests. <http://web.stanford.edu/~shayane/files/Contests.pdf>.
- Bimpikis, Kostas, Mihalis G. Markakis. 2015. Inventory pooling under heavy-tailed demand. *Management Sci.* **62**(6) 1800–1813.
- Boudreau, Kevin J., Nicola Lacetera, Karim R. Lakhani. 2011. Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management Sci.* **57**(5) 843–863.
- Chakraborty, Indranil. 1999. Bundling decisions for selling multiple objects. *Econom. Theory* **13**(3) 723–733.
- Che, Yeon-Koo, Ian Gale. 2003. Optimal design of research contests. *Amer. Econom. Rev.* **93**(3) 646–671.
- Choi, Jay P. 1991. Dynamic R&D competition under “hazard rate” uncertainty. *RAND J. Econom.* **22**(4) 596–610.
- D’aspremont, Claude, J. Jaskold Gabszewicz, J. F. Thisse. 1979. On Hotelling’s stability in competition”. *Econometrica: J. of the Econometric Society* **47**(5) 1145–1150.

- David, Herbert A., Haikady N. Nagaraja. 2003. *Order Statistics*. Wiley, New York.
- Dewan, Rajiv, Marshall Freimer, Abraham Seidmann. 2000. Organizing distribution channels for information goods on the internet. *Management Sci.* **46**(4) 483–495.
- DiPalantino, Dominic, Milan Vojnovic. 2009. Crowdsourcing and all-pay auctions. *Proceedings of the 10th ACM conference on Electronic commerce*. ACM, 119–128.
- Dixit, Avinash. 1987. Strategic behavior in contests. *Amer. Econom. Rev.* **77**(5) 891–898.
- Eckalbar, John C. 2010. Closed-form solutions to bundling problems. *J. of Economics & Management Strategy* **19**(2) 513–544.
- Fang, Hanming, Peter Norman. 2006. To bundle or not to bundle. *RAND J. Econom.* **37**(4) 946–963.
- Fey, Mark. 2008. Rent-seeking contests with incomplete information. *Public Choice* **135**(3-4) 225–236.
- Flint, Joe, Ben Fritz. 2015. Netflix viewership finally gets a yardstick. *The Wall Street Journal* (August 26).<https://www.wsj.com/articles/netflix-viewership-finally-gets-a-yardstick-1440630513/>.
- Fullerton, Richard L., R. Preston McAfee. 1999. Auction entry into tournaments. *J. Political Econom.* **107**(3) 573–605.
- Garcia, Stephen M., Avishalom Tor. 2009. The n-effect more competitors, less competition. *Psychological Science* **20**(7) 871–877.
- Geng, Xianjun, Maxwell B. Stinchcombe, Andrew B. Whinston. 2005. Bundling information goods of decreasing value. *Management Sci.* **51**(4) 662–667.
- Girotra, Karan, Christian Terwiesch, Karl T. Ulrich. 2010. Idea generation and the quality of the best idea. *Management Sci.* **56**(4) 591–605.
- Goltsman, Maria, Arijit Mukherjee. 2011. Interim performance feedback in multistage tournaments: The optimality of partial disclosure. *J. Labor Econom.* **29**(2) 229–265.
- Guo, Pengfei, Paul Zipkin. 2007. Analysis and comparison of queues with different levels of delay information. *Management Sci.* **53**(6) 962–970.
- Hadley, G., T. M. Whitin. 1961. Letter to the editor-useful properties of the poisson distribution. *Oper. Res.* **9**(3) 408–410.

- Halac, Marina, Navin Kartik, Qingmin Liu. 2016. Contests for experimentation. http://www.columbia.edu/~nk2339/Papers/HKL_contests.pdf.
- Hanson, Ward, R. Kipp Martin. 1990. Optimal bundle pricing. *Management Sci.* **36**(2) 155–174.
- Harris, Christopher, John Vickers. 1987. Racing with uncertainty. *Rev. Econom. Stud.* **54**(1) 1–21.
- Harsanyi, John C. 1968. Games with incomplete information played by “Bayesian” players part II. Bayesian equilibrium points. *Management Sci.* **14**(5) 320–334.
- Hassin, Refael. 2016. *Rational queueing*. CRC press.
- Hassin, Refael, Moshe Haviv. 2003. *To queue or not to queue: Equilibrium behavior in queueing systems*, vol. 59. Springer Science & Business Media.
- Hassin, Refael, Irit Nowik, Yair Y. Shaki. 2016. On the price of anarchy in a single server queue with heterogenous service valuations induced by travel costs. <https://arxiv.org/pdf/1610.04194v1.pdf>.
- Hausch, Donald B. 1986. Multi-object auctions: Sequential vs. simultaneous sales. *Management Sci.* **32**(12) 1599–1610.
- Hitt, Lorin M., Pei-yu Chen. 2005. Bundling with customer self-selection: A simple approach to bundling low-marginal-cost goods. *Management Sci.* **51**(10) 1481–1493.
- Hotelling, Harold. 1929. Stability in competition. *Econom. J.* **39**(153) 41–57.
- Hu, Ming, Mengze Shi, Jiahua Wu. 2013. Simultaneous vs. sequential group-buying mechanisms. *Management Sci.* **59**(12) 2805–2822.
- Jiang, Zhaohui Zoey, Yan Huang, Damian R. Beil. 2016. The role of feedback in dynamic crowdsourcing contests: A structural empirical analysis. https://deepblue.lib.umich.edu/bitstream/handle/2027.42/134692/1334_YHuang.pdf.
- Johnson, Justin P., David P. Myatt. 2006. On the simple economics of advertising, marketing, and product design. *Amer. Econom. Rev.* **96**(3) 756–784.
- Kalra, Ajay, Mengze Shi. 2001. Designing optimal sales contests: A theoretical perspective. *Marketing Sci.* **20**(2) 170–193.

