Title
Essays on Information Technology and Firms' Pricing Strategy

Permalink
https://escholarship.org/uc/item/8px6c8r4

Author
Zhang, Zhe

Publication Date
2014

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA, IRVINE

Essays on Information Technology and Firms’ Pricing Strategy

DISSERTATION

submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Management

by

Zhe Zhang

Dissertation Committee:
Professor Vidyanand Choudhary, Co-Chair
Professor Shivendu Shivendu, Co-Chair
Professor Sanjeev Dewan
Professor Mingdi Xin

2014
DEDICATION

This dissertation is dedicated to my parents

Who inspired me by practicing the virtues of hard work, honesty, commitment, and passion in

Their life.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>vii</td>
</tr>
<tr>
<td>CURRICULUM VITAE</td>
<td>viii</td>
</tr>
<tr>
<td>ABSTRACT OF THE DISSERTATION</td>
<td>x</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 1: Vertically Differentiated Markets with Heterogeneous Disutility From Under-Provisioning: The Case of Versioning in Software Industry</td>
<td>4</td>
</tr>
<tr>
<td>CHAPTER 2: Economics of Online Daily-Deal Business Model</td>
<td>88</td>
</tr>
<tr>
<td>CHAPTER 3: Recommender Systems and Consumer Product Search</td>
<td>151</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>User benefit from functionality, inconvenience from under-provisioning of</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>functionality, and the net utility</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>Optimal functionality level of the low version as the proportion of</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>high-type users’ changes</td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td>Profits under the three product strategies when user experience disutility</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>from under-provisioning of functionality</td>
<td></td>
</tr>
<tr>
<td>1.4</td>
<td>Utility curves for both types of users, and the optimal functionality level</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>of the low version as required functionality level of high-type users</td>
<td></td>
</tr>
<tr>
<td></td>
<td>increases where ( f(q) = q ), ( g(x,q) = (x-q)^2 ), ( c(q) = q^2 ),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>and ( x_{La} &lt; x_{Lb} ).</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>Utility curves for both types of users, and the optimal functionality level</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>of the low version as required functionality level for low-type users</td>
<td></td>
</tr>
<tr>
<td></td>
<td>increases when ( f(q) = q ), ( g(x,q) = (x-q)^2 ), ( c(q) = q^2 ),</td>
<td></td>
</tr>
<tr>
<td></td>
<td>and ( x_{La} &lt; x_{Lb} ).</td>
<td></td>
</tr>
<tr>
<td>1.6</td>
<td>Consumer surplus of high-type users</td>
<td>39</td>
</tr>
<tr>
<td>1.7</td>
<td>Profits under different product strategies when there is no disutility</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>from under-provisioning</td>
<td></td>
</tr>
<tr>
<td>1.8</td>
<td>Utility of high and low-type users when (a) ( x_L &lt; q^* &lt; q' &lt; q^* ),</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>(b) ( x_L &lt; q' &lt; q^* ), and (c) ( x_L &lt; q' = q^* )</td>
<td></td>
</tr>
<tr>
<td>1.9</td>
<td>Utility of high and low-type users when a) ( q^* &lt; q' &lt; x_L ), b) ( q'</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>&lt; q^* &lt; x_L ), c) ( q' &lt; q^* &lt; x_L ), d) ( q' &lt; q^* = x_L ).</td>
<td></td>
</tr>
<tr>
<td>1.10</td>
<td>Utility of high-type and low-type consumers when ( x_{LL} &lt; x_L )</td>
<td>47</td>
</tr>
<tr>
<td>1.11</td>
<td>Impact of consumer inconvenience on consumer utility</td>
<td>49</td>
</tr>
<tr>
<td>2.1</td>
<td>Sequence of the game</td>
<td>104</td>
</tr>
<tr>
<td>2.2</td>
<td>Optimal discount rate for merchants of different types</td>
<td>113</td>
</tr>
<tr>
<td>2.3</td>
<td>Impact on optimal discount rates when revenue sharing ratio increases</td>
<td>119</td>
</tr>
<tr>
<td>2.4</td>
<td>A merchant’s participation decision with revenue sharing ratios, ( s = 0 )</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>and ( s = 1 ).</td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td>Daily-deal publisher’s optimal revenue sharing strategy</td>
<td>129</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 3.1</td>
<td>Amazon product recommendation</td>
<td>152</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Consumer decision tree</td>
<td>158</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Consumer bucket ([a_j, a_j + r]) for a consumer at (x_i)</td>
<td>160</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Consumer’s acceptance range</td>
<td>162</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>The firm’s product recommendation and consumer’s decision when (a_j + \sqrt{c/t} \leq x_i(a_j) \leq a_j + r + \sqrt{c/t}) and (x_i(a_j) \geq a_j + r - \sqrt{c/t})</td>
<td>165</td>
</tr>
<tr>
<td>Figure 3.6</td>
<td>Impact of accuracy of recommender system on the firm’s expected revenue where (\bar{m} = 1), (c = 0.2), (t = 1), and (\mu = 0.6)</td>
<td>168</td>
</tr>
<tr>
<td>Figure 3.7</td>
<td>Impact of bucket range on the realized consumer surplus</td>
<td>173</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table 1.1</th>
<th>Summary of findings of literature on versioning of information goods for a monopolist firm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

vi
ACKNOWLEDGMENTS

I express the deepest appreciation to my committee co-chairs, Professor Vidyanand Choudhary and Professor Shivendu Shivendu at the Paul Merage Business School, UC Irvine, who introduced me to Economics of Information Systems and guided me through the course of my PhD research. Their work continually demonstrated to me the importance of economic implications of information technology and information goods. They are examples of excellence in academic life, and their dedication to good research has always been the source of inspiration for me. Without their guidance and persistent help this dissertation would not have been possible.

I thank my committee members, Professor Sanjeev Dewan and Professor Mingdi Xin, who provided valuable comments and suggestions to improve the model setup and analysis of my dissertation. In particular, I improved the positioning and other important aspects of my dissertation thank to their insights in software pricing and strategies of online intermediaries.

I also thank the Paul Merage Business School at UC Irvine for the financial support over the course of my PhD study.
CURRICULUM VITAE

Zhe Zhang

2006 Bachelor in Computing (honors), Department of Information Systems, the National University of Singapore, School of Computing, Singapore

2006-2009 Data Modeler, Department of Enterprise Information Service, Oversea Chinese Bank Corporate, Singapore

2009-2014 Teaching Assistant, Information Systems, Paul Merage Business School, University of California, Irvine

2009-2013 Research Assistant, Information Systems, Paul Merage Business School, University of California, Irvine

2013 Instructor of MGMT 107 -- Introduction to Management of Information Systems, Summer Session, Paul Merage Business School, University of California, Irvine

2014 Ph.D. in Management, University of California, Irvine

FIELD OF STUDY

Economic implications of software pricing and online intermediaries’ strategy

PUBLICATIONS


CONFERENCE PRESENTATIONS

   b. Conference on Information Systems and Technology (CIST), 2013, Minneapolis, MN.

   a. Conference on Information Systems and Technology (CIST), 2012, Phoenix, NV.

   a. INFORMS Annual Meeting, 2012, Phoenix, NV.
   b. Conference on Information Systems and Technology (CIST), 2011, Charlotte, NC.

   a. Hawaii International Conference on System Sciences (HICSS), 2012, Maui, HI.
   b. INFORMS Annual Meeting, 2011, Charlotte, NC.

   a. Workshop on Information Systems and Economics (WISE), 2011, Shanghai, China.
   b. INFORMS Annual Meeting, 2011, Charlotte, NC.
ABSTRACT OF THE DISSERTATION

Essays on Information Technology and Firms’ Pricing Strategy

By

Zhe Zhang

Doctor of Philosophy in Management

University of California, Irvine, 2014

Professor Vidyanand Choudhary, Co-Chair
Professor Shivendu Shivendu, Co-Chair

Why do software firms sell more than one version of software to consumers? In the first chapter, I develop an analytical game theory model to examine a monopolist software firm’s motivation of adopting a versioning strategy. We show that consumers’ heterogeneity in requirements for software functionality and consumers’ disutility from under-provisioning of software functionality are sufficient conditions for a software firm to adopt versioning strategy.

Why do deals have different discount rates on a daily-deal website? In the second chapter, I develop a two-period model to capture the strategic interaction between a daily-deal publisher and a merchant who wants to offer a deal to consumers on the publisher’s website. I show that a merchant’s optimal discount rate strategy and participation decision depends on the trade-offs between four effects – advertising, sampling, cannibalization and revenue sharing. Moreover, I recommend that a daily-deal publisher should take into account a merchant’s marginal cost, proportion of informed consumers and consumer characteristics in formulating an appropriately customized revenue sharing contract for the merchant.
Do consumers benefit from using an online retailer’s recommender system? In the third chapter, I develop an analytical framework to examine the optimal recommender system strategy of an online multi-product retailer and the effect of the recommender system on consumer surplus. I find that when the recommender system is relatively accurate, the retailer will recommend products that are never a perfect match to the recipients of the recommendations. I also find that improving the precision of a relatively accurate recommender system leads to a reduction in consumer surplus.
INTRODUCTION

The cost structure of information goods has been characterized by high fixed costs and low marginal costs. These characteristics present a unique and important phenomenon of study for Information Systems researchers. The value-based pricing is much more appropriate than the traditional cost-based pricing for information goods. Information goods provider firms can adopt a versioning strategy by selling versions of information goods at different quality levels in such a way that consumers can self-select among these versions.

Literature on versioning of information goods recommends that a software firm should adopt versioning only under certain narrowly defined contextual conditions such as the presence of network externalities, piracy, and non-convex cost structure. However, versioning is ubiquitous in the software industry. For example, Microsoft sells their on-premise operating system Windows 8 in four versions (Windows 8, Windows 8 Pro, Windows 8 Enterprise, and Windows RT). In order to bridge the gap between literature and practice, in Chapter 1, I analyze the versioning strategy from an alternative perspective of how users derive utility from software. I focus on ‘inconvenience’ or disutility experienced by users when software has lower functionality than what they require to accomplish their tasks. I show that in a vertically differentiated market of users who are heterogeneous in their valuation for functionality and their required levels of functionality, the utility functions of high-type and low-type users cross over at a positive utility level. In this setting, I find that versioning is an optimal product-pricing strategy for a software firm even in the absence of any contextual conditions identified in the existing literature.

This research makes two key contributions. First, this is the first paper that develops the utility function based on the conceptualization that consumers not only consider the benefit from
functionalities that are available in the software, but also take into account the disutility from under-provisioning of functionality. Second, this paper identifies users’ disutility from under-provisioning of functionality as a sufficient condition for optimality of versioning strategy.

Electronic commerce has grown at a fast and steady rate in the last two decades. Daily-deal websites, such as Groupon and LivingSocial, are part of the emerging trend of online social marketing, and have attracted significant amount of attention from both practitioners and researchers. These daily-deal websites follow an Online-to-Offline business model, that is, they attract consumers online and direct them to offline sellers. One distinct feature of the Online-to-Offline business model is that these websites have revenue sharing contracts with offline merchants. In Chapter 2, I develop an analytical model that captures merchants’ benefits and costs from offering deals on a daily-deal website. In a vertically differentiated market, consumers are heterogeneous in two dimensions: valuation for quality and expectation about the quality of merchants’ goods. In a leader-follower two-period game setting, a daily-deal website is the leader and first announces a revenue sharing ratio, then merchants decide discount rates and whether to offer deals on the website.

The contributions of this paper are threefold. First, I capture the strategic interactions between a daily-deal website and merchants, which is often ignored in the literature. Second, I identify and analyze four effects that are key to understanding Online-to-Offline business models – advertising, sampling, cannibalization, and revenue sharing – on a merchant’s profit from offering deals on the website. Third, I show how product and merchant characteristics, such as quality expectation and merchants’ marginal cost, affect the website’s optimal revenue sharing ratio.
A popular feature of online retailers is the recommender system which provides personalized product recommendations to individual consumers. Consumers can accept the recommended products by the retailer to avoid searching in the myriad of products in the context of online shopping. Thus, the use of recommender systems affects consumers’ product search process and reduces consumers’ search cost. There is a large body of work on designing recommender systems in computer science, but the economic implications of recommender systems are less understood. In Chapter 3, I develop an analytical framework to examine the optimal recommender system strategy of an online multi-product retailer and the effects of use of a recommender system on consumers.

Different from existing literature, I consider the case where a monopolist online retailer can strategically skew recommendations in favor of products with higher profit margin. In this sense, the retailer needs to balance the tradeoff between profit margin of the recommended product and the likelihood of consumers’ acceptance of the recommendation. I find that when the recommender system is relatively accurate, the retailer will recommend products that are purchased by consumers, although the recommended products are not a perfect match for the recipients of the recommendations. I also find that improving the precision of a relatively accurate recommender system leads to a reduction in consumer surplus.
CHAPTER 1

Vertically Differentiated Markets with Heterogeneous Disutility from Under-provisioning: The Case of Versioning in Software Industry

Abstract

Literature has identified factors such as piracy, network externality, or concave cost of producing quality as key drivers of software versioning. However, software firms adopt versioning strategies that are often invariant across different market settings. In order to explain widespread business practice of software versioning, we focus on ‘inconvenience’ or disutility that users continue to experience when software has lower functionality than what they require to accomplish tasks. In our model, users are heterogeneous on valuation for functionality and required level of functionality such that those with higher valuation have a higher required level of functionality. Users do not derive any additional utility if the software has more functionality than what they require. We show that heterogeneous disutility from under-supply of functionality is a sufficient condition for optimality of versioning under fairly general conditions. As the required level of functionality of high-type users increases, the firm increases the functionality level of the high version, but may decrease the functionality level of the low version when the proportion of high-type users is moderate. On the other hand, as the required level of functionality of low-type users increases, the firm may reduce functionality level of the low version when the proportion of high-type users is high, though the functionality level of the high version remains the same. Counter intuitively, an increase in the high- (low-) type users’ required level of functionality negatively (positively) impacts high-type users’ consumer surplus.
1. Introduction

Versioning\(^1\) as a product-price strategy is ubiquitous in information goods in general and in software in particular. A firm offers different versions of software at different prices such that different “types” of users self-select the version-price pair that is “targeted” to them (Mussa and Rosen 1978; Varian 1998; Shapiro and Varian 1998). Software firms first develop a flagship product with the highest level of functionality and then create different versions by strategically disabling some of the functionality of the flagship product (Ghose and Sundararajan, 2005; Gershoff et al., 2011; Wei and Nault, 2012; Dey and Lahiri, 2013). For example, Microsoft offers their on-premise operating system Windows 8 in three versions and their on-premise productivity software Office 2013 in five versions. Adobe Systems offers their graphics editing software Adobe Photoshop in two versions. IBM offers their statistical software IBM SPSS Statistics in three versions. Similarly, Intuit, which specializes in the tax-related software TurboTax, offers five versions of its federal and state tax filing software. In light of this commonly observed business practice of software firms, it is not surprising that study of versioning strategies is of central interest to the academics in Information Systems (IS). However, they generally recommend versioning only under certain market conditions such as the presence of piracy, product characteristics such as network externality or non-convex cost structure of producing functionality (see Table 1.1), and not as a universal strategy.

Often software firms’ versioning strategy does not appear to be determined by specific market conditions identified in the literature. For example, Microsoft has a worldwide product strategy for Windows operating systems and adopts versioning strategy in the US and the

---

\(^{1}\) Software versioning also refers to the practice of assigning unique version numbers to the unique states of the computer wherein these numbers are assigned in increasing order and correspond to new developments in the software (http://en.wikipedia.org/wiki/Software_versioning). This research refers to versioning as in Shapiro and Varian (1998), wherein firms “create different versions of the same core of information, based on different buyers’ needs.”
Chinese markets for Windows 8, even though piracy rates are much higher in China (BSA, 2011). Further, firms often adopt versioning strategy for software products with weak network externality. TurboTax is available in five versions with increasingly more complex capabilities, even though users of electronic tax filing software do not necessarily benefit from an increase in the installed base. Similarly, since it is almost costless for software firms to remove functionality from the flagship product to create a version with lower level of functionality, and marginal cost of additional copies is negligible (Varian, 1998; Bhargava and Choudhary, 2004; Chellappa and Shivendu, 2005; Wei and Nault, 2011), and hence cost structure may not be the driver of versioning strategy in many instances. Moreover, literature also recommends that versioning is not optimal if the marginal cost of producing different versions is the same (Anderson and Dana 2009; Salant 1989).

Table 1.1: Summary of findings of literature on versioning of information goods for a monopolist firm

<table>
<thead>
<tr>
<th>Paper</th>
<th>Model Specification</th>
<th>Conditions under which versioning is optimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bhargava and Choudhary (2001)</td>
<td>Vertical</td>
<td>When the highest quality product does not have the lowest cost-quality ratio</td>
</tr>
<tr>
<td>Wu et al. (2003), Chellappa and Shivendu (2005), Chen and Wu (2008)</td>
<td>Vertical</td>
<td>In the presence of piracy</td>
</tr>
<tr>
<td>Chen and Seshadri (2007)</td>
<td>Vertical</td>
<td>When users have convex value for outside options</td>
</tr>
<tr>
<td>Wei and Nault (2011)</td>
<td>Vertical and</td>
<td>When shared characteristics are not too valuable</td>
</tr>
</tbody>
</table>
The research objective of this paper is to bridge the gap between the recommendations of extant literature and universally adopted business practice of versioning by software firms. Our work belongs to the stream of literature that has analyzed software firms’ motivation for versioning strategy from a perspective of how users derive utility from software, in the absence of any specific product characteristics or market conditions (Raghunathan, 2000; Chen and Seshadri, 2007; Wei and Nault, 2011). In our abstraction, users consider buying software to accomplish specific set of tasks. In order to accomplish these tasks users require a set of functionality in software. If software has less functionality than what users require, then their willingness to pay is moderated by the ‘inconvenience’ or disutility that they continue to experience from under-provisioning. In this paper, we examine the role of ‘inconvenience’ or disutility from under-provisioning of functionality in a firm’s product-pricing strategy and show that it is a sufficient condition for the optimality of versioning strategy.
Software functionality consists of a set of features that enable users to perform certain tasks (Wilde and Scully, 1995). For example, Microsoft product development teams begin the process of new software development by first creating a “vision statement” that provides a list of “user activities that need to be supported by the product features” (Cusumano and Selby 1997). Following literature (Musa and Rosen, 1978; Bhargava and Choudhary, 2001; Chellappa and Shivendu, 2005; Wei and Nault, 2006; Dey and Lahiri, 2013), we take into account users’ heterogeneity in valuation for ‘features’ and, in our model, users derive heterogeneous benefit from functionality.

In addition, we also consider the situation where users are heterogeneous in requirements of functionality. High valuation or high-type users have higher requirements for functionality because they use the software in a complex task environment\(^2\). Furthermore, users derive no additional utility if the software has more functionality than what they require because they do not use functionality beyond their requirement.\(^3\) For example, a statistician (high valuation), who uses spreadsheet software like Excel for demand forecasting in a business firm, has a higher requirement of functionality than a high school student (low valuation) who uses spreadsheet software to do her school work. In this sense, the statistician’s required level of functionality is higher than that of the high school student. Similarly, the student may have no additional utility from spreadsheet software that has functionality for demand forecasting because she does not require it.

\(^2\) One observes that progressively higher versions of software are almost always forward compatible (higher version has all the functionality of lower version), and higher versions are sold at higher prices. If high valuation users were to require a lower set of functionality then it is natural to think that they would buy the lower version at lower price.\(^3\) Users may be inconvenienced by over-provisioning. For example, users may face hardware resource constraint when they adopt Microsoft Office 2010 that has lots of functionality (Methvin 2009). Since the hardware becomes less expensive and more powerful over time, we assume that the ‘inconvenience’ due to over-provisioning is negligible.
Users experience ‘inconvenience’ or disutility if they do not adopt the software or if the level of functionality in software is lower than their required level. This disutility from under-supply of functionality can be either due to users’ incurring time and effort to manually accomplish their subtasks, or from annoyance and frustration or even emotional distress from not being able to accomplish some of the subtasks. Continuing with the statistician and student example, if the spreadsheet software does not have the functionality to do demand forecasting, then statistician continues to use multiple computations to do the same as before she adopted this software. This takes time and effort, and thus her ‘inconvenience’ from performing multiple computations persists. However, the student does not experience any ‘inconvenience’ due to missing functionality for demand forecasting because she does not require it. However, if the software does not have the functionality to copy and paste cells, then both the statistician and the student continue to manually input the same value in different cells. This takes time and effort, and their ‘inconvenience’ persists.

There is anecdotal evidence to support our conceptualization of required level of functionality and ‘inconvenience’ that users continue to experience if software has a lower level of functionality than required level. It appears that there is a general agreement amongst various user groups that users evaluate the features of any particular version of software in relation to their requirements of ‘features’ or functionality and experience ‘inconvenience’ if software has lower level of functionality than what they require. Moreover, business reports support our abstraction that more functionality does not necessarily translate in more usefulness.

4 See “Do you need more than Windows 7 Home Premium?” by Ed Bott at http://www.zdnet.com/blog/bott/do-you-need-more-than-windows-7-home-premium/1128
5 See Discussion forum “Can we use Windows 7 Home Premium for work?” at http://superuser.com/questions/186146/can-we-use-windows-7-home-premium-for-work.
In our conceptualization, a user’s willingness to pay for different versions of software consists of two parts: benefit from the level of functionality in software, and remaining disutility from under-provisioning of functionality in relation to her required level of functionality. Therefore, a user’s utility from a particular version of software is additive in benefit from functionality and disutility from under-supply of functionality. Our conceptualization of how users derive utility from software is different from extant literature in versioning (see Table 1.1) in the following three key aspects: (1) different types of users have different required levels of functionality, (2) more functionality than the required level of functionality does not increase users’ utility, and (3) users continue to experience ‘inconvenience’ or disutility if software has less functionality than their required level. These three key differences imply that in our setup utility functions of high-type and low-type users cross over at some functionality level which we characterize as ‘indifferent level of functionality’, and high-type users have higher willingness to pay only if functionality level is higher than the utility cross-over point. Note that the extant literature assumes that high-type users have higher willingness to pay for any functionality level. We characterize this property of high-type and low-type users’ utility function to cross over at a positive utility level as ‘single-crossing property’ (Cooper, 1984).

Our work is close to Wei and Nault (2011) in the sense that in their hierarchical characteristics case users are heterogeneous in requirement for functionality beyond which their utility does not increase in functionality. But the key difference is that in their case users do not experience ‘inconvenience’ from under-provisioning of functionality. This key difference implies that while in their model users who have higher valuation for functionality have higher willingness-to-pay for all levels of functionality; in our model users with higher valuation for functionality (high-type) have lower willingness-to-pay than users with low valuation for
functionality (low-type) if there is severe under-supply of functionality. As a result, in our setting, even though the market is vertically differentiated and all users’ willingness to pay weakly increases in functionality, ordering of willingness to pay of high-type and low-type users is not maintained for all functionality levels.

One of our key findings is that when users experience heterogeneous ‘inconvenience’ if the software has less functionality than what they require, a versioning strategy is optimal even in the absence of other market conditions. This is so because when users experience ‘inconvenience’ from under-provisioning of functionality, then high-type users are less attracted to the low version and, therefore, the firm needs to provide them less incentive or information rent to stay with the high version. This saving in information rent makes versioning strategy profitable. This new finding identifies heterogeneous user ‘inconvenience’ or disutility from under-provisioning of functionality as a sufficient condition for optimality of versioning strategy. This conceptualization provides an alternative explanation for the widespread practice of versioning across software products and markets.

With fast changing computing environment and technologies, users’ required set of functionality is likely to increase over time period. We identify some interesting insights relating to the firm’s optimal versioning strategy as users’ required level of functionality increases. When high-type users require more functionality, they get a high version with a higher functionality level, but their surplus decreases. Nevertheless, whether the firm should increase or decrease the functionality level of the low version depends on the distribution of user types. When low-type users require more functionality, one might think that the firm should increase the functionality level of the low version to extract surplus. However, we find that the firm may lower the
functionality level of low version to economize on information rent when the market is dominated by high-type users.

Our contributions in this research are twofold. First, we identify heterogeneous disutility from under-provisioning as a sufficient condition for optimality of versioning. This allows us to provide an additional and new explanation for the widespread business practice of versioning in the software industry. We show that when users experience heterogeneous ‘inconvenience’ from under-supply of functionality, versioning strategy is optimal product-pricing strategy for software providers, under commonly observed conditions. Our conceptualization sheds light on conditions under which a firm provides low-type users software with less functionality than they require, even if the software firm has zero cost of providing them their required level of functionality. Our key managerial recommendation is that software firms should take users’ ‘inconvenience’ or disutility from under-provisioning of functionality into consideration in designing a versioning strategy.

Second, we propose a novel conceptualization of consumer utility that captures consumer heterogeneities in marginal valuation for functionality and required level of functionality. This allows us to analyze commonly observed situations in which the high-type users not only have higher marginal valuation for functionality but also have a higher requirement for functionality. Previous studies have considered only user marginal valuation for functionality or quality in either linear utility function (Chellappa and Shivendu, 2003; 2005; Bhargava and Choudhary, 2001; 2004; 2008; Jones and Mendelson, 1998; 2011; Wei and Nault, 2012; Dey and Lahiri, 2013) or quadratic utility function (Raghunathan, 2000, Ghose and Sundararajan, 2005). Our utility framework allows us to study the impact of changes in users’ requirements for
functionality on a versioning strategy even when marginal valuation for functionality does not change.

The remainder of this paper is organized as follows: In §2, we outline our conceptualization of consumer utility function and monopolist software company’s product strategies. In §3, we discuss a software firm’s three product strategies and show that a versioning strategy is optimal. In §4, we discuss the impact of changes in users’ required level of functionality on the software firm’s optimal versioning strategy. In §5, we analyze some extensions to our model and also discuss implications of relaxing some of the model assumptions. We conclude our analysis and provide managerial implications in §6.

2. Model

We consider a class of software which provides a set of functionality to users to perform their tasks to meet specific needs (Wilde and Scully, 1995). Users evaluate software in a task-oriented context akin to Garvin’s (1984) perspective of “user view” of quality which corresponds to “the totality of characteristics of an entity that bear on its ability to satisfy stated and implied needs” (ISO 1994). Software quality consists of multidimensional attributes like functionality, reliability, correctness, and usability (Kitchenham and Pfleeger 1996), and very often the only differentiating attribute among the different versions is functionality. In our model, users derive utility from a set of functionality of the software (Kekre et al., 1995) and recognize functionality as a measure of quality (Wei and Nault, 2011).

Users require a set of functionality in software because they use it to accomplish certain specific tasks. We refer to this set as the users’ required level of functionality. Users, who have higher valuation for functionality, use software in a more complex task-oriented environment and, therefore, require more functionality. If the software has less functionality than users’
required level of functionality, then users experience ‘inconvenience’. The source of this ‘inconvenience’ may be time or effort required to accomplish subtasks for which software has no functionality or it may stem from annoyance, frustration or emotional distress from not being able to accomplish subtasks. Note that users experience ‘inconvenience’ when they do not adopt software. We refer to this ‘inconvenience’ as disutility from under-provisioning of functionality.

Further, in our context, users do not use functionality which is beyond their required functionality level and hence, derive no additional benefit from over-provisioning of functionality.

We consider a market that is vertically differentiated with heterogeneous consumers. Without loss of generality we assume that there are two types of consumers, namely high-type and low-type and each type of users are denoted by a duplet \( \{ \theta, x \} \) where \( i \in \{ H, L \} \).\(^7\) The parameter \( \theta \) captures users’ heterogeneities in valuation for functionality, where \( \theta_H > \theta_L \), and the parameter \( x \) denotes users required levels of functionality, where \( x_H > x_L \). Note that in our conceptualization, high-type users have higher valuation for functionality as well as higher required level of functionality compared to low-type users. Elements of duplet \( \{ \theta, x \} \) can take any value as long as the rank ordering is maintained. This conceptualization of maintaining the rank ordering closely mimics the reality wherein one observes that progressively higher versions of software are almost always forward compatible (higher version has all the functionality of lower version), and higher versions are sold at higher prices\(^8\). Note that this conceptualization of two dimensional ordered user heterogeneities is similar to the hierarchal characteristics case of

---

\(^7\) A summary of our notation is provided in Appendix 1.A.

Wei and Nault (2011). For the sake of completeness, in §5.4, we discuss a firm’s versioning strategy when the rank ordering is not maintained.

Valuation for functionality ($\theta$), and required level of functionality ($x$), are users’ private information and the software firm only knows the distribution. The proportion of high-type users, denoted by $\{\theta_h, x_h\}$, in the market is $\alpha$ while the proportion of low-type users, denoted by $\{\theta_l, x_l\}$, is $1-\alpha$. The market consists of both types of users, that is, $\alpha \in (0,1)$. Our setting of a market with two types of users is sufficient to obtain useful insights of the monopolist product-pricing decisions (Laffont and Martimort 2002, pp.31, Chellappa and Shivendu 2005, Wu and Chen 2008). Later, in §5.2 we extend our analysis to a market with $N$ types of users.

A user evaluates software by taking into consideration not only the benefit from functionality which is present in the software, but also ‘inconvenience’ or disutility from those functionality which is required by her but is not provided in software. To that extent in our conceptualization, a user derives utility from software in two distinct parts. This conceptualization has some similarity to Chellappa and Shivendu (2010), in which consumers evaluate utility from personalization services in two parts, benefit from personalization services and disutility from loss of privacy as they use personalization services.

First part of a user’s utility is benefit derived from the functionality available in the software, which is given by $\theta f(q)$ where function $f(q)$ maps software functionality to benefit. We assume $f(q)$ is continuous and differentiable for all $\forall q > 0$, users’ benefit increases at a non-increasing rate as functionality level of the software increases: $\frac{\partial f(q)}{\partial q} > 0$ and $\frac{\partial^2 f(q)}{\partial q^2} \leq 0$, and $f(0) = 0$. Since in our set up users do not derive any additional benefit if the level of functionality in software is greater than their required level of functionality, we define two
benefit functions: \( f_H(q) \) for high-type user and \( f_L(q) \) for low-type users, where 
\[ f_H(q) = f(q) \quad \forall q \in [0,x_H) \quad \text{and} \quad f_H(q) = f(x_H) \quad \forall q \geq x_H, \] and 
\[ f_L(q) = f(q) \quad \forall q \in [0,x_L) \quad \text{and} \quad f_L(q) = f(x_L) \quad \forall q \geq x_L. \] Hence, the benefit from functionality to high-type users is given by 
\[ \theta_H f(q) \forall q \in [0,x_H) \quad \text{and} \quad \theta_H f(x_H) \forall q \geq x_H \] and the benefit from functionality to low-type users is given by 
\[ \theta_L f(q) \forall q \in [0,x_L) \quad \text{and} \quad \theta_L f(x_L) \forall q \geq x_L \] (see left plot, Figure 1.1). Note that literature has used linear utility function (Mussa and Rosen 1978; Moorthy and Png 1992; Chellappa and Shivendu 2003), that is \( f(q) = aq \), where \( a \) is a constant, which is a special case of our conceptualization of users’ benefit from software functionality without an upper bound.

Second part of a user’s utility function is ‘inconvenience’ or disutility that she continues to experience from the functionality that is required by her but is not available in the software. We capture user’s disutility from under-provisioning by a function \( g(x,q) \), where \( i \in \{H,L\} \). We assume that the function \( g(x,q) \) is continuous and differentiable \( \forall q \in (0,x_i) \), where \( i \in \{H,L\} \) and a user experiences zero disutility when the functionality level of the software is the same or greater than her required level, that is, \( g(x,q)|_{q=x_i}=0 \), \( i \in \{H,L\} \) (see left plot, Figure 1.1). Users experience decrease in disutility from under-provisioning at a non-decreasing rate as the software functionality level \( q \) increases: 
\[ \frac{\partial g(x,q)}{\partial q} < 0 \quad \text{and} \quad \frac{\partial^2 g(x,q)}{\partial q^2} \geq 0, \forall q \in (0,x_i), \quad \text{where} \quad i \in \{H,L\}. \] This implies that users experience less ‘inconvenience’ as the functionality increases and the rate of reduction of ‘inconvenience’ is lower with increasing functionality. In other words, users’ sensitivity to under-provisioning decreases as software functionality level gets closer to their required level of functionality. Note that users experience the greatest disutility when they do not adopt the software which is equivalent to \( q = 0 \).
On the other hand, users experience increase in disutility at a non-decreasing rate as their required level of functionality \( (x) \) increases: \( \frac{\partial g(x, q)}{\partial x} > 0 \) and \( \frac{\partial^2 g(x, q)}{\partial x^2} \geq 0 \), for \( \forall q \in (0, x_i) \), where \( i \in \{H, L\} \). This implies that given any functionality level \( q \in (0, x_i) \), as required level of functionality \( x_i \) increases, disutility from under-provisioning of functionality increases at a non-decreasing rate. This is because as the gap between the required level of functionality and functionality in the software increases, ‘inconvenience’ to users increases. In other words, users’ sensitivity to under-provisioning increases as their required level of functionality increases, given a functionality level \( (q \in (0, x_i)) \) in the software.

Furthermore, users who have higher required level of functionality experience greater reduction in ‘inconvenience’ as software functionality level increases: \( \frac{\partial^2 g(x_i, q)}{\partial x \partial q} < 0 \), for \( \forall q \in (0, x_i) \), where \( i \in \{H, L\} \). It implies that users who have higher required level of functionality experience greater reduction in ‘inconvenience’ due to under-provisioning as software functionality level increases. In other words, high-type users experience more ‘inconvenience’ compared to low-type users as software functionality decreases because they are more sensitive to under-provisioning. Now, we can write consumer utility as:

\[
U_i(\theta, x_i, q) = \begin{cases} 
\theta f(q) - g(x_i, q), & \text{if } q < x_i \\
\theta f(x_i), & \text{if } q \geq x_i 
\end{cases}, \quad i \in \{H, L\}
\] (1)

Note that utility function in (1), is similar to Chellappa and Shivendu (2010) in conceptualizing utility in two parts, namely benefit from functionality and disutility from under-provisioning, but it has two key differences from their non-monotonic concave utility function. First, in our setup, users with heterogeneous required level of functionality experience less disutility from under-provisioning as level of functionality in software gets closer to their.
required level of functionality. However, in their paper users with heterogeneous concern for privacy incur more disutility from loss of privacy as they use more personalization services. Second, in our setting users experience no disutility if level of functionality in software is equal to or more than their required level, while in their setting users incur more disutility from loss of privacy as they use more personalization services\(^9\). Third, in our setup users are heterogeneous in valuation for functionality (benefit from functionality), while in their setup users are homogeneous in valuation for services (benefit from services).

![Figure 1.1: User benefit from functionality, inconvenience from under-provisioning of functionality, and the net utility](image)

Figure 1.1 illustrates the conceptualization of our utility function (1). The left plot describes the increase in benefit and decrease in ‘inconvenience’ or disutility from increasing functionality for both types of users. And the right plot describes the net utility function \( U(\cdot) \) for both types of users. Since both the utility curves increase at a non-increasing rate in functionality, it implies that these two curves cross only once at \( q = q' \). A formal definition follows.

\(^9\) In Chellappa and Shivendu (2010), personalization services exhibit ‘No-Free Disposal (NFD)’ property, while in our paper users have ‘free disposal’ because they are indifferent to over-provisioning.
**Definition 1:** A functionality level \( q' \) is said to be ‘indifferent functionality level’ at which both high-type and low-type users derive the same positive utility \( U_n(q') = U_l(q') \).\(^{10}\)

Note three key characteristics of the user utility function in (1): (a) utility is monotonically increasing at a decreasing rate in functionality for high-type as well as for low-type users up to their respective required levels of functionality. This conceptualization has some similarity to quadratic utility functions employed in IS literature (Raghunathan, 2000; Sundararajan, 2004), (b) users’ utility remains constant for any functionality level higher than their required level which is similar to Ghose and Sundararajan (2005) and hierarchical case of Wei and Nault (2011), and (c) the Spence-Mirrlees Single-Crossing Condition is satisfied that is

\[
\frac{\partial^2 U_i(\theta, x, q)}{\partial \theta \partial q} > 0, \quad \frac{\partial^2 U_i(\theta, x, q)}{\partial x \partial q} > 0, \quad \text{for } q \in (0, x), \text{ where } i \in \{H, L\}.
\]

The key difference between our model and previous quadratic utility models in IS (Raghunathan, 2000; Sundararajan, 2004; Ghose and Sundararajan, 2005) is that in their models the utility maximizing functionality level depends only on the user’s marginal valuation for functionality, whereas, in our model, the utility maximizing functionality level depends on another parameter, that is, user’s required level of functionality \( x \).

Now we turn our attention to a software firm’s product-pricing strategy. Following versioning literature in IS (Bhargava and Choudhary, 2001; 2008; Chen and Seshadri, 2007; Chellappa and Shivendu, 2005; Wei and Nault, 2011; Chen and Wu, 2008; Lahiri et al., 2013) we consider a monopolistic market setting. A software firm may offer different versions of software which differ only in the level of functionality. High version software has higher level of functionality compared to low version which has a subset of functionality of high version and to that extent versions are upward compatible (Raghunathan, 2000). Without loss of generality, we

---

\(^{10}\) Please see Proof for Lemma 1 in online supplement.
assume that the high version has all the functionality that the software firm has developed and the low version has a subset of functionality of the high version.

Following the literature (Wei and Nault 2006; Chellappa and Shivendu 2005), we assume that the development cost of the first copy of the software, $c(q)$, is continuous, differentiable and convex in functionality level of the highest version of the software, $\frac{\partial c(q)}{\partial q} > 0$ and $\frac{\partial^2 c(q)}{\partial q^2} > 0$, for $\forall q$. The convex cost of producing functionality is consistent with the real world software development process where incorporating additional functionality in software requires more design and testing effort at an increasing rate (Chen and Seshadri 2007). The firm develops the high version of the software with the largest set of functionality and removes some functionality, without incurring any additional cost, to create the low version (Ghose and Sundrarajan 2005; Wei and Nault 2011). Further, there is no variable cost of producing additional copies of the high or low version software and all upfront RandD investment is sunk.

Before we analyze the firm’s optimal product-pricing strategy, we make the following two assumptions to focus our analysis to a market where high-type users are attracted to the low version and cost of producing functionality is not too high.

**Assumption 1:** (a) High-type users have higher utility than low-type users for the software that has required level of functionality of low-type users; (b) Functionality level at which high-type users derive zero utility is higher than the functionality level at which low-type users derive zero utility.

In other words, assumption 1(a) means that $U_H(\theta_H, x_H, x_L) > U_L(\theta_L, x_L, x_L)$. This assumption ensures that the software version that gives maximum utility to low-type users gives higher utility to high-type users. It allows us to exclude the trivial case where high-type users have no
incentive to buy the low version with functionality \( q = x_L \). Assumption 1(b) implies that \( U_{h}(0) > U_{l}(0) \). This ensures that there is some feasible version of software for which low-type users’ utility is higher than that of high-type users’.

**Assumption 2:** The marginal cost of developing functionality is relatively small such that the highest functionality level that the firm develops \( q^* \) is greater than the low-type users’ required level of functionality, that is, \( q^* > x_L \).  

Assumption 2 implies that irrespective of the firm’s product strategy, the rate of increase of net utility from functionality is greater than the rate of increase of cost of producing functionality at the low-type users’ required level of functionality. Since the rate of increase of net utility is higher than the rate of increase of cost at \( q = x_L \), optimal level of functionality developed by the firm is higher than \( x_L \). This assumption excludes the non-interesting case where the firm under-provides to low-type users because of high cost of producing functionality.

### 2.1 Firm’s product-price strategies

Monopolist software firm can adopt one of the two broad product-price strategies: sell only one version or sell two versions. We analyze these two strategies in the following subsections.

#### 2.1.1 Single version strategy

When the software firm sells a single version, it creates one version and can set a price such that (i) only the high-type users buy, or (ii) both types of users buy, or (iii) only the low-type users buy. Note that from Assumption 2, the firm produces functionality level which is higher than the required level of functionality of low-type, \( q^* > x_L \), (see Figure 1.1, right plot).

---

Note that \( q^* > x_L \) implies \( \min(\partial L \theta, f'(x_L), \alpha(\partial_h f'(x_L) - g'(x_H, q = x_L))) > c'(x_L) \). If this condition is satisfied at \( q > x_L \), then it is always satisfied for any \( q \leq x_L \).
From Assumption 1, we know $U_H(q = x_L) > U_L(q = x_L)$. This implies that for any functionality level which is higher or equal to the required level of functionality of low-type users, high-type users have higher utility than low-type users. Therefore, if the firm sets a price such that the low-type users buy, then high-type users always buy. Hence, selling a single version only to low-type users (strategy (iii)) is not feasible.

Under the strategy (i) the firm produces software of functionality level $q_H$ and sets price $p_H$ such that only the high-type users buy. The firm’s profit function is:

$$\pi_H(q_H, p_H) = \alpha p_H - c(q_H).$$

and the firm’s optimization problem is:

$$\max_{q_H, p_H} \pi_H(q_H, p_H)$$

subject to

IR (L: not buy): $U_L(\theta_L, x_L, q_H) - p_H < 0$

IR (H: buy): $U_H(\theta_H, x_H, q_H) - p_H \geq 0$

Note that IR (L) ensures that low-type users do not buy and IR (H) ensures that the high-type users buy the single version. Note that since there is only one version, we do not consider individual compatibility constraints (ICs). From Assumption 2, we know that highest optimal functionality $q^* > x_L$. This implies that $q_H > x_L$. Since the utility of high-type users increases but the utility of low-type users does not increase as functionality level increases beyond $x_L$, we have

$U_H(q = q_H) > U_H(q = x_L)$, and $U_L(q = q_H) = U_L(q = x_L)$. From Assumption 1, we know

$U_H(q = x_L) > U_L(q = x_L)$. Hence, $U_H(q = q_H) > U_L(q = q_H)$. Therefore, the IR (H) must be binding, i.e.,

$p_H = U_H(\theta_H, x_H, q_H)$ and this ensures that IR (L) is satisfied.

Under the strategy (ii) the firm produces software of functionality level $q_A$ and sets price $p_A$ such that both types of users buy. The firm’s profit function is:
\[ \pi_A(q_A, p_A) = p_A - c(q_A). \]

and the firm’s optimization problem is:

\[
\max_{q_A, p_A} \pi_A(q_A, p_A)
\]

subject to

IR (L: buy): \[ U_L(\theta_L, x_L, q_A) - p_A \geq 0 \]

IR (H: buy): \[ U_H(\theta_H, x_H, q_A) - p_A \geq 0 \]

Note that IR (L) and IR (H) ensure that both types of users buy. Since the utility of low-type user does not increase when the functionality level increases beyond \( x_L \), the maximum utility of low-type is \( U_L(\theta_L, x_L, q = x_L) \). From Assumption 2, we know that highest optimal functionality \( q^* > x_L \), therefore firm offers \( q_A = x_L \). When the firm sells to both types of users, the maximum price that it can set is \( U_L(\theta_L, x_L, q = x_L) \). And from Assumption 1, at that price, high-type users also buy. The firm sets the price such that IR (L) is binding and Assumption 1 ensures that IR (H) is satisfied.

