Welfare Properties of Recommender Systems

Submitted in partial fulfillment of the requirements for

for the degree of

Doctor of Philosophy

in

Engineering and Public Policy

Xiaochen Zhang

B.S. Telecom. Engineering, Beijing University of Posts & Telecom.

M.S. Electrical Engineering, Northwestern University

Carnegie Mellon University

Pittsburgh, PA

May 2017

To my father, Zhifeng Zhang

Committee Members

PEDRO FERREIRA (Chair)

Assistant Professor

the H. John Heinz III College & Engineering and Public Policy

Carnegie Mellon University

MARVIN SIRBU

Professor

Engineering and Public Policy & Tepper School of Business & Electrical and Computer Engineering

Carnegie Mellon University

RODRIGO BELO

Assistant Professor

Department of Technology and Operations Management

Rotterdam School of Management

Erasmus University Rotterdam

MIGUEL GODINHO DE MATOS

Assistant Professor

Information Systems

Catolica Lisbon School of Business & Economics

Acknowledgments

First and foremost, I would like to send my sincere gratitude to my advisor and committee chair Dr. Pedro Ferreira for his continuing support and guidance throughout the four years of my doctoral study. Pedro is a great mentor. He is always helpful in every milestone of my graduate study. He helped me practice talk and gave detailed feedbacks to slides and papers. I learned a lot from Pedro on writings and presentations. He also gave important advices on the courses I took to prepare me for analytic skills required for my research, and supported me to attend top-tier conferences and workshops. I could not finish my dissertation work without all of his help in these years.

My sincere thanks also go to my PhD committee members Dr. Rodrigo Belo, Dr. Miguel Godinho de Matos, and Dr. Marvin Sirbu for the questions and comments they gave during this process. Their feedbacks help immensely to improve my dissertation. I collaborated with Dr. Rodrigo Belo and Dr. Miguel Godinho de Matos for all three parts of my dissertation work and the accomplishment of my work is attributed largely to the regularly lively discussions with them. Dr. Marvin Sirbu provided me precious advice on my dissertation from the policy perspectives.

I would like to thank the lab members Baojiang Yang, Filipa Reis, and Qiwei Han for their help in my life and study both when I was in Pittsburgh and in Lisbon. I am so lucky to have such a united and friendly lab group with all these great people.

I would like to thank my home department EPP for all the departmental support that make all my qualifier, proposal and defense going so smoothly. Special thanks give to Adam Loucks, Victoria Finney, and Deborah Kuntz. I also want to thank Debora Porem from Catolica Lisbon School of Business & Economics for her help in the administrative work during my stay in Lisbon. Thanks all my friends, Fan Tong, Baojiang Yang, Nandi Zhang, Yu Gan, Daniel I. Posen, Brock Glasgo, Matt

Babcock, and Rebecca Ciez from EPP who provided me comments and feedback during my practices of the qualifier, proposal, and defense of the dissertation.

I took many great courses at CMU from which I acquired programming, mathematical, and modeling skills. These skills contribute a lot to the completion of all my dissertation work. I would like to thank the course professors Dr. Brian Kovak, Dr. David Krackhardt, Dr. Janusz Szczypula, Dr. Kaifu Zhang, Dr. Alex Davis, Dr. Shamena Anwar, and Dr. Amelia M. Haviland.

I would like to thank all my supporting friends at CMU. I feel so lucky to meet all of you in my life journey in these four years. I want to thank the four years of great friendship from Yingjie Zhang and Shanghang Zhang, my gorgeous roommate Dr. Vicky Chen, and my best friend Weiwei Liu.

I would like to acknowledge my parents, my father Zhifeng Zhang, and my mother Sumei Tian. They made me the person I am today. Thanks Dad for every word you gave to me since I could remember things. Those words lighted up the roads ahead many times during the darkness of my life and study. Thanks Mom for always believing that I can be a Ph.D. since I was five.

I met my boyfriend Yuanning Li at Carnegie Mellon and he is God's gift for me. He is always by my side setting a model for me and encouraging me all the way along. Thank you for all the support and the research passion you brought to me. I would never have found research and data science so interesting, and hard efforts so rewarding without you.