- Konrad, Kai A. 2007. Strategy in contests: An introduction. <https://ssrn.com/abstract=960458>.
- Konrad, Kai A., Dan Kovenock. 2009. Multi-battle contests. *Games and Econom. Behavior* **66**(1) 256–274.
- Körpeoğlu, Ersin, Soo-Haeng Cho. 2017. Incentives in contests with heterogeneous solvers. <https://ssrn.com/abstract=2575186>, forthcoming at *Management Sci.*
- Larsen, Christian. 1998. Investigating sensitivity and the impact of information on pricing decisions in an M/M/1/∞ queueing model. *International J. of Production Econom.* **56**(1) 365–377.
- Lazear, Edward P., Sherwin Rosen. 1981. Rank-order tournaments as optimum labor contracts. *J. Political Econom.* **89**(5) 841–64.
- Liu, Tracy Xiao, Jiang Yang, Lada A. Adamic, Yan Chen. 2014. Crowdsourcing with all-pay auctions: A field experiment on taskcn. *Management Sci.* **60**(8) 2020–2037.
- Lohr, Steve. 2012. Pentagon pushes crowdsourced manufacturing. *New York Times* (April 5). <http://bits.blogs.nytimes.com/2012/04/05/pentagon-pushes-crowdsourced-manufacturing/>.
- Malueg, David A., Shunichi O. Tsutsui. 1997. Dynamic R&D competition with learning. *RAND J. Econom.* **28**(4) 751–772.
- Marshall, Albert W., Ingram Olkin, Barry C. Arnold. 2016. *Inequalities: Theory of Majorization and its Applications..* Springer Series in Statistics.
- Mas-Colell, Andreu, Michael Dennis Whinston, Jerry R. Green, et al. 1995. *Microeconomic Theory.* Oxford University Press: New York.
- McAfee, R. Preston, John McMillan, Michael D. Whinston. 1989. Multiproduct monopoly, commodity bundling, and correlation of values. *Quarterly J. Econom.* **104**(2) 371–383.
- Mendelson, Haim, Seungjin Whang. 1990. Optimal incentive-compatible priority pricing for the M/M/1 queue. *Oper. Res.* **38**(5) 870–883.
- Miller, Bruce L, AG Buckman. 1987. Cost allocation and opportunity costs. *Management Sci.* **33**(5) 626–639.

- Moldovanu, Benny, Aner Sela. 2001. The optimal allocation of prizes in contests. *Amer. Econom. Rev.* **91**(3) 542–558.
- Moldovanu, Benny, Aner Sela. 2006. Contest architecture. *J. Econom. Theory* **126**(1) 70–96.
- Naor, Pinhas. 1969. The regulation of queue size by levying tolls. *Econometrica: J. of the Econometric Society* **37**(1) 15–24.
- Oh, Jaelynn, Xuanming Su. 2012. Pricing restaurant reservations: Dealing with no-shows. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2169567.
- Prasad, Ashutosh, R. Venkatesh, Vijay Mahajan. 2010. Optimal bundling of technological products with network externality. *Management Sci.* **56**(12) 2224–2236.
- Randhawa, Ramandeep S., Sunil Kumar. 2008. Usage restriction and subscription services: Operational benefits with rational users. *Manufacturing Service Oper. Management* **10**(3) 429–447.
- Rosen, Sherwin. 1986. Prizes and incentives in elimination tournaments. *Amer. Econom. Rev.* **76**(4) 701–715.
- Schmalensee, Richard. 1984. Gaussian demand and commodity bundling. *Journal of business* S211–S230.
- Shaked, Moshe, J. George Shanthikumar. 2007. *Stochastic Orders*. Springer Science & Business Media.
- Simhon, Eran. 2016. Advance reservations and information sharing in queues with strategic customers. <http://search.proquest.com/docview/1853473824>.
- Stigler, George J. 1963. United states v. loew's inc.: A note on block-booking. *Sup. Ct. Rev.* 152–157.
- Stouras, Konstantinos I., Karan Girotra, Serguei Netessine. 2016. First ranked first to serve: Strategic agents in a service contest. Working paper, University of Virginia and INSEAD.
- Taylor, Curtis R. 1995. Digging for golden carrots: An analysis of research tournaments. *Amer. Econom. Rev.* **85**(4) 872–890.
- Terwiesch, Christian, Yi Xu. 2008. Innovation contests, open innovation, and multiagent problem solving. *Management Sci.* **54**(9) 1529–1543.

- Wang, Ruxian, Maqbool Dada, Ozge Sahin. 2015. Pricing ancillary service subscriptions. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2671923.
- Watson, Ray, Ian Gordon. 1986. On quantiles of sums. *Austral. J. Statist.* **28**(2) 192–199.
- Yildirim, Huseyin. 2005. Contests with multiple rounds. *Games and Econom. Behavior* **51**(1) 213–227.