### 2.1.2 Versioning strategy

When the firm adopts a versioning strategy, it develops a high version with the highest functionality level and then creates a low version by disabling some functionality of the high version. The firm sells these two versions (high version and low version) at different prices and users self-select to buy the version of their choice. The firm’s profit function is:

\[
\pi_V(p_{VH}, p_{VL}, q_{VH}, q_{VL}) = \alpha p_{VH} + (1 - \alpha) p_{VL} - c(q_{VH}).
\]

And the firm’s optimization problem is:

\[
\max_{p_{VH}, p_{VL}, q_{VH}, q_{VL}} \pi_V(p_{VH}, p_{VL}, q_{VH}, q_{VL})
\]

subject to

IR (H: buy high version): \[ U_H(\theta_H, x_H, q_{VH}) - p_{VH} \geq 0 \]

IR (L: buy low version): \[ U_L(\theta_L, x_L, q_{VL}) - p_{VL} \geq 0 \]

IC (H: buy high version): \[ U_H(\theta_H, x_H, q_{VH}) - p_{VH} \geq U_H(\theta_H, x_H, q_{VL}) - p_{VL} \]
In the next section we analyze optimal functionality and prices set by the software firm under different product-pricing strategies and compare profits to identify the optimal strategy.

3. Product-pricing Strategies

Before we examine the firm’s profits under different product pricing strategies given in §2.1, we first consider some characteristics of the user utility function and the optimal functionality level under different strategies. Since high-type users have a higher required level of functionality \( (x_H > x_L) \), they are more sensitive to under-provisioning of functionality and their utility from software increases at a faster rate compared to low-type users as functionality level increases (see Figure 1.1). The following Lemma gives an important result relating to the indifferent level of functionality \( q' \) (see Definition 1).

**Lemma 1:** When users experience heterogeneous disutility from under-provisioning of functionality, the indifferent functionality level is always positive, that is, \( q' > 0 \).

For all proofs, see Appendix 1.B (online supplement).

When software has zero functionality, both types of users derive zero benefit from using the functionality to accomplish certain tasks. But high-type users experience greater ‘inconvenience’ from not being able to accomplish tasks, compared to low-type users because high-type users have higher required level of functionality and are more sensitive to under-provisioning. Therefore, the high-type users have lower net utility than the low-type users when \( q = 0 \), that is, \( U_H(q = 0) < U_L(q = 0) \). When the software has the same functionality as the high-type users’ required level of functionality \( (q = x_H) \), then neither type experiences any disutility because they can accomplish their tasks, that is, \( U_H = \theta_H f(x_H) \) and \( U_L = \theta_L f(x_L) \). It implies that high-type users derive higher utility from the software compared to the low-type users.
$U_H = \theta_H f(x_H) > U_L = \theta_L f(x_L)$, as $x_H > x_L$ and $\theta_H > \theta_L$. Further, since for both types of users net utility is monotonically non-decreasing in the level of functionality, there exists a functionality level $q^* > 0$, at which both types of users derive the same net utility from the software. Further, Assumption 1(b) $U_H^{-1}(0) > U_L^{-1}(0)$ implies that both types of users have positive utility at the indifferent functionality level, $q^*$.

Lemma 1 implies that when the software has functionality lower than the indifferent level, high-type users experience far more ‘inconvenience’ than low-type users. Even though high-type users derive higher benefit from the available functionality, their net utility is lower than the low-type users’ (see Figure 1.1, right plot, region between 0 and $q^*$). Continuing with the example of the statistician and the school student, both of them derive benefit from the functionality available in the spreadsheet software. The statistician has higher net utility than the school student if the spreadsheet has functionality for demand forecasting and complex modeling. On the other hand, if the software only has basic data processing functionality and does not have demand forecasting and complex modeling functionality, then the disutility experienced by the statistician may be so large that her willingness to pay for the spreadsheet software may be lower than the willingness to pay of the school student.

Before deciding the functionality-price menu under different product strategies, the firm determines the highest level of functionality to produce under each strategy. In our model, the utility derived by high-type users does not increase if the software’s functionality level is higher than the high-type user’s required level, that is, $q > x_H$. This implies that an increase in the level of functionality beyond the required level $x_H$ does not lead to any increase in high-type user’s utility. Nonetheless, the firm incurs a development cost for increasing functionality. Therefore, it
is never optimal for the firm to develop the level of functionality of the high version higher than $x_H$.

**LEMMA 2**: (A) When the firm adopts the product strategy of selling one version only to high-type users, the optimal level of functionality $q_h^*$ solves the first-order condition given by $\alpha \theta_H f'(q_h^*) - \alpha g'(x_H, q_H^*) - c'(q_H^*) = 0$. The optimal level of functionality is bound such that $x_L < q_h^* \leq x_H$ and it increases within this bound as the proportion of high-type users ($\alpha$) increases. 

(B) When the firm adopts the product strategy of selling one version to both types of users, then the optimal level of functionality is the same as the low-type users’ required functionality level, that is, $q_L^* = x_L$. 

(C) When the firm adopts a versioning strategy, then the highest level of functionality produced is the same as in (A), that is, $q_{vh}^* = q_h^*$.

When the software firm adopts the product strategy of selling only to high-type users, the optimal functionality is set at the level at which the marginal cost of functionality equals the marginal net utility of functionality for high-type users, weighted by the proportion of high-type users. As the proportion of high-type users increases, the firm increases the level of functionality of software because the increase in revenue outweighs the increase in development cost of additional functionality. When the firm adopts the product strategy of selling to both types of users, then the maximum price that the firm can charge is the maximum utility of low-type users. From Assumption 2, we know that highest optimal functionality $q^* > x_L$. Since the firm cannot charge a higher price by producing functionality beyond the low-type users’ required level of functionality, under this strategy, $q_{vh}^* = x_L$.

When the software firm adopts a versioning strategy, the optimal functionality level of the high version is the same as the optimal functionality level given in Lemma 2 (A). The economic intuition behind this result is as follows: When the firm adopts versioning strategy,
then the firm first develops the highest functionality level and targets it to the high-type users, and then removes some functionality from the high version to create low version and targets that to low-type users. The optimal level of functionality of the high version is determined by the tradeoff between the marginal benefit and the marginal cost of functionality. Since this tradeoff is the same under the option of selling only to high-type users or adopting a versioning strategy, the optimal functionality produced under both the options is the same.

Under versioning strategy, the firm has to ensure that each type of users buy the version targeted to them. Since for any functionality level greater than \( q' \), high-type users have higher utility, only they may consider switching to low version. In order to make them indifferent between buying the high version or switching to the low version, the firm needs to provide them some incentive or information rent.

**PROPOSITION 1:** Under a versioning strategy, the software firm’s optimal level of functionality of the low version is bounded between the indifferent functionality level and the required level of functionality of low-type user, such that, \( q' \leq q_{vl} \leq x_L \).

When the functionality level of low version is \( q' \), then the firm pays no information rent to high-type users because at that functionality level both types of users derive the same utility. If the firm offers the low version at a level of functionality lower than \( q' \), then the firm loses revenue from low-type users and also does not gain revenue from high-type users by saving more information rent, which is already zero at \( q' \). Therefore, the firm never lowers the functionality of the low version below the indifferent level, that is \( q_{vl} \geq q' \). On the other hand, if the firm offers the low version with functionality higher than the low-type users’ required level, then the firm has to pay higher information rent to high-type users, but does not gain any additional revenue from low-type users because their utility does not increase for functionality
level beyond $x_L$. Therefore, the firm never offers the low version with functionality level higher than the low-type users’ required level of functionality, that is, $q^*_{vl} \leq x_L$.

**LEMMA 3:** *When users experience heterogeneous disutility from under-provisioning of functionality, then under a versioning strategy, the optimal functionality level of the low version depends on the proportion of high-type users in the market.* (A) When $\alpha \in (0, \alpha_z)$, the optimal functionality level of the low version is $q^*_{vl} = x_L$. (B) When $\alpha \in [\alpha_z, 1]$, the optimal functionality level of the low version is bound within $[q', x_L]$, such that $q^*_{vl}$ solves the first order condition:

$$(\theta_L f'(q^*_{vl}) - g'(x_L, q^*_{vl})) - \alpha(\theta_H f'(q^*_{vl}) - g'(x_H, q^*_{vl})) = 0.$$ *(C) When $\alpha \in (\alpha_z, 1)$, the optimal functionality level of the low version is $q^*_{vl} = q'$. The two threshold values of proportion of high-type users are:

$$\alpha_1 = \frac{\theta_L f'(q') - g'(x_L, q')}{\theta_H f'(q') - g'(x_H, q')} \quad \text{and} \quad \alpha_2 = \frac{\theta_L f'(x_L)}{\theta_H f'(x_L) - g'(x_H, x_L)}.$$ 

Lemma 3 characterizes the optimal functionality level of the low version under a versioning strategy. This optimal functionality level depends on two critical proportions of high-type users in the market, namely the upper threshold proportion $\alpha_1$ and the lower threshold proportion $\alpha_2$. The optimal functionality level of the low version is $x_L$ as long as there are relatively fewer high-type users in the market, that is, $\alpha \in (0, \alpha_z]$. The economic intuition is that in this range, there are relatively more low-type users, thus the marginal benefit of functionality for the low-type users’ utility is greater than the marginal effect on information rent to maintain incentive compatibility. Therefore, it is optimal for the firm to offer the low version software at the highest possible functionality level for that version, $x_L$ (Figure 1.2). When the proportion of high-type users in the market is moderate, $\alpha \in [\alpha_z, \alpha_1]$, then the firm makes a tradeoff between the marginal revenue from low-type users and the marginal information rent to high-type users by adjusting functionality level of the low version. The firm reduces functionality of the low version...
from its upper bound $x_e$ as the proportion of high-type users increases (Figure 1.2). The economic intuition is easy to see: As the proportion of high-type users increases, functionality of the low version is distorted downwards to save on information rent. When the proportion of high-type users is relatively large in the market, $\alpha \in (\alpha_1, 1)$, the firm is better off by setting functionality level of the low version at its lower bound, that is, the indifferent functionality level $q'$ where information rent to high-type users is zero. Before we continue with our analysis, we present following two formal definitions.

**Definition 2:** When the proportion of high-type users in the market is higher than the upper threshold proportion $\alpha_1$, then under-provisioning of functionality in the low version of software is most severe ($q_{V_{L}} = q'$).

**Definition 3:** When the proportion of high-type users in the market is lower than the lower threshold proportion $\alpha_2$, then there is no under-provisioning of functionality in the low version of software ($q_{V_{L}} = x_L$).

Now we focus on comparing profits under different product strategies. Note that under the strategy of selling the high version only to high-type users, market coverage is partial. On the other hand, under the strategy of selling one version to both types of users and under the versioning strategy the market is covered. Under the versioning strategy, the firm’s ability to charge a high price for the high version is limited because of the potential of cannibalization between versions (Raghunathan 2000; Belleflamme 2002). In the presence of the low version, high-type users pay less for the high version as they get information rent, but low-type users, who previously did not find the high version attractive, buy the low version. The firm provides information rent to high-type users such that they receive the same net surplus from buying either of the versions, and therefore, do not switch from buying the high version to the low
version. As the level of functionality offered to low-type users decreases (from $x_L$), revenue from low-type users decreases, but revenue from high-type users increases because the firm pays less information rent. The firm determines an optimal level of functionality of the low version (Lemma 3) and the price of the high version by making a trade-off between these two opposite effects.

![Figure 1.2: Optimal functionality level of the low version as the proportion of high-type users’ changes](image)

**PROPOSITION 2: Sufficient Condition for Optimality of Versioning Strategy:** When users experience heterogeneous disutility from under-provisioning of functionality, versioning strategy dominates the product strategies of selling one version only to high-type users or selling to both types of users.

The economic intuition of Proposition 2 is as follows: when $\alpha \in (0, \alpha_2)$, then under a versioning strategy, the functionality level of low version is $x_L$. Therefore, in this region, compared with the product strategy of selling to both types of users, a versioning strategy has the same revenue from low-type users, but higher revenue from high-type users, after adjusting for higher development cost of high level of functionality. Hence, in this region, a versioning strategy dominates the strategy of selling to both types of users. While the optimal revenue for a strategy of selling to both types of users is constant, the optimal revenue for a versioning strategy increases as the proportion of high-type users increases in the market. Therefore, when the
proportion of high-type users is relatively higher, $\alpha \in [\alpha_2, 1)$, then a versioning strategy also dominates the strategy of selling to both types of users (Figure 1.3).

![Figure 1.3: Profits under the three product strategies when user experience disutility from under-provisioning of functionality](image)

Now we compare the optimal profit of a versioning strategy with the optimal profit of a strategy of selling single version only to high-type users. Since the firm provides some information rent to high-type users under a versioning strategy to incentivize them to refrain from switching to the low version, the price of the high version is lowered under versioning, even though the optimal functionality level of the high version is the same under both product strategies, that is, $q_{H H} = q_{H H}'$. Consequently, the revenue from high-type users under a versioning strategy is lower than under a strategy of selling a single version only to high-type users. But the gain in revenue from low-type users under versioning strategy more than compensates for this loss of revenue, when the proportion of high-type users is relatively low, $\alpha \in (0,\alpha_1)$. On the other hand, when the proportion of high-type users is relatively high, $\alpha \in (\alpha_2, 1)$, the firm sets the optimal functionality level of the low version at $q'$, and does not pay information rent to high-
type users. When $\alpha \in (\alpha_l, 1)$, the price of the high version is the same as under the strategy of selling a single version only to high-type users. It means the firm generates the same revenue from high-type users and some additional revenue from low-type users under a versioning strategy. Therefore, optimal profit is higher under a versioning strategy compared with a product strategy of selling only to high-type users. Figure 1.3 graphically shows profits under the three product-price strategies and the dominance of a versioning strategy.

4. Impact of increase in users’ required level of functionality

With the rapid technological changes, users’ requirements for software functionality and computing needs are changing at a fast pace. One would expect that the required level of functionality of both types of users will increase over time period, though users’ benefit from functionality may not change. In this section, we discuss the impact of increase in users’ required level of functionality, $x_H$ or $x_L$, one at a time, on the firm’s versioning strategy.

**Lemma 4:** As the required level of functionality of high-type users ($x_H$) increases, (i) the indifferent functionality level $q'$ increases, and (ii) the optimal level of functionality of the high version $q_{vh}$ increases.

The high-type users become more sensitive to under-provisioning of functionality as their required functionality level increases, because they experience more ‘inconvenience’ or disutility. Therefore, the utility of high-type users decreases for any functionality level lower than their previous required level of functionality. On the other hand, there is no impact on the low-type users’ utility when $x_H$ increases. This is illustrated in Figure 1.4, left plot, where the required level of functionality of high-type users increases from $x_{HA}$ to $x_{HB}$ ($x_{HA} < x_{HB}$). Since the high-type users are more sensitive to under-provisioning of functionality, they need more functionality to match the utility derived by the low-type users. This leads to shifting of the
intersection point of the two utility curves to the right. Therefore, the indifferent functionality level, \( q' \), increases, as \( x_h \) increases (see Figure 1.4, \( q'_a > q'_b \)).

For the rest of the analysis we assume that the increase in \( x_h \) is not large and Assumption 1(a) continues to hold, that is, \( U_H(\theta_H, x_{HB}, x_L) > U_L(\theta_L, x_L, x_L) \). Also note that since there is no change in the required level of functionality of low-type users, Assumption 2 (\( q^* > x_L \)) continues to hold. Moreover, since high-type users are more sensitive to disutility from under-provisioning, \( U^{-1}_H(0) \) increases, Assumption 1(b) (\( U^{-1}_H(0) > U^{-1}_L(0) \)) continues to hold.

The firm increases the high version’s functionality level, \( q^*_{VH} \), as high-type users’ required level of functionality increases. The economic intuition is easy to see: As \( x_h \) increases, the marginal effect of functionality on high-type users’ utility increases, though the marginal effect on development cost remains the same. Therefore, the firm increases the high version’s functionality level.

**PROPOSITION 3:** As the required level of functionality of high-type users \( x_h \) increases, (i) the lower threshold proportion \( \alpha_2 \) and the upper threshold proportion \( \alpha_1 \) decrease, (ii) the optimal functionality of the low version \( q^*_L \) (a) does not change if the proportion of high-type users is small, (b) decreases if proportion of high-type users is moderate, and (c) increases if the proportion of high-type users is large.

When high-type users’ required functionality level increases, they become less attracted to low version, because they experience more ‘inconvenience’ or more disutility from under-provisioning. Therefore, the firm needs to pay lower information rent to make high-type users indifferent to switching to low version. This implies that by reducing the low version functionality level from the low-type users’ required level, the marginal gain from saving on
information rent increases, but the marginal loss of revenue from low-type users remains the same. This leads to the firm starting to reduce the low version functionality level from the low-type users required functionality level, $x_L$, when there are relatively fewer high-type users in the market, that is, at a lower $\alpha_z$. This is illustrated in Figure 1.4, right plot, where $\alpha_{z_H} < \alpha_{z_A}$. Note that it also implies that the firm reduces the low version functionality at a faster rate as $\alpha$ increases. Combining these two factors - an increase in the indifferent functionality level, $q'_y > q'_s$ (Lemma 4), and an increase in sensitivity of high-type users’ to under-provisioning of functionality- the firm offers low version with functionality at $q'_s$ when there are relatively fewer high-type users in the market, that is, at a lower $\alpha_z$ (Figure 1.4, right plot, $\alpha_{H_H} < \alpha_{H_A}$).

As discussed above, when $x_H$ increases, both upper threshold proportion $\alpha_1$ and lower threshold proportion $\alpha_2$ decrease, but the impact on the functionality level of the low version is mixed. If there are few high-type users in the population (small $\alpha$), low-type users continue to receive the same level of functionality $x_L$, when $x_H$ increases. This occurs simply because there are not enough high-type users to make any impact on the trade-off between information rent and revenue from low-type users. When there are a relatively moderate number of high-type users, the firm lowers the functionality level of low version. This is because high-type users are now more sensitive to changes in functionality level of the low version, and the firm gains more revenue through reduction in information rent by decreasing functionality of the low version than loss of revenue from low-type users.
Figure 1.4: Utility curves for both types of users, and the optimal functionality level of the low version as required functionality level of high-type users increases where \( f(q) = q \), \( g(x,q) = (x-q)^2 \), \( c(q) = q^2 \), and

\[ x_{HA} < x_{HB}. \]

On the other hand, when the proportion of high-type users is relatively high, the functionality level of low version increases as \( x_h \) increases. Recall that when the proportion of high-type users is relatively high, low-type users get the low version at the indifferent functionality level \( q' \) because the firm economizes on information rent. Lemma 4 states that the indifferent quality level increases, \( q'_H > q'_L \) as \( x_h \) increases, and therefore, the firm increase the functionality level of low version when \( \alpha \) is relatively high. In Figure 1.4, low-type users get higher functionality level when \( \alpha_{HA} < \alpha < 1 \), lower functionality level when \( \alpha_{LB} < \alpha < \alpha_{HA} \), and no change in functionality level when \( 0 < \alpha < \alpha_{LB} \).

**Lemma 5:** As the required level of functionality of low-type users (\( x_L \)) increases, (i) the indifferent functionality level \( q' \) decreases, and (ii) there is no impact on the optimal level of functionality of high version \( q^*_VH \).

Increase in the low-type users’ required level of functionality has the opposite impact on the indifferent functionality level \( q' \) compared to the case described in Lemma 4. When \( x_L \)
increases, the high-type users’ utility remains the same, but the utility of low-type users decreases for any functionality level lower than their previous required level of functionality (see left plot, Figure 1.5). This is so because the low-type users become more sensitive to under-provisioning of functionality. But since they are always less sensitive to under-provisioning of functionality compared to the high-type users, they need less functionality to match the utility of the high-type users. This leads to the shifting of the intersection point of the two utility curves to the left. Hence, the indifferent functionality level, $q'$, decreases as $x_L$ increases (see Figure 1.5, left plot).

Note that as $x_{L\alpha}$ increases to $x_{L\beta}$, the utility of low-type users at $q = x_{L\alpha}$ increases, but the utility of high-type users at $q = x_{L\alpha}$ increases even more because they get more benefit and also experience less disutility. Hence, as $x_L$ increases, Assumption 1 continues to hold. Since increase in $x_L$ does not impact optimal functionality $q^*$, when $x_L$ increases, we are only considering the situations where Assumption 2 ($x_L < q^*$) and Assumption 1(b) ($U_H^0(0) > U_L^{-1}(0)$) continues to hold.

Recall that Lemma 2 states that the optimal functionality of the high version is independent of the characteristics of low-type users. Therefore, the optimal functionality of the high version does not change when required level of functionality of low-type users increases. However, one may expect that the firm will increase the functionality of the low version or reduce functionality distortion in low version as $x_L$ increases; since the firm increases the functionality level of the high version when high-type users have higher required level of functionality (Lemma 4). The following Proposition provides surprising results.

**PROPOSITION 4:** As the required level of functionality of low-type users ($x_L$) increases, (i) the lower threshold proportion $\alpha_z$ and the upper threshold proportion $\alpha_i$ increase, (ii) the optimal
functionality of the low version $q_{vl}$ (a) increases if proportion of high-type users is relatively low, and (b) decreases if proportion of high-type users is relatively high.

Low-type users become more sensitive to under-provisioning of functionality as their required functionality level increases. This implies that by reducing the low version functionality level from the low-type users’ required level of functionality, the marginal gain from saving on information rent remains the same, but the marginal loss of revenue from low-type users increases. This leads to the firm reducing the low version functionality level from the low-type users required functionality level $x_{L}$, when there are relatively more high-type users in the market, that is at higher $\alpha_{2}$ (Figure 1.5, right plot, $\alpha_{sla} < \alpha_{sll}$). Note that this also implies that the firm reduces the low version functionality at a slower rate as $\alpha$ increases. Furthermore, when $x_{L}$ increases the indifferent functionality level $q'$ decreases (Lemma 5). This results in the firm stopping to reduce the functionality level of the low version when there are relatively more high-type users in the market, that is at higher $\alpha_{1}$ (Figure 1.5, right plot, $\alpha_{sla} < \alpha_{sll}$).

**Figure 1.5:** Utility curves for both types of users, and the optimal functionality level of the low version as required functionality level for low-type users increases when $f(q) = q$, $g(x, q) = (x - q)^2$, $c(q) = q^2$, and $x_{La} < x_{Lb}$. 

37
The firm’s decision whether to reduce or increase functionality level of the low version when the low-type users’ required level of functionality increases depends on the proportion of high-type users in the market. And the impacts are opposite compared to an increase in the high-type users’ required level of functionality (Proportion 3). In Figure 1.5, right plot, low-type users get higher functionality level when $0 < \alpha < \alpha_{Lab}$, or lower functionality level when $\alpha_{Lab} < \alpha < 1$.

**PROPOSITION 5:** (1) As the required level of functionality of high-type users ($x_h$) increases, consumer surplus of those high-type users who get information rent decreases; (2) As the required level of functionality of low-type users ($x_l$) increases, consumer surplus of those high-type users who get information rent increases.

We know that the firm needs to provide information rent to high-type users only when the proportion of high-type is lower than the upper threshold proportion, $\alpha \in (0, \alpha_1)$ (Lemma 3). As high-type users’ required level of functionality increases, they become more sensitive to under-provisioning of functionality and less inclined to switch to the low version. Thus, the firm provides less information rent to make them indifferent to switching to the low version. Therefore, the consumer surplus of those high-type users who get information rent declines when $\alpha \in (0, \alpha_{lb})$ (see Figure 1.6(a)). Since the upper threshold proportion decreases (Proposition 3), $\alpha_{lb} < \alpha_{la}$, high-type users get zero rent when $\alpha \in [\alpha_{lb}, \alpha_{la})$. When the proportion of high-type users is large, that is, $\alpha \in [\alpha_{la}, 1)$, high-type users get zero information rent as before (see Figure 6a). Hence, counter to intuition, high-type users are not better off when their required level of functionality increases though they get a version with higher functionality level.

On the other hand, as the low-type users’ required level of functionality increases, they become more sensitive to under-provisioning. When $\alpha \in (0, \alpha_{la})$ (Figure 1.6(b)), the firm increases the functionality level of low version, and pays more information rent, leading to
increase in surplus of high-type users. When $\alpha \in [\alpha_{ILH}, \alpha_{ILB})$, then high-type users who did not get information rent get some information rent as $x_L$ increases since the upper threshold proportion of high-type users increase (Proposition 4). Therefore, when $\alpha \in (0, \alpha_{ILB})$, high-type users get more information rent leading to higher consumer surplus to high-type users as $x_L$ increases. When the proportion of high-type users is large, that is, $\alpha \in [\alpha_{ILB}, 1)$, high-type users get zero surplus because $q^*_L = q^*$ (see Figure 1.6(b)).

![Figure 1.6: Consumer surplus of high-type users](image)

5. Some extensions

In this section, we analyze the impact of following five settings on our results regarding optimal versioning strategy: (1) when users do not experience disutility from under-provisioning of functionality, (2) when the market consists of $N$ discrete types of users, (3) when the Assumption 1 or 2 is relaxed, one at a time, (4) when rank ordering of valuation and required level of functionality is not maintained, and (5) impact of relative strength of benefit function $f(q)$ and disutility function $g(x, q)$ on versioning strategy.
5.1 No disutility from under-provisioning of functionality

When users do not experience disutility from under-provisioning of functionality, then $g(x_i,q) = 0, \forall q$ for $i \in \{H,L\}$. The utility function in (1) can be modified as:

$$U_i(\theta_i, x_i, q) = \begin{cases} \theta_i f(q), & \text{if } q < x_i \\ \theta_i f(x_i), & \text{if } q \geq x_i \end{cases}, i \in \{H,L\}. $$

(5)

PROPOSITION 6: When users do not experience disutility from under-provisioning of functionality, (i) when the proportion of high-type users is small $\alpha < (\theta_H / \theta_L)$, the firm adopts a versioning strategy; and (ii) when the proportion of high-type users is large $\alpha \geq (\theta_L / \theta_H)$, the firm adopts single-version strategy and sells only to the high-type users.

Under a versioning strategy, the firm sets the optimal functionality level of the high version at the same level as when it sells a single version only to high-type users. Note that when there is no disutility from under-provisioning, the marginal gain from functionality decreases, but the marginal cost of producing functionality remains same as in §3. Hence the firm produces the high version with lower functionality level. The firm chooses the optimal functionality level of the low version by trading off decrease in revenue from high-type users due to payment of information rent in the presence of low version and increase in revenue from low-type users. Note that when $g = 0$, then the upper threshold proportion $\alpha_1$ and lower threshold proportion $\alpha_2$ values given in Lemma 3 converge to $\alpha_1 = \alpha_2 = \theta_L / \theta_H$. When there are relatively more high-type users in the market, $\alpha \in (\theta_L / \theta_H, 1)$, the firm sets the functionality level of the low version at zero to minimize information rent. It means that in this situation, the firm adopts a single-version strategy of selling only to high-type users.

When there are relatively less high-type users in the market, $\alpha \in (0, \theta_L / \theta_H)$, the firm sets the level of functionality of the low version at the required functionality level of low-type users,
that is \( x_L \), to extract the maximum revenue from low-type users. Though the firm pays information rent to high-type users, the revenue gain from low-type users is more than the revenue loss (due to information rent) from high-type users. Hence versioning strategy dominates single-version strategy of selling only to high-type users. Note that when the firm adopts a single-version strategy of selling to both types of users, the optimal functionality level is \( x_L \) as discussed in Lemma 2. Since under the versioning strategy, the firm sets the functionality of low version at \( x_L \), and gets more revenue from high-type users compared to selling a single version to both types of users, the profit under a versioning strategy dominates a single-version strategy of selling to both types of users when \( \alpha \in (0, \theta_L / \theta_H) \). Hence, the firm adopts versioning strategy. Profits under the three product strategies are shown in Figure 1.7.

**Figure 1.7: Profits under different product strategies when there is no disutility from under-provisioning**

When users do not experience disutility from under-provisioning, our model has close similarity to the hierarchical characteristics case of Wei and Nault (2011). Moreover, the condition under which a versioning strategy is optimal for the firm as given in Proposition 6 above is similar to their condition given in hierarchical characteristics case.
5.2 Market with \( N \) user segments

We extend our analysis of §3 with two user types to a market consisting of \( N \) user types. Let there be \( m_n \) users of each type where \( n \in \{1, 2, \ldots, N\} \). Each type of user type \( (n) \) is characterized by a duplet \( \{\theta_n, x_n\} \), \( n \in \{1, 2, \ldots, N\} \). Without loss of generality we assume that user segments are ordered such that \( \theta_n > \theta_{n-1} > \ldots > \theta_2 > \theta_1 \) and \( x_n > x_{n-1} > \ldots > x_2 > x_1 \). Note that this is consistent with our conceptualization in §2 that the ordering of users’ valuation for functionality and required level of functionality is maintained such that \( \theta_i < \theta_{i+1} \) and \( x_i < x_{i+1} \) for \( i = \{1, 2, \ldots, N-1\} \).

In this context, we rewrite Assumptions 1(a), 1(b) and 2 as equations (6), (7), and (8) respectively.

\[
U_i(\theta_i, x_i, q = x_{i-1}) > U_{i+1}(\theta_{i+1}, x_{i+1}, q = x_{i+1}) \quad \text{for} \quad \forall i \in \{2, \ldots, N\} \quad (6)
\]

\[
U_{i-1}^{-1}(0) > U_{i+1}^{-1}(0) \quad \text{for} \quad \forall i \in \{2, \ldots, N\} \quad (7)
\]

\[
q^* \geq x_{N-1} \quad (8)
\]

**PROPOSITION 7:** When users experience heterogeneous disutility from under-provisioning of functionality in a market consisting of \( N \) user segments, versioning strategy with customized version for each segment is optimal.

First, we start with two types of users, the \( N \)-type and the \( (N-1) \)-type. From Proposition 2 we know that it is optimal for the firm to adopt versioning strategy and target one version to \( N \)-type users and another version to \( (N-1) \)-type users. Then we consider the addition of a new market segment of \( (N-2) \)-type users. In this situation the firm has six possible product-price strategies: (1) sell one version so that only \( N \)-type buy, (2) sell one version so that only \( N \)- and \( (N-1) \)-types buy, (3) sell two versions so that \( N \)-type buy the high version, \( (N-1) \)-type buy the low version, and \( (N-2) \)-type do not buy, (4), sell one version so that \( N \)-, \( (N-1) \)-, and \( (N-2) \)-
types buy, (5) sell two versions so that \( N \)- and \((N-1)\)-types buy the high version, and \((N-2)\)-type buy the low version, and (6) sell three versions so that \( N \)-type buy the high version, \((N-1)\)-type buy the middle version and \((N-2)\)-type buy the low version. Comparing firm profits under these product strategies we show that offering three versions targeted to each of the three user segments is optimal.

Note that the results from two-type user segments in Section 3 (Proposition 2) extend to three-type segments because the firm gains more by providing another version targeted to the \((N-2)\)-type. The additional revenue from the \((N-2)\)-type is higher than the potential loss due to paying information rent to the \((N-1)\)-type and/or \( N \)-type consumers.

Assumptions in (6) and (7) above (extended form of Assumption 1 of two-type user segment case) implies that the indifferent functionality level between two adjacent user segments is higher than the required functionality level of the lower user segment, that is, \( q'_{i-1} < x_{i-1} \). This property implies that the ‘inconvenience’ or disutility experienced by the higher type user segment is such that the firm can offer a lower version at \( q'_{i-1} \) to the \( i-1 \)-type users without paying any information rent to the \( i \)-type and all user segments above \( i \)-type.

Now, if we add another user segment \((N-3)\) to our market such that \( \theta_{N-2} > \theta_{N-3} \) and \( x_{N-2} > x_{N-3} \), then similarly we can again show that the firm is better off by adding a new version targeted to the new \((N-3)\)-type users than to lower the price of the version targeted to higher type user segment to cover the new segment. Extending this analysis we show that when the market consists of \( N \) user segments then the firm’s optimal product-price strategy is to offer \( N \) versions and price them in such a way that each user segments self-selects version targeted to them.
5.3 Impact of model assumptions on versioning strategy

In this subsection we study the impact of relaxing Assumption 1(a), 1(b) and 2, one at a time, on the optimal versioning strategy of the firm. The Assumption 1(a) implies that \( U_H(q = x_L) > U_L(q = x_L) \). If this assumption is relaxed then there can be following two situations:

(i) \( U_H(q = x_L) = U_L(q = x_L) \), or (ii) \( U_H(q = x_L) < U_L(q = x_L) \). Situation (i) implies that the indifferent functionality level \( q' = x_L \). Under this situation, when the firm offers low version with functionality level at \( x_L \), then information rent to high-type users is zero. Therefore, we get a special case of Lemma 3, where \( q_{uu} = q^* \), \( q_{ul} = x_L \), and both the upper and the lower threshold proportions of high-type users are equal to 1, that is, \( \alpha_2 = \alpha_1 = 1 \). In this situation, it is optimal for the firm to adopt a versioning strategy.

Under situation (ii), the indifferent functionality level is higher than the low-type users’ required functionality level, that is \( q' > x_L \). In order to study the optimality of versioning strategy, we also need to take into consideration the relative position of highest functionality level developed by the firm (\( q^* \)) with respect to \( x_L \) and \( q' \). This leads us to three possible situations: (a) \( x_L < q^* < q' \), (b) \( x_L < q' < q^* \), and (c) \( x_L < q' = q^* \). (Figure 1.8(a), 8(b), and 8(c)).

![Figure 1.8: Utility of high and low-type users when (a) \( x_L < q^* < q' \), (b) \( x_L < q' < q^* \), and (c) \( x_L < q' = q^* \)](image1)

In situation (a), the optimal functionality level produced by the firm is such that at that functionality level, \( q = q^* \), high-type users have very high disutility from under-provisioning,
leading to their net utility being lower than low-type users. It implies that the firm needs to pay increasingly more information rent if it lowers the functionality of the low version from \( x_L \).

Therefore, in this situation, versioning is not optimal. On the other hand, in situation (b), the firm offers the low version with functionality level \( q_{vL} = x_L \) since there is no gain in revenue by offering low version with functionality more than \( x_L \). Furthermore, by providing the high version with \( q_{vH} = q^* \), the firm pays zero information rent to high-type users, and versioning is optimal. Moreover, in situation (c), low-type users have higher utility than high-type users for any functionality level \( q < q^* \). It is not optimal to offer a high version with functionality \( q_{vH} = q^* \) to low-type users and to offer a low version with functionality \( q_{vL} < q^* \) to high-type users, as the firm needs to pay increasingly more information rent to low-type users if it lowers the functionality of the low version from \( q^* \). Also, low-type users are indifferent from any functionality level from \( x_L \) to \( q' = q^* \). Therefore, the firm just offers a single version with \( q^* \) functionality.

If Assumption 1(b) were to be relaxed, then it implies that \( U_H^L(0) \leq U_L^L(0) \). This means that at \( q' \) (where the utility functions of both types of users intersect) both types of users derive either zero or negative utility. In this situation, the firm’s ability to distort the low version’s functionality to \( q' \) is limited (Lemma 3) and hence the versioning strategy will be optimal only when proportion of high-type users is smaller than some critical proportion \( \alpha < \hat{\alpha} \). The value of \( \hat{\alpha} \) depends on the properties of function \( f(.) \) and \( g(\ldots) \), and \( \hat{\alpha} \in [\alpha_2, \alpha_1] \), where \( \alpha_2 \) and \( \alpha_1 \) are given by Lemma 3.

Now, we consider the impact of relaxing Assumption 2 which is \( q^* > x_L \). If this assumption were to be relaxed, then we need to consider the relative position of optimal
functionality level with respect to indifferent functionality level $q'$. Therefore, following four situations may arise: (a) $q^* < q' < x_L$, (b) $q' = q^* < x_L$, (c) $q' < q^* < x_L$, and (d) $q' < q^* = x_L$ (Figure 1.9 (a), (b), (c) and (d)).

![Figure 1.9: Utility of high and low-type users when a) $q^* < q' < x_L$, b) $q' = q^* < x_L$, c) $q' < q^* < x_L$, d) $q' < q^* = x_L$.](image)

In situation (a) and (b), the optimal functionality level produced by the firm is such that at that functionality level, and the firm needs to pay increasingly more information rent if it lowers the functionality of the low version. Therefore, in these two situations (a) and (b), versioning is not optimal. In situation (c) and (d), the firm will produce $q'_{VH} = q^*$, and $q'_{VL} \in [q', q^*]$, and versioning is optimal, as the firm pays zero rent to high-type users when $q'_{VL} = q'$.

### 5.4 Impact of ordering of valuation and required level of functionality on versioning

In our conceptualization high-type users have higher valuation for functionality as well as higher required level of functionality compared to low-type users. This implies that the elements of duplet $\{\theta_i, x_i\}$ can take any value as long as the rank ordering is maintained. This conceptualization closely mimics the reality where one observes software versions with higher level of functionality are sold at higher prices. However, for the sake of completeness, in this subsection we study the impact on the versioning strategy of a software firm if the rank ordering
is not maintained, that is $\theta_H > \theta_L$ but $x_H \leq x_L$. Note that in this case, Assumption 1(b) does not hold.

Let us first consider the case where $x_H = x_L$. In this situation, both types of users experience the same disutility from under-provisioning, and thus, the utility curve of high-type users is above the utility curve of low-type users for all feasible functionality level. The implication of homogeneous disutility is that the versioning strategy is not optimal. This is because the revenue gain from selling two versions is always less than the information rent required to be paid to high-type users (formal analysis and proof is in Appendix 1.B).

When $x_H < x_L$, then three distinct sub-cases arise; (a) the maximum utility of high-type users is greater than maximum utility of low-type users, that is, $\theta_H f(x_H) > \theta_L f(x_L)$, (b) the maximum utility of high-type users is equal to maximum utility of low-type users, that is, $\theta_H f(x_H) = \theta_L f(x_L)$, and (c) the maximum utility of high-type users is lower than maximum utility of low-type users, that is, $\theta_H f(x_H) < \theta_L f(x_L)$. Figure 1.10 illustrates these three cases.

![Figure 1.10: Utility of high-type and low-type consumers when $x_H < x_L$](image)

In case (a), the utility function for high-type users is above the utility function of low-type users for all feasible functionality level. In this situation, when the firm sells one version, it considers selling $q = x_L$ to all users or to only high-type users. If firm adopts versioning
strategy, it will always sell the high version with \( q_{VH} = x_L \) and the low version with functionality level less than \( x_L \), that is \( q_{VL} \in [0, x_L) \). Since low-type users experience more disutility than high-type users, the information rent paid to high-type users will always be more than \((\theta_H - \theta_L)f(q_{VL})\). This implies that, when firm adopts versioning strategy, the loss in revenue due to information rent is always more than the gain in revenue from low-type users. Therefore, versioning strategy is not optimal (formal analysis and proof is in Appendix 1.B). Similar analysis applies to case (b) and leads to versioning being not optimal. Note that in case (a) and case (b) the utility functions of the low-type and high-type users do not cross-over. In case (c), the required level of functionality of low-type users \( x_L \) is so high that the cross-over point of utility function of the two types is to the right of the required level of functionality of high-type user \( x_H \). It is easy to see that in this case, since the utility functions cross over, the firm can save on the information rent by setting the functionality level of the high version at \( q^*_{VH} = x_L \) and the low version at \( q^*_{VL} = q' \). This leads to versioning being optimal strategy in this case.

5.5 Impact of relative strength of \( f(q) \) and \( g(x, q) \) on versioning strategy

In this subsection, we examine the impact of relative strengths of benefit function \( f(q) \) and disutility from under-provisioning function \( g(x, q) \) on the firm’s versioning strategy. Given a benefit function \( f(q) \), when the strength of disutility function \( g(x, q) \) increases (users experience more disutility from under-provisioning), it leads to steeper slope of utility curves of the both types of users. This results in shifting of the utility curves of the both types of users downward.
This implies that the indifferent functionality level \( q' \) shifts to the right which is illustrated in Figure 1.11 (left plot). Note that the strengthening of disutility function \( g(x,q) \) shifts the highest optimal functionality \( q^* \) to the right \( (q_2^* > q_1^*) \), because though the marginal cost of producing functionality remains the same the marginal benefit from functionality increases. Since, in this case the utility curves of the both types of users cross over, versioning strategy remains optimal. Therefore, as users’ disutility from under-provisioning becomes stronger, the firm adopts versioning and increases the level of functionality of the high version. The level of functionality of the low version may increase or decrease depending on the parameter values.

Given a disutility function \( g(x,q) \), when the strength of the benefit function \( f(q) \) increases \( (f_2(q) > f_1(q)) \), utility curves of both types of users become steeper, that is, the slopes of the utility curves increase. This shifts the utility curves of the both types of users upward. This implies that the indifferent functionality level shifts to the left \( (q_2' < q_1') \). Note that the strengthening of \( f(q) \) shifts the highest optimal functionality \( q^* \) to the right, \( q_2^* > q_1^* \). Since the utility curves of the both types of users cross over, versioning strategy remains optimal. The firm increases the level
of functionality of the high version, but may increase or decrease the level of functionality of the low version depending on the parameter values (Figure 1.11, right plot).

To summarize, the relative strengths of function $f(.)$ and $g(.,.)$ do not impact the existence of the cross-over point between utility functions of high-type and low-type users. This implies that the firm’s gain from offering two versions continues to be more than the information rent paid to high-type users. Therefore, versioning strategy remains optimal and only the functionality levels of high and low versions are impacted by the relatively strengths of function $f(.)$ and $g(.,.)$.

6. Discussion

This work belongs to the stream of literature in Information System that has studied versioning strategy from a perspective of users’ utility from software under general market conditions. In our abstraction, users derive utility from software in a task oriented context and require a certain set of functionality to accomplish their tasks. Users are heterogeneous on two dimensions: valuations for functionality and required levels of functionality; and high-type users require higher level of functionality. Users continue to experience disutility if functionality provided in software is lower than their required level. Moreover, users derive no additional utility if functionality provided in software is more than what they require. This conceptualization of utility function for software is unique and allows us to study software firm’s product-pricing strategies when users continue to experience disutility from under-provisioning.

Different parts of our utility function are informed by extant versioning literature; users derive benefit from using functionality (Moorthy and Png 1992; Bhargava and Choudhary 2001), and users derive no additional utility from over-provisioning beyond required level of functionality (Wei and Nault 2011). We integrate these parts with our novel abstraction that
users continue to experience ‘inconvenience’ or disutility if software provides less functionality than what they require. This leads to type-dependent utility functions which cross over at a positive utility level. We show that under our model framework and assumptions, heterogeneous disutility from under-provisioning of functionality is a sufficient condition for optimality of versioning strategy for a monopolist software firm.

The economic intuition for our finding is as follows: When a software firm adopts a single-version strategy of selling to all users, the firm’s profit is limited by the maximum utility derived by low-type users when software provides their required level of functionality. This is so because they do not derive additional utility from functionality beyond their required level. When the firm adopts a single-version strategy of selling only to high-type users, the firm does not serve low-type users. When the firm adopts a versioning strategy, it has to pay information rent to high-type users but can serve the entire market. When users experience disutility from under-provisioning, under a versioning strategy, high-type users are less attracted to the low version. This implies that the software firm has to pay less information rent to high-type users to incentivize them not to switch to the low version. Moreover, if the low version functionality is same as the indifferent functionality level (at which utility of both types of users is the same), the firm pays zero information rent. This saving in information rent, when users experience disutility from under-provisioning, results in the software firm’s profit under a versioning strategy dominating its profit under a single-version strategy.

When users continue to experience heterogeneous disutility from under-provisioning of functionality, then versioning is optimal irrespective of distribution of high-type users. On the other hand, when users do not experience any disutility from under-provisioning of functionality, then versioning strategy is sub-optimal when the proportion of high-type users is relatively high.
in the market. This highlights our key finding that heterogeneous disutility from under-provisioning of functionality is a sufficient condition for the optimality of versioning strategy.