I would like to acknowledge the funding support of my Ph.D. from Engineering and Public Policy department of Carnegie Mellon University, the NSF 12-503 Program on Secure and Trustworthy Cyberspace, and BI research fellowship of CUBE of Catolica Lisbon School of Business & Economics-UCP.

Abstract

Recommender systems are ubiquitously used by online vendors as profitable tools to boost sales and enhance the purchase experience of their consumers. In recent literature, the value created by recommender systems are discussed extensively. In contrast, few researchers look at the negative side of the recommender systems from the viewpoint of policymakers. To fill this gap, I critically investigate the welfare impact of recommender systems (RSs) during my Ph.D. study. The main focus of my Ph.D. dissertation is analyzing whether there exists a conflict of interest between the recommendations provider and its consumers in the electronic marketplace. My dissertation is composed of three parts. In Part I, I evaluate empirically whether in the real world, the profit-driven firm will choose a recommendation mechanism that hurts or is suboptimal to its consumers. In Part II, I analyze the role of personalization technology in the RSs from a unique perspective of how personalization resembles price discrimination as a profitable tool to exploit consumer surplus. In part III, I investigate the vendor's motivation to increase the level of personalization in two-period transactions.

As the RSs are designed by the firm, and the firm's objective is to maximize profits, the RSs might not maximize consumers' welfare. In Part I of my thesis work, I test the existence of such a conflict of interest between the firm and its consumers. I explore this question empirically with a concrete RS created by our industry collaborator for their Video-on-Demand (VoD) system. Using a large-scale dataset (300,000 users) from a randomized experiment on the VoD platform, I simulate seven RSs based on an exponential demand model with listed movie orders and prices as key inputs, estimated from the experimental dataset. The seven simulated RSs differ by the assignments of listed orders for selected recommended movies. Specifically, assignments are chosen to maximize profits, consumer surplus, social welfare, popularity (IMDB votes and IMDB ratings), and previous sales, as well as random assignments. As a result, the profit-driven recommender system generates 8% less consumer surplus than the consumer-driven RSs, providing evidence for a conflict of interest between the vendor and its consumers.

Major e-vendors personalize recommendations by different algorithms that depend on how much and types of consumer information obtained. Therefore, the welfare evaluations of personalized recommendation strategies by empirical methods are hard to generalize. In Part II of my thesis, I base my analysis of personalization in RSs on a conceptual approach. Under an analytic framework of horizontal product differentiation and heterogenous consumer preferences, the resemblance of personalization to price discrimination in welfare properties is presented. Personalization is beneficial to consumers when more personalization leads to more adoption of recommendations, since it decreases search costs for more consumers. However, when the level surpasses a threshold when all consumers adopt, a more personalized RS decreases consumer surplus and only helps the firm to exploit surplus from consumers. The extreme case of perfect personalization generates the same welfare

results as first-degree price discrimination where consumers get perfectly fit recommendations but are charged their willingness-to-pay.

As shown in Part II, personalization is always profitable for the monopoly seller. In Part III, I investigate the vendor's motivation to increase the level of personalization in a two-period transactions. In the first period, consumers do not observe the true quality of the recommendations and choose to accept recommended products or not based on their initial guesses. In the second period, consumers fully learn the quality. The settings of consumer uncertainty and consumer learning incentivize the firm to charge lower-than-exploiting price for recommendations to ensure consumers' first-period adoptions of the RS. Therefore, uncertainties mediate the conflicts of interest from the vendor's exploitive behavior even though the vendor might strategically elevate consumers' initial evaluation to reduce such effect.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview of Recommender Systems and Relevant Literature