With fast changing technological landscape and work environment, one would expect that software functionality required by users would increase over time period. When the high-type users’ required level of functionality increases, the firm increases the functionality level of high version, but decrease the functionality level of low version when the proportion of high-type users is relatively moderate in the market. It is so because high-type users become more sensitive to under-provisioning of functionality, and the firm gains more in saving of information rent from high-type users than loss of revenue from low-type users by downward distorting functionality of low version. On the other hand, when the low-type users’ required level of functionality increases, surprisingly, when the proportion of high-type users is relatively large, the firm lowers the functionality of low version. This is so because low-type users become more sensitive to under-provisioning of functionality and the indifferent functionality level (functionality level at which both type of users have same utility) deceases.

Extant literature in information systems has focused on either cost structure of developing quality or market characteristics like the presence of piracy or network externality or distribution of user types to study the profitability of a monopolistic firm’s versioning strategy. The general recommendation has been that it is suboptimal to offer versions of information goods like software where the development cost is convex and the marginal cost is negligible (Bhargava and Choudhary 2001, 2008; Jones and Mendelson, 2011). We recommend that since users derive utility from software in a task oriented context, firms need to take into account their required level of functionality and the remaining ‘inconvenience’ experienced by them from under-provisioning in determining optimal product-pricing strategies. Since our approach is
based on how users derive utility from software, there are some key similarities and differences between our work and the stream of literature in IS which has studied optimality of a versioning strategy from utility perspective, namely Chen and Seshadri (2007) and Wei and Nault (2011).

In Chen and Seshadri (2007), users have convex reservation utilities for outside options. It implies that high-type users have higher valuation of outside options than low-type users. This conceptualization has some similarity to our model where high-type users are more sensitive to under-provisioning of functionality, and hence, are less attracted to low version. The key difference is that in their model type-dependent outside-option is exogenous and is independent of quality of information good, while our type-dependent disutility from under-provisioning depends on level of functionality in software. This difference is key to our result that the entire market is covered, while in their case relatively high- and low-type users are excluded. Similarly, our abstraction of required level of functionality and users being indifferent to any functionality level beyond their required level is similar to the hierarchical characteristics case of Wei and Nault (2011). The key difference between our model and theirs is that in their model users do not experience any ‘inconvenience’ or disutility from under-provisioning while in our case they do. This difference is key to our result that optimality of versioning is independent of proportion of high-type users in the market, while in their hierarchical characteristics case versioning is optimal only under certain distribution of user types.

The managerial implication of our research for information goods firms is that careful investigation of functionality requirements of their different user segments is critical for an optimal versioning strategy. Inherent differences in users’ functionality requirements and ‘inconvenience’ from under-provisioning could potentially explain as why some providers sell a single version while others adopt versioning. It could explain differences in product-pricing
strategies even when firms have a similar cost structure of producing functionality, software has a similar level of network externality and markets have a similar piracy level. For example, Microsoft offers Windows 8 in three versions in the US because different user segments are not only heterogeneous in their valuation for functionality but also in their required level of functionality. On the other hand, Corel offers PhotoImpact in a single version only. Based on our analysis, one possible explanation for this may be that different user segments of PhotoImpact may have similar requirements for functionality though they may have heterogeneous valuation for functionality. This situation is similar to one discussed in §5.4 wherein both types of users have same required level of functionality. In that case, any under-provisioning would impact different user segments in the same way. This limits the ability of the firm to save information rent by distorting low version’s functionality. And the optimal strategy is likely to be a single-version product strategy. The same logic may explain cases where software firms simultaneously adopt versioning strategy for some products and single-version strategy for others. For example, Adobe offers Photoshop in five versions but sells a single version of Digital Publishing Suite.

This research also provides key insights to managers to develop optimal product-pricing strategies for information goods like software and information services from which users derive utility in a task-oriented context and users’ requirements for functionality increase over time. More specifically, when there is a shift in users’ required levels of functionality, firms may need to adjust the degree of functionality distortion of the low version to avoid either any loss of revenue from low-type users or any overpayment of information rent to high-type users. Managers also need to be aware that the optimal degree of functionality distortion in the low version depends on the proportion of high-type users in the market, and the firm may accordingly adjust the low version’s functionality level in different markets. For example, Microsoft offers
Windows 7 Home Basic only in some emerging markets. Moreover, since the relative strength of the benefit function from functionality provided in the software and the disutility function from under-provisioning of functionality is critical in determining extent of under-provisioning, managers should pay special attention to this in determining product-pricing strategies.

We made some assumptions regarding the users’ required level of functionality and cost of development to build the framework for analysis. Note that though we assume convex cost function, our results hold for any general cost function as long as Assumption 2 holds. Later in §5, we show that a versioning strategy often dominates other product strategies even when some of the assumptions are relaxed, one at a time. We also examine profitability of a versioning strategy when rank ordering is not maintained and find that versioning is optimal only under some conditions.

We assume that users have full information about the level of functionality in both the high and low versions, though in reality, users often may be uncertain about the true level of functionality and a potential future research project may be to allow uncertainty in the utility function. Moreover, an inter-temporal model that abstracts dynamics of changes in users’ requirements for functionality and analyzes multi-period product-pricing strategies of the software firm is an interesting avenue for future research. Further, the class of utility function with single-crossing property, developed in this research, can be used to study product-pricing strategy of other information goods that have characteristics similar to software.

We assume that over-provisioning is useless, though in some situation it may be utility decreasing when users have some resource constraints, as in Chellappa and Mehra (2011). Future research may examine versioning strategies under costly over-provisioning on the lines of Chellappa and Shivendu (2010), which may shed some light on theoretical understanding of the
widespread creation of bloatware\textsuperscript{12}. We assume monopolist market structure, and it may be fruitful to extend our setting to competitive market structure in the future. It may also be worthwhile to study a firm’s versioning strategy under different types of utility functions and analytically establish that only certain types of utility functions lend themselves to versioning.

References


Belleflamme, P. (2002). Pricing information goods in the presence of copying, mimeo, Queen Mary University of London.


\textsuperscript{12} Software bloat is a process whereby successive versions of a computer program include an increasing proportion of unnecessary features that the end users do not access, or that generally use more system resources than necessary, while offering little or no benefit to users (http://en.wikipedia.org/wiki/Bloatware).


Appendix 1.A: Summary of Notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_h$</td>
<td>Marginal valuation for functionality of high-type users</td>
</tr>
<tr>
<td>$\theta_l$</td>
<td>Marginal valuation for functionality of low-type users</td>
</tr>
<tr>
<td>$x_h$</td>
<td>High-type users required level of functionality</td>
</tr>
<tr>
<td>$x_l$</td>
<td>Low-type users required level of functionality</td>
</tr>
<tr>
<td>$q_h$</td>
<td>Functionality under a strategy of selling a single version only to high-type users</td>
</tr>
<tr>
<td>$q_s$</td>
<td>Functionality under a strategy of selling a single version to all users</td>
</tr>
<tr>
<td>$q_v$</td>
<td>Functionality of low version under a versioning strategy</td>
</tr>
<tr>
<td>$q_vH$</td>
<td>Functionality of high version under a versioning strategy</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Proportion of high-type users in the market</td>
</tr>
<tr>
<td>$\superscript{N}$</td>
<td>Superscript: case where users do not experience disutility from under-provisioning of functionality</td>
</tr>
<tr>
<td>$p_s$</td>
<td>Price under a product strategy of selling a single version to all users</td>
</tr>
<tr>
<td>$p_H$</td>
<td>Price under a product strategy of selling a single version only to high-type users</td>
</tr>
<tr>
<td>$p_{vH}$</td>
<td>Price of high version under a versioning strategy</td>
</tr>
<tr>
<td>$p_vL$</td>
<td>Price of low version under a versioning strategy</td>
</tr>
<tr>
<td>$\pi_s$</td>
<td>Profit under a product strategy of selling a single version to all users</td>
</tr>
<tr>
<td>$\pi_H$</td>
<td>Profit under a product strategy of selling a single version only to high-type users</td>
</tr>
<tr>
<td>$\pi_v$</td>
<td>Profit under a versioning strategy</td>
</tr>
<tr>
<td>$q'$</td>
<td>Functionality level where both types of users have the same utility</td>
</tr>
<tr>
<td>$\alpha_2$, $\alpha_1$</td>
<td>Critical values of proportion of high-type users in the market where the firm starts and stops distorting functionality of the low version respectively.</td>
</tr>
</tbody>
</table>
Appendix 1.B: Proofs of Lemmas and Propositions

**Lemma 1:** When users experience heterogeneous disutility from under-provisioning of functionality, the indifferent functionality level is always positive, that is, \( q' > 0 \).

**Proof of Lemma 1:**

Consumer utility is given by equation (1):

\[
U_i(\theta, x, q) = \begin{cases} 
\theta f(q) - g(x, q), & \text{if } q < x_i \\
\theta f(x_i), & \text{if } q \geq x_i
\end{cases}, \quad i \in \{H, L\}
\]

1. At \( q = x_H \), \( U_H(\theta_H, x_H, q = x_H) = \theta_H f(q = x_H) \) and \( U_L(\theta_L, x_L, q = x_H) = \theta_L f(q = x_L) \), and therefore, \( U_H(\theta_H, x_H, q = x_H) > U_L(\theta_L, x_L, q = x_H) \).

2. At \( q = 0 \), since \( f(0) = 0 \), we have \( U_i(\theta, x, q) = -g(x, q) \). And since \( \frac{\partial g(x, q)}{\partial q} < 0 \), this implies \( g(x_H, 0) > g(x_L, 0) \). Hence, \( U_H(\theta_H, x_H, q = 0) < U_L(\theta_L, x_L, q = 0) \).

3. Since \( \frac{\partial f(q)}{\partial q} > 0 \), \( \frac{\partial^2 f(q)}{\partial q^2} \leq 0 \), \( \frac{\partial g(x, q)}{\partial q} < 0 \), and \( \frac{\partial^2 g(x, q)}{\partial q^2} \geq 0 \), we have \( \frac{\partial U_i(q)}{\partial q} > 0 \), \( \frac{\partial^2 U_i(q)}{\partial q^2} \leq 0 \), \( i \in \{H, L\} \), for \( q \in (0, x_L) \).

From 1, 2 and 3 above, we know that utility of high-type and low-type users is monotonically non decreasing at a non-increasing rate in functionality level \( q \) in the interval \((0, x_L)\). This implies that two utility functions have the same value at some functionality level, \( q = q' \), where \( q' > 0 \).

\[ \blacksquare \]

**Lemma 2:** (A) When the firm adopts the product strategy of selling one version only to high-type users, the optimal level of functionality \( q_H^* \) solves the first-order condition given by \( \alpha \theta_H f'(q_H^*) - \alpha g'(x_H, q_H^*) - c'(q_H^*) = 0 \). The optimal level of functionality is bound such that \( x_L < q_H^* \leq x_H \) and it increases within this bound as the proportion of high-type users \( (\alpha) \) increases.

(B) When the firm adopts the product strategy of selling one version to both types of users, then the optimal level of functionality is the same as the low-type users’ required functionality level.
that is, \( q^*_x = x_L \). (C) When the firm adopts a versioning strategy, then the highest level of functionality produced is the same as in (A), that is, \( q^*_H = q^*_H \).

**Proof of Lemma 2:**

(A) Product strategy of selling one version to high-type users only:

When the firm adopts single version product strategy of selling to high-type consumers, then the firm set the price high version at \( U_H(\theta_H, x_H, q = q_H) \). The firm’s profit function in Equation (2) is:

\[
\pi_H = \alpha p_H - c(q_H) = \alpha(\theta_H f(q_H) - g(x_H, q_H)) - c(q_H).
\]

Let \( q_H = q_H^* \) be the solution to the first order condition:

\[
\frac{\partial \pi_H}{\partial q_H} = \alpha \theta_H f'(q_H) - \alpha g'(x_H, q_H^*) - c'(q_H^*) = 0
\]

(A1)

The optimal quality level \( q_H^* \leq x_H \), because high-type users do not derive higher utility for \( q > x_H \).

Since \( \frac{\partial^2 f(q)}{\partial q^2} \leq 0 \), \( \frac{\partial^2 g(x, q)}{\partial q^2} \geq 0 \), and \( \frac{\partial^2 c(q)}{\partial q^2} > 0 \), so \( \frac{\partial^2 \pi_H}{\partial q_H^2} = \alpha \theta_H f''(q_H) - \alpha g''(x_H, q_H) - c''(q_H) < 0 \), for all \( q_H \leq x_H \). Hence, \( \frac{\partial^2 \pi_H}{\partial q_H^2} \big|_{q_H=q_H^*} < 0 \).

Therefore, \( q_H = q_H^* \) is the optimal functionality level for the single version product strategy of only selling to high-type users. Hence, we have the first part of Lemma 2(A).

Since from Assumption 2, we have \( q_H^* > x_L \), and since \( \frac{\partial U_H(q)}{\partial q} = 0 \), for \( q > x_H \), hence, \( x_L < q_H^* \leq x_H \) and we have the second part of Lemma 2(A).

At \( q_H = q_H^* \), the first order condition of \( \pi_H \) with respect to \( q_H \) is satisfied, therefore, we have the identity equation:

\[
\alpha \theta_H f'(q_H^*) - \alpha g'(x_H, q_H^*) - c'(q_H^*) = 0
\]

(A2)

Taking the full derivative of both sides of the identity (A2) with respect to \( \alpha \), we have:

\[
\frac{\partial}{\partial \alpha} \left( \alpha \theta_H f'(q_H^*) - \alpha g'(x_H, q_H^*) - c'(q_H^*) \right) = 0
\]

By simple algebraic manipulation,

\[
\frac{\partial q_H^*}{\partial \alpha} = \frac{\theta_H f'(q_H^*)}{-\alpha \theta_H f''(q_H^*) + \alpha g''(x_H, q_H^*) + c''(q_H^*)} > 0 \text{ because } f' > 0, f^* \leq 0, g^* \geq 0, \text{ and } c^* > 0.
\]

Hence we have the third part of Lemma 2(A).

(B) Single version product strategy of selling to all consumers:
From Assumption 1, \( U_H(q = x_L) > U_L(q = x_L) \) and this implies \( q' < x_L \). From Assumption 2, \( q^* > x_L \).

Hence we have \( q' > x_L > q' \). The firm sets the price at \( U_L(\theta_L, x_L, q = q_H) \) when sells single version to all users. Since, \( \frac{\partial U_L(q)}{\partial q} = 0 \), for \( q > x_L \), firm never produces functionality greater than \( x_L \).

Therefore, \( q_H^* = x_L \).

Hence we have Lemma 2(B).

(C) Versioning strategy:

Under versioning strategy firm sets prices of the high version and low version such that high-type buy high version and low-type users buy the low version. Let the price of functionality and price of high version be \( q_{vh} \) and \( p_{vh} \), and of low version be \( q_{vl} \) and \( p_{vl} \). From Assumption 1 and 2, we have \( q^* > x_L > q' \). The optimal highest functionality is \( q_{vh}^* > x_L \), and \( x_L > q' \), thus \( q_{vh}^* > q' \).

The prices are set such that high-type consumers buy the high version and low-type consumers buy the low version. After satisfying the ICs and IRs, the prices for the two versions are:

\[
p_{vh} = U_H(q_{vh}) - (U_H(q_{vl}) - U_L(q_{vl})), \quad \text{and} \quad p_{vl} = U_L(q_{vl}).
\]

Demand for high version is \( \alpha \) and demand for low version is \( 1 - \alpha \). Firm’s profit function as in equation can be written as:

\[
\pi_v = \alpha p_{vh} + (1 - \alpha) p_{vl} - c(q_{vh})
= \alpha(U_H(q_{vh}) - (U_H(q_{vl}) - U_L(q_{vl}))) + (1 - \alpha)U_L(q_{vl}) - c(q_{vh})
= (\theta_H f(q_{vh}) - g(x_h, q_{vh})) + \alpha[(\theta_H f(q_{vh}) - g(x_h, q_{vh})) - (\theta_H f(q_{vh}) - g(x_h, q_{vh}))] - c(q_{vh})
= \theta_H f(q_{vh}) - g(x_h, q_{vh}) + \alpha(\theta_H f(q_{vh}) - g(x_h, q_{vh}) - \theta_H f(q_{vh}) + g(x_h, q_{vh})) - c(q_{vh})
\]

Let \( q_{vh}^* = q_{vh}^* \) be the solution to the first order condition:

\[
\frac{\partial \pi_v}{\partial q_{vh}} = \alpha \theta_H f'(q_{vh}^*) - \alpha g'(x_h, q_{vh}^*) - c'(q_{vh}^*) = 0 \tag{A3}
\]

Now, checking for the second order condition:

\[
\frac{\partial^2 \pi_v}{\partial q_{vh}^2} = \alpha \theta_H f''(q_{vh}^*) - \alpha g''(x_h, q_{vh}^*) - c''(q_{vh}^*) < 0,
\]

because \( f'' \leq 0 \), \( g'' \geq 0 \), and \( c'' > 0 \).

Since the first order condition in (A1) and (A3) are identical, \( q_{vh}^* = q_H^* \).

Hence, we have Lemma 2(C).
**PROPOSITION 1:** Under a versioning strategy, the software firm’s optimal level of functionality of the low version is bounded between the indifferent functionality level and the required level of functionality of low-type user, such that, \( q' \leq q_{vl}^* \leq x_L \).

**Proof of Proposition 1:**

Let the optimal functionality of low version be \( q_{vl}^* \). Since low-type users’ utility does not increase for any functionality \( q_{vl} > x_L \), the firm will never set the functionality level of the low version greater than \( x_L \). Hence \( q_{vl}^* \leq x_L \). We know that as proportion of high-type users increases, the firm may reduce the functionality of low version to economies on information rent. Therefore, there is some value of \( \alpha = \alpha_z \) above which it is optimal for the firm to lower the functionality level of the low version \( q_{vl}^* \leq x_L \). As \( \alpha \) keep increasing, the firm distorts functionality of the low version further till \( q_{vl}^* = q' \) at some value of \( \alpha = \alpha_z \) when the information rent to high-type users is zero, that is, \( \theta_H f(q') - g(x_H, q') - \theta_L f(q') + g(x_L, q') = 0 \). Therefore, the firm will never set the functionality level of low version lower than \( q' \), that is, \( q_{vl}^* \geq q' \). Hence, we have \( q' \leq q_{vl}^* \leq x_L \).

**LEMMA 3:** When users experience heterogeneous disutility from under-provisioning of functionality, then under a versioning strategy, the optimal functionality level of the low version depends on the proportion of high-type users in the market. (A) When \( \alpha \in (0, \alpha_z) \), the optimal functionality level of the low version is \( q_{vl}^* = x_L \). (B) When \( \alpha \in [\alpha_z, \alpha_1] \), the optimal functionality level of the low version is bound within \( [q', x_L] \), such that \( q_{vl}^* \) solves the first-order condition given by \( (\theta_H f'(q_{vl}^*) - g'(x_L, q_{vl}^*)) - \alpha(\theta_H f'(q_{vl}^*) - g'(x_H, q_{vl}^*)) = 0 \). (C) When \( \alpha \in (\alpha_1, 1) \), the optimal functionality level of the low version is \( q_{vl}^* = q' \). The two threshold values of proportion of high-type users are: \( \alpha_z = \frac{\theta_H f'(x_L)}{\theta_H f'(x_L) - g'(x_H, q')} \) and \( \alpha_1 = \frac{\theta_L f'(x_L)}{\theta_L f'(x_L) - g'(x_H, x_L)} \).

**Proof of Lemma 3:**

We first provide the definitions of upper threshold level and lower threshold level, given as Definition 2 and 3 in the paper.
Definition 2: When the proportion of high-type users in the market is higher than the upper threshold proportion $\alpha$, then under-provisioning of functionality in the low version of software is most severe ($q_{L}^{*} = q'$).

Definition 3: When the proportion of high-type users in the market is lower than the lower threshold proportion $\alpha$, then there is no under-provisioning of functionality in the low version of software ($q_{L}^{*} = q_x$).

Let $q_{H}^{*} = q'$ when $\alpha \in [\alpha_1, 1]$ and $q_{L}^{*} = x_L$ when $\alpha \in (0, \alpha_1]$, from Definition 1 and Definition 2.

From Proposition 1, optimal functionality level of the low version is $q_{L}^{*} = q_x$ when $\alpha \leq \alpha \leq \alpha_1$, and $q_{L}^{*} = x_L$ when $\alpha \leq \alpha \leq \alpha_1$.

At $q_{H}^{*} = x_L$ and for $\alpha_2 \leq \alpha \leq \alpha_1$, we have the identity equation from the first order condition of the profit function:

$$(\theta - \alpha g'_{H})(q_{L}^{*}) - g'(x_L, q_{L}^{*}) + \alpha g'(x_H, q_{L}^{*}) = 0$$

(A5)

Taking first derivative of (A5) with respect to $\alpha$, and we have:

$$\frac{\partial q_{L}^{*}}{\partial \alpha}(\theta - \alpha g'_{H})(q_{L}^{*}) - \theta g'(x_L, q_{L}^{*}) + \frac{\partial q_{L}^{*}}{\partial \alpha} \frac{\partial g'(x_L, q_{L}^{*})}{\partial \alpha} + \alpha \frac{\partial g'(x_H, q_{L}^{*})}{\partial \alpha} = 0$$

After simplifying we have:

$$\frac{\partial q_{L}^{*}}{\partial \alpha} = \frac{\theta g'(x_L, q_{L}^{*}) - g'(x_H, q_{L}^{*})}{\theta g'(x_L, q_{L}^{*}) - g'(x_H, q_{L}^{*}) - \alpha g'(x_H, q_{L}^{*})}.$$  

Since the second derivative is negative,

$$\frac{\partial^2 \pi_{V}}{\partial q_{L}^{*}^2} = \theta g'(x_L, q_{L}^{*}) - g'(x_H, q_{L}^{*}) - \alpha g'(x_H, q_{L}^{*}) < 0$$

for $\alpha_2 \leq \alpha \leq \alpha_1$, $f' > 0$, and $g' < 0$, then, $rac{\partial q_{L}^{*}}{\partial \alpha} < 0$ for $\alpha_2 \leq \alpha \leq \alpha_1$, and $q' \leq q_{L}^{*} \leq x_L$.

This implies that the optimal functionality level of low version decreases, as proportion of high-type users in the market increases.

Since the optimal functionality level of the low version is between $q'$ and $x_L$ (Proposition 1), and the optimal functionality level of low version decreases as the proportion of high-type users in the market increases, we find the corresponding threshold values of $\alpha$ when $q_{H}^{*} = x_L$ and $q_{L}^{*} = q'$.

(A) From Lemma 1, we know that the optimal functionality level decreases within in $x_L \geq q_{H}^{*} \geq q'$.

The lower threshold value for the proportion of high-type users, $\alpha = \alpha_2$, corresponds to $q_{H}^{*} = x_L$.

By substituting $q_{H}^{*} = x_L$ and $\alpha = \alpha_2$ into the first order condition:

$$\frac{\partial \pi_{V}}{\partial q_{L}^{*}} = \theta f'(x_L) - g'(x_L, x_L) - \alpha_2(\theta f'(x_L) - g'(x_H, x_L)) = 0,$$

we have:
\[ \alpha_z = \frac{\theta_L f'(x_L) - g'(x_L, x_L)}{\theta_H f'(x_H) - g'(x_H, x_L)} = \frac{\theta_L f'(x_L)}{\theta_H f'(x_H) - g'(x_H, x_L)}. \]

Therefore, the optimal functionality level of the low version is at \( x_L \) where \( \alpha \in (0, \alpha_z) \). Hence, we have Lemma 3 (A).

(B) When \( \alpha \in [\alpha_z, 1] \), the optimal functionality level, \( q_{vz}^* \), is the solution to the first order condition

\[ \frac{\partial \pi_v}{\partial q_{vz}} = (\theta_L f'(q_{vz}^*) - g'(x_L, q_{vz}^*)) - \alpha(\theta_H f'(q_{vz}^*) - g'(x_H, q_{vz}^*)) = 0. \]

It is bounded by \([q', x_L]\) and decreases as the proportion of high-type users increases, \( \frac{\partial q_{vz}^*}{\partial \alpha} < 0 \) as shown at the start of the proof of Lemma 3.

Hence we have Lemma 3 (B).

(C) At the upper threshold value for the proportion of high-type users, \( \alpha = \alpha_i \), the optimal functionality level of the low version is at the indifference functionality level, \( q_{vz}^* = q' \). By substituting \( q_{vz}^* = q' \) and \( \alpha = \alpha_i \) into the first order condition:

\[ \frac{\partial \pi_v}{\partial q_{vz}} = \theta_L f'(q') - g'(x_L, q') - \alpha_i(\theta_H f'(q') - g'(x_H, q')) = 0, \]

we have: \( \alpha_i = \frac{\theta_L f'(q') - g'(x_L, q')}{\theta_H f'(q') - g'(x_H, q')} \). Therefore, the optimal functionality level of the low version is set at \( q' \) where \( \alpha \in (\alpha_i, 1) \). Hence, we have Lemma 3 (C).

Now we check that the upper threshold value (\( \alpha_i \)) is always higher or equal to the lower threshold value (\( \alpha_z \)), that is \( \alpha_i \geq \alpha_z \),

\[ \alpha_i - \alpha_z = \frac{\theta_L f'(q') - g'(x_L, q')}{\theta_H f'(q') - g'(x_H, q')} - \frac{\theta_L f'(x_L)}{\theta_H f'(x_L) - g'(x_H, x_L)} \]
\[ = \frac{(\theta_L f'(q') - g'(x_L, q'))(\theta_H f'(x_L) - g'(x_H, x_L)) - \theta_L f'(x_L)(\theta_H f'(q') - g'(x_H, q'))}{(\theta_H f'(q') - g'(x_H, q'))(\theta_H f'(x_L) - g'(x_H, x_L))} \]
\[ = \frac{\theta_L f'(x_L)g'(x_H, q') - \theta_L f'(x_L)(\theta_H f'(x_L) - g'(x_H, x_L))}{(\theta_H f'(q') - g'(x_H, q'))(\theta_H f'(x_L) - g'(x_H, x_L))} \]

(A6)

From model setup and Lemma 1, \( f'(q') \leq f'(x_L) \) and \( g'(x_H, x_L) \leq g'(x_H, q') \), therefore, RHS of A6 is always greater than or equal to zero. Hence, \( \alpha_i - \alpha_z \geq 0 \). Therefore, \( \alpha_i \geq \alpha_z \).
PROPOSITION 2: Sufficient Condition for Optimality of Versioning Strategy: When users experience heterogeneous disutility from under-provisioning of functionality, versioning strategy dominates the product strategies of selling one version only to high-type users or selling to both types of users.

Proof of Proposition 2:

First, we derive the optimal functionality levels and profits under the three product strategies and the effect of the proportion of high-type users in the market on optimal profits.

From Lemma 2 (A), the optimal profit of a single version strategy of selling to high-type users only is:

$$\pi_h^* = \alpha \theta_h f(q_h^*) - g(x_h, q_h^*) - c(q_h^*),$$

where the optimal functionality level satisfies the first order condition:

$$(\alpha \theta_h f'(q_h^*) - \alpha g'(x_h, q_h^*) - c'(q_h^*)) = 0 \text{ and } x_h < q_h^* \leq x_H.$$  

We take the first order derivative of the optimal profit with respect to $\alpha$, and according to the Envelope Theorem at optimality:

$$\frac{d\pi_h^*}{d\alpha} = \theta_h f(q_h^*) - g(x_h, q_h^*) = U_h(q_h^*) > 0.$$  

It implies that as the proportion of high-type users in the market increases, the optimal profit under a single version strategy of selling to high-type users increases.

From Lemma 2 (B), the optimal functionality level under a strategy of selling one version to all consumers is $q_L^* = x_L$ and the optimal profit is: $\pi_L^* = \theta_L f(x_L) - c(x_L)$. And we have: $\frac{d\pi_L^*}{d\alpha} = 0$.

From Lemma 2 (C) and Lemma 3, the optimal profit of the versioning strategy is:

$$\pi_v^* = \theta_L f(q_v^*) - g(x_L, q_v^*) + \alpha \theta_H f(q_v^*) - g(x_H, q_v^*) - \theta_H f(q_v^*) + g(x_H, q_v^*) - c(q_v^*).$$

Taking the first order derivative with respect to $\alpha$, and according to the Envelope theorem at optimality:

$$\frac{d\pi_v^*}{d\alpha} = \frac{d\theta_L f(q_v^*)}{d\alpha} - g'(x_L, q_v^*) - \theta_L f'(q_v^*) - \alpha \theta_H f'(q_v^*) - g'(x_H, q_v^*) + \theta_H f'(q_v^*) - g(x_H, q_v^*) - \theta_H f(q_v^*) + g(x_H, q_v^*) - c'(q_v^*).$$

From Proposition 1 and Lemma 3, we have $q' \leq q_v^* \leq x_L$, and $q_v^*$ satisfies the first order condition, $\theta_L f(q_v^*) - g'(x_L, q_v^*) - \alpha \theta_H f(q_v^*) - g'(x_H, q_v^*) = 0$, where $\alpha \in [\alpha_2, \alpha_1]$; and $\frac{d\theta_L f(q_v^*)}{d\alpha} = 0$, where $\alpha \in (0, \alpha_2)$ or $\alpha \in (\alpha_1, 1)$.

Therefore, $\frac{d\pi_v^*}{d\alpha} = \theta_L f(q_v^*) - g(x_L, q_v^*) - \theta_H f(q_v^*) + g(x_H, q_v^*)$, where $\alpha \in (0.1)$.  

66
From Proposition 1 and Lemma 2, \( q_{vl} \leq x_L \) and \( x_L < q_{vl}^* \), therefore, we have \( q_{vl}^* < q_{vh}^* \) and, hence, \( \theta_H f(q_{vh}^*) - g(x_H, q_{vh}^*) - \theta_H f(q^*) + g(x_H, q^*_H) > 0 \). Therefore, \( \frac{\partial \pi_v^*}{\partial \alpha} > 0 \), where \( \alpha \in (0,1) \).

Now we compare the optimal profits under the three strategies. The optimal functionality of the low version is \( q_{vl}^* = x_L \), and the optimal functionality of the high version is \( q_{vh}^* = q_H^* \) under a versioning strategy when \( \alpha \in (0, \alpha_1) \). The optimal profit under a versioning strategy is:

\[
\pi_v^* = \theta_L f(x_L) - g(x_L, x_L) + \alpha(\theta_H f(q_{vh}^*) - g(x_H, q_{vh}^*) - \theta_H f(x_L) + g(x_H, x_L)) - c(q_{vh}^*) = \theta_L f(x_L) - g(x_L, x_L) + \alpha(\theta_H f(q_{vh}^*) - g(x_H, q_{vh}^*) - \theta_H f(x_L) + g(x_H, x_L)) - c(q_{vh}^*)
\]

and

\[
\pi_v^* = \theta_L f(x_L) - g(x_L, x_L) + \alpha(\theta_H f(q_{vh}^*) - g(x_H, q_{vh}^*) - \theta_H f(x_L) + g(x_H, x_L)) - c(q_{vh}^*) = \theta_L f(x_L) - g(x_L, x_L) + \alpha(\theta_H f(q_{vh}^*) - g(x_H, q_{vh}^*) - \theta_H f(x_L) + g(x_H, x_L)) - c(q_{vh}^*)
\]

Therefore, \( \pi_v^* > \pi_h^* \) and \( \pi_v^* > \pi_h^* \) when \( \alpha \in (0, \alpha_1) \).

We then compare the optimal profits under the strategies when \( \alpha \in (\alpha_1, 1) \). The optimal functionality of the low version is \( q_{vl}^* = q' \), and the optimal functionality of the high version is \( q_{vh}^* = q_h^* \) under a versioning strategy when \( \alpha \in (\alpha_1, 1) \). The optimal profit under a versioning strategy is:

\[
\pi_v^* = \theta_L f(q') - g(x_L, q') + \alpha(\theta_H f(q_{vh}^*) - g(x_H, q_{vh}^*) - \theta_H f(q') + g(x_H, q')) - c(q_{vh}^*) = \theta_L f(q') - g(x_L, q') + \alpha(\theta_H f(q_{vh}^*) - g(x_H, q_{vh}^*) - \theta_H f(q') + g(x_H, q')) - c(q_{vh}^*)
\]

Therefore, we show that \( \pi_v^* > \pi_h^* \), where \( \alpha \in (\alpha_1, 1) \).

As we shown before that \( \frac{\partial \pi_v^*}{\partial \alpha} = 0 \) and \( \frac{\partial \pi_v^*}{\partial \alpha} > 0 \), for \( \alpha \in (0,1) \), and also since \( \pi_v^* > \pi_h^* \) where \( \alpha \in (0, \alpha_2) \), and \( \alpha_2 \leq \alpha_1 \). Hence, \( \pi_v^* > \pi_h^* \) where \( \alpha \in (0,1) \).

And since
\[
\frac{\partial \pi_h^*}{\partial \alpha} = \theta_H f(q_{vh}^*) - g(x_H, q_{vh}^*) > 0 \quad \text{for} \quad \alpha \in (0,1),
\]
and also since \( \pi_v^* > \pi_h^* \), where \( \alpha \in (0,\alpha_1) \) and \( \alpha \in (\alpha_1,1) \), then, \( \pi_v^* > \pi_h^* \), for \( \alpha \in (0,1) \).

Therefore, \( \pi_v^* > \pi_h^* \) and \( \pi_v^* > \pi_h^* \) for \( \alpha \in (0,1) \). Hence, we have Proposition 2.
**Lemma 4:** As the required level of functionality of high-type users \( x_H \) increases, (i) the indifferent functionality level \( q' \) increases, and (ii) the optimal level of functionality of the high version \( q'_{vh} \) increases.

**Proof of Lemma 4:**

(i) From Lemma 1, at \( q = q' \), \( U_H(\theta_H, x_H, q = q') = U_L(\theta_L, x_L, q = q') \). We take first order derivative with respect to \( x_H \) and get FOC as:

\[
(\theta_H - \theta_L) f'(q') \frac{\partial q'}{\partial x_H} - (g'(x_H, q') \frac{\partial g(x_H, q')}{\partial x_H} - g'(x_L, q') \frac{\partial q'}{\partial x_H}) = 0.
\]

After simplifying, we have:

\[
\frac{\partial g(x_H, q')}{\partial x_H} = \frac{\theta_H f'(q') - g'(x_H, q') - \theta_L f'(q') + g'(x_L, q')}{(\theta_H - \theta_L)}.
\]

From model setup, we have:

\[
\frac{\partial^2 g(x_H, q)}{\partial x \partial q} < 0, \quad \theta_H f'(q') - g'(x_H, q') - \theta_L f'(q') + g'(x_L, q') > 0, \text{ and}
\]

\[
\frac{\partial g(x_H, q)}{\partial x} > 0.
\]

Therefore, we have \( \frac{\partial q'}{\partial x_H} > 0 \). Hence, we have Lemma 4 (i).

(ii) From Lemma 2, at \( q_{vh} = q_{vh}^* \), the first order condition is satisfied:

\[
a \theta_H f'(q_{vh}^*) - \alpha g'(x_H, q_{vh}^*) - c'(q_{vh}^*) = 0.
\]

Taking the first order derivative of both sides, we get:

\[
\frac{\partial q_{vh}^*}{\partial x_H} (a \theta_H f'(q_{vh}^*) - \alpha g'(x_H, q_{vh}^*) - c'(q_{vh}^*)) - a \frac{\partial^2 g(x_H, q = q_H^*)}{\partial x_H \partial q} = 0.
\]

After simplifying we have

\[
\frac{\partial q_{vh}^*}{\partial x_H} = \frac{a \frac{\partial g(x_H, q)}{\partial x} \frac{\partial q_{vh}^*}{\partial x_H}}{a \theta_H f'(q_{vh}^*) - \alpha g'(x_H, q_{vh}^*) - c'(q_{vh}^*)}.
\]

From Lemma 2, \( a \theta_H f'(q_{vh}^*) - \alpha g'(x_H, q_{vh}^*) - c'(q_{vh}^*) < 0 \), and from model setup, \( \frac{\partial^2 g(x_H, q)}{\partial x \partial q} < 0 \).

Therefore, we have \( \frac{\partial q_{vh}^*}{\partial x_H} > 0 \). Hence, we have Lemma 4 (ii).

**Proposition 3:** As the required level of functionality of high-type users \( x_H \) increases, (i) the lower threshold proportion \( \alpha_z \) and the upper threshold proportion \( \alpha_l \) decrease, (ii) the optimal functionality of the low version \( q_{vl}^* \) (a) does not change if the proportion of high-type users is small, (b) decreases if proportion of high-type users is moderate, and (c) increases if the proportion of high-type users is large.
Proof of Proposition 3:

(i) At the lower threshold proportion of high-type users in the market, \( \alpha = \alpha_z \), the optimal functionality level of the low version is, \( q^*_v = x_L \), and we have
\[
\theta_L f'(x_L) - g'(x_L, x_L) - \alpha_z(\theta_H f'(x_L) - g'(x_H, x_L)) = 0.
\]
We take its first order derivative with respect to \( x_H \):
\[
\alpha_z \frac{\partial}{\partial x_H} \left[ \frac{\partial^2 g(x_H, q = x_L)}{\partial x_H \partial q} \right] - \frac{\partial}{\partial x_H} \left[ \frac{\partial^2 g(x_H, q = x_L)}{\partial x_H \partial q} \right] = 0.
\]
After simplifying we have, \( \alpha_z \frac{\partial^2 g(x_H, q = x_L)}{\partial x_H \partial q} < 0 \), we have \( \frac{\partial^2 g(x_H, q = x_L)}{\partial x_H \partial q} < 0 \).

Hence, we have the first part of Proposition 3(i).

At the upper threshold proportion of high-type users in the market \( \alpha = \alpha_i \), the optimal functionality level of the low version, \( q^*_v = q_i' \), and then we have
\[
\theta_L f'(q_i') - g'(x_L, q_i') - \alpha_i(\theta_H f'(q_i') - g'(x_H, q_i')) = 0.
\]
We take its first order derivative with respect to \( x_H \) and have:
\[
\alpha_i \frac{\partial}{\partial x_H} \left[ \frac{\partial^2 g(x_H, q = q_i')}{\partial x_H \partial q} \right] - \frac{\partial}{\partial x_H} \left[ \frac{\partial^2 g(x_H, q = q_i')}{\partial x_H \partial q} \right] = 0.
\]
From Lemma 4(i), we have \( \alpha_i > 0 \), and since \( \theta_L f'(q_i') - g'(x_L, q_i') - \alpha_i \theta_H f'(q_i') + \alpha_i g'(x_H, q_i') < 0 \), \( f' > 0 \), \( g' < 0 \), and \( \frac{\partial^2 g(x_H, q = q_i')}{\partial x_H \partial q} < 0 \), we have \( \frac{\partial^2 g(x_H, q = q_i')}{\partial x_H \partial q} < 0 \).

Hence, we have the second part of Proposition 3(i).

(ii) For \( q_{iv} = q_{iv}' \), the first order condition, \( \theta_L f'(q_{iv}') - g'(x_L, q_{iv}') - \alpha(\theta_H f'(q_{iv}') - g'(x_H, q_{iv}')) = 0 \), is satisfied, where \( q' \leq q_{iv}' \leq x_L \). We take its first order derivative with respect to \( x_H \) and get:
\[
\frac{\partial}{\partial x_H} \left[ \frac{\partial^2 g(x_H, q = q_{iv}')}{\partial x_H \partial q} \right] - \frac{\partial}{\partial x_H} \left[ \frac{\partial^2 g(x_H, q = q_{iv}')}{\partial x_H \partial q} \right] = 0.
\]
After simplifying we have:
\[
\frac{\partial}{\partial x_H} \left[ \frac{\partial^2 g(x_H, q = q_{iv}')}{\partial x_H \partial q} \right] = \frac{\alpha \frac{\partial^2 g(x_H, q = q_{iv}')}{\partial x_H \partial q} - \theta_L f'(q_{iv}') + \alpha g'(x_H, q_{iv}')}{\theta_L f'(q_{iv}') - g'(x_L, q_{iv}') - \alpha \theta_H f'(q_{iv}') + \alpha g'(x_H, q_{iv}')}.
\]
Since \( \theta_L f'(q) - g'(x_L, q) - \alpha \theta_H f'(q) + \alpha g'(x_H, q) < 0 \), and \( \frac{\partial^2 g(x_H, q = q_{iv}')}{\partial x_H \partial q} < 0 \), we have \( \frac{\partial}{\partial x_H} \left[ \frac{\partial^2 g(x_H, q = q_{iv}')}{\partial x_H \partial q} \right] < 0 \).
(a) We have shown that $\frac{\partial \alpha_z}{\partial x_H} < 0$ in the first part of Proposition 3(i). From Lemma 3 have and that the firm’s optimal functionality level of the low version is $q_{vl}^* = x_L$, when $\alpha \in (0, \alpha_z)$. Since $x_L$ does not change, the firm does not change the functionality level of the low version when there is relatively low proportion of high-type users in the market. Hence, Proposition 3 (ii) (a).

(ii) We have shown that $\frac{\partial q'}{\partial x_H} > 0$ in Lemma 4, $\frac{\partial \alpha_z}{\partial x_H} < 0$ and $\frac{\partial \alpha_z}{\partial x_H} < 0$ in Proposition 3 (i), and also $\frac{\partial q_{vl}}{\partial x_H} < 0$ for $q' \leq q_{vl}^* \leq x_L$ from the first part of Proposition 3(ii).

It implies that the firm starts to reduce functionality level of low version at a smaller $\alpha_z$ at $q_{vl}^* = x_L$, and stops reducing functionality level of low version at a smaller $\alpha_z$ at $q_{vl}^* = q'$. It also implies the condition for the firm to provide functionality level of the low version at a higher $q'$ is less restrictive. Therefore, for some large proportion of high-type users in the market, the firm provides less functionality distortion as $\frac{\partial q'}{\partial x_H} > 0$, but for some moderate proportion of high-type users in the market, the firm provide more functionality distortion as $\frac{\partial q_{vl}}{\partial x_H} < 0$.

Hence, we prove Proposition 3 (ii) (b) and (c).

\[\text{Lemma 5: As the required level of functionality of low-type users (} x_L \text{) increases, (i) the indifferent functionality level } q' \text{ decreases, and (ii) there is no impact on the optimal level of functionality of high version } q_{vh}^*.\]

Proof of Lemma 5:

(i) From Lemma 1, at $q = q'$, $U_H(\theta_H, x_H, q = q') = U_L(\theta_L, x_L, q = q')$. Taking the first order derivative with respect to $x_L$ we get FOC as: $(\theta_H - \theta_L)f'(q)\frac{\partial q'}{\partial x_L} - g'(x_H, q')\frac{\partial q'}{\partial x_H} + g'(x_L, q')\frac{\partial q'}{\partial x_L} + \frac{\partial g(x_H, q')}{\partial x_L} = 0$. After
simplifying, we have \( \frac{\partial q'}{\partial L} = \frac{-g'(x_L,q)\,\partial x_L}{\theta_H f'(q')-\theta_L f'(q')-g'(x_H,q') + g'(x_L,q')} \). From model setup, \( \frac{\partial g(x,L)}{\partial x_L} > 0 \), and \( \theta_H f'(q')-\theta_L f'(q')-g'(x_H,q')+g'(x_L,q) > 0 \), thus \( \frac{\partial q'}{\partial x_L} < 0 \). Hence we have Lemma 5 (i).

(ii) From Lemma 2 (c), the optimal functionality of the high version is \( q_{vh} = q_{vh}' \), which satisfies the first order condition: \( \alpha \theta_H f'(q_{vh}') - \alpha g'(x_H,q_{vh}) - c'(q_{vh}') = 0 \). Thus, it is clear that optimal functionality of the high version, \( q_{vh} = q_{vh}' \), is independent of \( x_L \). Hence, we have Lemma 5 (ii).

**PROPOSITION 4:** As the required level of functionality of low-type users \((x_L)\) increases, (i) the lower threshold proportion \( \alpha_z \) and the upper threshold proportion \( \alpha_i \) increase, (ii) the optimal functionality of the low version \( q_{vl}' \) (a) increases if proportion of high-type users is relatively low, and (b) decreases if proportion of high-type users is relatively high.

**Proof of Proposition 4:**

(i) By Definition 3, \( \alpha_z = \frac{\theta_z f'(x_L)}{\theta_H f'(x_L)-g'(x_H,x_L)} \). We divide numerator and denominator by \( f'(x_L) \), and we get: \( \alpha_z = \frac{\theta_z}{\theta_H - g'(x_H,x_L)/f'(x_L)} \).

From model setup, we have \( g^* \geq 0, f^* \leq 0 \) and \( g' < 0 \), therefore, as \( x_L \) increases, the term \( -\frac{g'(x_H,x_L)}{f'(x_L)} \) decreases. Thus, \( \alpha_z \) increases.

By Definition 2, at \( \alpha = \alpha_i \), the optimal functionality level of the low version is: \( q_{vl}' = q' \), and the FOC is \( \theta_L f'(q')-g'(x_L,q')-\alpha_i(\theta_H f'(q')-g'(x_H,q')) = 0 \). We take first order derivative of this identity with respect to \( x_L \) and have:

\[
\theta_L f''(q') \frac{\partial q'}{\partial x_L} - g''(x_L,q') \frac{\partial q'}{\partial x_L} - (\theta_H f'(q')-g'(x_H,q')) \frac{\partial \alpha_i}{\partial x_L} - \alpha_i(\theta_H f'(q') - g'(x_H,q')) \frac{\partial q'}{\partial x_L} - g''(x_H,q') \frac{\partial q'}{\partial x_L} - \frac{\partial^2 g(x_L,q = q')}{\partial x_L \partial q} = 0
\]

After simplifying, we have \( \frac{\partial \alpha_i}{\partial x_L} = \frac{\theta_L f''(q') - g''(x_L,q') - \alpha_i \theta_H f'(q') + \alpha_i g''(x_H,q')}{{\theta_H f'(q') - g'(x_H,q')}} \). We know that \( \theta_L f''(q_{vl}') - g''(x_L,q_{vl}') - \alpha \theta_H f'(q_{vl}) + \alpha g''(x_H,q_{vl}'') < 0 \), where \( \alpha = \alpha_i \) and \( q_{vl}' = q' \), therefore,
\[ \theta_l f'(q) - g'(x_l, q) - \alpha \theta_H f'(q) + \alpha g''(x_H, q) < 0. \] From model setup, \( f' > 0 \) and \( g' < 0 \) and
\[ \frac{\partial^2 g(x, q)}{\partial x \partial q} < 0, \] then,
\[ \frac{\partial^2 g(x, q = q)}{\partial x \partial q} < 0 \] and \( \theta_H f'(q) - g'(x_H, q) > 0. \) And since also \( \frac{\partial q'}{\partial x_l} < 0 \) (Lemma 5), therefore, we have \( \frac{\partial \alpha_l}{\partial x_l} > 0. \)

Hence we have Proposition 4 (i).

(ii) (a) We have shown that \( \frac{\partial \alpha_l}{\partial x_l} > 0 \) and \( \frac{\partial \alpha_z}{\partial x_l} > 0 \) in part (i). And where \( \alpha \in (0, \alpha_z) \), the firm’s optimal functionality level of the low version, \( q_{vl} = x_l \), increases as \( x_l \) increases. It implies that the firm will increase the functionality level of the low version if proportion of high-type users is relatively low. Hence we have Proposition 4 (ii) (a).

(ii)(b): When \( \alpha \in [\alpha_z, \alpha_z] \), then the firm set \( q_{vl} = q_{vl}^* \), which satisfies the first order condition:
\[ \theta_l f'(q_{vl}^*) - g'(x_l, q_{vl}^*) - \alpha \theta_H f'(q_{vl}^*) - g'(x_H, q_{vl}^*) = 0, \] and \( q' \leq q_{vl}^* \leq x_l \). Taking its derivative with respect to \( x_l \) and we have:
\[ \frac{\partial q_{vl}^*}{\partial x_l} \left( (\theta_l - \alpha \theta_H) f''(q_{vl}^*) - g''(x_l, q_{vl}^*) + \alpha g''(x_H, q_{vl}^*) \right) = 0. \] After simplifying, we have:
\[ \frac{\partial q_{vl}^*}{\partial x_l} = \frac{\partial^2 g(x, q = q_{vl}^*)}{\partial x \partial q} \left( (\theta_l - \alpha \theta_H) f''(q_{vl}^*) - g''(x_l, q_{vl}^*) + \alpha g''(x_H, q_{vl}^*) \right). \] From model setup,
\[ \frac{\partial^2 g(x, q)}{\partial x \partial q} < 0, \] and \( \theta_l f''(q_{vl}^*) - g''(x_l, q_{vl}^*) - \alpha \theta_H f''(q_{vl}^*) + \alpha g''(x_H, q_{vl}^*) < 0 \) where \( \alpha \in [\alpha_z, \alpha_z] \) and \( q_{vl} = q_{vl}^* \).

Therefore, we have \( \frac{\partial q_{vl}^*}{\partial x_l} > 0 \), where \( \alpha \in [\alpha_z, \alpha_z] \).

We have shown that \( \frac{\partial \alpha_l}{\partial x_l} > 0, \frac{\partial \alpha_z}{\partial x_l} > 0, \) and \( \frac{\partial q'}{\partial x_l} < 0 \) in Lemma 5, and also \( \frac{\partial q_{vl}^*}{\partial x_l} > 0 \) for \( q' \leq q_{vl}^* \leq x_l \).

It implies that the firm provides the low version at higher functionality level, \( q_{vl}^* = x_l \), and starts to reduce functionality level of low version at a greater \( \alpha_z \), and stops reducing functionality level of low version at a greater \( \alpha_z \) at smaller functionality level, \( q_{vl} = q' \). It implies that for a relatively low proportion of high-type users in the market, the firm increases functionality level of low version as the optimal functionality level increases as \( x_l \) increases when \( \alpha < \alpha_{lab} \) at which the functionality qualities of low version are the same before and after the increase in \( x_l \). It also implies that the firm decreases the functionality level of low version if the proportion of high-type users in the market is relatively high, \( \alpha_{lab} < \alpha < 1 \). Hence, we have Proposition 4 (ii) (b).
**PROPOSITION 5:** (1) As the required level of functionality of high-type users \( x_H \) increases, consumer surplus of those high-type users who get information rent decreases (2); As the required level of functionality of low-type users \( x_L \) increases, consumer surplus of those high-type users who get information rent increases.

**Proof of Proposition 5:**

(1) High-type users’ surplus is the information rent paid to them by the firm,

\[
CS_H = U_H(\theta_H, x_H, q_{vl}) - U_L(\theta_L, x_L, q_{vl}).
\]

When \( \alpha \in (0, \alpha_z) \), then \( q_{vl}^* = x_L \), and surplus is \( CS_H = \theta_H f(x_L) - g(x_H, q = x_L) - \theta_L f(x_L) + g(x_L, q = q_{vl}) \). We take first order derivative with respect to \( x_H \), and get \( \frac{\partial CS_H}{\partial x_H} = \frac{\partial g(x_H, q = x_L)}{\partial x_H} \). From model setup,

\[
\frac{\partial g(x_L, q)}{\partial x_H} > 0, \text{ therefore, we have } \frac{\partial CS_H}{\partial x_H} < 0, \text{ for } \alpha \in (0, \alpha_z) \text{ and } q_{vl}^* = x_L.
\]

Where \( \alpha \in [\alpha_z, \alpha_i] \), then \( q_{vl} = q_{vl}^z \), and we showed \( \frac{\partial q_{vl}^z}{\partial x_H} < 0 \) at the beginning of part (ii) Proposition 3.

Then high-type users’ surplus is \( CS_H = \theta_H f(q_{vl}^z) - g(x_H, q = q_{vl}^z) - \theta_L f(q_{vl}^z) + g(x_L, q = q_{vl}) \). We take its first order derivative with respect to \( x_H \), and we get:

\[
\frac{\partial CS_H}{\partial x_H} = \frac{\partial q_{vl}^z}{\partial x_H}(\theta_H f'(q_{vl}^z) - g'(x_H, q = q_{vl}^z)) - \frac{\partial g(x_H, q = q_{vl})}{\partial x_H} - \frac{\partial q_{vl}^z}{\partial x_H}(\theta_L f'(q_{vl}^z) + g'(x_L, q = q_{vl})) + \frac{\partial q_{vl}^z}{\partial x_H} g'(x_L, q = q_{vl}).
\]

After simplifying we have \( \frac{\partial CS_H}{\partial x_H} = \frac{\partial q_{vl}^z}{\partial x_H}(\theta_H f'(q_{vl}^z) - g'(x_H, q = q_{vl}^z) - \theta_L f'(q_{vl}^z) + g'(x_L, q = q_{vl})) - \frac{\partial g(x_H, q)}{\partial x_H}. \)

From model setup, \( \theta_H f'(q_{vl}^z) - g'(x_H, q = q_{vl}^z) - \theta_L f'(q_{vl}^z) + g'(x_L, q = q_{vl}) > 0, \) and \( \frac{\partial g(x_H, q)}{\partial x_H} > 0, \) therefore, we have \( \frac{\partial CS_H}{\partial x_H} < 0, \) when \( \alpha \in [\alpha_z, \alpha_i] \).

When \( \alpha \in (\alpha_i, 1) \), then \( q_{vl} = q' \), and by Lemma 1, the information rent is 0, thus high-type users’ surplus is zero.
Considering $\frac{\partial CS_H}{\partial x_H} < 0$ where $\alpha \in (0, \alpha_1]$ and $CS_H = 0$ where $\alpha \in (\alpha_1, 1)$, and also from Lemma 4 (ii) that $\frac{\partial \alpha_H}{\partial x_H} < 0$ and $\frac{\partial \alpha_L}{\partial x_H} < 0$, the high-type users’ surplus deceases as the required functionality level of high-type users increases. Hence, we have Proposition 5 (1).

(2) We have shown that the high-type users’ surplus is: $CS_H = U_H(\theta_H, x_H, q_{i_H}^*) - U_L(\theta_L, x_L, q_{i_L}^*)$. When $\alpha \in (0, \alpha_1)$, then $q_{i_H}^* = x_L$, and $g(x_L, x_L) = 0$, thus surplus is $CS_H = \theta_H f(x_L) - g(x_L, q = x_L) - \theta_L f(x_L)$. We take the first derivative with respect to $x_L$, and we have:

$\frac{\partial CS_H}{\partial x_L} = (\theta_H - \theta_L) f'(x_L) - g'(x_H, q = x_L)$. From model setup, $f' > 0$ and $g' < 0$, then we have $\frac{\partial CS_H}{\partial x_L} > 0$, for $\alpha \in (0, \alpha_1)$.

When $\alpha \in [\alpha_2, \alpha_1]$, then $q_{i_H}^* = q_{i_L}^*$, and we showed $\frac{\partial q_{i_H}^*}{\partial x_L} > 0$ in Proposition 4. Then high-type users’ surplus is $CS_H = \theta_H f(q_{i_H}^*) - g(x_H, q = q_{i_H}^*) - \theta_L f(q_{i_H}^*) + g(x_L, q = q_{i_L}^*)$, when $\alpha \in [\alpha_2, \alpha_1]$. We take it derivative with respect to $x_L$ and we get:

$\frac{\partial CS_H}{\partial x_L} = \frac{\partial q_{i_H}^*}{\partial x_L} (\theta_H f'(q_{i_H}^*) - g'(x_H, q = q_{i_H}^*) - \theta_L f'(q_{i_H}^*) + g'(x_L, q = q_{i_L}^*)) + \frac{\partial g(x_L, q = q_{i_L}^*)}{\partial x_L}$. From model setup, $\frac{\partial g(x_L, q)}{\partial x_L} > 0$, and also $\theta_H f'(q) - g'(x_H, q) - \theta_L f'(q) + g'(x_L, q) > 0$, and $\frac{\partial q_{i_H}^*}{\partial x_L} > 0$, then we have $\frac{\partial CS_H}{\partial x_L} > 0$, when $\alpha \in [\alpha_2, \alpha_1]$.

When $\alpha \in (\alpha_1, 1)$, then $q_{i_H} = q'$, and by Lemma 1, the information rent is 0, thus the high-type users’ surplus is zero. Considering $\frac{\partial CS_H}{\partial x_L} > 0$ when $\alpha \in (0, \alpha_1]$ and $CS_H = 0$ when $\alpha \in (\alpha_1, 1)$, and also from Proposition 4 that $\frac{\partial \alpha_H}{\partial x_L} > 0$ and $\frac{\partial \alpha_L}{\partial x_L} > 0$, the high-type users’ surplus increases as the required functionality level of low-type users increases. Hence we have Proposition 5 (ii).

PROPOSITION 6: When users do not experience disutility from under-provisioning of functionality, (i) when the proportion of high-type users is small $\alpha < (\theta_H / \theta_L)$, the firm adopts a
versioning strategy; and (ii) when the proportion of high-type users is large $\alpha \geq (\theta_L / \theta_H)$, the firm adopts single version strategy and sells only to the high-type users.

Proof of Proposition 6:

(A) Under the product strategy of selling only to high-type users: the firm will provide $q_H^N$, and price at $p_H^N = \theta_H f(q_H^N)$, and the profit function is: $\pi_H^N = \alpha p_H^N - c(q_H^N) = \alpha \theta_H f(q_H^N) - c(q_H^N)$. Thus, the optimal functionality level is $q_H^N = q_H^N$, which is the solution to $\frac{\partial \pi_H^N}{\partial q_H^N} = \alpha \theta_H f'(q_H^N) - c'(q_H^N) = 0$.

Compared to $q_H^*$ where $\alpha \theta_H f'(q_H^*) - g'(x_H, q_H^*) - c'(q_H^*) = 0$, we have: $q_H^N < q_H^*$, since $g'(x_H, q_H^*) < 0$.

Moreover, from Assumption 2, $x_L < q_H^N \leq x_H$. Also from Assumption 1 that $q' < x_L$, therefore, $p_H^N = U_H(q_H^N) = \theta_H f(q_H^N) > \theta_L f(q_H^N) = U_L(q_H^N)$. It implies high-type users buy, but low-type users do not buy.

Second order condition: since $\frac{\partial^2 f(q)}{\partial q^2} \leq 0$ and $\frac{\partial^2 c(q)}{\partial q^2} > 0$, therefore, $\frac{\partial^2 \pi_H^N}{\partial q_H^N} = \alpha \theta_H f'(q_H^N) - c'(q_H^N) = 0$.

Therefore, $q_H^N = q_H^N$, maximizes the profit function. The optimal profit function is:

$\pi_H^N = \alpha \theta_H f(q_H^N) - c(q_H^N)$, where $\alpha \theta_H f'(q_H^N) - c'(q_H^N) = 0$.

Similar to Lemma 2, we have: $\alpha \theta_H f'(q_H^N) \frac{\partial q_H^N}{\partial \alpha} + \theta_H f'(q_H^N) - c'(q_H^N) \frac{\partial q_H^N}{\partial \alpha} = 0$. By simple algebraic manipulation, we have: $\frac{\partial q_H^N}{\partial \alpha} = \frac{\theta_H f'(q_H^N) - \alpha \theta_H f'(q_H^N) + c'(q_H^N)}{\theta_H f'(q_H^N) - \alpha \theta_H f'(q_H^N) + c'(q_H^N)} > 0$, because $f'' \leq 0$ and $c'' > 0$.

(B) Product strategy of selling to both types of users: from Assumption 1 and 2, and for the similar intuition as in Lemma 2, we have: $q_L^N = x_L$. The firm’s profit function is: $\pi_L^N = \theta_L f(x_L) - c(x_L)$.

(C) Versioning strategy: the prices are set such that high-type buy high version and therefore, demand for high version is $\alpha$ and demand for low version is $1 - \alpha$. Let the functionality and price of high version be $q_H^N$ and $p_H^N$, and that of low version be $q_L^N$ and $p_L^N$. Firm’s profit function in equation (6) can be written as: $\pi_N^N = \alpha p_H^N + (1 - \alpha) p_L^N - c(q_H^N)$. ICs are:

$U_L(\theta_L, x_L, q_L^N) - p_L^N \geq 0$, $U_H(\theta_H, x_H, q_H^N) - p_H^N \geq 0$, ICs are:

$U_L(\theta_L, x_L, q_L^N) - p_L^N \geq U_H(\theta_H, x_H, q_H^N) - p_H^N$, $U_L(\theta_L, x_L, q_L^N) - p_L^N \geq U_L(\theta_L, x_L, q_L^N) - p_L^N$. From Assumption 1 and Assumption 2, $p_L^N = U_L(\theta_L, x_L, q_L^N)$, $p_L^N = U_L(\theta_L, x_L, q_L^N) - U_H(\theta_H, x_H, q_H^N) + U_L(\theta_L, x_L, q_L^N)$. Thus, the profit function
is: $\pi^*_V = \theta_L f(q^*_V) + \alpha(\theta_H f(q^*_H) - \theta_H f(q^*_H)) - c(q^*_H)$. Let $q^*_V = q^*_V$ be the solution to the first order condition: $\frac{\partial \pi^*_V}{\partial q^*_V} = \alpha \theta_H f'(q^*_H) - c'(q^*_H) = 0$. Second order condition is checked:

$$\frac{\partial^2 \pi^*_V}{\partial q^*_V^2} = \alpha \theta_H f''(q) - c''(q) < 0,$$ since $f^* \leq 0$, and $c^* > 0$. Since the first order condition in (A) are identical, we have: $q^*_V = q^*_V > x_L$.

Let $q^*_L = q^*_L$ be the solution to the first order condition: $\frac{\partial \pi^*_L}{\partial q^*_L} = (\theta_L - \alpha \theta_H) f'(q^*_L) = 0$.

If $\alpha > \frac{\theta_L}{\theta_H}$, then $\frac{\partial \pi^*_L}{\partial q^*_L} < 0$, and thus, then the low version is of the lowest possible functionality and

$p^*_L = 0$. If $\alpha < \frac{\theta_L}{\theta_H}$, then $\frac{\partial \pi^*_L}{\partial q^*_L} > 0$, and thus, the optimal is the highest possible value, $q^*_L = x_L$ and

$p^*_L = \theta_L f(x_L)$. If $\alpha = \frac{\theta_L}{\theta_H}$, then $\frac{\partial \pi^*_L}{\partial q^*_L} = 0$, and $\frac{\partial^2 \pi^*_L}{\partial q^*_L^2} = 0$, the firm is indifferent between the optimal functionality to 0 or $x_L$.

Now we compare the optimal profit under the three product strategies.

From (A), the optimal profit of selling only to high-type users: $\pi^*_H = \alpha \theta_H f(q^*_H) - c(q^*_H)$, where

$q^*_H = q^*_H$ satisfies the first order condition: $\alpha \theta_H f'(q^*_H) - c'(q^*_H) = 0$ and $x_L < q^*_H \leq x_H$.

Take the first order derivative of the optimal profit with respect to $\alpha$, and we get:

$$\frac{\partial \pi^*_H}{\partial \alpha} = \theta_H f(q^*_H) + \frac{\partial q^*_H}{\partial \alpha} (\alpha \theta_H f'(q^*_H) - c'(q^*_H)) > 0.$$ It implies that as the proportion of high-type users in the market increases, the optimal profit under a strategy of selling to high-type users increases.

From (B), the optimal profit of selling one version to both types of users is: $\pi^*_A = \theta_L f(x_L) - c(x_L)$

And we have $\frac{\partial \pi^*_A}{\partial \alpha} = 0$. Hence, as proportion of high-type users increases optimal profit under the product strategy of selling to all remains constant.

From (C), the optimal profit under a versioning strategy when $\alpha \in (0, \frac{\theta_L}{\theta_H})$:

$$\pi^*_V = \theta_L f(x_L) + \alpha \theta_H f(q^*_V) - \alpha \theta_H f(x_L) - c(q^*_H) = \theta_L f(x_L) - c(x_L) + [\alpha \theta_H f(q^*_V) - c(q^*_H) - (\alpha \theta_H f(x_L) - c(x_L))].$$

Since $q^*_V > x_L$, and $\frac{\partial q^*_V}{\partial \alpha} > 0$, then, $\alpha \theta_H f(q^*_V) - c(q^*_V) > (\alpha \theta_H f(x_L) - c(x_L)) > 0$. Therefore,

$$\pi^*_V > \theta_L f(x_L) - c(x_L) = \pi^*_A,$$ where $\alpha \in (0, \frac{\theta_L}{\theta_H})$. 

76
Since \( q^{N^*}_{VH} = q^{N^*}_{H} \), and \( \alpha \in (0, \frac{\theta_L}{\theta_H}) \), then:

\[
\pi_{v}^{N^*} = \alpha \theta_H f(q^{N^*}_{VH}) - c(q^{N^*}_{VH}) + \theta_L f(x_L) - \alpha \theta_H f(q^{N^*}_{VH}) - c(q^{N^*}_{VH}) = \alpha \theta_H f(q^{N^*}_{VH}) - c(q^{N^*}_{VH}) = \pi_{H}^{N^*} \]

Therefore,

\[
\pi_{v}^{N^*} > \pi_{A}^{N^*} \quad \text{and} \quad \pi_{v}^{N^*} > \pi_{H}^{N^*} , \quad \text{where} \quad \alpha \in (0, \frac{\theta_L}{\theta_H}) .
\]

The optimal profit under a versioning strategy when \( \alpha \in [\frac{\theta_L}{\theta_H}, 1) \) is: \( \pi_{v}^{N^*} = \alpha \theta_H f(q^{N^*}_{VH}) - c(q^{N^*}_{VH}) \).

Since \( q^{N^*}_{VH} = q^{N^*}_{H} \), then \( \pi_{v}^{N^*} = \alpha \theta_H f(q^{N^*}_{VH}) - c(q^{N^*}_{VH}) = \alpha \theta_H f(q^{N^*}_{H}) - c(q^{N^*}_{H}) = \pi_{H}^{N^*} \).

Also because \( q^* = q^{N^*}_{VH} = q^{N^*}_{H} \), which maximizes \( \alpha \theta_H f(q) - c(q) \) and \( q^{N^*}_{VH} > x_L \), then

\[
\alpha \theta_H f(q^{N^*}_{VH}) - c(q^{N^*}_{VH}) > \alpha \theta_H f(x_L) - c(x_L) .
\]

Thus, when

\[
\alpha \in [\frac{\theta_L}{\theta_H}, 1) , \quad \pi_{v}^{N^*} = \alpha \theta_H f(q^{N^*}_{VH}) - c(q^{N^*}_{VH}) \geq \alpha \theta_H f(x_L) - c(x_L) > \theta_L f(x_L) - c(x_L) = \pi_{A}^{N^*} . \]

Therefore, \( \pi_{v}^{N^*} = \pi_{A}^{N^*} \) and \( \pi_{v}^{N^*} > \pi_{H}^{N^*} , \quad \text{where} \quad \alpha \in [\frac{\theta_L}{\theta_H}, 1) .
\]

Hence we have Proposition 6. ■

**PROPOSITION 7:** When users experience heterogeneous disutility from under-provisioning of functionality in a market consisting of \( N \) user segments, versioning strategy with customized version for each segment is optimal.

**Proof of Proposition 7:**

In the case of market consisting of \( N \) consumer types, the \( n \)th type segment has \( m_n \) consumers who are characterized by \( \{ \theta_n, x_n \} \), \( n \in \{ 1, 2, ..., N \} \). We know that \( \theta_i < \theta_{i+1} \) and \( x_i < x_{i+1} \), for \( i \in \{ 1, 2, ..., N-1 \} \).

We start by study the firm’s product pricing strategies when only two user segments, \( N, N-1 \), are targeted:

1) Sell one version only to the \( N \)-type:

The firm’s profit function is: \( \pi_N = m_N p_N - c(q_N) . \) The firm’s profit maximizing problem is:

\[
\max_{p_N} \pi_N , \quad \text{with subject to:} \quad (IR) \quad U_N(q_N) - p_N \geq 0 , \quad \text{and} \quad U_{N-1}(q_N) - p_N < 0 .
\]

Therefore, \( p_N = U_N(q_N) \) binds, for \( q_N^* \geq q_{N,N-1}^* \). And the profit function is:
\[ \pi_N = m_N U_N(q_N) - c(q_N) . \] We get \( q = q_N^* \) maximizes \( m_N U_N(q) - c(q) \). By assumption in Equation (8) that \( q^* > x_{N-1} \), we get \( q_N^* = q^* > x_{N-1} > q^*_{N-1} \). The optimal profit function is:

\[ \pi_N^* = m_N U_N(q_N^*) - c(q_N^*) . \]

2) Sell one version so both \( N \) and \( N-1 \) type:

The firm’s profit function is: \( \pi_{N-1}^* = (m_N + m_{N-1}) p_{N-1} - c(q_{N-1}) \). The firm’s profit maximizing problem is:

\[
\max_{q_{N-1}} \pi_{N-1}^* \text{, with subject to: (IR) } U_N(q_{N-1}) - p_{N-1} \geq 0 \text{, and } U_{N-1}(q_{N-1}) - p_{N-1} \geq 0 .
\]

The profit function is: \( \pi_{N-1}^* = (m_N + m_{N-1}) U_{N-1}(q_{N-1}) - c(q_{N-1}) \), as \( p_{N-1} = U_{N-1}(q_{N-1}) \) binds. We get \( q = q_{N-1}^* \) which maximizes \( (m_N + m_{N-1}) U_{N-1}(q) - c(q) \), and since \( q^* \geq x_{N-1} \), \( U_{N-1}(q) \) does not increase for any \( q > x_{N-1} \). Therefore, we have \( q_{N-1}^* = x_{N-1} \), and since the extend assumptions in Equation (6) and (7), \( U_N(x_{N-1}) \geq p_{N-1} = U_{N-1}(x_{N-1}) \). The optimal profit is:

\[ \pi_{N-1}^* = (m_N + m_{N-1}) U_{N-1}(x_{N-1}) - c(x_{N-1}) . \]

3) Sell two versions, high version to \( N \) type and low version to \( N-1 \) type:

The firm’s profit function is similar to the versioning strategy in the earlier analysis when we consider high and low-type users, but with additional constraints. The profit function is: \( \pi_{V,(N,N-1)} = m_N p_{V,N} + m_{N-1} p_{V,N-1} - c(q_{V,N}) \). The firm’s maximizing problem is:

\[
\max_{q_{V,N},q_{V,N-1}} \pi_{V,(N,N-1)} \text{, subject to constraints:}
\]

For \( N \) type:

\[
(\text{IR}) \ U_N(q_{V,N}) - p_{V,N} \geq 0 \text{, and (IC) } U_N(q_{V,N}) - p_{V,N} \geq U_N(q_{V,N-1}) - p_{V,N-1}.
\]

For \( N-1 \) type:

\[
(\text{IR}) \ U_{N-1}(q_{V,N-1}) - p_{V,N-1} \geq 0 \text{, and (IC) } U_{N-1}(q_{V,N-1}) - p_{V,N-1} \geq U_{N-1}(q_{V,N}) - p_{V,N}.
\]

From the previous analysis on the versioning strategy for high and low-type users, the (IR) for the \( N-1 \) type and the (IC) for the \( N \) type bind. We obtain \( q_{V,N}^* \) maximizes \( m_N U_N(q) - c(q) \), thus, \( q_{V,N}^* = q_N^* \); \( q_{V,N-1}^* \) maximizes \( m_N(U_N(q) - U_N(q)) + m_{N-1}U_{N-1}(q) \) and \( q_{V,N-1}^* \in [q_{V,N-1}^*, x_{N-1}] \). The optimal profit is:

\[ \pi_{V,(N,N-1)}^* = m_N(U_N(q_N^*) - U_N(q_{V,N-1}^*)) + m_{N-1}U_{N-1}(q_{V,N-1}^*) - c(q_N^*) \]

From Proposition 2, we know that \( \pi_{V,(N,N-1)}^* > \pi_N^* \), and \( \pi_{V,(N,N-1)}^* > \pi_{N-1}^* \).
Now we consider the case where the firm considers targeting to another user segment, $N-2$. Therefore, the firm has six possible product-price strategies: (1) sell one version so that only $N$-type buys, (2) sell one version so that only $N$- and $N-I$- types buy, but $N-2$-type do not buy, (3) sell two versions so that $N$-type buy the high version, $N-I$-type buy the low version, but $N-2$-type do not buy, (4) sell one version so that $N$, $N-I$, and $N-2$- types buy, (5) sell two versions so that $N$- and $N-I$- type buy the high version, and $N-2$- type buy the low version, and (6) sell three versions so that $N$-type buy the high version, $N-I$-type buy the middle version and $N-2$-type buy the low version. Note that because assumption in Equation (6), (7) and (8), there is no change in the firm’s optimal product-pricing strategy in (1), (2) and (3) that are discussed previously in the case where the firm only targets to $N$ and $N-I$-type users, and we know that (3) is the optimal strategy. Now we just need to compare the profitability of strategies (3), (4), (5), and (6) to derive the optimal strategy for the firm. The strategies (4), (5) and (6) are:

4) Sell one version to all $N$, $N-I$, and $N-2$ type:

Similar to the previous strategy which sells one version to $N$ and $N-I$, we have the optimal profit:

$$\pi_{N-2}^* = (m_N + m_{N-1} + m_{N-2})U_{N-2}(x_{N-2}) - c(x_{N-2}).$$

5) Sell two versions, high version to $N$ and $N-I$ type, and low version to $N-2$ type:

The firm’s profit function is similar to the versioning strategy in the previous strategy. The profit function is:

$$\pi_{v,(N,N-2)} = (m_N + m_{N-1})p_{v,N-1} + m_{N-2}p_{v,N-2} - c(q_{v,N-1}).$$

The firm’s maximizing problem is:

$$\max_{q_{v,N-1}, q_{v,N-2}} \pi_{v,(N-1,N-2)}, \text{ subject to constraints:}$$

For $N$ type:

(IR) $U_N(q_{v,N-1}) - p_{v,N-1} \geq 0$, and (IC) $U_N(q_{v,N-1}) - p_{v,N-1} \geq U_N(q_{v,N-2}) - p_{v,N-2}$

For $N-1$ type:

(IR) $U_{N-1}(q_{v,N-1}) - p_{v,N-1} \geq 0$, and (IC) $U_{N-1}(q_{v,N-1}) - p_{v,N-1} \geq U_{N-1}(q_{v,N-2}) - p_{v,N-2}$

For $N-2$ type:

(IR) $U_{N-2}(q_{v,N-2}) - p_{v,N-2} \geq 0$, and (IC) $U_{N-2}(q_{v,N-2}) - p_{v,N-2} < U_{N-2}(q_{v,N-1}) - p_{v,N-1}$
From assumptions in Equation (6), (7) and (8), and the previous analysis on the versioning strategy for high and low-type users, both the (IR) for the \(N\)-2 type and the (IC) for the \(N\)-1 type bind.

We obtain that \(q^*_v, N = 1\) maximizes \((m_N + m_{N-1})U_{N-1}(q) - c(q)\), and since \(q^*_v > x_{N-1}\), and also, \(U_{N-1}(q)\) is constant for \(q > x_{N-1}\), therefore, \(q^*_v = x_{N-1}\); and \(q = q^*_v, N - 2\) maximize \((m_N + m_{N-1})(U_{N-2}(q) - U_{N-1}(q)) + m_{N-2}U_{N-2}(q)\), and \(q^*_v, N - 2 \in [q_{1, N-2}, x_{N-2}]\). Since \(q^*_v, N = x_{N-1} > q^*_v, N - 1\), (IR) for \(N\)-1 type and the (IR) for the \(N\) type is satisfied; and since (IC) for the \(N\)-1 type binds, the (IC) for the \(N\) type is also satisfied. The optimal profit is:

\[
\pi_{v, (N, N-1, N-2)}^* = (m_N + m_{N-1})(U_{N-1}(x_{N-1}) - U_{N-1}(q^*_v, N - 2) + U_{N-2}(q^*_v, N - 2)) + m_{N-2}U_{N-2}(q^*_v, N - 2) - c(x_{N-1})
\]

6) Sell three versions, high version to \(N\) type, middle version to \(N\)-1 type, and low version to \(N\)-2 type.

The profit function is: \(\pi_{v, (N, N-1, N-2)} = m_N \tilde{p}_{v, N} + m_{N-1} \tilde{p}_{v, N-1} + m_{N-2} \tilde{p}_{v, N-2} - c(q_v, N)\). The firm’s maximizing problem is:

\[
\max_{q_v, N, \tilde{q}_{v, N-1}, \tilde{q}_{v, N-2}} \pi_{v, (N, N-1, N-2)}, \text{ subject to constraints}
\]

For \(N\) type: (IR) \(U_N(q^*_v, N) - \tilde{p}_{v, N} \geq 0\),
(IC) \(U_N(q^*_v, N) - \tilde{p}_{v, N} \geq U_N(q^*_v, N - 1) - \tilde{p}_{v, N - 1}\),
(IC) \(U_N(q^*_v, N) - \tilde{p}_{v, N} \geq U_N(q^*_v, N - 2) - \tilde{p}_{v, N - 2}\);

For \(N\)-1 type: (IR) \(U_{N-1}(q^*_v, N - 1) - \tilde{p}_{v, N - 1} \geq 0\),
(IC) \(U_{N-1}(q^*_v, N - 1) - \tilde{p}_{v, N - 1} \geq U_{N-1}(q^*_v, N) - \tilde{p}_{v, N}\),
(IC) \(U_{N-1}(q^*_v, N - 1) - \tilde{p}_{v, N - 1} \geq U_{N-1}(q^*_v, N - 2) - \tilde{p}_{v, N - 2}\);

For \(N\)-2 type: (IR) \(U_{N-2}(q^*_v, N - 2) - \tilde{p}_{v, N - 2} \geq 0\),
(IC) \(U_{N-2}(q^*_v, N - 2) - \tilde{p}_{v, N - 2} \geq U_{N-2}(q^*_v, N) - \tilde{p}_{v, N}\),
(IC) \(U_{N-2}(q^*_v, N - 2) - \tilde{p}_{v, N - 2} \geq U_{N-2}(q^*_v, N - 1) - \tilde{p}_{v, N - 1}\).

From Assumption (A) and (B), after analysis, the (IR) for the \(N\)-2 type binds, the 2\textsuperscript{nd} (IC) for the \(N\)-1 type binds, and the 1\textsuperscript{st} (IC) for the \(N\) type binds, and other constrains are also satisfied, for \(q^*_v, N, N - 1 \geq q^*_v, N, N - 2\). Thus, the profit function is:
\[ \pi_{v,(N=1, N-1)} = m_{q,N}(U_{q,N} - U_{q,N}) + m_{n,N}(U_{n,N} - U_{n,N}) + m_{x,N}U_{N-1}(q_{v,N-1}) - U_{N-1}(q_{v,N}) \]

We obtain that \( q_{v,N} \) maximizes \( m_{q,N}U_{q,N} - c(q) \), thus, \( q_{v,N} = q_{v,N} = q_{N} > x_{N-1} \); \( q_{v,N-1} \) maximizes \( m_{n,N}(U_{n,N}(q) - U_{n,N}(q)) + m_{n,N}U_{N-1}(q) \), thus, \( q_{v,N-1} = q_{v,N-1} \) and \( q_{v,N-1} \in [q_{v,N-1}, x_{N-1}] \); \( q_{v,N-2} \) maximizes \( m_{n,N}(U_{n,N}(q) - U_{n,N}(q)) + m_{n,N}U_{N-2}(q) \), thus, \( q_{v,N-2} = q_{v,N-2} \) and \( q_{v,N-2} \in [q_{v,N-2}, x_{N-2}] \). Therefore, all constraints are satisfied. The optimal profit is:

\[
\pi_{v,(N=1, N-1)} = m_{q,N}(U_{q,N} - U_{q,N}) + m_{n,N}(U_{n,N} - U_{n,N}) + m_{x,N}U_{N-1}(q_{v,N-1}) - U_{N-1}(q_{v,N})
\]

Now, we can compare profitability of these strategies.

We already know that \( \pi_{v,(N=1)} > \pi_{v,N} \) and \( \pi_{v,(N-1)} > \pi_{v,N} \). And from Proposition 2, we know that the firm gets higher profitability by selling two versions to target \( N-I \) and \( N-2 \)-types of users than selling single version to both types of users. Moreover, the \( N \)-type users will buy \( q_{v,N-1} \) at the price \( U_{N-1}(x_{N-1}) - U_{N-1}(q_{v,N-1}) \) under the strategy (5) or buy \( x_{N-2} \) at \( U_{N-2}(x_{N-2}) \) under the strategy (4), therefore, the firm generates more revenue from selling \( q_{v,N-1} \) than selling \( x_{N-2} \) to \( N \)-type users. Thus, \( \pi_{v,(N=1, N-1)} > \pi_{v,N-2} \).

To compare the optimal profits, \( \pi_{v,(N=1, N-1)} \) (strategy (3)) and \( \pi_{v,(N=1, N-1)} \) (strategy (6)):

\[ \pi_{v,(N=1, N-1)} - \pi_{v,(N=1)} = (m_{x,N}U_{X,N} - U_{X,N}) + m_{n,N}U_{N-1}(q_{v,N-1} - U_{N-1}(q_{v,N-1})) + m_{n,N}U_{N-2}(q_{v,N-2} - U_{N-2}(q_{v,N-2})) \]

and we also show that \( q = q_{v,N-2} \) maximizes \( m_{n,N}(U_{n,N}(q) - U_{n,N}(q)) + m_{n,N}U_{N-2}(q) \), and \( q_{v,N-2} \in [q_{v,N-2}, x_{N-2}] \). At \( q = q_{v,N-2} \), \( U_{N-2}(q) - U_{N-1}(q) = 0 \) and \( \pi_{v,(N=1, N-1)} - \pi_{v,(N=1)} = U_{N-2}(q_{v,N-2}) > 0 \), therefore, \( \pi_{v,(N=1, N-1)} > \pi_{v,(N=1)} \).

To compare the optimal profits, \( \pi_{v,(N=1, N-2)} \) (strategy (5)) and \( \pi_{v,(N=1, N-2)} \) (strategy (6)):

\[ \pi_{v,(N=1, N-2)} - \pi_{v,(N=1, N-2)} = m_{n,N}(U_{N}(q_{N}) + U_{N-1}(q_{v,N-1}) - U_{v,N-1}(q_{v,N-1})) + m_{n,N}U_{N-1}(q_{v,N-1}) - c(q_{N}) \]

As \( \pi_{v,(N=1, N-2)} = m_{n,N}(U_{N}(q_{N}) + U_{N-1}(q_{v,N-1}) - U_{v,N-1}(q_{v,N-1})) + m_{n,N}U_{N-1}(q_{v,N-1}) - c(q_{N}) \), \( \pi_{N-1} = (m_{n,N}U_{N-1}(x_{N-1}) - c(x_{N-1})) \), and we also show that \( \pi_{v,(N=1, N-2)} > \pi_{N-1} \), thus, \( m_{n,N}(U_{N}(q_{N}) + U_{N-1}(q_{v,N-1}) - U_{v,N-1}(q_{v,N-1})) + m_{n,N}U_{N-1}(q_{v,N-1}) - c(q_{N}) > (m_{n,N}U_{N-1}(x_{N-1}) - c(x_{N-1})) \). Therefore, we have \( \pi_{v,(N=1, N-2)} > \pi_{v,(N=1, N-2)} \).
We show $\pi_{V,N,N-1,N-2}^* > \pi_{V,N-1,N-1,N-2}^*$, $\pi_{V,N,N-1,N-3}^* > \pi_{V,N-1,N-2,N-3}^*$, $\pi_{V,N,N-4}^* > \pi_N$, $\pi_{V,N-1,N-4}^* > \pi_{N-1,N-4}^*$, and $\pi_{V,N-1,N-2,N-3} > \pi_{N-2,N-3}^*$, thus, the firm’s optimal strategy is to offer three customized versions for each type of users. And we can also extend these results by considering firm’s offering a new customized version to target a lower ordered user type each time, as it generates more revenue from the new user segment than loss from sequentially higher segments. It implies that versioning strategy of selling $N$ versions to $N$ types of users is the optimal strategy for the firm. Hence, we prove that versioning strategy is always optimal for $N$ types of users in the market.

**Proof for Subsection 5.4:**

In this proof, we discuss the optimality of versioning strategy when the rank order is not maintained. This means that the required level of functionality of high-type users is not higher than the required level of functionality of low-type users, that is $\theta_H > \theta_L$ but $x_H \leq x_L$.

1. When $x_H = x_L$, high-type and low-type users incur the same disutility for any given functionality level, $g(x_L, q) = g(x_H, q)$.

When the software firm sells only to high-type users, the firm will set the functionality at $q_H = x_L$ (Assumption 2), and price $p_H = U_H(x_L) = \theta_H f(x_L)$, and the profit is $\pi_H = \alpha p_H - c(x_L)$. Since $x_H = x_L$, then $\pi_H = \alpha \theta_H f(x_H) - c(x_H)(9)$.

When the software firm sells one version to all users, the optimal functionality is $q_A = x_L$, and the price $p_A = U_L(q_L)$, and the profit is $\pi_A = p_A - c(q_A) = \theta_L f(x_L) - c(x_L)(10)$.

When the software firm adopts versioning strategy, $p_{VH} = U_H(q_{VH}) - (U_H(q_{VL}) - U_L(q_{VL}))$ and $p_{VL} = U_L(q_{VL})$, and the profit is:

$$\pi_V = \alpha p_{VH} + (1 - \alpha)p_{VL} - c(q_{VH})$$
$$= \alpha(U_H(q_{VH}) - (U_H(q_{VL}) - U_L(q_{VL}))) + (1 - \alpha)U_L(q_{VL}) - c(q_{VH})$$
$$= \alpha U_H(q_{VH}) + U_L(q_{VL}) - \alpha U_H(q_{VL}) - c(q_{VH})$$

The firm will set $q_{VH} = x_L$. Also since $x_H = x_L$, and $g(x_L, q) = g(x_H, q)$, then, the profit can be rewritten as:
\[ \pi_V = \alpha \theta_H f(x_H) + (\theta_L - \alpha \theta_H) f(q_{VL}) - (1 - \alpha) g(x_L, q_{VL}) - c(x_H). \]

Then, we have the first derivative of profit function of versioning strategy with respect to functionality of low version:

\[ \frac{d\pi_V}{dq_{VL}} = (\theta_L - \alpha \theta_H) f'(q_{VL}) - (1 - \alpha) g'(x_L, q_{VL}). \]

When \( \alpha < \frac{\theta_L}{\theta_H} \), then \( \frac{d\pi_V}{dq_{VL}} > 0 \), which implies \( q_{VL}^* = x_L \). It also means that the versioning strategy converges to the strategy of selling one version to all users when \( \alpha < \frac{\theta_L}{\theta_H} \).

When \( \alpha \geq \frac{\theta_L}{\theta_H} \), then \( \frac{d\pi_V}{dq_{VL}} \leq 0 \), and assume \( q_{VL} = q_{VL}^* \) is the solution to the first order condition: \((\theta_L - \alpha \theta_H) f'(q_{VL}) - (1 - \alpha) g'(x_L, q_{VL}) = 0\), and \( q_{VL}^* \in [0, x_L] \).

Hence, when \( \alpha \geq \frac{\theta_L}{\theta_H} \), then \( (\theta_L - \alpha \theta_H) \leq 0 \), and from (9), it is easy to see:

\[ \pi_V \leq \alpha \theta_H f(x_H) - (1 - \alpha) g(x_L, q_{VL}^*) - c(x_H) < \alpha \theta_H f(x_H) - c(x_H) = \pi_H. \]

Hence, the profit of versioning strategy is lower than the strategy of selling one version only to high-type users, when \( \alpha \geq \frac{\theta_L}{\theta_H} \).

Therefore, versioning is never optimal when \( x_H = x_L \).

2. Let’s consider the case where \( x_H < x_L \), and there are three sub-cases, (a) \( \theta_H f(x_H) > \theta_L f(x_L) \), (b) \( \theta_H f(x_H) = \theta_L f(x_L) \), and (c) \( \theta_H f(x_H) < \theta_L f(x_L) \).

In case (a) \( x_H < x_L \) and \( \theta_H f(x_H) > \theta_L f(x_L) \):

The firm sells only to high-type users with \( q_H = x_L \) (Assumption 2) and sets the price \( p_H = U_H f(x_L) \). The profit is: \( \pi_H = \alpha U_H(x_L) - c(x_L) = \alpha \theta_H f(x_H) - c(x_L) \) (11).
When the software firm sells to all users, then the firm offers \( q_A = x_L \), and price \( p_A = U_L(x_L) = \theta_L f(x_L) \), and the profit is \( \pi_A = p_A - c(q_A) = \theta_L f(x_L) - c(x_L) \) (12).

When the software firm adopts a versioning strategy, \( p_{vH} = U_H(q_{vH}) - (U_H(q_{vL}) - U_L(q_{vL})) \) and \( p_{vL} = U_L(q_{vL}) \). Since \( U_H(q) > U_L(q) \), the firm sells \( q_{vH} \) to high-type users and \( q_{vL} \) to low-type users, and the firm’s profit is:

\[
\pi_V = \alpha p_{vH} + (1 - \alpha)p_{vL} - c(q_{vH}) \\
= \alpha(U_H(q_{vH}) - (U_H(q_{vL}) - U_L(q_{vL}))) + (1 - \alpha)U_L(q_{vL}) - c(q_{vH}) \\
= \alpha U_H(q_{vH}) + U_L(q_{vL}) - \alpha U_H(q_{vL}) - c(q_{vH})
\]

The firm sets \( q_{vH} = x_L \), and \( q_{vL} < x_L \).

Since high-type users do not incur disutility if \( q > x_H \), we consider two cases here. The profit of versioning strategy:

\[
\pi_V = \theta_L f(q_{vL}) - g(x_L, q_{vL}) - c(x_L) \text{ (13), when } q_{vL} \in (x_H, x_L) \text{; or} \\
\pi_V = \alpha \theta_H f(x_H) + \theta_L f(q_{vL}) - g(x_L, q_{vL}) - \alpha(\theta_H f(q_{vL}) - g(x_H, q_{vL})) - c(x_L) \text{ (14), when } q_{vL} \in [0, x_H].
\]

When \( q_{vL} \in (x_H, x_L) \), from (13) and (12), it is easy to see that \( \pi_V = \theta_L f(q_{vL}) - g(x_L, q_{vL}) - c(x_L) < \theta_L f(x_L) - c(x_L) = \pi_A \).

Hence, it is never optimal for the firm to offer low version at a functionality level in the range \( q_{vL} \in [x_H, x_L] \).

When \( q_{vL} \in [0, x_H] \), taking the first derivative with the profit (14) with respect to functionality level of low version, we have:

\[
\frac{d\pi_V}{dq_{vL}} = (\theta_L - \alpha \theta_H) f'(q_{vL}) - g'(x_L, q_{vL}) + \alpha g'(x_H, q_{vL}).
\]

When \( \alpha \leq \frac{\theta_L}{\theta_H} \), then \( \frac{d\pi_V}{dq_{vL}} > 0 \), since \( \alpha g'(x_H, q_{vL}) > g'(x_L, q_{vL}) \). It implies, \( q_{vL} = x_H \), and from (14)

\[
\pi_V = \theta_L f(x_H) - g(x_L, x_H) - c(x_L). \text{ Hence, we can compare it with (12),}
\]
\[ \pi_V < \theta_L f(x_L) - c(x_L) = \pi_A. \]
It implies that profit of versioning strategy is lower than the profit of selling one version to all users when \( \alpha \leq \frac{\theta_L}{\theta_H}. \)

When \( \alpha > \frac{\theta_L}{\theta_H}, \) then \( \frac{d\pi_v}{dq_{VL}} \leq 0. \) Assume \( q_{VL} = q_{VL}^* \) is the solution to the first order condition:

\[
(\theta_L - \alpha \theta_H) f'(q_{VL}) - g'(x_L, q_{VL}) + \alpha g(x_L, q_{VL}) = 0, \quad \text{and} \quad q_{VL}^* \in [0, x_H].
\]
Hence, we can rewrite the profit function from (14) as,

\[
\pi_V = \alpha \theta_H f(x_H) - c(x_L) + \theta_L f(q_{VL}^*) - g(x_L, q_{VL}^*) - \alpha (\theta_H f(q_{VL}^*) - g(x_H, q_{VL}^*)).
\]
Also because, when \( \alpha \geq \frac{\theta_L}{\theta_H}, \) then \( \theta_L f(q_{VL}^*) - g(x_L, q_{VL}^*) - \alpha (\theta_H f(q_{VL}^*) - g(x_H, q_{VL}^*)) < 0. \) Hence, we can compare it with (11), and \( \pi_V < \alpha \theta_H f(x_H) - c(x_L) = \pi_H. \) This implies that profit of versioning strategy is lower than the strategy of selling only to high-type users when \( \alpha > \frac{\theta_L}{\theta_H}. \)

Therefore, versioning is never optimal in case (a).

In case (b) \( x_H < x_L \) and \( \theta_H f(x_H) = \theta_L f(x_L): \)

When the software firm sells one version, then the firm offers \( q_A = x_L \) to all users, and price \( p_A = U_L(x_L) = \theta_L f(x_L), \) and the profit is \( \pi_A = p_A - c(q_A) = \theta_L f(x_L) - c(x_L) \) (15).

When the firm sells two versions, it is similar to the case (a), and we have \( q_{VH} = x_L, \quad q_{VL} \in [0, x_H], \) and the profit function is:

\[ \pi_V = \alpha \theta_H f(x_H) + \theta_L f(q_{VL}) - g(x_L, q_{VL}) - \alpha (\theta_H f(q_{VL}) - g(x_H, q_{VL})) - c(x_L) \] (16).

When \( \alpha \leq \frac{\theta_L}{\theta_H}, \) then \( \frac{d\pi_v}{dq_{VL}} > 0, \) which implies, \( q_{VL} = x_H. \) Hence, by comparing (15) and (16), we have \( \pi_V < \theta_L f(x_L) - c(x_L) = \pi_A. \) This implies that profit of versioning strategy is lower than the strategy of selling one version to all users when \( \alpha \leq \frac{\theta_L}{\theta_H}. \)
When \( \alpha > \frac{\theta_L}{\theta_H} \), then \( \frac{d\pi_V}{dq_V} \leq 0 \). We assume \( q^*_V \in [0, x_H] \), and we have \( \pi_V = \alpha \theta_H f(x_H) - c(x_L) + \theta_L f(q^*_V) - g(x_L, q^*_V) - \alpha(\theta_H f(q^*_V) - g(x_H, q^*_V)) \). By comparing it with (15),
\[
\pi_V < \alpha \theta_H f(x_H) - c(x_L) = \alpha \theta_L f(x_L) - c(x_L) < \theta_L f(x_L) - c(x_L) = \pi_A.
\]
This implies that profit of versioning is lower than the strategy of selling one version to all users when \( \alpha > \frac{\theta_L}{\theta_H} \).

Therefore, versioning is never optimal in case (b).

In the case (c) \( x_H < x_L \), and \( \theta_H f(x_H) < \theta_L f(x_L) \):

The strategy of selling one version to one type users is that the firm will choose \( q_H = x_L \) (Assumption 2), and sell only to low-type users, and \( \pi_L = (1 - \alpha)\theta_L f(x_L) - c(x_L) \) (17).

The strategy of selling one version to all users is the firm will set \( q_A = x_L \) and \( \pi_A = \theta_H f(x_L) - c(x_L) \) (18).

When the software firm adopts a versioning strategy, since \( \theta_H f(x_H) < \theta_L f(x_L) \), the firm sells \( q_{VH} \) to low-type users and \( q_{VL} \) to high-type users, \( p_{VH} = U_L(q_{VH}) - (U_L(q_{HL}) - U_H(q_{VL})) \) and \( p_{VL} = U_H(q_{VL}) \), and the firm’s profit is:
\[
\pi_V = \alpha p_{VL} + (1 - \alpha)p_{VH} - c(q_{VH})
\]
\[
= (1 - \alpha)(U_L(q_{VH}) - (U_L(q_{VL}) - U_H(q_{VL}))) + \alpha U_H(q_{VL}) - c(q_{VH})
\]
\[
= (1 - \alpha)U_L(q_{VH}) + U_H(q_{VL}) - (1 - \alpha)U_L(q_{VL}) - c(q_{VH})
\]
Thus, the firm will set \( q_{VH} = x_L \), and \( q_{VL} = q' \). And since \( q_{VL} = q' > x_H \), we have:
\[
\pi_V = (1 - \alpha)U_L(q_{VH}) + U_H(q_{VL}) - (1 - \alpha)U_L(q_{VL}) - c(q_{VH})
\]
\[
= (1 - \alpha)\theta_L f(x_L) + \alpha \theta_H f(q') - c(x_L)
\]
(19)

Hence, by comparing (17) and (19), we have \( \pi_V > \pi_L = (1 - \alpha)\theta_L f(x_L) - c(x_L) \). This implies that profit of versioning strategy is higher than the profit of offering one version only to the low-type users.
Also since $\theta_H f(x_H') = \theta_H f(x_L)$, and $\theta_L f(x_L') > \theta_H f(x_H)$, by comparing (18) and (19), we have $\pi_V - \pi_A = (1 - \alpha)\theta_L f(x_L) + \alpha \theta_H f(x_H) - \theta_H f(x_L) = (1 - \alpha)(\theta_L f(x_L) - \theta_H f(x_H)) > 0$. This implies that profit of versioning strategy is higher than the profit of offering one version to all users.

Therefore, versioning strategy is optimal in the case (c).
CHAPTER 2

Economics of Online Daily-deal Business Model

Abstract

Daily-deal publishing is a type of performance-based advertising model wherein a publisher provides advertising space to a merchant who pays a proportion of revenue generated on the website to the publisher. We develop a two-period model to capture the strategic interaction between a publisher and a merchant where consumers are heterogeneous in their willingness to pay for quality. In our Stackelberg game, the daily-deal publisher is the leader who decides the revenue sharing ratio and the merchant is the follower who responds by deciding whether to offer a deal on the publisher’s website and choosing an appropriate discount rate. Our analysis shows that a merchant’s optimal discount rate strategy and participation decision depends on the trade-offs between four effects – advertising, sampling, cannibalization and revenue sharing. Different from the literature on performance-based advertising, we show that the impact of the publisher’s revenue sharing ratio on a merchant’s optimal discount rate is non-monotonic. Counter-intuitively, a merchant increases the discount rate (decreases price) on the daily-deal website when the publisher increases the revenue sharing ratio under some conditions. Moreover, we recommend that a daily-deal publisher should take into account the merchant’s marginal cost, proportion of informed consumers and consumer characteristics in formulating appropriate customized revenue sharing contract for each merchant.
1. Introduction

A daily-deal publisher provides a platform for local businesses or merchants who offer services, like restaurants, yoga classes, and tanning salons to advertise to online users by offering deals. Moreover, given that 58% of consumers in the US research products and services online before purchase (Jansen, 2010), it is becoming increasingly critical for local brick-and-mortar businesses like restaurants to advertise their offerings on daily-deal publisher’s mobile platforms. A merchant offers price discount as “daily-deal” through a daily-deal publisher to encourage first-time consumers to try the good, and some of these first-time consumers may become repeat consumers in future. U.S. consumers’ spending on daily-deals is expected to grow from $873 million in 2010 to $4.2 billion in 2015 (BIA/Kelsey, 2011).

The popularity of daily-deal publishing is also fueled by the rapid emergence of technologies enabling Location Based Service (LBS) and Social, Local, Mobile (SOLOMO) services. For example, Groupon claims that its deals are inherently local and their distinguishing feature is that they are designed for mobile commerce platforms. The mobile commerce is expected to grow to $31 Billion in 2016 from a level of around $3 Billion in 2010 (Forrester, 2011). This presents unprecedented opportunities for the local businesses to connect to consumers who are online 24x7 through their mobile devices.

What differentiates a daily-deal business model from the traditional ecommerce is that though buyers purchase goods online, they receive or consume them offline. Therefore, this

---


14 A merchant may also consider offering a deal on the daily-deal website due to competitive consideration, and/or due to operational reasons, such as capacity utilization/planning, and demand management. In order to keep our analysis tractable, we do not consider competitive or operational dimensions of daily-deals in this paper.

15 See Groupon CEO interview on CNBC on 10/23/2013: [http://www.cnbc.com/id/101136815](http://www.cnbc.com/id/101136815)

business model is essentially based on first attracting consumers online and then directing them to offline local stores. Since service is an experience good, some consumers who have not experienced the good before may not know the quality of the good, and may have a lower quality estimate of the good (Shapiro, 1983). Therefore, a merchant may offer price discount to incentivize consumers with low quality estimate to buy (Bagwell, 1990).

Another notable aspect of daily-deal business model is that consumers complete the purchases of daily-deals on the website\textsuperscript{17}. This means that a publisher, like Groupon or Living Social, can monitor revenue generated, and, hence, can offer a revenue sharing contract to merchants in which a merchant pays a proportion of revenue (revenue sharing ratio) to the daily-deal publisher for the advertising services only when a transaction has been completed. In that sense the daily-deal business model can be considered as a performance-based online advertising model. Note that a revenue sharing scheme may not be viable in the case of traditional advertising because it is almost impossible for the publisher to track revenue impact of advertising stimulus.

One critical feature of daily-deal business model is that a publisher often offers a standard revenue sharing scheme to all merchants\textsuperscript{18}. However, it is not clear if offering one revenue sharing ratio is optimal for the publisher given that merchants are often heterogeneous in term of marginal costs, market size, and first-time consumers’ quality perception. Furthermore, there is little information available in industry reports which provide insights about how do heterogeneous merchants decide discount rate given a revenue sharing ratio, if they choose to participate. Even though the daily-deal advertising model is becoming increasingly important for

\textsuperscript{17} A daily-deal publisher can reach consumers via different outlets, like website and mobile app.

\textsuperscript{18} For example, the standard revenue sharing ratio offered by Groupon and LivingSocial is 50%, and 40% respectively (Patel, 2011).
offline merchants who provide services, there is very little academic research relating to the ecosystem of daily-deal business model.

In this paper, we develop a two-period analytical model to abstract the ecosystem of daily-deal publishing where a participating merchant sells an experience good at the regular price directly to consumers or in the direct channel, and offers deals on the publisher’s website or in the publisher channel. Consumers have heterogeneous quality estimate of the experience good in the publisher channel, and decide whether to buy from the merchant in each period. We model the strategic interaction between a daily-deal publisher and an advertising merchant as a leader-follower game where the daily-deal publisher is the leader and the merchant is the follower. More specifically, we answer the following questions: (i) What are the benefits and costs to a merchant in offering a deal on a daily-deal publisher’s website? (ii) What is the optimal discount rate strategy of a merchant? (iii) How does a daily-deal publisher’s revenue sharing ratio affect a merchant’s optimal discount rate strategy? (iv) Which merchant type should offer deal on a daily-deal publisher’s website? and (v) What is the optimal revenue sharing scheme of a daily-deal publisher?

In the context of daily-deal business model, a merchant faces a market consisting of two consumer segments, namely, *informed consumers* and *uninformed consumers*. While the *informed consumers* are aware about the merchant, the *uninformed consumers* are not aware about the merchant. The daily-deal website attracts both these types of consumers, wherein *uninformed consumers* who are exposed to a merchant’s offerings for the first time on the daily-deal website underestimate the quality of the experience good. This setting has some similarity to two-period model of Shapiro (1983), wherein pessimistic consumers are exposed to the merchant
offerings for the first time, and underestimate the quality of the experience good in the first period.

A local merchant can reach out to these uninformed consumers on the publisher’s website by offering a price discount, as known as daily-deal. Exposure to these uninformed consumers is the advertising effect of daily-deal publishing to the merchant. Moreover, uninformed consumers on the publisher’s website are incentivized to purchase the merchant’s experience good through an attractive deal which offers discount. Those uninformed consumers, who buy the deal, realize the true quality of the experience good, and their willingness to pay increases in the next period. Hence, some of these uninformed consumers may buy the experience good at the regular price in the second period (Dholakia, 2012). This aspect of daily-deal publishing is akin to the introductory pricing strategy of underestimated experience goods (Shapiro 1983), and is the sampling effect of daily-deal publishing. Moreover, some of the informed consumers, who would have bought the experience good at the regular price in the direct channel, may purchase the deal at a discounted price in the publisher channel (Dholakia, 2012). This leads to the cannibalization effect, which is the revenue loss to a merchant from informed consumers. In addition, a merchant pays to the daily-deal publisher a proportion of revenue generated on the website according to the revenue sharing scheme. This payment is the merchant’s cost of offering a deal on the publisher’s website.

Merchants value these effects of daily-deal publishing differently because of their inherent heterogeneities in proportion of informed consumers, marginal cost, and the quality estimate of the uninformed consumers on the publisher’ website. Therefore, a merchant’s profit of offering a deal on a daily-deal publisher’s website depends on the trade-offs between the positive effects of advertising and sampling and the negative effects of cannibalization and
revenue sharing. Therefore, a merchant’s optimal discount rate strategy is driven by these trade-offs.

We find that the impact of the publisher’s revenue sharing ratio on the merchant’s optimal discount rate is non-monotonic. More specifically, a merchant, who has low marginal cost and/or has low proportion of informed consumers, increases discount rate in the publisher channel as the publisher increases revenue sharing ratio. On the other hand, a merchant, who has higher marginal cost and/or high proportion of informed consumers, decreases discount rate in the publisher channel as the publisher increases the revenue sharing ratio. This is because a merchant does not try to optimize the profit in the first period, but optimizes profit over the two periods by balancing the loss of revenue from informed consumers in the publisher channel in the first period and gain of revenue from uninformed consumers in both the periods.

In turn, a merchant’s participation strategy is driven by the merchant’s optimal discount rate strategy which is impacted by the daily-deal publisher’s revenue sharing ratio. A merchant with low proportion of informed consumers and low marginal cost may participate even if the publisher has a very high revenue sharing ratio. This is because that the merchant has relatively large positive advertising benefit and relatively small negative cannibalization effect. Finally, we find that a publisher’s revenue depends on a merchant’s characteristics, such as the proportion of informed consumers, marginal cost, and uninformed consumers’ quality estimate. The real world implication of this is that the current industry practice of one size fits all – one revenue sharing ratio for all merchants - is suboptimal. We develop an optimal revenue sharing scheme for a daily-deal publisher who offers customized revenue sharing contracts.

To our knowledge, this is the first paper which develops an analytical model to study the dynamics of the daily-deal publishing and the strategic interaction between a merchant and a
daily-deal publisher. In doing so, we make three significant contributions to literature. First, while the extent literature in pricing of experience goods (Nelson, 1974; Shapiro, 1983; Bils, 1989) has considered merchants as the only strategic players, our setup has two strategic players, namely, a merchant and a daily-deal publisher, and we specifically model their strategic interaction. We find that a merchant does not always participate, and a merchant with high marginal cost or high proportion of informed consumers may offer higher introductory price compared to that recommended by the extant literature (Shapiro, 1983).

Our second contribution is related to the performance-based advertising literature (Dellarocas, 2012), which models only one period setting. Our two-period model allows us to study not only the advertising effect, but also the sampling effect. Our research studies the trade-offs that a merchant makes between inter-temporal profits, while single-period models of performance-based advertising are unable to do so. Our analysis of two-period model finds that a merchant’s pricing decision in the first period in the publisher channel is non-monotonic in the publisher’s pricing decision. This finding is new because the extant literature finds that optimal response of a merchant is to increase price as publisher increases per action advertising fee (Dellarocas, 2012). Models employed by key word search advertising literature (Dellarocas, 2012, Edelman et al., 2007; Liu et al. 2009) do not consider revenue sharing contract which is key to daily-deal publishing. Hence, this paper sheds new insights on merchant strategy in the context of performance-based advertising taking into consideration the revenue sharing effect.

Third, we contribute to the literature of daily-deal publishing business model by showing that a merchant’s adoption of daily-deal advertising largely depends on the merchant’s characteristics, and is impacted by the publisher’s revenue sharing contract. We find that the
publisher should offer higher (lower) revenue sharing ratio for a merchant with lower (higher) proportion of informed consumers and lower (higher) marginal cost.

1.1 Literature review

There are four broad streams of literature that are relevant to our paper: pricing strategy of experience goods, product sampling, performance-based advertising, and daily-deal publishing business model. The first stream of literature - pricing strategy of experience goods- posits that the first-time consumers (in our setup, uninformed consumers on the publisher’s website) learn the true quality of an experience good only upon consumption (Nelson 1974). Shapiro (1983) and Goering (1986) develop two-period models wherein consumers have lower quality estimate of an experience good in the first period, and recommend that a merchant should charge a lower price in the first period to encourage consumers to purchase the good. After consumption, consumers learn the true quality which leads to increase in their willingness to pay, and hence, the firm should charge a higher price in the second period.

This introductory pricing strategy is similar to a merchant offering price discount on a daily-deal publisher channel to incentivize uninformed consumers to purchase in the first period. However, in the context of daily-deal publishing model, a merchant faces a mix of informed and uninformed consumers in the publisher’s channel in the first period while the extant literature (Shapiro, 1983; Goering, 1986) consider a case of one consumer segment – uninformed consumers, in the first period. Our setting of two consumer segments in the first period is similar to Bils (1989) which studies a market with a mix of consumers – first-time consumers who are uncertain about the product quality and repeat consumers who know the product quality, and finds that an optimal pricing strategy should consider the trade-offs between exploiting repeat consumers and attracting first-time consumers. However, papers in this stream of literature do
not model a setting in which a merchant needs a third party - a daily-deal publisher- to access the mass of uninformed consumers. In other words, though the pricing of experience goods literature informs our research, it does not model a setting where a merchant requires a daily-deal publisher to access first-time consumers.

The second stream of literature - product sampling - has studied the strategy of offering free samples for introduction and promotion of new products (Bettinger et al. 1979; Jain et al, 1995; Heiman et al., 2001). Offering free samples to new consumers has a direct experiential effect that reduces the risk of product uncertainty (Heiman et al., 2001). This is akin to introductory pricing strategy for uninformed consumers, in our setup, in the first period in the publisher channel. Information Systems researchers have specifically studied sampling and pricing strategy related to digital experience goods (Chellappa and Shivendu, 2005; Dey and Lahiri, 2013; Dou et al, 2013; Wei and Nault, 2013). While digital goods are characterized by zero marginal cost, in the context of daily-deal publishing, a merchant may sell an experience good with non-negligible marginal cost. Though our research is informed by this sampling literature, a key difference is that, in the context of daily-deal publishing, a merchant faces consumers with heterogeneous quality estimates in the publisher channel in the first period. Moreover, in our setup, a merchant can sell in the direct channel and in the publisher channel in the first period, and has full control over the pricing decisions in both channels. Note this setting is different from the dual channel distribution literature where a manufacturer has access to two channels - direct channel and retailer channel - but has control over the pricing decision only in the direct channel, and the retailer decides the price in the retailer channel (Chiang et al., 2003; Tsay and Agrawal, 2004).
The third relevant stream of literature that informs our research relates to online performance-based advertising models. This literature has primarily focused on online keyword search advertising in general, and on auction design (Liu et al., 2009; Edelman et al., 2007), a publisher’s effort to improve conversion rates (Hu, 2004; Asdemir et al., 2012), and advertising merchants’ strategy (Chen and He 2011; Feng and Xie 2012; Dellarocas, 2012) in particular. Keyword search advertising and the daily-deal publishing are both performance-based because an advertiser pays to the publisher based on the advertising performance. However, there are three key differences between the two. The first difference is that while in the keyword search advertising model a merchant can access consumers in the publisher channel only, in the context of daily-deal publishing, a merchant accesses consumers both in the direct and the publisher channels. The second difference is that while in the keyword search advertising models, there is only one period and consumers do not update their quality belief, in our model, there are two periods, and uninformed consumers update quality belief upon consumption. The third difference is that while in the auction-based keyword advertising model, a publisher does not determine a merchant’s advertising cost per user action, in the daily-deal publishing model, the publisher can strategically decide the revenue sharing ratio which directly determines a merchant’s cost per action (purchase).

The fourth relevant literature stream relates to the ecosystem of daily-deal business models which studies the relationship between design and profitability of daily-deals (Byers et al., 2011), and profitability of advertising merchants (Edelman et al., 2011; Dholakia, 2012; Kumar and Rajan, 2012). Some earlier studies (Anand and Aron 2003; Kauffman and Wang 2001) have examined other online group buying schemes that arose before emergence of daily-deal websites. The closest paper to our research is Edelman et al. (2011) wherein authors study
the price discrimination and advertising effects of a daily-deal publishing business. However, they only focus on the merchants’ advertising strategy, and assume away the strategic role played by the publisher. Moreover, in their setting, a merchant’s discount rate is exogenously determined, and discount coupons only reach to the consumer segment with lower valuation, so they neither study merchant’s optimal discount rate strategy nor the cannibalization effect.

The rest of the paper is organized as follows. In §2, we present the model. In §3, we examine a merchant’s tradeoffs in offering a deal in the publisher channel. In §4, we present the optimal discount rate strategy of a merchant given the revenue sharing ratio announced by the daily-deal publisher. In §5 we analyze the participation decision of a merchant and in §6 we examine the daily-deal publisher’s optimal revenue sharing scheme given a merchant’s characteristics. In §7, we discuss our results and identify appropriate theoretical and managerial implications.

2. Model

We consider a market consisting of three types of players, a monopolist online daily-deal advertising publisher\(^{19}\) who publishes only daily-deals, a merchant, and a set of consumers. In our setting, a merchant sells an experience good of quality \(q\) which may be a product or service, and the marginal cost of production of the experience good is \(c\). The merchant faces two consumer segments that differ in terms of information about the merchant’s offering. The first segment consists of consumers who are aware about the merchant’s offering, and we refer to this segment as informed consumers. The second segment consists of consumers who are not aware about the merchant’s offering, and we refer to this segment as uninformed consumers. We denote

\(^{19}\) This conceptualization on monopolistic website offering daily-deals is similar to (Edelman et al, 2011). We discuss the implications of daily-deal websites competitive setting on our result in §7.
a merchant’s proportion of *informed consumers* in the market as $\delta$, and we refer to this as merchant type. More specifically, for a merchant of type $\delta$, the proportion of *informed consumers* is $\delta$ and the proportion of *uninformed consumers* is $1-\delta$. The merchant’s type $\delta$ is exogenously given, such that $\delta \in [0,1]$, and is common knowledge.

The market consists of a unit mass of consumers who have heterogeneous valuation for quality. The quality valuation parameter, $\theta$, is consumers’ private information, though the distribution $\theta \sim U[0,1]$ is common knowledge. In our model, *informed consumer* segment of a merchant knows the true quality of the experience good ($q$). In the absence of daily-deal publisher’s website, a merchant sells the experience good directly or in the direct channel to *informed consumer* segment, since *uninformed consumer* segment is not aware about the merchant. An *informed consumer’s* willingness to pay for one unit of the experience good in the direct channel is $\theta q$. It implies that consumers of higher $\theta$ have higher willingness to pay for the experience good.

We consider a two-period game. In the presence of the daily-deal publisher, a merchant can offer the experience good through two channels in the first period: direct channel and publisher channel or daily-deal website. A merchant does not offer deal in the second period. It means that a merchant can access consumers through dual channel (direct channel and publisher channel) in the first period but can access consumers only in the direct channel in the second period.\(^{20}\) We assume $\mu$ proportion of consumers in the market join the publisher’s website, where $\mu$ is exogenously given and $\mu \in (0,1)$. Note that the $\mu$ mass of consumers who join the website consists of a mix of *informed* and *uninformed* consumers. This implies that the merchant

\(^{20}\) “Upon execution of the Agreement, Merchant agrees that Merchant will not promote an online offer with respect to the products or services described in this Agreement of similar or greater value for a period up to 90 days from the Effective Date, plus a minimum of 90 days following the Merchant's date of feature on the Groupon Website.”- Groupon Merchant Terms, Clause 2- Terms and Termination: http://www.groupon.com/pages/merchant-terms#
has access to both the consumer segments in the publisher channel. An informed consumer’s willingness to pay for one unit of the experience good in the publisher channel is the same as in the direct channel ($\theta q$).

On the other hand, uninformed consumers are not aware about the true quality of the experience good. We assume that uninformed consumers have a quality estimate $R$, where $0 < R < 0.82$ (Shapiro, 1983; Jain et al., 1995). This implies that uninformed consumers uniformly agree upon the quality estimate of the experience goods offered on the publisher’s website (Chellappa and Shivendu, 2005). The value of this quality estimate depends on characteristics of the merchant’s experience good. We assume that $R$ is a point expectation, and hence an uninformed consumer’s willingness to pay for one unit of the experience good in the publisher is $\theta R$.

A daily-deal publisher offers a merchant a revenue sharing scheme wherein the merchant pays a proportion of revenue generated on the website to the publisher for the access to consumers in the publisher channel. Since the merchant pays a proportion of revenue generated in the publisher channel to the publisher, the total payment to the publisher depends on the performance of the merchant’s deal. Therefore, this revenue sharing scheme is a performance-based contract. Note that if the daily-deal publisher were to charge an upfront fixed fee from the merchant for accessing consumers on the website, it would be a nonperformance-based contract. The daily-deal publisher decides the proportion of revenue that a merchant pays for offering a deal, and we refer to this proportion as the revenue sharing ratio $s \in [0,1]$. We assume that the daily-deal publisher and the merchant have the same information about consumer characteristics, market demand and the cost structure.

---

21 The directionality of results in this paper does not change if some of the uninformed consumers underestimate while others overestimate the quality of experience goods in the publisher channel in the first period. Please see Appendix 2.B.
2.1 Purchase decision by consumers

In the first period, there are four types of consumers. First, there are $\delta\mu$ mass of consumers who are informed consumers on the publisher’s website. Second, there are $\delta(1-\mu)$ mass of informed consumers who are not on the publisher’s website. Third, there are $(1-\delta)\mu$ mass of uninformed consumers who are on the publisher’s website. Fourth, there are $(1-\delta)(1-\mu)$ mass of uninformed consumers who are not on the publisher’s website, and they do not buy in either period.

We assume that the merchant commits to a price in the direct channel for both the periods (Edelman et al. 2011) and we refer to this price as the regular price $p$. A merchant offers discount rate $d \in (0,1)$ on the publisher’s website and hence price in the publisher channel is $p(1-d)$ in the first period. To keep our exposition simple, we normalize the quality of the experience good $q=1$ in the subsequent analysis. A consumer’s surplus from buying the experience good in the first period is:

$$U_i(\theta) = \begin{cases} 
\theta - p & \text{informed consumers in the direct channel} \\
\theta - p(1-d) & \text{informed consumers in the publisher channel} \\
\theta R - p(1-d) & \text{uninformed consumers in the publisher channel} 
\end{cases} \quad (1)$$

Consumers are risk neutral, and consider purchasing one unit of good or nothing in each period. Consumers purchase only if their surplus is non-negative. Note that uninformed consumers who are not in the publisher channel ($(1-\delta)(1-\mu)$ mass of consumers) do not consider buying the good in the first period because they are not aware about product offerings of the merchant.

Uninformed consumers who buy the experience good in the first period in the publisher channel update their willingness to pay from $\theta R$ to $\theta$ upon consumption. A consumer’s surplus from buying the experience good in the direct channel in the second period is:
\[ U_\theta(\theta) = \begin{cases} \theta - p & \text{informed consumers, and uninformed consumer who bought in first period} \\ \theta R - p & \text{uniformed consumer in the publisher channel who did not buy in first period} \end{cases} \]  

Note that in the second period, the mass of consumers who know the true quality of the experience goods consists of informed consumers and those uninformed consumers who bought in the first period in the publisher channel.

### 2.2 Strategic interaction between daily-deal website and merchant

The strategic interaction between a publisher and a merchant takes place in a leader-follower setting. The publisher is the leader, and first announces the revenue sharing ratio \(s\). The merchant is the follower who takes into consideration this revenue sharing ratio in deciding whether to offer a deal on the publisher’s website and, if so, the optimal discount rate \(d\) to offer on the website. Moreover, while the publisher and the merchant make strategic decisions about revenue sharing ratio \(s\) and discount rate \(d\) respectively; consumers are price-takers.

In this setting, if a merchant decides not to offer a deal in the publisher channel, then, in the first period, she offers the experience good only in the direct channel. We refer to this case as the benchmark case. In this case the merchant decides the optimal price \(p\) in the direct channel to maximize profit over two periods. We denote the demand proportions from the informed consumers in the first period as \(D^i\) and as \(D^u\) in the second period. Therefore, in the benchmark case, the profit function of a merchant of type \(\delta\) over the two periods is:

\[
\pi^\delta(p, \delta) = \delta(D^i + D^u)(p - c)
\]  

The marginal consumer \((\hat{\theta})\) who is indifferent from purchasing, and hence, \(\hat{\theta} = p\). All consumers whose \(\theta \geq \hat{\theta}\) purchase the experience good. Since informed consumers’ valuation for quality is \(\theta \sim U[0,1]\), the demand proportions from informed consumers in each period is:

\[ D^i = D^u = 1 - p. \]
When a merchant decides to offer a deal in the publisher channel, she offers a deal with a regular price, $p$, and a discount rate $d$, on the publisher’s website, and hence, the effective price on the publisher’s website is $p(1-d)$. Let the demand proportions of informed and uninformed consumers on the publisher’s website in the first period be $D_{hi}^w$ and $D_{iu}^w$ respectively. Let the demand proportion of informed consumers in the direct channel in the first period be $D_{hi}^d$. The first period profit function of a merchant of type $\delta$ is:

$$\pi_1 = \mu \delta D_{hi}^w (p(1-d)(1-s)-c) + \mu(1-\delta) D_{iu}^w (p(1-d)(1-s)-c) + (1-\mu) \delta D_{hi}^d (p-c),$$

(4)

where $D_{hi}^w = 1-p(1-d)$, $D_{iu}^w = (1-p(1-d)/R)$, and $D_{hi}^d = 1-p$.

In the second period, the merchant offer the experience good only in the direct channel regular price $p$. uninformed consumers, who buy in the first period in the publisher channel, update their quality estimate to true quality ($q$), and update their willingness to pay from $\theta R$ to $\theta$. We denote the proportion of uninformed consumers who buy in the second period as $D_{iu}^o$. Some of these uninformed consumers who bought the experience good in the first period may also buy at the regular price in the second period. Moreover, it is easy to see that an uninformed consumer, who did not buy at a discounted price in the first period in the publisher channel, will not buy in the second period at the regular price in the direct channel, which implies $D_{iu}^o \leq D_{iu}^w$.

Let the demand proportion from informed consumers in the second period be $D_{hi}^d$. The profit function of a $\delta$ type merchant in the second period is:

$$\pi_2 = \delta D_{hi}^o (p-c) + \mu(1-\delta) D_{iu}^o (p-c),$$

(5)

where $D_{hi}^o = 1-p$ and $D_{iu}^o = \min\{1-p(1-d)/R, 1-p\}$.
Now we can write the two-period profit function of a merchant of type $\delta$, who offers a deal in the publisher channel as:

$$\pi_M(d;\delta) = \pi_1 + \pi_2 = \mu(\delta D_{1w}^w + (1-\delta)D_{1w}^w)(p(1-d)(1-s)-c) + (1-\mu)(\delta D_{1w}^e(p-c) + (\delta D_{1w}^w + \mu(1-\delta)D_{1w}^e)(p-c)$$

(6)

Thus, a merchant’s profit maximization problem is $\max_{d} \pi_M(d;\delta)$. Let $d^*$ be the solution to this profit maximization problem, and the optimal profit for a merchant of type $\delta$ be $\pi_M(d^*;\delta)$. A merchant offers a deal on the publisher’s website only if the profit from offering a deal is not less than the profit from the benchmark case, that is $\pi_M(d^*;\delta) \geq \pi_M^b(\delta)$.

Now, we can write the profit function of a daily-deal publisher who offers a revenue sharing contract with a revenue sharing ratio ($s$) to a merchant of type $\delta$ as:

$$\pi_p(s) = \begin{cases} s\mu(\delta D_{1w}^w p(1-d) + (1-\delta)D_{1w}^w p(1-d)) & \text{if } \pi_M(d^*;\delta) \geq \pi_M^b(\delta) \\ 0 & \text{if } \pi_M(d^*;\delta) < \pi_M^b(\delta) \end{cases}$$

(7)

where $\mu(\delta D_{1w}^w p(1-d) + (1-\delta)D_{1w}^w p(1-d))$ is the revenue generated on the publisher’s website when a merchant of type $\delta$ offers a deal. The daily-deal publisher chooses a revenue sharing ratio ($s = s^*$) which maximizes the profit in (7).

Figure 2.1: Sequence of the game

Figure 2.1 describes the sequence of the two-period game. The summary of notations is provided in Appendix 2.A.
3. A merchant’s trade-offs in offering a deal

Before we discuss the optimal strategies of a daily-deal publisher and a merchant, we first establish the theoretical framework to study a merchant’s trade-offs in offering a deal on a daily-deal publisher’s website. First, a merchant offers a deal on the daily-deal publisher’s website to acquire new consumers, that is, uninformed consumers, who are not aware about the merchant in the direct channel. Second, uninformed consumers who purchased the experience good in the publisher channel update their willingness to pay in the second period. Third, informed consumers who are in the publisher channel may also buy the deal at a discounted price, though they have high willingness to pay for the experience good. In addition, the merchant shares a proportion of revenue generated in the publisher channel with the publisher. We explain these aspects that impact the merchant’s profit of offering a deal on the publisher’s website in the following paragraphs.

When a merchant offers a deal on the daily-deal publisher’s website, uninformed consumers get information about the merchant’s experience good. Note that these uninformed consumers do not know about the merchant in the direct channel. Therefore, by offering a deal on the daily-deal publisher’s website, a merchant is able to reach out to uninformed consumers who are on the website. We characterize this impact of reaching out to uninformed consumers on the publisher’s website as the advertising effect. This effect is positive for the merchant, and it may increase the merchant’s market size in both the periods. If the merchant were to offer the experience good on the website at the regular price \( p \), then in the first period, the demand from uninformed consumers who are on the website is \( \mu(1-\delta)(1-\frac{p}{R}) \). Those uninformed consumers who buy in the first period, then they update their willingness to pay and also buy in the second
period. The merchant’s advertising effect is the gain in profit from having access to uninform
ed consumers on the publisher’s website. Therefore, advertising effect is:

\[ F_a = (1-\frac{P}{R})\mu(1-\delta)(p-c)+(1-\frac{P}{R})\mu(1-\delta)(p-c). \]  (8)

Note that the advertising effect is stronger, (1) if the merchant has lower proportion of informed consumers (\(\delta\)), or (2) if uninformed consumers have a larger quality estimate (\(R\)), or (3) if merchant has a smaller marginal cost (\(c\)), or (4) if the mass of consumers on the publisher’s website (\(\mu\)) is larger.

When a merchant offers a deal on the daily-deal publisher’s website, the merchant offers a price discount (\(d\)) in the first period, hence, the price of the experience good on the publisher’s website is \(p(1-d)\). Due to this lower price on the publisher’s website, some informed consumers on the website, who do not buy at the regular price may buy at the discounted price. The demand proportion for these informed consumers, who take the advantage of price discount in the publisher channel, is \(\mu\delta(pd)\). However, these informed consumers do not buy at the regular price in the second period. Furthermore, uninformed consumers who are on the website have lower willingness to pay because they have lower quality estimate than the true quality of the experience good (\(R < q = 1\)). Some uninformed consumers who would have not bought the experience good at the regularly price in the first period in the publisher channel, buy at the discounted price. The demand proportion of these uninformed consumers is \(\mu(1-\delta)(\frac{pd}{R})\). Recall that advertising effect in (8) accounts for those uninformed consumers who buy in the first period in the publisher channel even if discount rate is zero, and also buy in the direct channel in the second period. This implies that the demand proportion of uninformed consumers in the second
period who buy because they update their quality belief, is \( D_{2u}^c (1 - \frac{P}{R}) \). We characterize the impact of offering price discount in the publisher channel on the merchant’s profit as sampling effect. More formally, the sampling effect is:

\[
F_s = \mu \delta (p_d)(p(1-d)-c) + (\frac{pd}{R}) \mu (1-\delta)(p(1-d)-c) + \mu (1-\delta)(D_{2u}^c (1 - \frac{P}{R}) (p-c)).
\]  

Note that the sampling effect is stronger, (1) if the merchant has lower proportion of informed consumers (\( \delta \)), or (2) if uninformed consumers have a smaller quality estimate (\( R \)), or (3) if merchant has a smaller marginal cost (\( c \)), or (4) if the mass of consumers on the publisher’s website (\( \mu \)) is larger.

The advertising and sampling effects result in increase in the profit of a merchant who offers a deal on the daily-deal publisher’s website, and hence, both are beneficial effects. However, offering a deal on the daily-deal website also imposes two types of costs on a participating merchant.

When a merchant offers a discount on the daily-deal publisher’s website, she incurs loss of revenue from those consumers who have high valuation for quality (\( \theta \)), because they would have bought the experience good at the regular price. More formally, the cannibalization effect is:

\[
F_c = (1-p) \mu \delta p_d + (1-\frac{P}{R}) \mu (1-\delta)p_d,
\]  

where the first and second terms of equation (10) are the loss of revenue from informed consumers and uninformed consumers respectively due to discount in the publisher channel. The cannibalization effect is stronger, (1) if a merchant has higher proportion of informed consumers (\( \delta \)), or (2) if the merchant offers higher discount rates (\( d \)), or (3) if uninformed consumers have
a smaller quality estimate ($R$), or (4) if the mass of consumers on the publisher’s website ($\mu$) is larger. Note that the magnitude of the cannibalization effect does not depend on the merchant’s marginal cost ($c$).

A merchant also shares $s$ proportion of revenue generated on the daily-deal website with the publisher in the first period. This is the cost to the merchant for getting access to uninformed consumers on the publisher’s website. We characterize this as the revenue sharing effect, and more formally, it is:

$$F_R = (1 - \frac{p(1-d)}{R})(1-\delta)\mu(p(1-d)s) + (1 - p(1-d))\delta\mu(p(1-d)s),$$

(11)

The first term of equation (11) is the revenue from uninformed consumers in the publisher channel that a merchant pays to the publisher in the first period. The second term of equation 11 is the revenue from informed consumers in the publisher channel that a merchant pays to the publisher in the first period. The revenue sharing effect is stronger, (1) if a merchant has higher proportion of informed consumers ($\delta$), or (2) if uninformed consumers have a larger quality estimate ($R$), or (3) if the mass of consumers on the publisher’s website ($\mu$) is larger, or (4) if the revenue sharing ratio ($s$) is larger. Note that the magnitude of the revenue sharing effect does not depend on the merchant’s marginal cost ($c$).

A merchant’s discount rate strategy is driven by the trade-off between these four effects. In Section 4, we analyze a merchant’s optimal discount rate strategy.

### 4. A merchant’s discount rate strategy

Before we present our analysis regarding a merchant’s optimal discount rate strategy in offering a deal on the daily-deal publisher’s website, we establish the regular price, the demand
proportion and optimal profit under the benchmark case where a merchant sells the experience good only in the direct channel.

4.1 Benchmark case: a merchant sells only in the direct channel

When a merchant of type $\delta$ sells the experience good only in the direct channel, she has access only to the informed consumers who know the true quality of the experience good. An informed consumer who has $\theta$ valuation for quality buys the experience good in each period if $\theta - p \geq 0$. More formally since $\theta \sim U[0,1]$, we have the demand in each period as $D^i = D^r = 1 - p$. Now, we can rewrite the merchant’s benchmark profit function over the two periods as:

$$\pi^b(p) = 2\delta(1 - p)(p - c)$$  \hspace{1cm} (12)

A merchant of type $\delta$ maximizes her profit in (12) by choosing the optimal price $p^*$. Solution to the merchant’s optimization problem gives the optimal price, $p^* = (1 + c) / 2$, demand proportion in each period, $D^i = D^r = (1 - c) / 2$, and profit in the benchmark case as $\pi^b = \delta(1 - c)^2 / 2$.

4.2 Merchant’s optimal discount rate strategy in offering a deal

A merchant considers the trade-offs between the positive and negative effects of offering a deal on the daily-deal publisher’s website in the first period to determine the optimal discount rate ($d^*$). Note that the regular price offered by a merchant in the direct channel is the merchant’s optimal price in the benchmark case, that is $p^* = (1 + c) / 2$. It is so because in the direct channel, the merchant sells only to the informed consumer segment. This also implies that a merchant who offers a deal with discount rate $d$ sells the experience good at the discounted price $p^*(1 - d)$ in the publisher channel, where $p^* = (1 + c) / 2$. 
**Uninformed consumers**, who are in the daily-deal publisher channel (proportion $\mu(1-\delta)$), underestimate the quality of the experience good ($R<1$), and, the demand proportion for the experience good from the *uninformed consumers* in the publisher channel is $D_{1u}^p = 1 - \frac{p^*(1-d)}{R}$.

Note that this demand proportion increases when merchant’s discount rate increases, as well as when *uninformed consumers*’ quality estimate ($R$) increases. Before we discuss the optimal discount rate strategy of a merchant, Proposition 1 reports the range within which the optimal discount rate is bounded.

**PROPOSITION 1:** A merchant’s optimal discount rate $d^*$ is bounded such that

$$\max\{d, 0\} \leq d^* \leq 1 - R,$$

where $d = 1 - \frac{2R}{1 + c}$.

When a merchant offers discount rate $d = 1 - R$, the demand proportion from *uninformed consumers* in the publisher channel is $D_{1u}^w = \frac{p^*(1-d)}{R} |_{d=1-R} = 1 - p^*$. Those *uninformed consumers*, who buy at the discounted price in the first period in the publisher channel, also buy at the regular price in the second period in the direct channel ($D_{1u}^p = 1 - p^*$, and $D_{1u}^w = D_{1u}^p$). As mentioned in §2.2, the demand proportion of *uninformed consumers* on the website in the first period is always weakly higher than the demand proportion of *uninformed consumers* in the direct channel in the second period, that is, $D_{1u}^w \geq D_{1u}^p$. Hence, if a merchant offers a discount rate $d > 1 - R$, then some *uninformed consumers*, who buy at discounted price in the first period, do not buy at the regular price in the direct channel in the second period ($D_{1u}^w > D_{1u}^p$). This implies that the benefit to the merchant due to the sampling effect does not increase when the discount rate is $d > 1 - R$.

On the other hand, by offering a higher discount rate, the cost to the merchant due to the cannibalization effect increases. Hence, the highest discount rate that a merchant considers offering in the publisher channel is $d = 1 - R$. When a merchant offers the highest discount rate
the merchant gets the same demand proportion of uninformed consumers in both the periods, that is, $D^w_{au} = D^c_{au} = \frac{1-c}{2}$.

A merchant’s sampling effect is zero when the discount rate is $d \leq d$, since no uninformed consumers buy at the discounted price in the publisher channel ($D^w_{au} = 0$), where $d = 1 - \frac{2R}{1+c}$. However, the merchant still incurs cost due to the cannibalization effect, as some of informed consumers, who would have bought at the regular price in the direct channel, buy at the discounted price in the publisher channel. Therefore, the lowest discount rate that a merchant consider offering in the publisher channel is $d = \max\{d, 0\}$. Next, Lemma 1 reports the optimal discount rate offered by a merchant in the daily-deal publisher’s channel.

**LEMMA 1:** There exists a lower merchant type $\delta$ and a higher merchant type $\bar{\delta}$ such that (i) for a merchant of type $\delta \in [0, \underline{\delta})$ the optimal discount rate is $d^* = 1-R$; (ii) for a merchant of type $\delta \in (\underline{\delta}, \bar{\delta})$ the optimal discount rate is $d^* = \hat{d} = \frac{(1-c + 2(1-R)(1-s))(1-\delta) - 2cs(1-(1-R)\delta)}{2(1+c)(1-s)(1-(1-R)\delta)}$; and (iii) for a merchant of type $\delta \in [\bar{\delta}, 1]$ the optimal discount rate is $d^* = 0$. The lower merchant type $\underline{\delta}$ and the higher merchant type $\bar{\delta}$ are: 

\[
\underline{\delta} = \begin{cases} 
\frac{1-c(3-2R(1-s))}{1-c+2(1-R)(1-s(1+c))} & \text{for } s \leq s', \\
0 & \text{for } s' < s 
\end{cases}
\]

and 

\[
\bar{\delta} = \begin{cases} 
\frac{2csR}{1-c+2(1-R)(1-s(1+c))} & 0 \leq c \leq c', \\
\frac{1-c+2R(1-s)}{1-3c+2R(1-s)} & c' < c < 1 
\end{cases}
\]

\[\text{where } s' = \frac{1-3c+2cR}{2cR}, \text{ and } c' = 2R-1.\]

Lemma 1 outlines the optimal discount rate offered by a merchant of type $\delta$. Recall that for a merchant of type $\delta$, the proportion of informed consumer segment is $\delta$, and proportion of uninformed consumer segment is $1-\delta$. First, there are two threshold merchant types, the lower merchant type $\underline{\delta}$ and the higher merchant type $\bar{\delta}$, which influence merchant’s optimal discount
rate strategy. In turn, while, the value of \( \delta \) critically depends on the daily-deal publisher’s choice of revenue sharing ratio \( s \), the value of \( \delta^* \) critically depends on the merchant’s marginal cost. Second, as discount rate increases in the publisher channel, a merchant with low \( \delta \in [0, \bar{\delta}) \) gets larger benefit from the sampling effect compared to increase in loss due to the cannibalization effect. Hence, a merchant of type \( \delta \in [0, \bar{\delta}) \) offers the highest discount rate \( d^* = 1 - R \). Third, as a merchant’s informed consumer segment \( (\delta) \) increases, the loss due to the cannibalization effect increases, and thus, the merchant lowers the discount rate as \( \delta \) increases beyond \( \bar{\delta} \). Hence, for a merchant with moderate \( \delta \in [\bar{\delta}, \bar{\delta}) \), the optimal discount rate \( (d^* = \hat{d}) \) is lower than highest discount rate, and decreases in \( \delta \). This optimal discount rate \( (\hat{d}) \) depends on uninformed consumers’ quality estimate \( R \), marginal cost \( c \), and the daily-deal publisher’s revenue sharing ratio \( s \). Fourth, for a merchant with high proportion of informed consumers \( \delta \in [\bar{\delta}, 1] \), loss due to cannibalization effect dominates the net benefits from the advertising and sampling effects. Hence, a merchant of type \( \delta \in [\bar{\delta}, 1] \) offers zero discount rate.
Figure 2.2 illustrates the discount rate strategy of a merchant (reported in Lemma 1) with reference to the merchant’s marginal cost and the publisher’s revenue sharing ratio. In the Region I and II, the publisher’s revenue sharing ratio is relatively low, \( s \leq s' \). In this case, for a merchant with relatively low \( \delta \in [0, \bar{\delta}] \), the loss due to the cannibalization and the revenue sharing effects is low. Hence, a merchant with relatively low \( \delta \in [0, \bar{\delta}] \) offers the highest discount rate \( d' = 1 - R \). On the other hand, when the publisher’s revenue sharing ratio is relatively high, \( s > s' \), the loss due to revenue sharing effect increases, therefore, a merchant decreases the discount rate to reduce the loss due to the revenue sharing and the cannibalization effects. Hence, in the Region III and IV, a merchant of any type offers discount rate lower than the highest discount rate, that is \( d' < 1 - R \).
In Region III and IV, the highest discount rate offered by a merchant is denoted by \( d' = \bar{d} \) and is lower than \( 1-R \). In this case, the lower merchant type is \( \bar{\xi} = 0 \). The value of \( \bar{d} \) can be derived from Lemma 1, and is given as \( \bar{d} = \frac{3-c-2R-2(1+c-R)s}{2(1+c)(1-s)} \). (Proof is in the Appendix). Note that \( \bar{d} \) decreases, as uninform ed consumers’ quality estimate \( R \) increases, or the merchant’s marginal cost \( c \) increases.

In the Region I and III, a merchant of type \( \delta = \bar{\delta} \) offers the lowest discount rate \( d' = 0 \). Since the marginal cost is relatively low in these two regions \( 0 \leq c \leq c' \), the merchant’s regular price in the direct channel is low, and hence, she attracts some uninform ed consumers to buy in the publisher channel even when the discount rate is zero. On the other hand, when the merchant’s marginal cost is relatively high \( c' < c < 1 \) in the Region II and IV, the regular price in the direct channel is high. If a merchant offers a discount rate lower than \( d = 1 - \frac{2R}{1+c} \) (Proposition 1), no uninform ed consumer buys in the publisher channel. Hence, by offering a discount rate than \( \bar{d} \), a merchant has no benefit from sampling effect, but incurs loss due to the cannibalization effect. Therefore, it is never optimal for a merchant to offer discount rate lower than \( \bar{d} \) if marginal cost is relatively high \( c > c' \).

4.3 Impact of revenue sharing ratio on a merchant’s discount rate strategy

From Lemma 1, we know that a merchant trade-offs the relative benefits and costs (four effects described in §3) of offering a deal on the publisher’s website to determine the optimal discount rate. Since the revenue sharing ratio \( s \) announced by the daily-deal publisher determines the revenue sharing effect, it affects a merchant’s optimal discount rate strategy in the following way. When the daily-deal publisher increases revenue sharing ratio \( s \), the merchant pays a larger proportion of revenue generated in the publisher channel to the publisher, hence, the
merchant’s cost due to the revenue sharing effect increases. The merchant can respond to the increase in this cost by adjusting the discount rate in the publisher channel, which in turn impacts the magnitude of sampling and cannibalization. Note that advertising effect (Equation (8)) does not depend on discount rate offered by the merchant in the publisher channel.

Recall from Lemma 1, a merchant of type $\delta \in [0, \bar{\delta})$ offers optimal discount rate $d^* = 1 - R$, a merchant of type $\delta \in [\bar{\delta}, \tilde{\delta})$ offers optimal discount rate $d^* = \hat{d}$ and a merchant of type $\delta \in [\tilde{\delta}, 1]$ offer optimal discount rate $d^* = 0$. Before we analyze the impact of the revenue sharing ratio on the merchant’s optimal discount rate, we first analyze the impact of the revenue sharing ratio on the two critical threshold values of merchant type, that is the lower merchant type ($\hat{\delta}$) and the higher merchant type ($\tilde{\delta}$).

**PROPOSITION 2 (a):** When the revenue sharing ratio ($s$) increases, the lower merchant type $\hat{\delta}$ (i) increases when marginal cost is relatively low, $c < \bar{c}$, (ii) decreases when marginal cost is relatively high, $c > \bar{c}$, and (iii) does not change when marginal cost $c = \bar{c}$, where $\bar{c} = \frac{3 - 2R - \sqrt{9 - 8(2 - R)R}}{2R}$.

Note that an increase in $\hat{\delta}$ implies that some merchants, who offered discount rate $d^* = \hat{d}$ (lower than $1 - R$), may offer the highest discount rate $d^* = 1 - R$. Conversely, a decrease in $\hat{\delta}$ implies that some merchants, who offer highest discount rate $d^* = 1 - R$, may offer lower discount rate $d^* = \hat{d} < 1 - R$. Proposition 2 reports that the impact of revenue sharing ratio ($s$) on lower merchant type ($\hat{\delta}$) is mediated by merchant’s marginal cost ($c$). When the marginal cost is relatively low ($c < \bar{c}$), the merchant’s benefit from each uninformed consumer who buys the experience good in either period is high. This implies that for such a merchant, the advertising and sampling effects are strong. Note that while the revenue gain to the merchant from the
sampling and advertising effects high is in both the periods, revenue loss to the merchant due to the cannibalization and revenue sharing effects is limited only to the first period. Therefore, when revenue sharing ratio ($s$) increases, it is optimal for a merchant, whose type is not much higher than the lower merchant type, to increases the discount rate, if marginal cost is low. In this circumstance, the merchant can gain more from the sampling effect than the loss due to the increase in the cannibalization effect. Hence, lower merchant type ($\delta$) increases as revenue sharing ratio ($s$) increases. On the other hand, when the marginal cost is relatively high ($c>\bar{c}$), the revenue gain to a merchant from sampling and advertising effects is small in both the periods. In this situation, when revenue sharing ratio ($s$) increases, a merchant whose type is not much lower than the lower merchant type, decreases the discount rate to reduce cannibalization effect. This leads to decrease in the lower merchant type ($\delta$).

**PROPOSITION 2 (b):** When the revenue sharing ratio ($s$) increases, the higher merchant type $\bar{\delta}$ (i) increases when the uninformed consumers’ quality estimate is small, $R<1/2$, and the marginal cost is sufficiently small, $c<\bar{c}$, and (ii) decreases when the uninformed consumers’ quality estimate is large, $R\geq1/2$, where $\bar{c} = \frac{1-2R}{3-4R}$.

Proposition 2 (b) describes the effect of an increase in the publisher’s revenue sharing ratio on higher merchant type $\bar{\delta}$. Recall that a merchant of type $\delta \in [0, \bar{\delta})$ offers positive optimal discount rate and a merchant type $\delta \geq \bar{\delta}$ offers zero discount rate. This implies that for a merchant of type $\bar{\delta}$, the cannibalization effect and sampling effect is zero. As the revenue sharing ratio increase, the loss to a merchant due to the revenue sharing effect increases. Recall that the sampling effect is strong and the cannibalization effect is weak when the *uninformed consumers* have relatively low quality estimate of the experience good (low $R$) and the marginal cost is sufficiently small. This implies that if a merchant increases discount rate, then her revenue
sharing cost decreases, and she has net gain from the sampling effect and cannibalization effect.

Therefore, it is optimal for a merchant, whose type is not much higher than the higher merchant type ($\tilde{\delta}$), to offer a positive discount rate $\hat{d} > d$, when marginal cost is sufficiently low. In this situation, the higher merchant type ($\tilde{\delta}$) increases as revenue sharing ratio ($s$) increases. On the other hand, when uninformed consumers have relatively high quality estimate of the experience good (high $R$), the sampling effect is low. Hence, in this case, it is optimal for a merchant, whose type is not much lower than the $\tilde{\delta}$, to offer zero discount rate. In this way, the merchant can balance off the cannibalization, sampling, and revenue sharing effect when the revenue sharing ratio increases. Therefore, in this case, the higher merchant type ($\tilde{\delta}$) decreases as revenue sharing ratio ($s$) increases.

For a merchant of type $\delta \in [0, \tilde{\delta})$ or $\delta \in [\tilde{\delta}, 1]$, the optimal discount rate is independent of the revenue sharing ratio. Therefore, the merchant does not adjust discount rate as revenue sharing revenue increases. On the other hand, the optimal discount rate for a merchant of type $\delta \in [0, \tilde{\delta})$ is $\hat{d}$, which is a function of the daily-deal publisher’s revenue sharing ratio. Hence, a merchant of type $\delta \in [\tilde{\delta}, 1]$ adjusts discount rate as revenue sharing ratio increases. After establishing the impact of revenue sharing ratio on the two critical merchant types, the lower merchant type ($\hat{\delta}$) and the higher merchant type ($\tilde{\delta}$), now we discuss the impact of revenue sharing ratio on the optimal discount rate strategy of a merchant type $\delta \in [\hat{\delta}, \tilde{\delta})$.

**Proposition 3 (a):** When marginal cost is relatively small ($c < \bar{c}$) and uninformed consumers’ quality estimate is small ($R < 1/2$), the optimal discount rate offered by a merchant of type $\delta \in [\hat{\delta}, \tilde{\delta})$ increases as revenue sharing ratio ($s$) increases.

A merchant’s cost due to the revenue sharing effect increases as the daily-deal publisher increases the revenue sharing ratio. To mitigate the increase in the revenue sharing cost, a
merchant may respond by increasing the discount rate in the publisher channel. This increases the cannibalization effect; however, if uninformed consumers have relatively low quality estimate of the experience good \((R < 1/2)\) and marginal cost is sufficiently small \((c < \hat{c})\), then the sampling effect increases at a faster rate compared with the cannibalization effect. This leads to the result in Proposition 3(a) that it is optimal for a merchant of type \(\delta \in [\hat{\delta}, \delta)\) to increase the discount rate as the revenue sharing ratio increases when marginal cost is relatively low \((c < \hat{c})\).

**PROPOSITION 3 (b):** If marginal cost is moderate \((\bar{c} \leq c \leq \hat{c})\), then as the revenue sharing ratio \((s)\) increases, (i) the optimal discount rate increases for a merchant of type \(\delta \in [\hat{\delta}, \delta)\); (ii) the optimal discount rate decreases for a merchant of type \(\delta \in (\bar{\delta}, \bar{\delta})\); and (iii) the optimal discount rate does not change for a merchant of type \(\delta = \bar{\delta}\), where \(\bar{\delta} = \frac{1-3c}{1-3c+2cR}\).

When marginal cost is moderate \((\bar{c} \leq c \leq \hat{c})\) and merchant type is relatively low \((\delta \in [\hat{\delta}, \bar{\delta})\), a merchant has a strong sampling effect and a weak cannibalization effect. Hence, when the publisher increases the revenue sharing ratio, a merchant can respond by increasing the discount rate to decrease the revenue sharing costs, while the net result of the sampling, cannibalization and revenue sharing effects is positive. However, this does not hold when the merchant type is relatively high \((\delta \in (\bar{\delta}, \bar{\delta})\), since the cannibalization effect is stronger than the sampling effect. As a result, if the merchant increases discount rate, the gain from the reduction in the revenue sharing cost is lower than the net loss of the sampling and cannibalization effect. Hence, when marginal cost is moderate \((\bar{c} \leq c \leq \hat{c})\) and merchant type is relatively high \((\delta \in (\bar{\delta}, \bar{\delta}))\), a merchant decreases the optimal discount rate as the revenue sharing ratio increases.

**PROPOSITION 3 (c):** When marginal cost is relatively high, \(c > \hat{c}\), as the revenue sharing ratio \((s)\) increases, the optimal discount rate of a merchant of type \(\delta \in [\hat{\delta}, \bar{\delta})\) decreases.
When marginal cost is relatively high \((c > \bar{c})\), a merchant of type \(\delta \in [\underline{\delta}, \bar{\delta})\) has a weak sampling effect and a strong cannibalization effect. Hence, such a merchant responds to an increase in the revenue sharing ratio by decreasing the discount rate in the publisher channel. This mitigates the increase in the revenue sharing cost, and the cannibalization effect decreases at a faster rate than the sampling effect. Therefore, it is optimal for a merchant of type \(\delta \in [\underline{\delta}, \bar{\delta})\) to decrease the discount rate as the revenue sharing ratio increases when marginal cost is relatively high \((c > \bar{c})\).

Propositions 3 (a), 3 (b) and 3 (c) highlight that the impact of the daily-deal publisher’s revenue sharing ratio is non-monotonic on the optimal discount rate strategy for a merchant of type \(\delta \in [\underline{\delta}, \bar{\delta})\). The left panel of Figure 2.3 illustrates Proposition 3 (a), the middle panel of Figure 2.3 illustrates Proposition 3 (b) and the right panel of Figure 2.3 illustrates Proposition 3 (c).

**Figure 2.3: Impact on optimal discount rates when revenue sharing ratio increases**

The qualitative nature of the impact of an increase in the revenue sharing ratio on the optimal discount rate critically depends on marginal cost and merchant type. This is so because the sampling effect is dominant when marginal cost is relatively low and merchant type is relatively low, while the cannibalization effect is dominant when marginal cost is relatively
higher and merchant type is relatively high. Therefore, a merchant decreases the discount rate when the cannibalization effect is dominant because the rate of decrease in the net effect of cannibalization and revenue sharing is faster than the decrease in the sampling effect. On the other hand, a merchant increase the discount rate when the sampling effect is dominant because the rate of increase in sampling effect is faster than the increase in the net effect of cannibalization and revenue sharing.

In the next sub-section, we study a merchant’s optimal discount rate strategy when a publisher adopts a nonperformance-based scheme, and compare this with the optimal discount rate strategy under a performance-based scheme.

### 4.4 Merchant’s discount rate strategy under nonperformance-based scheme

When the daily-deal publisher adopts a nonperformance-based scheme, the publisher charges a membership fee from each participating merchant who offers a deal on the publisher’s website. In this situation, the membership fee paid by a merchant to the publisher is independent of the revenue generated in the publisher channel. If the membership fee charged by the publisher is \( f \), then the two-period profit function of a merchant of type \( \delta \) who offers a deal with discount rate \( d^{np} \) in the publisher channel is:

\[
\pi^{np}_M = \mu(\delta D^w + (1-\delta)D^w_\mu)(p(1-d^{np}) - c) + (1-\mu)\delta D^w_\mu(p - c) + (\delta D^w_\mu + \mu(1-\delta)D^w_\mu)(p - c) - f
\]  

(13)

where \( D^w_\mu = (1 - \frac{p(1-d^{np})}{q}) \), \( D^w = D^w_\mu = (1 - \frac{p(1-d^{np})}{R}) \), and \( D^w_\mu = D^w_\mu = (1 - \frac{p}{q}) \).

Therefore, the profit maximization problem of a merchant who participates in a nonperformance-based scheme is \( \max_{d^{np}} \pi^{np}_M \). Comparing a merchant’s profit function under a nonperformance-based scheme given in (13) with that under a performance-based scheme given in (6), it is easy to see that the solution of the profit maximization problem under a
nonperformance-based scheme is \( d^{NP} = d^* (s = 0) \). This implies that the optimal discount rate offered by a merchant under the nonperformance-based scheme is a special case of the optimal discount under the performance-based scheme with zero revenue sharing ratio.

**PROPOSITION 4 (a):** When the daily-deal publisher adopts a nonperformance-based scheme, compared to a performance-based scheme, a merchant offers a higher discount rate (\( d^{NP} > d^* \)), if the marginal cost is relatively low (\( c < \bar{c} \)) and type is \( \delta \in [\bar{\delta}^{NP}, \delta^{NP}) \), or the marginal cost is moderate (\( \bar{c} \leq c \leq \bar{c} \)) and type is relatively low (\( \delta \in [\bar{\delta}^{NP}, \delta) \)).

**PROPOSITION 4 (b):** When the daily-deal publisher adopts a nonperformance-based scheme, compared to a performance-based scheme, a merchant offers a lower discount rate (\( d^{NP} < d^* \)), if the marginal cost is relatively high (\( c > \bar{c} \)) and type is \( \delta \in [\bar{\delta}^{NP}, \delta^{NP}) \), or the marginal cost is moderate and type is relatively high (\( \delta \in (\bar{\delta}, \delta^{NP}) \)).

First note the merchant of type \( \delta \) is defined in Proposition 3(b), and two critical merchant types in the case of nonperformance-based scheme, lower merchant type (\( \bar{\delta}^{NP} \)) and higher merchant type (\( \delta^{NP} \)), are derived from Lemma 1, such as \( \bar{\delta}^{NP} = \bar{\delta} (s = 0) \) and \( \delta^{NP} = \bar{\delta} (s = 0) \). Moreover, the two critical threshold of marginal cost, \( \bar{c} \) is defined in Proposition 2 (a) and \( \bar{c} \) is defined in Proposition 2(b). The optimal discount rate under a performance-based scheme (Lemma 1) may be higher or lower compared to a nonperformance-based scheme, depending on marginal cost (\( c \)) and merchant type (\( \delta \)). This implies that adoption of nonperformance-based scheme has non-monotonic impact on a merchant’s optimal discount rate. The economic intuition is as follows. Note that under a nonperformance-based scheme, a merchant does not incur the cost due to the revenue sharing effect. Moreover, the advertising effect is independent of discount rate. Therefore, when the publisher changes scheme from performance-based to nonperformance-based, a merchant can respond either by increasing discount rate to increase the
gain from sampling effect, or by decreasing discount rate to reduce the loss due to cannibalization effect.

The optimal strategy for a merchant with lower marginal cost \((c < \bar{c})\) and type \(\delta \in [\delta^{np}, \delta^{np}]\) is to increase the discount rate because she gains more from the increased sampling effect than loss due to the increased cannibalization effect. This dynamic remains the same when the marginal cost is moderate \((c \leq \bar{c})\) and merchant type is relatively low \(\delta \in [\delta^{np}, \delta^{np}]\). On the other hand, the optimal strategy for a merchant with relatively high marginal cost \((c > \bar{c})\) and type \(\delta \in [\delta^{np}, \delta^{np}]\) is to decrease the discount rate because she gains more from the decreased cannibalization effect than the loss from decreased sampling effect. This dynamic remains the same when the marginal cost is moderate \((c \leq \bar{c})\) and merchant type is relatively high \(\delta \in [\delta^{np}, \delta^{np}]\).

5. Merchant’s participation decision

A merchant offers a deal only if the optimal profit in participating in the publisher channel is higher than the profit in the benchmark case, that is \(\pi^*_m > \pi^*_b\). It is easy to see that a merchant’s optimal profit in the benchmark case, that is, \(\pi^*_b = \delta(1-c)^2/2 \ (§4.1)\), is linearly increasing in merchant type \(\delta\), and \(\pi^*_b(\delta = 0) = 0\). From Lemma 1, we know that a merchant’s optimal profit in offering a deal in the publisher channel is driven by the optimal discount rate strategy, which depends on merchant type \(\delta\), marginal cost \(c\) and the publisher’s revenue sharing ratio \(s\). We define a merchant type as the indifferent merchant type, \(\hat{\delta} \in (0,1)\), for whom the optimal
profit of offering a deal in the publisher channel is the same as the optimal profit in the benchmark case, \( \pi^*_m(\hat{\delta}) = \pi^*_p(\hat{\delta}) \).22

**LEMMA 3**: The value of the indifferent merchant type (\( \check{\delta} \)) decreases as the publisher’s revenue sharing ratio (\( s \)) increases, \( \frac{\partial \hat{\delta}}{\partial s} < 0 \).

In the first period, in the direct channel, a merchant of higher type has higher proportion of informed consumers who consider buying. Therefore, a merchant of higher type also has higher profit in the direct channel in the first period. Moreover, a merchant of higher type has higher proportion of informed consumers with higher willingness to pay for the experience good in the first period in the publisher channel. In the publisher channel, given a discount rate, a merchant of higher type has higher demand proportion, and hence, has a higher profit. Furthermore, a merchant of higher type, who offers the optimal discount rate, has higher profit in the second period in the direct channel. Therefore, as merchant type increases, the merchant’s optimal profit in offering a deal in the publisher channel increases.

When the revenue sharing ratio increases, the cost due to the revenue sharing effect increases, thus, the net gain to a merchant from offering a deal decreases. This implies that an increase in the publisher’s revenue sharing ratio leads to lowering of a merchant’s optimal profit in offering a deal in the publisher channel irrespective of merchant type. Hence, the value of indifferent merchant type (\( \check{\delta} \)) decreases as the revenue sharing ratio increases.

From the analysis in §4, we know that for a higher merchant type, the benefits from sampling and advertising effects are lower, and the cost due to cannibalization effect is higher. This implies that the net gain to a merchant from offering a deal in the publisher channel increases.

---

22 For a closed form expression of the indifferent merchant type (\( \check{\delta} \)), please see the Proof of Lemma 3 in Appendix 2.B.
decreases as the merchant type increases. Since a merchant of type $\delta > \hat{\delta}$ has lower profit in offering a deal in the publisher channel than the benchmark profit, $\pi^*_m(\delta) < \pi^*_b(\delta) \forall \delta > \hat{\delta}$, such a merchant does not offer a deal in the publisher channel. Conversely, a merchant of type $\delta \leq \hat{\delta}$ has higher or equal profit in offering a deal in the publisher channel than the benchmark profit, $\pi^*_m(\delta) \geq \pi^*_b(\delta) \forall \delta \leq \hat{\delta}$. Hence, a merchant of type $\delta \leq \hat{\delta}$ offers a deal in the publisher channel.

**PROPOSITION 5:** (a) If the marginal cost is relatively low, $c < 1/3$, then there exists a feasible indifferent merchant type $\hat{\delta}$ for any value of revenue sharing ratio. (b) If the marginal cost is relatively moderate, $1/3 \leq c < (1+2R)/3$, then the existence of a feasible indifferent merchant type $\hat{\delta}$ depends on the value of revenue sharing ratio. (c) If the marginal cost is relatively high, $c \geq (1+2R)/3$, then there is no feasible indifferent merchant type $\hat{\delta}$ for any value of revenue sharing ratio.

The value of the indifferent merchant type $\hat{\delta}$ critically depends on the marginal cost and the publisher’s revenue sharing ratio. The net gain to any merchant from offering a deal decreases as revenue sharing ratio increases. However, the net gain to any merchant from offering a deal increases as marginal cost decreases. When a daily-deal publisher charges a revenue sharing ratio close to one, the cost to the merchant due to revenue sharing effect is very high. However, if the marginal cost is low ($c < 1/3$), then even when revenue sharing ratio is close to one, for a merchant of sufficiently low type, the benefits of sampling and advertising effects are higher than the cost of revenue sharing effect. Therefore, when the marginal cost is low ($c < 1/3$), there exists some merchant types who are better off by offering a deal for any value of revenue sharing ratio.

When the marginal cost is moderate ($1/3 \leq c < (1+2R)/3$), the benefits of sampling and advertising effects are moderate to a merchant. If the revenue sharing ratio is high, then for a
merchant who has a higher value of $\delta$, the cost due to revenue sharing effect may be more than the benefits of offering a deal on the website. Therefore, for a very high value of revenue sharing ratio, no merchant will offer a deal in the publisher channel. When the marginal cost is relatively high ($c \geq (1+2R)/3$), no merchant finds it beneficial to offer a deal in the publisher channel, even if the revenue sharing ratio is zero. This is so because, for a merchant of any type with high marginal cost, the cost due to cannibalization effect is higher than the benefits from sampling and advertising effects.

From Lemma 1, we know that the merchants’ discount rate strategy and the optimal profit depend on *uninformed consumers’* quality estimate of a merchant’s experience goods ($R$) and marginal cost ($c$). Hence, a merchants’ participation decision is also impacted by values of $R$ and $c$. Note that as the value of $R$ increases, the value of indifferent merchant type ($\hat{\delta}$) also increases. It occurs because a merchant’s optimal discount rate decrease as the value of $R$ increases (Lemma 1). This leads to a decrease in the cannibalization effect and an increase in the optimal profit for any merchant type. On the other hand, as the marginal cost $c$ increases, benefits due to sampling and advertising effects decrease, and the optimal profit of offering a deal in the publisher channel decreases for any merchant type. Therefore, the value of indifferent merchant type decreases as marginal cost increases.

6. **Revenue sharing scheme of a daily-deal publisher**

The publisher’s optimal revenue sharing strategy takes into consideration a merchant’s response in term of optimal discount rate strategy ($\S4$) and the participation decision ($\S5$) which in turn depends on the publisher’s revenue sharing ratio. Before we analyze the publisher’s optimal revenue sharing ratio, we report some properties of publisher’s profit function.
**Lemma 4**: The daily-deal publisher’s profit from a participating merchant of type \( \delta \), who has marginal cost \( c \) and uninformed consumers’ quality estimate \( R \), increases in revenue sharing ratio \( s \), that is, \( \frac{\partial \pi_r(s; \delta, c, R)}{\partial s} > 0 \forall s \in \{ s : \delta \leq \hat{\delta}(s; c, R) \} \).

Lemma 4 reports that the daily-deal publisher’s revenue increases in revenue sharing ratio from a merchant as long as the revenue sharing ratio is such that the merchant participates, that is, the merchant’s optimal profit of offering a deal in the publisher channel is not less than the optimal profit in the benchmark case. The economic intuition for this is as follows. First consider a merchant of type \( \delta < \min\{ \hat{\delta}, \hat{\delta} \} \). The optimal discount rate for this merchant is \( d^* = 1 - R \) (Lemma 1) and this merchant does not reduce the discount rate as revenue sharing ratio increases as long as \( \delta < \hat{\delta} \) holds. From such a merchant, the publisher extracts more revenue by increasing revenue sharing ratio \( (s) \). Now consider the other case where the merchant type \( \delta \) is such that \( \delta \in [\hat{\delta}, \hat{\delta}] \). Even though such a merchant changes the optimal discount rate as publisher increases the revenue sharing ratio (Propositions 3(a), 3(b) and 3(c)), the publisher’s revenue increases in \( s \) as long as the merchant participates.

Before we analyze the optimal revenue sharing ratio of the publisher, we discuss two special cases, a) revenue sharing ratio \( s = 1 \), and b) revenue sharing ratio \( s = 0 \) (Figure 2.4). When the revenue sharing ratio is set to 1 \( (s = 1) \), any merchant who has marginal cost \( c > 1/3 \) does not participate. Similarly, any merchant of type \( \delta > 1/2 \) does not participate. This is so because the cost due to the revenue sharing effect is more than the benefits from offering a deal in the publisher channel. A merchant of type \( \delta \leq \hat{\delta}(s = 1; c, R) = \frac{1 - 4c + 3c^2}{2 - 2c(1 + R) + c^2(4 - 2R)} \) finds it beneficial to offer a deal (under the curve in Figure 2.4 left plot). As the uninformed consumers’ quality estimate \( (R) \) increases, the set of merchant types who offer deals increases.
Figure 2.4: A merchant’s participation decision with revenue sharing ratios, $s=0$ and $s=1$.

On the other hand, when the revenue sharing ratio is set to 0 ($s=0$), no merchant will participate if marginal cost is relatively large ($c > \frac{2R+1}{3}$). Moreover, when *uninformed consumers*’ quality estimate is relatively low ($R < 1/2$), a merchant of relatively large type ($\delta > \frac{(1+2R)^2}{1+12(1-R)R}$) does not offer a deal. A merchant of type $\delta \leq \delta(s=0; c, R) = \frac{1-3c(2-3c)+8R-4c(3+c)R-4(1-2c)R^2-4(1-c)(1+c-2R)R}{(1-3c)^2+4(3-c(3+2c))R-4(3-4c)R^2}$ finds it beneficial to offer a deal (under the curve in Figure 2.4 right plot). As the *uninformed consumers*’ quality estimate ($R$) increases, the set of merchant types who offer deals increases.

Now we provide a general approach that a publisher should follow to offer the optimal revenue sharing ratio to a merchant. Let the merchant type be $\delta_m$, the marginal cost be $c_m$, and the *uninformed consumers*’ quality estimate be $R_m$. The approach is as follows:
**Step 1:** Check $\delta_m > \hat{\delta}(s=0; c_m, R_m)$, if yes, then this merchant will not participate; if no, then go to Step 2.

**Step 2:** Check $\delta_m \leq \hat{\delta}(s=1; c_m, R_m)$, if yes then the optimal revenue sharing ratio $s^* = 1$; if no, go to Step 3.

**Step 3:** Offer the optimal revenue sharing ratio $s^* = \hat{\delta}^{-1}(\delta_m; c_m, R_m)$, where $\hat{\delta}(s; c, R)$ is given in Lemma 3.

From the above approach, it is recommended that the publisher should offer a customized revenue sharing ratio to merchant whose type is $\delta_m \in (\hat{\delta}(s=1; c_m, R_m), \hat{\delta}(s=0; c_m, R_m))$. Due to operational reasons, this general approach of customized revenue sharing ratio for each merchant may not be feasible. Therefore, to provide some managerial guidelines, we consider a setting where the daily-deal publisher offers two distinct revenue sharing ratios, $s \in \{s_L, s_H\}$, where $0 < s_L < s_H < 1$. The following proposition provides the publisher's optimal revenue sharing ratio strategy.

**PROPOSITION 6:** Given a merchant type $\delta_m$, marginal cost $c_m$ and uninformed consumers’ quality estimate $R_m$, the daily-deal publisher’s optimal revenue sharing strategy is to offer (a) a revenue sharing ratio $s = s_H$ to any merchant of type $\delta_m \leq \hat{\delta}(s=s_H; c_m, R_m)$; and (b) revenue sharing ratio, $s = s_L$, to any merchant of type, $\hat{\delta}(s=s_H; c_m, k_m) < \delta_m \leq \hat{\delta}(s=s_L; c_m, R_m)$.

The optimal revenue sharing strategy of the publisher is illustrated in Figure 2.5. The daily-deal publisher’s optimal revenue sharing ratio is $s_H$ for any merchant of type $\delta_m \leq \hat{\delta}(s=s_H; c_m, R_m)$ (highlighted in the Region 3 in Figure 2.5). Note that any merchant in Region 3 has relatively low merchant type and low marginal cost. These types of merchants have strong benefits from advertising and sampling effects and low cost due to cannibalization effect in
offering a deal in the publisher channel. Therefore, these merchant types are willing to participate even if the publisher charges a relatively high revenue sharing ratio \((s_H)\) leading to high cost to these merchants due to revenue sharing effect.

![Figure 2.5: Daily-deal publisher’s optimal revenue sharing strategy](image)

The daily-deal publisher’s optimal revenue sharing ratio is \(s_L\) for any merchant of type \(\hat{ \delta}(s_H; c_m, R_m) < \delta_m \leq \hat{ \delta}(s_L; c_m, R_m)\) (highlighted in the Region 2 in Figure 2.5). Note that any merchant in Region 2 has relatively moderate merchant type. These types of merchants have moderate benefits from advertising and sampling effects which decreases as marginal cost increases. Moreover, these merchant types have moderate cost due to cannibalization effect. Therefore, these merchant types do not participate if the publisher charges a relatively high revenue sharing ratio \((s_H)\), but are better off by participating when the publisher offers them lower revenue sharing ratio \((s_L)\). On the other hand, a merchant of relatively high type \(\delta_m > \hat{ \delta}(s_L; c_m, R_m)\) has low benefits, but high cost due to cannibalization effect. Therefore, these merchant types (Region 1 in Figure 2.5) do not participate.
7. Discussion

Internet technologies have made it increasingly feasible for advertisers and merchants to track the impact of advertising on consumers’ purchase actions. Performance-based advertising, like pay-per-purchase, allows merchants to pay to a publisher only if consumers’ purchases can be credited to a specific advertising stimulus. Merchants view this business model of performance-based advertising favorably because it limits their risks when investing in new and often untested advertising technologies (Mahdian and Tomak, 2008). This has led to popularity of daily-deal business models wherein a merchant pays to a daily-deal publisher only when consumers buy goods at discounted price in the publisher channel.

In this paper, we develop a theoretical framework to study the strategic interaction between a merchant and a monopolist daily-deal publisher where consumers are heterogeneous in their willingness to pay for quality of an experience good. In the context of daily-deal publishing, the market consists of two consumer segments; informed consumers, who are aware about the merchant and the quality of the experience good, and uninformed consumers who are not aware about the merchant in the direct channel and underestimate the quality in the publisher channel. Moreover, ours is a two-period model where a merchant offers the experience good in the direct channel at the regular price in both the periods and may offer the experience good in the publisher channel at a discounted price in the first period. Those uninformed consumers, who buy in the first period in the publisher channel, update their quality belief in the second period. In this framework, we analyze how a daily-deal publisher’s performance-based revenue sharing scheme impacts a merchant’s optimal discount rate strategy and participation decision.

In the context of daily-deal publishing, we identify four effects - two benefits and two costs - for a merchant who offers a deal in the publisher channel. While the advertising effect is
the benefit of reaching out to the merchant’s uninformed consumer segment in the publisher channel, the sampling effect is the benefit of incentivizing uninformed consumers to purchase the experience good at a discounted price in the publisher channel in the first period who subsequently update their belief about quality to true quality and may buy at the regular price in the second period. On the other hand, while the cannibalization effect is the loss of revenue from informed consumers, who have high willingness to pay but buy at the discounted price in the publisher channel, the revenue sharing effect is the payment by the merchant of a proportion of revenue generated on the website to the publisher based on revenue sharing scheme. While the value of discount rate affects the magnitude of the sampling, cannibalization and revenue sharing effects, it does not impact the magnitude of advertising effect. Moreover, marginal cost and merchant type also affect the magnitude of all the four effects. Hence, a merchant decides the optimal discount rate in the publisher channel based on the trade-offs of these four effects, which also affect the merchant’s participation decision.

In this framework, we find that a daily-deal publisher’s choice of revenue sharing ratio has a non-monotonic impact on a merchant’s optimal discount rate. When the publisher increases the revenue sharing ratio, while a merchant with low marginal cost and/or relatively low merchant type increases the optimal discount rate, a merchant with relatively high marginal cost and/or relatively high merchant type decreases the optimal discount rate. The economic intuition for this result is that when both marginal cost and/or merchant type are relatively low, a merchant can benefit from strong advertising and sampling effects by increasing the discount rate which leads to higher demand from informed and uninformed consumers in both the periods. As the publisher increases the revenue sharing ratio, the merchant’s profit decreases in the first period due to higher revenue sharing cost, as a result the merchant has higher incentive to
increase discount rate to acquire more uninformed consumers in the first period. On the other hand, if marginal cost and/or merchant type are relatively high, then advertising and sampling effects are not strong. Therefore a merchant decreases the optimal discount rate to reduce cannibalization effect as well as revenue sharing effect as the publisher increases the revenue sharing ratio. Nonetheless, in the context of daily-deal publishing (performance-based advertising), a merchant deviates from the optimal discount rate that she would offer in a nonperformance-based advertising context.

We also find that a merchant with low marginal cost and low merchant type benefits the most from offering a deal on a daily-deal publisher’s website. Hence, marginal cost, merchant type and the daily-deal publisher’s revenue sharing scheme affect the merchant’s participation decision. When a daily-deal publisher increases revenue sharing ratio, a merchant’s revenue sharing effect increases, and thus, her profit of offering a deal on the website decreases. Since a merchant participates and offers a deal on the website only if the associated profit is more than the benchmark profit, fewer merchant types will participate if the publisher increases the revenue sharing ratio.

The literature on performance-based advertising posits that a merchant’s pricing decision can be affected by an advertising publisher’s per action fee. Dellarocas (2012) shows that a merchant increases price when the advertiser moves from a nonperformance-based advertising scheme to a performance-based advertising scheme due to decentralization of the pricing decision. In the case of daily-deal advertising, a merchant’s choice of discount rate (and therefore, discounted price in the publisher channel) is also influenced by the daily-deal publisher’s choice of revenue sharing ratio. Hence, though the decentralized nature of a merchant’s decision of discount rate in the context of daily-deal publishing is similar to the one
in Dellarocas (2012), our model is distinct in three ways—two consumer segments, two channels and two periods. We show that in the context of daily-deal publishing a merchant may increase (decrease price) or decrease (increase price) the optimal discount rate under certain conditions when publisher moves from a nonperformance-based advertising scheme to a performance-based advertising scheme. This is different from Dellarocas (2012) where the impact is monotonic. This is because a merchant’s long-term and balanced considerations over two periods, two consumer segments and two channels change the dynamics of decentralized pricing decision.

Our paper provides a daily-deal publisher a set of guidelines about how a merchant would react to a performance-based revenue sharing contract. More importantly, we show that it is optimal for a daily-deal publisher to offer a customized revenue sharing ratio based on characteristics of the merchant and his consumer base. While it is optimal for the publisher to offer a higher revenue sharing ratio to a merchant with low marginal cost and low merchant type, the publisher should offer a lower revenue sharing ratio to a merchant with high marginal cost and high merchant type. Further, the publisher can ignore some merchants with very high marginal cost, as they will always find it unprofitable to offer a deal on the publisher’s website.

Our results also offer some insights about the economic dynamics of daily-deal websites. One of managerial recommendation is that the daily-deal website should reduce the new consumers’ uncertainty about the quality of the experience goods for which deals are offered. This may be achieved by providing consumer reviews about the merchants, or through some quality assurance certification. Another managerial recommendation is that the daily-deal website should design customized revenue sharing scheme for merchants taking into account expected quality parameter of the new consumers, marginal cost and the existing consumer base of the merchants. For example, certain professional services like dental and chiropractic services
require professional certifications, and the quality estimate of new consumers for these services is likely to be higher than a yoga classes. Hence, a daily-deal publisher should offer a lower revenue sharing ratio to a dental service provider compared to a yoga class provider, all else being similar. Similarly, the cost of serving an additional customer of a tanning service may be much smaller than that of a restaurant. Hence, a daily-deal publisher should offer a higher revenue sharing ratio to a tanning parlor than to a restaurant holding all else constant.

Our stylized two-period model has some limitations. We do not consider the effect of competition on a daily-deal website’s strategy. Competition between rival daily-deal websites will potentially exert downward pressure on revenue sharing ratio. We also assume uninformed consumers are rational that they buy in the direct channel in the second period if their updated willingness to pay is higher than the regular price. This may not be true in the reality where many consumers on the daily-deal website are deal seeking (Dholakia, 2012). This implies that these deal seeking consumers will return to the merchant in future only when the merchant offers another deal. Moreover, there may be other motivations for a merchant to offer a deal on the daily-deal website. For example a merchant may have operational considerations, like excessive capacity utilization, and/or demand smoothing. These operational considerations present different dynamics, like channel coordination and production planning, other than pricing/advertising considered in this paper. We can also study the impact of different types of payment schemes, like fixed fee or two-part tariff fee, on a merchant’s discount rate strategy. We leave these possible extensions for future research.
References


## Appendix 2.A: Summary of Notations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>proportion of consumers who are present on the daily-deal website</td>
</tr>
<tr>
<td>$\theta$</td>
<td>consumer valuation for quality</td>
</tr>
<tr>
<td>$\delta$</td>
<td>proportion of informed consumers for a merchant</td>
</tr>
<tr>
<td>$\delta^L$</td>
<td>lower merchant type</td>
</tr>
<tr>
<td>$\delta^H$</td>
<td>higher merchant type</td>
</tr>
<tr>
<td>$\hat{\delta}$</td>
<td>merchant type who is indifferent between offering a deal in the publisher channel and not offering a deal in the publisher channel.</td>
</tr>
<tr>
<td>$R$</td>
<td>uninformed consumers’ quality estimate</td>
</tr>
<tr>
<td>$q$</td>
<td>true quality of experience goods</td>
</tr>
<tr>
<td>$p$</td>
<td>merchant’s regular price</td>
</tr>
<tr>
<td>$d$</td>
<td>discount rate on the website in the first period</td>
</tr>
<tr>
<td>$D^w_{1i}$</td>
<td>demand proportion from uninformed consumers in the publisher channel in the first period</td>
</tr>
<tr>
<td>$D^w_{2i}$</td>
<td>demand proportion uninformed consumers in the direct channel in the first period</td>
</tr>
<tr>
<td>$D^w_{1u}$</td>
<td>demand proportion from uninformed consumers in the publisher channel in the first period</td>
</tr>
<tr>
<td>$D^w_{2u}$</td>
<td>demand proportion uninformed consumers in the direct channel in the second period</td>
</tr>
<tr>
<td>$c$</td>
<td>marginal cost</td>
</tr>
<tr>
<td>$s$</td>
<td>daily-deal website’s revenue sharing ratio</td>
</tr>
<tr>
<td>$\pi_p$</td>
<td>merchant’s profit in the benchmark case where the merchant does not offer a deal</td>
</tr>
<tr>
<td>$\pi_{p,w}$</td>
<td>merchant’s profit in offering a deal in the performance-based advertising publisher channel</td>
</tr>
<tr>
<td>$\pi_{w}$</td>
<td>publisher’s profit from a merchant who offers a deal in the publisher channel</td>
</tr>
<tr>
<td>$\pi_{w,m}$</td>
<td>merchant’s profit in offering a deal in the nonperformance-based advertising publisher channel</td>
</tr>
<tr>
<td>$\superscript{NP}$</td>
<td>the superscript indicates the case of nonperformance-based advertising scheme.</td>
</tr>
</tbody>
</table>
Appendix 2.B: Proofs of Lemmas and Propositions

PROPOSITION 1: A merchant’s optimal discount rate $d^*$ is bounded such that
\[
\max\{d, 0\} \leq d^* \leq 1 - R, \text{ where } d = 1 - \frac{2R}{1 + c}.
\]

Proof of Proposition 1:

(a) The optimal demand proportion from uninformed consumers in the first period in the publisher channel is the same as the demand from uninformed consumers in the direct channel in the second period, that is $D_{u1}^w = D_{su}^0$.

(b) Moreover, since the price in the direct channel is $p$ for both informed consumers as well as uninformed consumers, and some uninformed consumers who do not buy in the publisher channel in the first period may have lower quality estimate ($R < q$) in the second period, the demand proportions from uninformed consumers and informed consumers must be such that $D_{u1}^w \leq D_{si}^0 = 1 - p$.

From (a) and (b) above, we have $\max D_{u1}^w = 1 - p$.

If the merchant offers discount $d$ in the publisher channel, then $D_{u1}^w = 1 - \frac{p(1-d)}{R}$. Substituting $p = \frac{1+c}{2}$ and solving $1 - \frac{p(1-d)}{R} = 1 - p$, we get the discount rate that corresponds to this maximum demand proportion $D_{u1}^w$ as: $d = 1 - R$. Hence, this is the maximum discount rate that a merchant may offer in the publisher channel.

By solving $D_{u1}^w = 1 - \frac{p(1-d)}{R} = 0$, we get
\[
d = 1 - \frac{2R}{1+c} > 0, \text{ if } c > 2R - 1; \text{ and } d = 0, \text{ if } c \leq 2R - 1.
\]

Therefore, the minimum discount rate is $\max\{1 - \frac{2R}{1+c}, 0\}$.

Hence, we have Proposition 1. ■

LEMMA 1: There exists a lower merchant type $\underline{\delta}$ and a higher merchant type $\overline{\delta}$ such that (i) for any merchant of type $\delta \in [0, \overline{\delta})$ the optimal discount rate is $d^* = 1 - R$; (ii) for any merchant of
type $\delta \in [\delta, \bar{\delta})$ the optimal discount rate is $d' = \hat{d} = \frac{(1-c+2(1-R)(1-s))(1-\delta) - 2cs(1-(1-R)\delta))}{2(1+c)(1-s)(1-(1-R)\delta)}$; and

(iii) for any merchant of type $\delta \in [\delta, 1]$ the optimal discount rate is $d' = 0$. The lower merchant type $\underline{\delta}$ and the higher merchant type $\bar{\delta}$ are:

$$\underline{\delta} = \begin{cases} \frac{1-c(3-2R(1-s))}{1-c-2(1-R)(c-R(1+c)(1-s))} & \text{if } 0 \le s \le s' \\ 0 & \text{if } s' < s < 1 \end{cases}, \quad \bar{\delta} = \begin{cases} \frac{2csR}{1-c+2(1-R)(1-s(1+c))} & 0 \le c \le c' \\ \frac{1-3c+2R(1-s)}{1-c-2(1-R)(c-R(1+c)(1-s))} & c' < c < 1 \end{cases},$$

where $s' = \frac{1-3c+2cR}{2cR}$, and $c' = 2R-1$.

**Proof of Lemma 1:**

Substitute $D_{\mu}^U = 1-p(1-d)$, $D_{\mu}^W = D_{\mu}^G = 1-p(1-d)/R$, $D_{\nu}^O = D_{\nu}^G = 1-p$, and $p = \frac{1+c}{2}$, into the merchant’s profit function in Equation (6), we get:

$$\pi_M = \mu(1-c(\frac{1+c}{2}(1-d))) + (1-\delta)(1-\frac{1+c}{2R}(1-d))(\frac{1+c}{2}(1-d)(1-s-c)) + (1-\mu)\delta(1-\frac{1+c}{2}) - (\frac{1+c}{2} + (\delta(\frac{1+c}{2} + \mu(1-\delta))(1-\frac{1+c}{2}(1-d)) - \frac{1-c}{2})$$

Differentiate $\pi_M$ with respect to discount rate $d$, and solving the first order condition, and we have the optimal discount rate $\hat{d} = \frac{(1-c)(1-\beta) + 2(1-R)(1-s)(1-\delta) - 2cs(1-(1-R)\delta)}{2(1+c)(1-s)(1-(1-R)\delta)}$.

We know from Proposition 1 that the maximum discount rate is $1-R$ and minimum discount rate is $\max\{1-\frac{2R}{1+c}, 0\}$. Now, we get the merchant types that correspond to the boundaries of the discount rate $\hat{d}$ wherein $\underline{\delta}$ corresponds to the value of $\delta$ for which $\hat{d}(\delta = \underline{\delta}) = 1-R$, and $\bar{\delta}$ corresponds to the value of $\delta$ for which $\hat{d}(\delta = \bar{\delta}) = \max\{1-\frac{2R}{1+c}, 0\}$.

We define $c' = 2R-1$.

Solving $\hat{d}(\delta = \underline{\delta}) = 1-R$, we get $\underline{\delta} = \frac{1-c(3-2R(1-s))}{1-c-2(1-R)(c-R(1+c)(1-s))}$. Now, $\underline{\delta} \ge 0$ if $s \le \frac{1-3c+2cR}{2cR}$, and $\underline{\delta} = 0$ if $s > \frac{1-3c+2cR}{2cR}$. We define $s' = \frac{1-3c+2cR}{2cR}$.

From Proposition 1, if $0 \le c \le 2R-1 = c'$, then minimum discount rate is $d=0$. By solving $\hat{d}(\delta = \bar{\delta}) = 0$, we get $\bar{\delta} = 1-\frac{2csR}{1-c+2(1-R)(1-s(1+c))}$.
From Proposition 1, if \( 2R-1=c'<c<1 \), the minimum discount rate is \( d=1-\frac{2R}{1+c} \). By solving \( \hat{d}(\delta) = 1-\frac{2R}{1+c} \), we get \( \delta = \frac{1-3c+2R(1-s)}{1-c-2(1-R)(c-2R(1-s))} \).

When \( s'<s<1 \), the lower merchant type \( \delta = 0 \), and the highest discount rate is \( \bar{d} = \hat{d}(\delta=0) = \frac{3-2c-2R-2(c-R)s+2(1+c)(1-s)}{2(1+c)(1-s)} \). Moreover, since \( s'<s<1 \), \( \bar{d} < 1-R \).

Hence, we have Lemma 1. ■

**PROPOSITION 2 (a):** When the revenue sharing ratio \( s \) increases, the lower merchant type \( \delta \) (i) increases when marginal cost is relatively low, \( c<\bar{c} \), (ii) decreases when marginal cost is relatively high, \( c>\bar{c} \), and (iii) does not change when marginal cost \( c=\bar{c} \), where \( \bar{c} = \frac{3-2R-\sqrt{9-8(2-R)R}}{2R} \).

**Proof of Proposition 2 (a):**

From Lemma 1, \( \delta = \frac{1-c(3-2R(1-s))}{1-c-2(1-R)(c-2R(1+c)(1-s))} \), when \( s \leq s' \), equivalently \( c \leq \frac{1}{3-2R(1-s)} \).

Taking the first derivative of \( \delta \), we get \( \frac{\partial \delta}{\partial s} = \frac{2R(1-3c-(1-c(2+c))R)}{(1-c(3-2R(2-R(1-s)-s)+2(1-R)(1-R)(1-s))^2} \).

By solving \( \frac{\partial \delta}{\partial s} = 0 \), we have \( c = \bar{c} = \frac{3-2R-\sqrt{9-8(2-R)R}}{2R} \).

It is easy to see that, \( \frac{\partial \delta}{\partial s} > 0 \) when \( c<\bar{c} \); \( \frac{\partial \delta}{\partial s} = 0 \) when \( c=\bar{c} \); and \( \frac{\partial \delta}{\partial s} < 0 \) when \( \bar{c} < c \).

Hence, we have Proposition 2 (a). ■

**PROPOSITION 2 (b):** When the revenue sharing ratio \( s \) increases, higher merchant type \( \bar{\delta} \) (i) increases when the uninformed consumers’ quality estimate is small, \( R<1/2 \), and the marginal cost is sufficiently small, \( c<\bar{c} \), and (ii) decreases when the uninformed consumers’ quality estimate is large, \( R \geq 1/2 \), where \( \bar{c} = \frac{1-2R}{3-4R} \).

**Proof of Proposition 2 (b):**

From Lemma 1:

When \( 0 \leq c \leq c' \), \( \bar{\delta} = 1-\frac{2csR}{1-c+2(1-R)(1-s(1+c))} \), and we get \( \frac{\partial \bar{\delta}}{\partial s} = \frac{-2cR(3-c-2R)}{(c+2R-3+2(1+c)(1-R)s)^2} < 0 \).
When $c' < c < 1$, $\delta = \frac{1 - 3c + 2R(1-s)}{1 - c - 2(1-R)(c - 2R(1-s))}$, and we get $\frac{\partial \delta}{\partial s} = \frac{2R(1-2R-c(2-4R)-c)}{(1-c(3-2R)+4(1-R)R(1-s))^2}$.

When $R \leq 1/2$, if $c < \frac{1-2R}{3-4R}$, then $\frac{\partial \delta}{\partial s} > 0$; if $c = \frac{1-2R}{3-4R}$, then $\frac{\partial \delta}{\partial s} = 0$; if $c > \frac{1-2R}{3-4R}$, then $\frac{\partial \delta}{\partial s} < 0$.

When $R > 1/2$, $\frac{1-2R}{3-4R} < 0$. Hence, $c > \frac{1-2R}{3-4R}$ and, therefore, $\frac{\partial \delta}{\partial s} < 0$.

Hence, we have Proposition 2 (b). ■

**PROPOSITION 3 (a):** When marginal cost is relatively small ($c < \tilde{c}$) and uninformed consumers’ quality estimate is small ($R < 1/2$), the optimal discount rate offered by a merchant of type $\delta \in [\underline{\delta}, \overline{\delta})$ increases as revenue sharing ratio ($s$) increases.

**Proof of Proposition 3 (a):**

For merchant type $\delta \in [\underline{\delta}, \overline{\delta})$, the optimal discount rate is $d = \hat{d}$ (From Lemma 1).

Taking the first derivative of $\hat{d}$ with respect to $s$, we get: $\frac{\partial \hat{d}}{\partial s} = \frac{1 - \delta + c(-3 + (3-2R)\delta)}{2(1+c)(-1+s)^2(1+(1+R)\delta)}$.

By solving for $\frac{\partial \hat{d}}{\partial s} = 0$, we have $\delta = \tilde{\delta} = \frac{1-3c}{1-3c+2cR}$.

This implies if $\delta < \tilde{\delta}$, then $\frac{\partial \hat{d}}{\partial s} > 0$.

By setting $\delta = \tilde{\delta}$, and solving we get $c = \hat{c} = \frac{1-2R}{3-4R}$ (From Proposition 2 (b))

If $R < 1/2$, then $\hat{c} > 0$. Therefore, if $c < \hat{c}$, then $\tilde{\delta} < \delta$. This implies that for all merchant type $\delta \in [\underline{\delta}, \tilde{\delta})$, $\frac{\partial \hat{d}}{\partial s} > 0$.

Hence, we have Proposition 3 (a). ■

**PROPOSITION 3 (b):** If marginal cost is moderate ($\hat{c} \leq c \leq \bar{c}$), then as the revenue sharing ratio ($s$) increases, (i) the optimal discount rate increases for a merchant of type $\delta \in [\underline{\delta}, \tilde{\delta})$; (ii) the optimal discount rate decreases for a merchant of type $\delta \in (\tilde{\delta}, \overline{\delta})$; and (iii) the optimal discount rate does not change for a merchant of type $\delta = \overline{\delta}$, where $\tilde{\delta} = \frac{1-3c}{1-3c+2cR}$.

**Proof of Proposition 3 (b):**

Continuing from the proof of Proposition 3 (a),
By setting $\tilde{\delta} = \hat{\delta}$, and solving we get $c = \tilde{c} = \frac{3 - 2R - \sqrt{9 - 8(2 - R)R}}{2R}$. This implies that if $\tilde{c} \leq c \leq \bar{c}$, $\tilde{\delta} \in [\tilde{\delta}, \bar{\delta}]$.

Therefore, when $\delta < \tilde{\delta}$, $\frac{\partial \tilde{d}}{\partial s} > 0$; when $\delta = \tilde{\delta}$, $\frac{\partial \tilde{d}}{\partial s} = 0$; and when $\delta > \tilde{\delta}$, $\frac{\partial \tilde{d}}{\partial s} < 0$.

Hence, we have Proposition 3 (b). □

**PROPOSITION 3 (c):** When marginal cost is relatively high, $c > \bar{c}$, as revenue sharing ratio ($s$) increases, the optimal discount rate of a merchant of type $\delta \in [\tilde{\delta}, \bar{\delta})$ decreases.

**Proof of Proposition 3(c):**

Continuing from the proof of Proposition 3 (b);

When $c > \bar{c}$, the critical merchant type $\tilde{\delta} < \hat{\delta}$. This implies that for any $\delta \in [\tilde{\delta}, \bar{\delta})$, $\delta > \tilde{\delta}$

Therefore, for any $\delta \in [\tilde{\delta}, \bar{\delta})$, $\frac{\partial \tilde{d}}{\partial s} < 0$.

Hence, we have Proposition 3 (c). □

**PROPOSITION 4 (a):** When the daily-deal publisher adopts a nonperformance-based scheme, compared to a performance-based scheme, a merchant offers a higher discount rate ($d^{NP} > d^*$), if the marginal cost is relatively low ($c < \bar{c}$) and type is $\delta \in [\tilde{\delta}^{NP}, \bar{\delta}^{NP})$, or the marginal cost is moderate ($\tilde{c} \leq c \leq \bar{c}$) and type is relatively low ($\delta \in [\tilde{\delta}^{NP}, \bar{\delta}]$).

**PROPOSITION 4 (b):** When the daily-deal publisher adopts a nonperformance-based scheme, compared to a performance-based scheme, a merchant offers a lower discount rate ($d^{NP} < d^*$), if the marginal cost is relatively high ($c > \bar{c}$) and type is $\delta \in [\tilde{\delta}^{NP}, \bar{\delta}^{NP})$, or the marginal cost is moderate and type is relatively high ($\delta \in (\tilde{\delta}, \bar{\delta}^{NP})$).

**Proof of Propositions 4 (a) and 4(b):**

In the case of nonperformance-based advertising, a merchant’s payment to a publisher for offering a deal does not depend on the merchant’s revenue. Hence, the merchant’s discount rate strategy is equivalent to the case where the revenue sharing ratio is zero in the case of performance-based advertising, that is, $d^{NP} = d^*(s = 0)$. 

143
On the other hand, in the case of performance-based advertising, the revenue sharing ratio is greater than zero \((s > 0)\). Hence, to compare a merchant’s optimal discount rate in the case where the daily-deal publisher adopts a nonperformance-based scheme to that of a performance-based scheme, it is equivalent to compare the change in a merchant’s discount rate strategy due to decrease in the revenue sharing ratio.

From Propositions 3 (a), the optimal discount rate increases as \(s\) increases, when \(c < \hat{c}\), \(R < 1/2\), and \(\delta \in [\hat{\delta}, \overline{\delta})\).

From Proposition 3 (b) (i), the optimal discount rate increases as \(s\) increases, when \(\hat{c} \leq c \leq \bar{c}\) and \(\delta \in [\hat{\delta}, \bar{\delta})\).

Therefore, a merchant offers a higher discount rate in the case of nonperformance-based advertising than in the case of performance-based advertising \((d^{np} > d^\ast)\), if the marginal cost is relatively low \((c < \hat{c})\) and type is \(\delta \in [\hat{\delta}^{np}, \overline{\delta}^{np})\), or the marginal cost is moderate \((\hat{c} \leq c \leq \bar{c})\) and type is relatively low \((\delta \in [\hat{\delta}^{np}, \bar{\delta}^{np})\).

Hence, we have Proposition 4 (a).

From Proposition 3 (b) (ii), the optimal discount rate increases as \(s\) decreases, when \(\hat{c} \leq c \leq \bar{c}\) and \(\delta \in (\hat{\delta}, \bar{\delta})\).

From Proposition 3 (c), the optimal discount rate decreases as \(s\) decreases, when \(c > \hat{c}\) and \(\delta \in [\hat{\delta}, \overline{\delta})\).

Therefore, a merchant offers a lower discount rate in the case of nonperformance-based advertising than in the case of performance-based advertising \((d^{np} < d^\ast)\), if the marginal cost is relatively high \((c > \hat{c})\) and type is \(\delta \in [\hat{\delta}^{np}, \overline{\delta}^{np})\), or the marginal cost is moderate \((\hat{c} \leq c \leq \bar{c})\) and type is relatively high \((\delta \in (\hat{\delta}, \overline{\delta}^{np}))\).

Hence, we have Proposition 4 (b).

**Lemma 3:** The value of the indifferent merchant type \((\hat{\delta})\) decreases as the publisher’s revenue sharing ratio \((s)\) increases, \(\frac{\partial \hat{\delta}}{\partial s} < 0\).
Proof of Lemma 3:

From Lemma 1, we get the optimal profit of a merchant of type $\delta$ who offers a deal in the publisher channel and adopts optimal discount rate strategy. When $s \leq s'$ (From Lemma 1), the optimal profit of offering a deal, for a merchant $\delta \in \mathbb{D}$, is

$$\pi_M^* = \pi_M(d = 1-R) = \frac{1}{4}((1-c)\mu(1-c(3-R(1-s)) + R(1-s)) + 2(1-c)^2 \delta + \mu((1-R)R(1-s)) + 2c(1+R(2-R(1-s))-4c^2(4-R(3-R(1-s)-s))(1-2\delta))$$

Taking first derivative of $\pi_M^*$ with respect to $s$, we get

$$\frac{d\pi_M^*}{ds} = \frac{1}{16(1-s)(1-R)\delta}\mu((\delta - 1)(1-3c) + 2R\delta c)^2 - 4R^2(1-s)^3$$

By solving $\frac{d\pi_M^*}{ds} = 0$, we get that if $\delta > \frac{1-3c-2R(1-s)}{1-3c+2cR}$, then $\frac{d\pi_M^*}{ds} < 0$.

Since $\frac{1-3c-2R(1-s)}{1-3c+2cR} < \hat{\delta}$, therefore, for $\delta \in [\hat{\delta}, \bar{\delta})$, $\frac{d\pi_M^*}{ds} < 0$.

Solving $\pi_M^*(\hat{\delta} | \bar{\delta}) = \pi_M^*(\bar{\delta})$, we have:

$$\hat{\delta} = \frac{1}{4(1+c)^2(1-R)R}((1-c)(1+(16-c)c)R-2(c(7+c(2+c))-2)-4(1+c)^2 R^2 \sqrt{4(6+c(5-14-c))c(21-(2-c)c)-8})R+(1+c)^2(9-(22-c)c)R^2$$

When $s > \hat{s}$, $\hat{\delta} (s) \in [0, \bar{\delta})$. Solving $\pi_M^* = \pi_M(d = 1-R) = \pi_M^*$, we have:

$$\hat{\delta} (s) = \hat{\delta} = \frac{4c - R(1-s) - c^2(3-R+sR)-1}{2c(1+R(2-R(1-s)))-c^2(4-R(3-R(1-s)-s))+(1-R)R(1-s)-2}$$
When $s \leq \hat{s}$, $\hat{s} (s) \in (\hat{\delta}, \overline{\delta})$. Solving $\pi^*_m = \pi^*_m (d = \hat{d}) = \pi^*_b$, we have:

$$
\hat{\delta} (s) = \hat{\delta}^*_2 = (1 - 2c(3 + R(6 - 4(1 - s) - 7R))) + 4(2 - R)R(1 - s) + c^2(9 - 2R(2 + s)) - \\
2\sqrt{(1 - c)R^2(4 + c(-4 + s) - 3s)(1 - s)((1 + c - 2R)^2 + 4(1 - R)Rs)) / \\
(1 - 2c(3 + 2R(3 - 4(1 - s) - 4s)) + 12(1 - R)R(1 - s) + c^2(9 - 8R - 4(1 - R)Rs))}
$$

Since the profit of offering a deal in the publisher channel decreases as revenue sharing ratio increase, that is $\frac{\partial \pi^*_m}{\partial s} < 0$, and the benchmark profit $\pi^*_b$ is independent of revenue sharing ratio, the value of $\hat{\delta}$ decreases, that is $\frac{\partial \hat{\delta}}{\partial s} < 0$.

Hence, we have Lemma 3. ■

PROPOSITION 5: If the marginal cost is relatively low, $c < 1/3$, then there exists a feasible indifferent merchant type $\hat{\delta}$ for any value of revenue sharing ratio. If the marginal cost is relatively moderate, $1/3 \leq c < (1 + 2R)/3$, then the existence of a feasible indifferent merchant type $\hat{\delta}$ depends on the value of revenue sharing ratio. If the marginal cost is relatively high, $c \geq (1 + 2R)/3$, then there is no feasible indifferent merchant type $\hat{\delta}$ for any value of revenue sharing ratio.

Proof of Proposition 5:

From Lemma 3, when the revenue sharing ratio is very close to 1 from below, $s \to 1$, we have

$$
\hat{\delta}_{\tau = 1} = \hat{\delta}_1 (s = 1) = \frac{1 - 4c + 3c^2}{2 - 2c(1 + R) + 2c^2(2 - R)}.
$$

By solving $\hat{\delta}_{\tau = 1} = 0$, we find that if $c < \frac{1}{3}$, then $\hat{\delta}_{\tau = 1} > 0$.

When the revenue sharing ratio is zero, $s = 0$, we have

$$
\hat{\delta}_{\tau = 0} = \hat{\delta}_2 (s = 0) = \frac{1 - 3c(2 - 3c) + 8R - 4c(3 + c)R - 4(1 - 2c)R^2 - 4(1 - c)(1 + c - 2R)R}{(1 - 3c)^2 + 4(3 - c(3 + 2c))R - 4(3 - 4c)R^2}.
$$

By solving $\hat{\delta}_2 (s = 0) = 0$, we find that if $c \geq \frac{1 + 2R}{3}$, then $\hat{\delta}_{\tau = 0} \leq 0$. This implies that if $c \geq \frac{1 + 2R}{3}$, then no merchant type will participate.
From the proof of Lemma 3, if $1/3 \leq c < (1 + 2R)/3$, then the existence of a feasible indifferent merchant type $\delta$ depends on the value of revenue sharing ratio.

Hence, we have Proposition 5. ■

**LEMMA 4:** The daily-deal publisher’s profit from a participating merchant of type $\delta$, who has marginal cost $c$ and uninformed consumers’ quality estimate $R$, increases in revenue sharing ratio $s$, that is, $\frac{\partial \pi_w(s|\delta,c,R)}{\partial s} > 0 \forall s \in \{s: \delta \leq \hat{\delta}(c,R|s)\}$.

**Proof of Lemma 4:**

The profit function of a daily-deal publisher who offers a revenue sharing contract with a revenue sharing ratio $(s)$ to a merchant of type $\delta$ is:

$$\pi_w(s) = \begin{cases} s\mu(\delta D^W_\mu p(1-d) + (1-\delta)D^W_\mu p(1-d)) & \text{if } \pi_w(\delta,d') \geq \pi_w(\delta) \\ 0 & \text{if } \pi_w(\delta,d') < \pi_w(\delta) \end{cases}$$ (20)

From Equation 1 (Equation 7 in the paper), the publisher’s profit from a participating merchant is:

$$\pi_w(s) = s\mu(\delta D^W_\mu p(1-d) + (1-\delta)D^W_\mu p(1-d)).$$

If $\delta \in [0,\hat{\delta})$ and $\delta \leq \hat{\delta}$, then $d = 1 - R$, and the publisher’s profit is

$$\pi_w = s \left\{ \frac{1}{4}(1+c)R\mu(1-c+(1+c)(1-k)\delta) \right\}.$$ And it follows that $\frac{\partial \pi_w}{\partial s} > 0$, as long as $\delta \leq \hat{\delta}$.

If $\delta \in [\hat{\delta},\bar{\delta})$ and $\delta \leq \hat{\delta}$, then $d = \hat{\delta}$ and the publisher’s profit is

$$\pi_w = \mu s \frac{4R^2(1-s)^2 - ((1-\delta)(1-c)-2R\delta c)^2}{16(1-s)^2 R(1 -(1-R)\delta)}.$$ Solving $\frac{\partial \pi_w}{\partial s} = 0$,

we have that if $\delta > \frac{1 + s - 2c(3-R)(1+s) + 3c^2(3-2R)(1+s) - 2R(1-s)\sqrt{(1-c(3-2R))^2(1-s^2)} - c^2(1+s)(3-2R)\delta^2}{(1-c(3-2R))^2(1+s)}$,

then $\frac{\partial \pi_w}{\partial s} > 0$. Since $\frac{1 + s - 2c(3-R)(1+s) + 3c^2(3-2R)(1+s) - 2R(1-s)\sqrt{(1-c(3-2R))^2(1-s^2)} - c^2(1+s)(3-2R)\delta^2}{(1-c(3-2R))^2(1+s)} < \delta$,
therefore, for \( \delta \in [\tilde{\delta}, \delta_0] \), \( \frac{\partial \pi_w}{\partial \delta} > 0 \).

Hence, we have Lemma 4. ■

**PROPOSITION 6:** Given a merchant type \( \delta_m \), marginal cost \( c_m \) and uninformed consumers’ quality estimate \( R_m \), the daily-deal publisher’s optimal revenue sharing strategy is to offer (a) a revenue sharing ratio \( s = s_H \) to any merchant of type \( \delta_m \leq \hat{\delta}(s = s_H; c_m, R_m) \); and (b) revenue sharing ratio, \( s = s_L \), to any merchant of type, \( \hat{\delta}(s = s_H; c_m, k_m) < \delta_m \leq \hat{\delta}(s = s_L; c_m, R_m) \).

**Proof of Proposition 6:**

From Lemma 3, we have that merchant participation decreases as revenue sharing ratio increases.

From Lemma 4, we have that the publisher’s revenue increases as revenue sharing ratio increases.

Therefore, from Lemma 3 and Lemma 4 we have that it is optimal for a publisher to offer the highest revenue sharing ratio such that the merchant participates.

Hence, the optimal revenue sharing ratio the publisher offers is the high revenue sharing ratio \( (s_H) \) to a merchant who is willing to participate, \( \delta_m \leq \hat{\delta}(s = s_H; c_m, R_m) \), and offers the low revenue sharing ratio \( (s_L) \) to a merchant who is willing to participate, but does not want to participate if given the high sharing ratio, that is \( \hat{\delta}(s = s_H; c_m, k_m) < \delta_m \leq \hat{\delta}(s = s_L; c_m, R_m) \).

Hence, we have Proposition 6. ■

**Merchant’s optimal discount rate strategy when some of the uninformed consumers underestimate while others overestimate the quality of experience goods in the publisher channel in the first period.**

Let’s assume that \( \beta \) proportion of the uninformed consumers overestimate the quality, \( R_H > 1 \), and \( 1 - \beta \) proportion of the uninformed consumers underestimate the quality, \( R_L < 1 \).

A consumer’s surplus from buying the experience good in the first period is:

\[
U_i(\theta) = \begin{cases} 
\theta - p & \text{informed consumers in the direct channel} \\
\theta - p(1-d) & \text{informed consumers in the publisher channel} \\
\theta R_i - p(1-d), i \in \{H, L\} & \text{uninformed consumers in the publisher channel}
\end{cases}
\]
A consumer’s surplus from buying the experience good in the direct channel in the second period is:

\[ U_2(\theta) = \begin{cases} 
\theta - p & \text{informed consumers, and uninformed consumer who bought in first period} \\
\theta R_i - p, i \in \{H, L\} & \text{uniformed consumer in the publisher channel who did not buy in first period}
\end{cases} \]

The uninformed consumers who overestimate the quality of the experience goods, have surplus \( U_1(\theta) = \theta R_i - p(1-d) \) in the first period in the publisher channel, and upon consumption of the experience goods, their surplus in the second period is \( U_2(\theta) = \theta - p \).

Since \( R_i > 1 \), some of the uninformed consumers who buy in the first period, do not buy in the second period, and therefore the demand proportion of uninformed consumers who overestimate the quality of the experience good is \( \beta \) in the second period.

The merchant’s two-period profit in offering a deal in publisher channel is:

\[ \pi^m(\beta, d) = \mu(\delta D^R_i + (1-\delta)D^L_i)(p(1-d)(1-s)-c) + (1-\mu)(\delta D^L_i - p-c) + (\beta(1-d) + (1-\beta)(1-p(1-d)/R_i)), \]

where \( D^R_i = 1-p(1-d) \), \( D^L_i = \beta(1-p(1-d)/R_i) + (1-\beta)(1-p(1-d)/R_i) \), \( D^0_i = D^L_i = 1-p \).

\[ D^0_i = \beta(1-p) + (1-\beta)(1-p(1-d)/R_i). \]

Note that when \( \beta = 0 \) and \( R_i = R \), the profit in (A1) is the same as in (6) in the paper.

It is easy to see that the boundaries for the optimal discount rates in Proposition 1 hold in the case of heterogeneous estimate of quality of the experience among uninformed consumers.

Taking derivative of the profit function in (A1) and solving the first order condition, we get:

\[ \hat{\beta} = \frac{2R_i(1-s-cs)\beta(1-\delta) - R_i((1-\delta)(2R_i(1-s) - 3 - 2s)(1-\beta)) - c((1-\beta)(1-\delta) + 2s(1-\beta(1-\delta) - (1-R_i)\delta)))}{2(1+c)(1-s)(R_i\beta(1-\delta) + R_i(1-\beta(1-\delta) - (1-R_i)\delta))} \]

We take the derivative of \( \beta \), and we get

\[ \frac{\hat{\beta}}{\hat{\beta}} = -\frac{R_i R_i(1-\delta)(1-c) + 2(1-s)(R_i - R_i) + (c)(1-R_i)\delta}{2(1+c)(1-s)(R_i\beta(1-\delta) + R_i(1-\beta(1-\delta) - (1-R_i)\delta))^2} < 0. \]

This implies that the optimal discount rate \( \hat{\beta}(\delta, \beta) \) decreases as the proportion of the uninformed consumers who overestimate the quality of the experience goods ( \( \beta \) ) increases.

The sampling effect decreases and the revenue sharing effect increases as the proportion of the uninformed consumers, who overestimate the quality of the experience goods, increases. The
advertising and cannibalization effects remain unchanged. It implies that the overall benefit of offering a deal by a merchant decreases. Hence, the optimal discount rate of a merchant ( \( \hat{d} \)) is lower compared to the case where \( \beta = 0 \), that is, all uninformed consumers underestimate the quality of experience good (the model setup in the paper).

The merchant’s participation decision is also negatively affected by the proportion of the uninformed consumers, who overestimate the quality of the experience goods. Thus, the marginal merchant type ( \( \hat{\delta} \)) is lower compared to the case where \( \beta = 0 \), that is, all uninformed consumers underestimate the quality of experience good (the model setup in the paper).

Having a higher value of \( R_H \) has similar effect as having a large \( \beta \). The benefits of offering a deal by a merchant is lower compared to the case where \( R_H = R_L < q \), that is, all uninformed consumers underestimate the quality of experience good (the model setup in the paper).

Therefore, when some proportion of uninformed consumers overestimate the quality of the experience good, the directionality of four effects of offering a deal by a merchant does not alter. In turn, adding this consumer segment does not change the directionality of a publisher’s optimal revenue sharing scheme.
CHAPTER 3

Recommender Systems and Consumer Product Search

Abstract

Recommender systems have become increasingly popular on online retailers, such as those featuring music, movies, and other online content. The use of recommender systems can reduce consumers’ cost of product search and online retailers can use this technology to entice consumers to buy the recommended products. In this paper, we develop an analytical framework to examine the optimal recommender system strategy of an online multi-product retailer. We show that the retailer’s optimal recommender system strategy is driven by consumers’ search cost and misfit cost, and the probability that a consumer continues to search if she rejects the recommendation. We find that when the recommender system is relatively accurate, the retailer will recommend products that are purchased by consumers, although the recommended product is not a perfect match for the recipient of the recommendation. On the other hand, when the recommender system is relatively less accurate, the firm recommends products are a perfect match for some consumers, but other consumers may reject the recommendation. We also find that improving the precision of a relatively accurate recommender system leads to a reduction in consumer surplus.
1. **Introduction**

Product recommender systems are personalized sale assistance tools provided by online retailers to make product search easier for consumers before making their purchase decisions. Recommender systems have become extremely popular in recent years in a variety of applications, such as movies, music and books, where personalized recommender systems are useful for consumers who choose from a myriad of products.

![Figure 3.1: Amazon product recommendation](image)

In general, online retailers or intermediaries use personalized recommender systems to provide products or contents that individual consumers are interested in. Figure 3.1 illustrates the personalized product recommendation provided by Amazon’s recommender system. A consumer sees this set of recommended products at the home page immediately after her logging in, and she can end her product search if she finds what she wants among these products. From the consumers’ perspective, recommender systems make their product search easier and save product search time. The question is how a recommender system affects a consumer’s product search process; and what is an online retailer’s optimal recommender system strategy.

The value proposition of recommender systems is that they provide consumption experience that is personalized to consumers’ tastes (Hosanagar, Fleder, Lee and Buja, 2014).
Giving individual consumers what they want reduces their product search cost in a context where consumers face many choices. In Figure 3.1, Amazon identifies this consumer as someone who is looking for books about business intelligence or data mining, and thus, recommend to her three books, *Principles of Data Mining*, *Python in a Nutshell*, and *Introductory Statistics with R*. She would buy a book if it was the book she was looking for and finish the product search; otherwise, she might start searching for other books using the in-store search engine.

As a result, a consumer has a better experience with Amazon if the recommended products are the ones she is interested in. Thus, from the online retailers’ perspective, these tools are used to build loyalty and turn browsers into buyers. Google News reports that the use of recommendation increases articles viewed by 38% (Das et al. 2007); Amazon reports that 35% of sales originate from recommendations (Lamere and Green 2008); and Netflix reports that 60% of their movie rentals originate from recommendations (Thompson 2008). Thus recommender systems are highly relevant and becoming increasingly powerful tools for retailers.

On the other hand, online retailers can manipulate the product search process of consumers. Online advertising-supported intermediaries, like online news sites and search engines, expose advertisements to viewers in more prominent places than the content that is of interest to consumers. This is referred to as search diversion by Hagiu and Jullien (2014). One recent account is that Google allegedly manipulated search results—promoting its own services and suppressing competitors. In addition, Amazon deliberately removed pages promoting books by Hachette in the midst of a dispute over e-book pricing.

---

25 http://bits.blogs.nytimes.com/2014/05/23/amazon-escalates-its-battle-against-hachette/?_php=true&_type=blogs&_r=0
In this paper, we examine a context where consumers first use the recommender system and then they may use a product search engine to gather information before making their final purchase decision. The research questions are: 1) What is the effect of product recommendation on consumers’ product search and purchasing decisions? 2) How does an online retailer strategically deploy a recommender system? 3) What is the effect of product and consumer characteristics on the online retailer’s optimal product recommendation strategy? 4) What is the impact of recommender systems on consumer surplus?

Online retailers or intermediaries use recommender systems to guide consumers to locate content, products, or services by aggregating product and user information from different sources. Recommender systems can be broadly classified into two types: collaborative filtering and content-based filtering. Both types involve techniques, like data mining, to extract consumer and product matching knowledge from a large quantity of data, and analytically calculate the likelihood that a consumer will purchase one of the products. These systems are used in cases where there are large number of differentiated products, like books, movies, and music.

Understanding consumer information search and acquisition has been an important theme in Economics and Marketing, and Information Systems. In Bakos (1997), the electronic market system provides product search functions that lower buyer search costs and improve market efficiency. In recent years, online retailers have not only improved product search functions, but also provided recommender systems to consumers to improve their product search experience. Due to its popularity, recommender systems have become a major theme of research (Murthi and Sarkar 2003, Dellarocas 2009, Hosanagar et al. 2014), and its impact on consumer product search and market structure is being studied (Kim, Albuquerque, and Bronnenberg, 2010).
Prior research has typically assumed that recommender systems provide an accurate match between buyers and products, and firms do not strategically manage their recommender systems. In this study, we develop a framework in a context where products are differentiated and an online multi-product monopolist can skew recommendations in favor of products with higher profit margin. Consumers can accept or reject the product recommendation depending on their preferences. If consumers reject the recommendation, then they may continue to search for alternative products, thus incurring search costs. The firm’s decision about product recommendation depends on product profit margin, consumer search cost and misfit cost. We find that the firm can benefit from strategically recommending products with higher profit margin. We also find that the optimal accuracy of the recommender system depends on factors such as consumers’ search and misfit cost, and the profit margin of products.

1.1 Literature review

A simplified taxonomy of recommender systems divides them into two types: content-based and collaborative filtering-based systems. Content-based systems use product information (e.g., genre, mood, author) as criteria to recommend items similar to those a user rated highly. On the other hand, collaborative filtering-based systems recommend products that other consumers with similar tastes and preferences bought in the past. There is a large body of work on designing recommender systems, and Adomavicius and Tuzhilin (2005) provide an extensive review in the information systems literature.

The economic implications of recommender systems and how they affect society are not well understood. There is a growing stream of research focusing on the economic implications of electronic commerce in the online market. Studies have examined the impact of use of

Another relevant stream of literature examines the impact of the reduction of consumer search cost on firms’ profit and social welfare. Bakos (1997) studies the implication of consumer search cost on online sellers’ pricing and product strategies, and the resulting reduction on market inefficiency. Branco, Sun and Villas-Boas (2012) develop a framework to study the role between gradual learning, search cost, and consumers’ purchase decisions. They discuss the impact of consumer search on profits and social welfare, and how the seller chooses its price to strategically influence the extent of the consumers’ search.

We consider recommender systems as an integral part of the consumer product search process, and consumers use recommender systems in combination with other product search features, such as search engines, before making a purchase decision. We consider the case where recommender systems can be used strategically to favor higher margin products. The firm needs to balance the tradeoff between profit margin of the recommended product and the likelihood of consumers’ acceptance of the recommendation. In our framework, product recommendation is a substitute for the use of search engines as consumers can learn about the product recommendation without incurring search cost. Therefore, this paper also contributes to the literature on consumer search.

We describe the model setup in §2 and the results in §3. We conclude with a discussion in §4.
2. **Model Setup**

2.1 **Consumers**

In a world with differentiated product offerings, we consider a market of consumers with heterogeneous preferences and one monopolist firm selling products that are spatially differentiated along a single dimension. In our horizontal differentiation model, consumer preference or location is denoted by, \( x_i \) and it is distributed on the interval \([0,1]\). The monopolist’s products are differentiated along the same dimension so that product locations are evenly distributed on \([0,1]\).

A consumer incurs a “misfit” cost \( t \) per unit distance between her ideal location and product location. This cost represents a consumer’s loss due to obtaining a product at a location that is different from her ideal location per unit distance. All consumers have identical unit demand, subject to a reservation utility \( V \), and they are risk neutral. Utility of a consumer at \( x_i \) from buying a product at \( x \) is:

\[
U(x_i, x) = V - |x_i - x|t
\]  

(1)

All products seem ex ante identical to consumers. This implies that consumers have no prior knowledge of the location of any product. Each consumer can learn a product’s location by searching and evaluating the product. The time and effort to search and analyze the product is the search cost, represented by parameter \( c \).

A consumer sequentially searches and examines one product at a time. She decides whether to purchase the examined products or to continue the search. She will stop searching and accept a product if the expected gain in utility from an additional search is less or equal to the cost of another search. It follows that a consumer will purchase a product when it is within a
certain distance to her ideal location. We refer to this distance as the acceptance range, and we formally define this range in the next section.

![Diagram of Consumer decision tree]

**Figure 3.2: Consumer decision tree**

There are two types of consumers: a proportion of consumers, $\mu$, who are willing to search and the remaining consumers, $1-\mu$, who do not search. Consumers of the first type will search if they reject the recommendation. They will continue searching till they find a product $x$ in their acceptance range. Consumers of the second type will never search for products and will not buy any product if they reject the recommendation. This may occur for several reasons for example some consumers may have an outside option in the form of an existing or substitute product. We assume that the acceptance range is identical for both types of consumer. Figure 3.2 illustrates the flow of consumer decisions.

### 2.2 Firm and product profit margin

The monopolist firm sells a continuum of horizontally differentiated products at the same price. There are several examples that fit this setting in the online movie or music subscription models (Netflix and Spotify), where online users pay monthly subscription fee and do not pay for individual movie or music titles. This setting is also applicable to the revenue models where online intermediaries such as search engines, newspapers that charge little or no fee from users, and instead obtain revenue from advertisers. The firm gains heterogeneous profit margin from sales of different products. This can happen due to revenue and cost factors. For example, in the
case of Netflix and Spotify, the quality of digital media (non-HD and HD) impacts the price that retailer can charge as well as the retailer’s marginal costs. In the case of the Google search engine, the revenue from keyword advertising varies across different keywords. In general, the profit margin can vary across products because they are provided by different suppliers with different contract terms. To simplify our model, we assume that products have heterogeneous profit margins and all products have the same exogenously determined price. We assume a mapping from product location to profit margin for each product as \( v(x) = mx \), where \( x \) is product location and \( m \) is the highest profit margin.

The firm knows the distribution of consumer locations, but not the exact location of a consumer. The firm can invest time and effort to gather information about consumers and such analysis can better inform the firm about the consumer’s location. In our model, the firm is able to narrow a consumer’s possible location to a window. We refer to this window of possible consumer locations as a bucket. The firm knows that the consumer’s true location is within the bucket so that the probability of the bucket containing the consumer’s location is 1. The range of a bucket is denoted by \( r \). In this way, instead of a consumer’s exact location, the firm identifies a bucket that the consumer belongs to, which spans from \( a_j \) to \( a_j + r \). Figure 3.1 illustrates that the firm observes the bucket \( [a_j, a_j + r] \) associated with a consumer without knowing her ideal location, \( x_j \). We assume that the consumer’s location is equally likely to be at any point within the bucket and this is known to the firm. This implies that the firm has better information about the location of consumers when the range of the bucket \( (r) \) is small. At the extreme values of \( r \), when \( r = 1 \) the firm has no information about consumers, and when \( r = 0 \), the firm has precise information about consumers’ locations.
Figure 3.3: Consumer bucket $[a_j, a_j + r]$ for a consumer at $x_i$

When a consumer arrives at the retailer, the firm does not observe the true location of the consumer. Instead, he observes a bucket. In aggregate, the firm observes buckets whose left border ($a_j$) are uniformly distributed on the interval $[0, 1-r]$ (Figure 3.3). When the firm recommends a product at $x_i$ to a consumer associated with a bucket $[a_j, a_j + r]$, there are three possible outcomes: 1) the consumer buys the recommended product and the firm gets the profit margin $x_i \bar{m}$; 2) the consumer rejects the recommendation, and searches on her own, to purchase a product at $x_i$, resulting in expected profit of $E[x, \bar{m}]$ to the firm; and 3) the consumer rejects the recommendation, and does not buy any product, resulting in the firm getting zero profit margin. The probabilities associated with these three outcomes depend on the location of the product recommendation ($x_i$), the left border of the bucket ($a_j$), and the range of bucket ($r$). The firm’s expected revenue from a bucket whose left border is at $a_j$ is $E[R_{a_j}(x_i | r)]$. Thus, the firm’s total expected revenue consists of the expected revenue from all buckets and it is:

$$E[R] = \int_0^{1-r} E[R_{a_j}(x_i | r)] da$$

(2)

In this paper, we make two assumptions. 1) We assume the price of all products is zero so that consumers’ purchasing decisions depend on their baseline utility, misfit and search costs. 2) We assume that for consumers, whose locations are close to the ends of the line (0 or 1), they can...
always find the same number of products located on the left and on the right of their individual locations.

3. Analysis

A consumer sequentially searches for product information on the retailer firm’s website. In a sequential search, a consumer decides to stop or continue a search each time after having searched a product. The theory of optimal sequential search states that a consumer continues a search only if the marginal benefit of doing so outweighs the marginal cost. This leads to a stopping criterion that we refer to as the Acceptance Range of a consumer.

3.1 Acceptance range

A consumer incurs cost $c$ per search when she looks for a product and learns its location. Moreover, a consumer incurs misfit cost $t$ per unit distance between her ideal location and the location of the purchased product. A consumer will stop searching and purchase the currently examined product if the expected gain in utility of an additional search is less than or equal to the search cost $c$. We define a critical distance $x'$ for a consumer who is indifferent between buying a product located at $x'$ distance from her ideal location $x_i$ and her continuing searching. This critical distance $x_i$ is obtained from the following equation:

$$2\int_{x_i}^{x_i+x'} (x-x_i)dx = c$$  \hspace{1cm} (3)

The distribution of the firm’s products is common knowledge, thus, a consumer knows there is a product with probability $dx$ along any differentiated part of the unit line. On the left-hand side of equation (3) is the expected gain of utility due to finding a product that has a distance less than $x'$ to the consumer’ location, and on the right-hand side of the equation is the search cost. Solving equation (3), we obtain the acceptance range $x' = \sqrt{c/t}$. This implies that a
consumer will stop searching for products and buy the currently examined product if it is closer than or equal to \( x' = \sqrt{c/t} \) distance from her ideal location.

**PROPOSITION 1**: When the search costs increase \( \frac{dx'}{dc} > 0 \) or the misfit costs decrease \( \frac{dx'}{dt} < 0 \) then the expected distance between the consumer’s ideal location and the location of the purchased product will increase.

Proof is in the Appendix 3.A.

Proposition 1 establishes the stopping rule of an individual consumer who strategically balances the trade-offs between her misfit cost and search cost. When her search cost is high, having examined a product, the cost of finding a fit product (closer to her ideal location) is likely to be high. Thus, she will stop searching when she finds a product even if the product is distant from her ideal location when she has high search cost. Figure 3.4 illustrates the concept of this stopping rule for a consumer, whose location is at \( x_i \). She will buy a product if it is within the acceptance range \( [x_i - \sqrt{c/t}, x_i + \sqrt{c/t}] \), otherwise, she may continue to search for alternative products.

![Figure 3.4: Consumer's acceptance range](image)

### 3.2 In the absence of a recommender system

Some consumers will search for products when the firm does not have a recommender system. If a consumer at \( x_i \) searches, she incurs cost \( c \) for each search, and she keeps searching until she finds a product that is within her acceptance range \( [x_i - \sqrt{c/t}, x_i + \sqrt{c/t}] \). Hence, the
probability of finding products outside the acceptance range on the first $k-1$ searches and finding a product within the acceptance range on the $k^{th}$ search is $P(k) = p(1-p)^{k-1}$, where $p$ is the probability of a random draw of a product that falls in the acceptance range. Therefore, the number of searches till a product within the acceptance range is discovered by a consumer follows a geometric distribution, and the expected number of searches is $1/p$. For a consumer at $x_i$, the probability of a random draw of products being within the acceptance range is $p = 2\sqrt{ct}/t$ since she would buy any product within $[x_i - \sqrt{ct}, x_i + \sqrt{ct}]$. Hence, the expected number of searches is $E[k] = 1/p = \sqrt{ct}/2$, and the expected cost of searches is the product of the expected number of searches and the cost per search: $E[kc] = \sqrt{ct}/2$.

A consumer will continue to search for products until she discovers a product $x_i$ that is within the acceptance range $[x_i - \sqrt{ct}, x_i + \sqrt{ct}]$. Since the products are evenly distributed, the expected value of product $x_i$ is $E[x_i] = \frac{x_i + \sqrt{ct} + x_i - \sqrt{ct}}{2} = x_i$. Moreover, a $\mu$ proportion of consumers will search and a $1-\mu$ proportion of consumers will not search. This implies that the firm’s expected profit from a consumer at $x_i$ in the absence of a recommender system is:

$$\mu E[x_i \bar{m}] = \mu x_i \bar{m}$$

### 3.3 Product recommendation

The firm can recommend a product to each individual consumer when she arrives at the website. Depending on the bucket $[a_j, a_j + r]$ associated with the consumer $x_i$, the firm recommends a product at $x_r(a_j)$. This consumer will accept the recommended product $x_r$ if it is within her acceptance range $[x_r - \sqrt{ct}, x_r + \sqrt{ct}]$ as we have shown in §3.1. If so, then the firm has profit margin $x_r \bar{m}$. The consumer at $x_i$ will reject the recommended product $x_r$ if it is outside the
acceptance range \([x_i - \sqrt{c/t}, x_i + \sqrt{c/t}]\). The consumer will continue to search with a probability \(\mu\) until she finds an acceptable product, and thus the firm’s expected profit margin is \(\mu x \bar{m}\) as shown in (4).

The firm prefers to recommend products with larger \(x\), as the profit margin is \(v(x) = mx\), which implies larger profit margin for larger \(x\). Figure 3.5 illustrates the process of the firm’s product recommendation. The first step is that the firm understands the true location of the consumer can be at any location within the range of the bucket \([a_j, a_j + r]\) with equal probability. The second step is that the firm decides to recommend a product at \(x,(a_j)\), and expects three outcomes which depend on the relative position of the left border of the bucket \((a_j)\), the range of a bucket \((r)\), and the product recommendation \((x,(a_j))\). When the product recommendation is within the range of \(a_j + \sqrt{c/t} \leq x,(a_j) \leq a_j + r + \sqrt{c/t}\) and \(x,(a_j) \geq a_j + r - \sqrt{c/t}\), the consumer will accept the recommendation \(x_r\) if \(x,(a_j) - \sqrt{c/t} \leq x_i \leq a_j + r\). The likelihood of a consumer accepting the recommendation \(x_r\) is: \(L_{accept}(x_r) = \frac{a_j + r - (x_r - \sqrt{c/t})}{r}\).

If \(a_j \leq x_i < x,(a_j) - \sqrt{c/t}\), then the consumer does not accept the recommendation. If she rejects the recommendation, then with a probability \(\mu\), she will search, and with a probability \(1 - \mu\), she will leave the retailer’s website. Although the range of the buckets seems to be greater than the acceptance range \((r > \sqrt{c/t})\) in Figure 3.5, the concept is applicable to the cases where \(r \leq \sqrt{c/t}\).
Figure 3.5: The firm’s product recommendation and consumer’s decision

when \( a_j + \sqrt{c/t} \leq x, (a_j) \leq a_j + r + \sqrt{c/t} \) and \( x, (a_j) \geq a_j + r - \sqrt{c/t} \)

When \( a_j + \sqrt{c/t} \leq x, (a_j) \leq a_j + r + \sqrt{c/t} \) and \( x, (a_j) \geq a_j + r - \sqrt{c/t} \), the firm’s expected revenue of recommending product at \( x \) is:

\[
E[R_j(x)] = \int_{x, -\sqrt{c/t}}^{\infty} x, \mu / rd x \cdot \mu \int_{x, -\sqrt{c/t}}^{\infty} x, \mu / rd x
\]

(5)

The first term is the expected profit margin if the recommendation was within the consumer’s acceptance range (\( x, \in [x, -\sqrt{c/t}, a_j + r] \)). The second term is the expected profit margin if the recommendation was outside her acceptance range (\( x, \in [a_j, x, -\sqrt{c/t}] \)). When \( x > a_j + r + \sqrt{c/t} \), the firm gets nothing from the recommendation. When \( x < a_j + r + \sqrt{c/t} \), the second term is zero; while when \( x < a_j + r - \sqrt{c/t} \), the consumer may reject the recommendation and give up the product search if the consumer location is greater than \( x, + \sqrt{c/t} < x, \leq a_j + r \). Thus, the firm’s maximization problem is:

\[
\max x E[R_j(x,)]
\]

subject to \( x \geq a_j + \sqrt{c/t} \)

\( x \leq a_j + r + \sqrt{c/t} \)

\( x \geq a_j + r - \sqrt{c/t} \)
We solve the maximization problem and obtain \( \frac{\bar{m}(\sqrt{c}(1-\mu)+\sqrt{r}(a_j+r-(2-\mu)x_j))}{r^{\sqrt{t}}} = 0 \) as first order condition (FOC). The value of recommendation \( x_i = \sqrt{c/t}(1-\mu) + \frac{a_j + r}{2-\mu} \) satisfies the FOC. Since \( 0 < \mu < 1 \), the optimal product recommendation \( x_i = a_j + r + \sqrt{c/t} \), the constraints on the optimal product recommendation are only \( x_i \geq a_j + r - \sqrt{c/t} \) and \( x_i \geq a_j + \sqrt{c/t} \). When \( x_i < a_j + \sqrt{c/t} \), the firm can increase the value of \( x_i \) to \( x_i = a_j + \sqrt{c/t} \), which leads to higher expected profit margin from recommendation (greater value of the first term in (5) and the second term remains as 0), and thus, the greater expected revenue. Note that when \( r \leq 2\sqrt{c/t} \), \( a_j + r - \sqrt{c/t} \leq a_j + \sqrt{c/t} \); \( x_i = a_j + \sqrt{c/t} \) is the only boundary solution.

When \( r > 2\sqrt{c/t} \), \( a_j + r - \sqrt{c/t} > a_j + \sqrt{c/t} \), it is possible that \( a_j + \sqrt{c/t} \leq x_i(a_j) < a_j + r - \sqrt{c/t} \)

which means that a consumer, whose location is either extremely large or small within the bucket, may reject the recommended product \( x_i \). In this case, the firm’s expected revenue of recommending a product at \( x_i \) is:

\[
E[R_{a_j}(x_i)] = \int_{x_i - \sqrt{c/t}}^{x_i + \sqrt{c/t}} x m / r dx_i + \mu \int_{x_i - \sqrt{c/t}}^{x_i + \sqrt{c/t}} x m / r dx_i + \mu \int_{x_i - \sqrt{c/t}}^{a_j + \sqrt{c/t}} x m / r dx_i \] (6)

The first order derivative of the expected revenue in (6) is \( \frac{dE[R_{a_j}(x_i)]}{dx_i} = \frac{2(1-\mu)\bar{m}}{r} > 0 \).

This implies that the optimal solution is the highest possible value \( x_i = a_j + r - \sqrt{c/t} \). Thus, when \( r \geq 2\sqrt{c/t} \), the constraint is \( x_i \geq a_j + r - \sqrt{c/t} \), and \( x_i = a_j + r - \sqrt{c/t} \) is a boundary solution.

We further assume that the search cost is low compared to the misfit cost, that is \( \frac{t}{9} < c < \frac{t}{4} \). The optimal product recommendation strategy is summarized in Lemma 1.
**LEMMA 1:** 1) When \(0 < r \leq \sqrt{c/t} \), the optimal product recommendation is \(x = a_j + \sqrt{c/t} \) for buckets \(a_j \in [0,1 - \sqrt{c/t}] \), and \(x = 1 \) for buckets \(a_j \in (1 - \sqrt{c/t}, 1 - r] \); 2) When \(\sqrt{c/t} < r < r_1 \), the optimal product recommendation is \(x = x_i \) for buckets \(a_j \in [0,a_i] \), and \(x = a_j + \sqrt{c/t} \) for buckets \(a_j \in (a_i, 1 - r] \); 3) When \(r_1 \leq r < 1 \), the optimal product recommendation is \(x = x_i \) for all buckets \(a_j \in [0,1 - r] \); where \(x_i = \frac{\sqrt{c/t} - \frac{1}{\mu} + a_j + r}{\frac{2}{\mu}} \), \(r_1 = \frac{1 - \mu + \sqrt{c/t}}{\frac{2}{\mu}} \) and \(a_i = \frac{r - \sqrt{c/t}}{1 - \mu} \).

It is obvious to see that it is more likely that a consumer will accept the recommended product if the range of a bucket \((r)\) decreases. Moreover, when every consumer who rejects the recommended product continues to search for alternative products \((\mu = 1)\), the firm recommends products that do not match with the consumers’ preferences. The firm recommends products outside the bucket if the range of buckets is relatively small, or recommends the product at right-most right border of the bucket when the range is large. When consumers may give up the product search all together, the firm recommends products that are inside the bucket so that the recommendation may be a perfect match for the consumer.

**PROPOSITION 2:** The location of the product recommended by the recommender system \(x\), is independent of the profit margin parameter \(\bar{m}\).

The firm’s strategic recommendation strategy is driven by heterogeneity among products’ profit margin \(x\bar{m}\) and the profit margin increases as either \(x\) or \(\bar{m}\) increases. Therefore, one may think that the firm’s product recommendation strategy will shift to the right when profit margin \(\bar{m}\) is higher. However, counter-intuitively, Proposition 2 says that the product recommendation does not depend on \(\bar{m}\).
3.4 Firm’s total expected revenue

Lemma 1 says that the firm’s recommendation strategy for each consumer bucket depends on the range of a bucket \( r \), the consumer’s search cost \( c \), misfit cost \( t \), and the probability that a consumer will continue to search after rejecting the recommendation \( \mu \). Thus, the firm’s expected revenue of a bucket over \([a_j, a_j + r]\) also depends on these factors, \( E[R_j(x^*_j | r, \mu, c, t)] \). Without considering the cost of the recommender system, we show that the firm’s expected total revenue is:

\[
E[R(r)] = \int_0^{1-r} E[R_j(x_j | r)]da_j,
\]
and this decreases as the bucket range increases (Figure 3.6).

![Figure 3.6: Impact of accuracy of recommender system on the firm's expected revenue where \( \bar{m} = 1, c = 0.2, t = 1, \) and \( \mu = 0.6 \)]

**PROPOSITION 3:** The firm’s expected total revenue increases as the range of consumer buckets decreases (lower \( r \)), \( \frac{dE[R(r)]}{dr} < 0 \).

Decrease in the range of buckets means the firm can more accurately identify each consumer, and thus it is more likely for a consumer to accept a recommended product that has a higher profit margin. Thus, the firm’s total revenue increases as \( r \) decreases.
**PROPOSITION 4:** When the range of buckets is relatively small, \( r \leq \sqrt{c/t} \), the firm’s optimal product recommendation is outside the bucket \( x^* \not\in [a_j, a_j + r] \).

When the range of the buckets is \( r' \leq \sqrt{c/t} \), the firm’s optimal product recommendation is \( x^* = a_j + \sqrt{c/t} \). Proposition 4 states that the firm will never provide a recommendation that perfectly matches the consumer’s preference if the recommender system is sufficiently accurate (\( r \leq \sqrt{c/t} \)). This implies that the consumers will accept the recommendation, but they never get a perfect match to their ideal locations. This is intuitive as the firm will try to sell products that are profitable and acceptable to consumers. This implies that the firm sells the product that is acceptable to consumers but he chooses to recommend products that are outside the consumers’ buckets. Such a recommendation strategy yields higher profit margin to the retailer. Proposition 5 explores the firm’s product recommendation strategy when the accuracy of the recommender system is relatively low.

**PROPOSITION 5:** When the range of buckets is relatively large, \( r' > \sqrt{c/t} \), (1) the firm’s optimal product recommendation is inside the bucket \( x^* \in [a_j, a_j + r] \); and (2) a consumer is more likely to accept the recommendation if \( \mu \) is smaller \( \frac{dL_{\text{accept}}(x_c)}{d\mu} < 0 \).

When \( \sqrt{c/t} < r < r_1 \), the firm’s optimal product recommendation is \( x_c = x_1 \) for buckets \( a_j \in [0, a_1] \), or \( x_c = a_j + \sqrt{c/t} \) for buckets \( a_j \in (a_1, 1-r] \); or when \( r_1 \leq r < 1 \), the firm’s optimal product recommendation is \( x_c = x_i \) for all buckets (Lemma 1). In either case, the product recommendation is inside the range of buckets \( x^* \in [a_j, a_j + r] \), which implies that a consumer may get a perfect match with the product recommendation. However, in this case a consumer may reject the recommendation if her location is \( x_c \in [a_j, x_c - \sqrt{c/t}] \), since the optimal range of bucket is
relatively large, $r' > \sqrt{c/t}$. Proposition 5 states that when the accuracy of the recommender system is relatively low, some consumers may reject the recommendation, but other consumers may get a perfect match with the product recommendation.

Moreover, it is more likely that she accepts the recommendation when it is less likely for consumers to continue to search if they reject the product recommendation (smaller $\mu$), or the acceptance range is greater. Intuitively, when there is a probability that a consumer will give up the product search and do not buy any product if she rejects the recommendation ($0 < \mu < 1$), the firm needs to sacrifice the expected profit margin from recommending the product (by recommending a lower $x_i$) as he may lose a consumer if she rejects the recommendation. Therefore, the chance that a consumer will accept the product recommendation is higher, if there is a higher chance the firm loses a consumer.

### 3.5 Realized consumer surplus

The expected misfit cost for a consumer at $x_i$ who search products till she finds an acceptable product is the expected loss from products that are within the range $[x_i - \sqrt{c/t}, x_i + \sqrt{c/t}]$:

$$2 \int_{x_i - \sqrt{c/t}}^{x_i + \sqrt{c/t}} (x - x_i) dx = \sqrt{ct}/2$$

The expected number of attempts for a consumer who searches products until she finds an acceptable product is $\sqrt{ct}/2$, and the expected search cost is $\sqrt{ct}/2$ from §3.2.

Therefore, the consumer surplus of a consumer who rejects project the recommendation, and continues to search for alternative product until an acceptable product is the baseline utility minus the expected search cost and misfit costs:

$$CS_i' = V - \sqrt{cs}/2 - \sqrt{cs}/2 = V - \sqrt{cs}$$
The consumer surplus of a consumer at $x_i$ who accepts the product recommendation $x$, is:

$$CS_i' = V - |x - x_i|t$$

In order to examine the impact of the firm’s recommender system strategy on consumer surplus, we need to assume a mapping from consumer location to consumer bucket, $a_j = x_j/(1 - r)$. In this subsection, we assume consumer locations $x_i$ are distributed uniformly on the interval [0,1], and the left borders of buckets $a_j$ are distributed uniformly on the interval [0,1 - r].

Lemma 1 states that when the range of buckets is relatively small, $r \leq \sqrt{c/t}$, the firm’s product recommendation is $x_i^* = a_j + \sqrt{c/t}$ for all buckets. Since all consumers accept the recommendation and no one searches for products, consumers only incur misfit cost (Proposition 4). For consumers whose locations are close to 1, $x_i > (1 - \sqrt{c/t})/(1 - r)$, the product recommendation is $x_i^* = 1$ as it is the most profitable product; while for consumers whose location are $x_i \leq (1 - \sqrt{c/t})/(1 - r)$, the product recommendation is $x_i^* = a_j + \sqrt{c/t}$. The aggregate consumer surplus is:

$$CS = \int_{0}^{(1 - \sqrt{c/t})/(1 - r)} CS_i'(x_i = a_j + \sqrt{c/t})dx_i + \int_{(1 - \sqrt{c/t})/(1 - r)}^{1} CS_i'(x_i = 1)dx_i \quad (7)$$

When the range of buckets is $\sqrt{c/t} < r < 1$, the firm’s optimal product recommendation is $x_i^* = x_i = \sqrt{c/t} \frac{1 - \mu}{2 - \mu} + \frac{a_j + r}{2 - \mu}$ for buckets $a_j \in [0, a_1]$, or $x_i^* = a_j + \sqrt{c/t}$ for buckets $a_j \in (a_1, 1 - r]$. Some consumers ($x_i < x_i - \sqrt{c/t}$) reject the product recommendation because it is outside their acceptance ranges, and they continue to search and buy products that are acceptable. Their expected misfit cost is $\sqrt{c/t}/2$ and search cost is $\sqrt{c/t}/2$. Other consumers ($x_i \geq x_i - \sqrt{c/t}$) accept the product recommendation, and they do not search for other products. Their search cost is zero.
Moreover, for consumers $x_i \in [x_i - \sqrt{c/t}, a_i/(1-r)]$, whose corresponding buckets are $a_j \in [0,a_i]$, the product recommendation is $x_i^* = x_i$; while for consumers $x_i \in [a_i/(1-r),1]$, the product recommendation is $x_i^* = a_j + \sqrt{c/t}$. Thus, the aggregate consumer surplus is:

$$CS = \mu \int_0^{\sqrt{c/t}} CS_i dx_i + \int_{x_i = x_i}^{a_i/(1-r)} CS_i'(x_i = x_i) dx_i + \int_{a_i/(1-r)}^1 CS_i'(x_i = a_j + \sqrt{c/t}) dx_i$$

(8)

Furthermore, when the range of buckets is relatively large, $r_1 \leq r < 1$, the optimal product recommendation is $x_i = x_i$ for all consumers. Similar to the case where $\sqrt{c/t} < r < r_1$, some consumers ($x_i < x_i - \sqrt{c/t}$) reject the product recommendation, while other consumers ($x_i \geq x_i - \sqrt{c/t}$) accept the product recommendation. For consumers who reject the product recommendation, they may continue to search for products, and they incur search and misfit costs; while for consumers who accept the product recommendation, they only incur misfit cost. The aggregate consumer surplus is:

$$CS = \mu \int_0^{\sqrt{c/t}} CS_i dx_i + \int_{x_i = x_i}^{a_i/(1-r)} CS_i'(x_i = x_i) dx_i$$

(9)

We further assume that the unit misfit cost is $t = 1$, and thus, $\frac{1}{9} < c < \frac{1}{4}$, following the assumption in §3.2.

**PROPOSITION 6:** (1) The aggregate consumer surplus increases as the range of buckets increases, $\frac{dCS}{dr} > 0$, when the range of the bucket is relatively small $r \leq \sqrt{c}$; (2) aggregate consumer surplus decreases as the range of the buckets decreases, $\frac{dCS}{dr} < 0$, when the range of the bucket is relatively large, $r > \frac{1 - \mu + \sqrt{c}}{2 - \mu}$.
Proposition 6 states that when the accuracy of the recommender system is not accurate \( r > \frac{1-\mu+\sqrt{c}}{2-\mu} \), consumer surplus increases if the range of the buckets decreases. This implies that consumers benefit as the recommender system becomes more accurate in this range. On the other hand, when the recommender system is already accurate \( r < \sqrt{c} \), consumer surplus decreases if the range of the buckets decreases. This means that consumers become worse off if the recommender system becomes more accurate (Figure 3.7).

![Figure 3.7: Impact of bucket range on the realized consumer surplus](image)

4. Conclusion

In this paper, we develop a framework to examine the optimal recommender system strategy of an online monopolist multi-product retailer. In the context of online products, we consider consumers use of recommender systems before they search for product information. The use of recommender system reduces search cost as consumers can accept the recommendation without incurring search cost.

We find that the firm’s expected revenue increases when the accuracy of the recommender system is higher. We also find that the firm’s optimal recommender system strategy is driven by consumers’ search cost and misfit cost, and the probability that a consumer
does not buy any product if she rejects recommendation. We show that the firm recommends products that do not match consumers’ ideal locations when the recommender system is relatively accurate. On the other hand, the firm recommends products that perfectly match some consumers’ preference when the recommender system is relatively less accurate.

References


Appendix 3.A: Proof

PROPOSITION 1: When search costs increase \( \frac{dx'}{dc} > 0 \) or misfit costs decrease \( \frac{dx'}{dt} < 0 \) then the expected distance between the consumer’s ideal location and the location of the purchased product will increase.

Proof of Proposition 1:

Since \( x' = \sqrt{c/t} \), \( \frac{dx'}{dc} = \frac{1}{2\sqrt{ct}} \), and \( \frac{dx'}{dt} = -\frac{1}{2\sqrt{ct}} \). Since \( c > 0 \) and \( t > 0 \), \( \frac{dx'}{dc} > 0 \) and \( \frac{dx'}{dt} < 0 \).

Thus, we prove Proposition 1.

LEMMA 1: 1) When \( 0 < r \leq \sqrt{c/t} \), the optimal product recommendation is \( x_r = a_j + \sqrt{c/t} \) for buckets start from \( a_j \in [0,1-\sqrt{c/t}] \), or \( x_r = 1 \) for buckets start from \( a_j \in (1-\sqrt{c/t}, 1-r] \); 2) When \( \sqrt{c/t} < r < r_1 \), the optimal product recommendation is \( x_r = x_i \) for buckets start from \( a_j \in [0,a_i] \), or \( x_r = a_j + \sqrt{c/t} \) for buckets start from \( a_j \in (a_i,1-r] \); 3) When \( r_1 \leq r < 1 \), the optimal product recommendation is \( x_r = x_i \) for all buckets \( a_j \in [0,1-r] \); where \( x_i = \sqrt{c/t} \left[ 1 - \mu + \frac{a_i + r}{2 - \mu} \right] \), \( r_i = \frac{1 - \mu + \sqrt{c/t}}{2 - \mu} \) and \( a_i = \frac{r - \sqrt{c/t}}{1 - \mu} \).

Proof of Lemma 1:

When \( 0 < r \leq \sqrt{c/t} \), then \( x_r = \sqrt{c/t} \left[ 1 - \mu + \frac{a_i + r}{2 - \mu} \right] = a_j + 2\sqrt{c/t} - \sqrt{c/t} \mu + (r - \sqrt{c/t}) \) is lower than \( a_j + \sqrt{c/t} = a_j (2 - \mu) + 2\sqrt{c/t} - \sqrt{c/t} \mu \), thus, the optimal product recommendation is the boundary solution \( x_r = a_j + \sqrt{c/t} \).

When \( \sqrt{c/t} < r < r_1 \), then solve the value of \( a_i \) where \( a_i + \sqrt{c/t} = x_i (a = a_i) \), and we obtain \( a_i = \frac{r - \sqrt{c/t}}{1 - \mu} \).

Since the maximum value of \( a_j \) is \( 1-r \), then solving \( a_i = \frac{r - \sqrt{c/t}}{1 - \mu} = 1-r_i \), we get \( r_i = \frac{1 - \mu + \sqrt{c/t}}{2 - \mu} \).

This implies that when \( \sqrt{c/t} < r < r_1 \), if \( a_j \leq a_i \), then \( x_i = a_i + \sqrt{c/t} \), thus the optimal product recommendation is \( x_r = x_i \); and if \( a_j > a_i \), then \( x_i < a_i + \sqrt{c/t} \), thus the optimal product recommendation is \( x_r = a_j + \sqrt{c/t} \).

When \( r_1 \leq r < 1 \), \( x_r > a_i + \sqrt{c/t} \), thus, the optimal product recommendation is \( x_r = x_i \).

Thus, we prove Lemma 1.

PROPOSITION 2: The firm’s product recommendation \( x_r \) is independent of profit margin \( \bar{m} \).
Proof of Proposition 2:
Given the accuracy of the recommender system, it is obvious to see that the product recommendation, \( x_i = a_i + \sqrt{c/t} \) or \( x_i = \sqrt{c/t} \frac{1-\mu}{2-\mu} a_i + \frac{r}{2-\mu} \), is independent of \( \bar{m} \).
Thus, we prove Proposition 2. ■

PROPOSITION 3: The firm’s expected total revenue increases as the range of consumer buckets decreases (lower \( r \)), \( \frac{dE[R(r)]}{dr} < 0 \).

Proof of Proposition 3:
From Lemma 1, we show the firm’s product recommendation strategy depends on the range of buckets. And there are three regions for the range of buckets: 1) \( 0 < r \leq \sqrt{c/t} \), 2) \( \sqrt{c/t} < r < \epsilon \), and 3) \( \epsilon \leq r < 1 \).

1) When \( 0 < r \leq \sqrt{c/t} \), the product recommendation is \( x_i = a_i + \sqrt{c/t} \) for buckets \( a_i \in [0, 1-\sqrt{c/t}] \), or \( x_i = 1 \) for buckets \( a_i \in (1-\sqrt{c/t}, 1-r] \). Thus, the optimal expected total revenue is
\[
E[R(r)] = \int_0^{\sqrt{c/t}} E[R_{u_i}(x_i = a_i + \sqrt{c/t})]da_i + \int_{\sqrt{c/t}}^{1-r} E[R_{u_i}(x_i = 1)]da_j
\]
\[
= \bar{m}(\sqrt{c/t} - r) + \bar{m}(t-c) \frac{1}{2t}.
\]
We take the first order derivative of the optimal expected total revenue with respect to the range of buckets, and we obtain: \( \frac{dE[R(r)]}{dr} = -\bar{m} < 0 \).

2) When \( \sqrt{c/t} < r < \epsilon \), the optimal product recommendation is \( x_i = x_i \) for buckets start from \( a_i \in [0, a_i] \), or \( x_i = a_i + \sqrt{c/t} \) for buckets start from \( a_i \in (a_i, 1-r] \). Thus, the optimal expected total revenue is
\[
E[R(r)] = \int_0^{a_i} E[R_{u_i}(x_i = x_i)]da_i + \int_{a_i}^{1-r} E[R_{u_i}(x_i = a_i + \sqrt{c/t})]da_j
\]
We take the first order derivative of the optimal expected total revenue with respect to the range of buckets, and we obtain: \( \frac{dE[R(r)]}{dr} = \bar{m}(c^{3/2} - 3\sqrt{c}(5 - 6\mu + 2\mu^2)r^{2}t + 2r^2(9\mu(1-r) - 3\mu^2(1-r) + 7r - 6)\mu^{3/2})
\]
\[
6(2-\mu)(1-\mu)r^2t^{3/2}.
\]
We let \( A = c^{3/2} - 3\sqrt{c}(5 - 6\mu + 2\mu^2)r^{2}t + 2r^2(9\mu(1-r) - 3\mu^2(1-r) + 7r - 6)\mu^{3/2} \), so the numerator is \( \bar{m}A \). We then take the derivative of \( A \) with respect to \( \mu \), and we find that \( \frac{dA}{d\mu} = 6(3-2\mu)r^2(\sqrt{c} + (1-r)\sqrt{t})t \).
and \( \frac{d^2A}{d\mu^2} = -12r^2(\sqrt{c} + (1-r)\sqrt{t})t < 0 \). This implies that \( A \) is at the maximum when \( \mu = 2/3 \). We substitute \( \mu = 2/3 \) into \( A \), and we have \( A(\mu = 2/3) = \frac{1}{3}(3c^{3/2} - 17\sqrt{c}r^2t - 2r^2(4 - 7r)\mu^{3/2}) \). Since \( \sqrt{c/t} < r < \epsilon(\mu = 2/3) \rightarrow \sqrt{c/t} < r < \frac{3}{4}(1 + \sqrt{c/t}) \), and \( 2\sqrt{c/t} < 1 < 3\sqrt{c/t} \), we show that \( A(\mu = 2/3) < 0 \). This implies that \( A < 0 \) for all \( \mu \in (0, 1) \).
Since $\frac{dE[R(r)]}{dr} = \frac{\tilde{m}A}{6(2-\mu)(1-\mu)r^3t^{5/2}}$, and $6(2-\mu)(1-\mu)r^3t^{5/2} > 0$, we show that $\frac{dE[R(r)]}{dr} < 0$ for all $\mu \in (0,1)$.

3) When $\varepsilon \leq r < 1$, the optimal product recommendation is $x_i = x_i$ for all buckets $a_j \in [0,1-r]$.

Thus, the optimal expected total revenue is $E[R(r)] = \int_0^{1-r} E[R(x_i=x_i)] da_j$.

We take the first order derivative of the optimal expected total revenue with respect to the range of buckets, and we obtain:

$$\frac{dE[R(r)]}{dr} = \frac{\tilde{m}(3c + 3\sqrt{c}(1-\mu)(1+r^2)\sqrt{t} + ((1-\mu)^2 + 3(2-\mu)\mu r^2 + 2(1-\mu)^2 r^3)t)}{-6(2-\mu)r^3t}.$$ 

Since $0 < \mu < 1$, the numerator is positive, but the denominator is negative, and we have $\frac{dE[R(r)]}{dr} < 0$.

Thus, we show that $\frac{dE[R(r)]}{dr} < 0$. ■

**PROPOSITION 4:** When the range of buckets is relatively small, $r \leq \sqrt{c/t}$, the firm’s optimal product recommendation is outside the bucket $x_i^* \not\in [a_j, a_j + r]$.

**PROPOSITION 5:** When the range of buckets is relatively large, $r > \sqrt{c/t}$, (1) the firm’s optimal product recommendation is inside the bucket $x_i^* \in [a_j, a_j + r]$; and (2) a consumer is more likely to accept the recommendation if $\mu$ is smaller $\frac{dL_{\text{accept}}(x_i)}{d\mu} < 0$.

**Proof of Propositions 4 and 5:** When the range of buckets is relatively small, $r \leq \sqrt{c/t}$, $x_i = a_j + \sqrt{c/t} \geq a_j + r$. When the range of buckets is relatively large, $r > \sqrt{c/t}$, $x_i = x_i < a_j + r$, and the likelihood of a consumer may accept the recommendation is:

$$L_{\text{accept}}(x_i) = \frac{a_j + r - (x_i - \sqrt{c/t})}{r} = \frac{a_j + r + \sqrt{c/t} - (1-\mu)\sqrt{c/t} + (a_j + r)}{(2-\mu)r},$$

which decreases in $\mu$, so $\frac{dL_{\text{accept}}(x_i)}{d\mu} < 0$.

Thus, we prove Proposition 4 and 5. ■

**PROPOSITION 6:** (1) The aggregate consumer surplus increases as the range of buckets increases, $\frac{dCS}{dr} > 0$, when the range of bucket is relatively small $r \leq \sqrt{c}$; (2) aggregate consumer surplus decreases as the range of buckets decreases, $\frac{dCS}{dr} < 0$, when the range of bucket is relatively large $r > \frac{1-\mu + \sqrt{c}}{2-\mu}$.

**Proof of Proposition 6:**

When $r \leq \sqrt{c/t}$, the product recommendation is $x_i = a_j + \sqrt{c/t}$ for buckets $a_j \in [0,1-\sqrt{c/t}]$, or $x_i = 1$ for buckets $a_j \in (1-\sqrt{c/t},1-r]$. The aggregate consumer surplus from equation (7) is
\[
CS = \int_0^{1/\sqrt{c(1-r)}} (V - (a_j + \sqrt{c/t} - x_i)\,dx_i + \int_{1/\sqrt{c(1-r)}}^1 (V - (1-x_i)\,dx_i .
\]
We get \( CS = \frac{c-2\sqrt{ct+tr}}{2(1-r)} + V . \) Take the derivative with respect to \( r , \) and we obtain \( \frac{dCS}{dr} = \frac{(\sqrt{c} - \sqrt{t})^2}{2(1-r)} > 0 . \)

When \( \sqrt{c/t} < r < r \), the optimal product recommendation is \( x_i = x_i \) for buckets start from \( a_j \in [0,a_j] , \) or \( x_i = a_j + \sqrt{c/t} \) for buckets start from \( a_j \in (a_j,1-r] . \) We first solve for \( a_{i1} \) that satisfies \( a_{i1}/(1-r) = a_i + \sqrt{c/t} \). If \( r \leq \frac{c^{1/4}(c^{1/4} \mu + \sqrt{c\mu^2 + 4(1-\mu)/\sqrt{t}})}{2\sqrt{t}} \), then \( a_i, a_{i1} \). This implies when \( \sqrt{c/t} < r \leq \frac{c^{1/4}(c^{1/4} \mu + \sqrt{c\mu^2 + 4(1-\mu)/\sqrt{t}})}{2\sqrt{t}} \), consumers \( a_{i1}/(1-r) < x_i, a_{i1}/(1-r) \), will be given the product recommendation that is smaller than the true location \( x_i = a_j + \sqrt{c/t} < x_i \) ; and consumers \( x_i > a_{i1}/(1-r) \), will be given the product recommendation larger than the true location \( x_i = a_i + \sqrt{c/t} > x_i \). Thus, the aggregate consumer surplus from equation (8) is:

\[
CS = \mu \int_0^{1/\sqrt{c(1-r)}} (V - \sqrt{c} \,dx_i + \int_{1/\sqrt{c(1-r)}} (V - (x_i - x_i)\,dx_i + \int_{a_{i1}/(1-r)}^1 (V - (a_j + \sqrt{c/t} - x_i)\,dx_i + \int_t^{a_{i2}/(1-r)} (V - (x_i - a_j - \sqrt{c/t} r)\,dx_i
\]

If \( \frac{c^{1/4}(c^{1/4} \mu + \sqrt{c\mu^2 + 4(1-\mu)/\sqrt{t}})}{2\sqrt{t}} < r \), then \( a_{i2} < a_i \). We first solve for \( a_{i2} \) that satisfies \( a_{i2}/(1-r) = x_i(a_j = a_{i2}) \). When \( \frac{c^{1/4}(c^{1/4} \mu + \sqrt{c\mu^2 + 4(1-\mu)/\sqrt{t}})}{2\sqrt{t}} < r \leq r \), consumers \( a_{i2}/(1-r) < a_i/1-r \), will be given the product recommendation that is greater than the true location \( x_i = a_i > x_i \) ; and consumers \( a_{i2}/1-r) < x_i, a_{i2}/1-r) \), will be given the product recommendation that is greater than the true location \( x_i = a_j + \sqrt{c/t} > x_i \). Thus, the aggregate consumer surplus from equation (8) is:

\[
CS = \mu \int_0^{1/\sqrt{c(1-r)}} (V - \sqrt{c} \,dx_i + \int_{1/\sqrt{c(1-r)}} (V - (x_i - x_i)\,dx_i + \int_{a_{i2}/(1-r)}^1 (V - (x_i - a_j - \sqrt{c/t} r)\,dx_i
\]

When \( r \leq r < 1 \), the optimal product recommendation is \( x_i = x_i \) for all buckets \( a_j \in [0,1-r] \). Thus, the aggregate consumer surplus from equation (9) is

\[
CS = \mu \int_0^{1/\sqrt{c(1-r)}} (V - \sqrt{c} \,dx_i + \int_{1/\sqrt{c(1-r)}} (V - (x_i - x_i)\,dx_i + \int_{a_{i2}/1-r) (V - (x_i - a_j - \sqrt{c/t} r)\,dx_i
\]

Since we assume \( t = 1 \), we take the derivative with respect to \( r , \) and we obtain

\[
\frac{dCS}{dr} = \frac{c(5-10\mu+4\mu^2)-2\sqrt{c}(1-\mu)(1-\mu^2+\mu(1-V)+2V)+(1-\mu)^2(1-2(2-\mu)V)}{2(2-\mu)(1-\mu+r)^2} . \]

Also, since \( 1/9 < c < 1/4 , 0 < \mu < 1 \) and \( V > \sqrt{c} \), the numerator of the derivative is negative and the denominator is positive, thus, \( \frac{dCS}{dr} < 0 . \)

Thus, we prove Proposition 6.