Recommender Systems (RSs) are mechanisms designed to offer suggestions of items to consumers. The recommended item could be a movie, a book, a mobile app, or even a physician for a patient or an appropriate college for a new high school graduate. Especially, RSs are widely used in E-commerce websites such as eBay, Amazon, Netflix, and Hulu because of the significant business value created for online retailers (Gomez-Uribe and Hunt, 2015). Amazon, for example, boosted sales by 29% after incorporating its recommendation mechanism (Mangalindan, 2012b). By adopting RSs, the E-commerce giant is able to increase the overall purchase experience since RSs solve the information overload problems created by the explosive growth of product varieties available online (Brynjolfsson et al., 2003). Improvement in purchase experience adds a premium on the unit price sold and increases revenue (Garfinkel et al., 2006). Additionally, availability of massive historical user online behavior data and the rapid development of collaborative filtering algorithms allow the vendor to increase sales by predicting and recommending to each consumer the products they are most likely to purchase (predictive accuracy) (Linden et al., 2003). As a result, RSs become powerful business tools to capture market share and increase profits for firms and therefore, are commonly adopted in the electronic marketplace.

My dissertation work investigates the RSs from the perspective of policymakers. It is a standard practice in literature of policy research to evaluate the impact of a specific technology by measuring the welfare impacts on the stakeholders of this technology. The stakeholders of RSs can be the vendors (suppliers/intermediaries) who design the RSs, and buyers who purchase from recommendations. The vendor's welfare is measured by the total expected profit, and the individual buyer's welfare is measured by consumer surplus (CS). A majority of the previous works present the benefits of RSs not only for sellers (Schafer et al., 1999; Fleder and Hosanagar, 2009), but also for buyers (Brynjolfsson et al., 2006; Bakos, 1997). Nonetheless, there are limited discussions about the concerns and issues caused by RSs. Stead and Gilbert (2001) review potential ethical issues of E-commerce in general, such as privacy, spamming, cybersquatters, and conflicts of interest. But the paper does not talk about RSs specifically. Another example is Pollock (2010) which investigates whether the market concentration as a result of current search engine technologies is beneficial to consumers. This paper presents evidence of a conflict of interest in the search engine market. RSs in E-commerce are similar to search engines such that they both show products listed in orders and can impact sales concentration. Inspired by a notion of conflict in the ranking algorithms, I explore related issues to concern caused by RSs with the following rationale.

In many previous works, RSs are considered beneficial to consumers. On one hand, this is because RSs reduce search costs, and on the other hand, companies are thought of doing their best to recommend the most "favorable" products to consumers. Each recommendation algorithm has an objective function that defines the fitness of the recommended products to consumers. For a period of time, numerous machine learning algorithms sprung up as a result of heated discussions on which objective function derives the most social optimal RSs, such as diversity, novelty, serendipity, and coverage (Adamopoulos and Tuzhilin, 2014; Iaquinta et al., 2008a; McNee et al., 2006a; Zhang et al., 2012).

However, there is an essential question that is often ignored: do companies actually choose the aforementioned social optimal objective functions in designing the real RSs? They are ultimately profit maximizers, so if the RSs that maximize profits do not maximize consumer welfare or total welfare, then the RSs offered by firms are sub-optimal or even hurting consumers, especially when the firm is a monopoly provider and thus has more power to manipulate demand. Furthermore, as RSs facilitate consumer search, introduction of RSs increase the manipulating power of the firms potentially. It's likely that the recommendations offered to the consumers make them reluctant to search for other items since they already face more convenient choices.

Therefore, to evaluate the welfare impact of real RSs, it is necessary to model the firm as a provider of profit-maximizing recommendations. Yet, the impact of the profit-driven RSs on the consumers' welfare has been largely unstudied. My thesis work is among the first to bridge this gap and explore the conflict of interest between the firm and its consumers in the RSs.

1.2 Problem Formulation and Introduction to My Dissertation Work

To evaluate the welfare properties of the RSs, the first step is to understand the decision processes and the interactions of the firm and its consumers in the RSs. The firm, as the provider of recommendations, decides the core attributes of RSs, such as prices, brands, and layouts of recommendations. On the other hand, the consumer decides on whether and which product to purchase, such that their consumer surplus is higher. Analyzing the past purchase data, the firm predicts the potential outcomes and chooses the design of RS that maximizes the expected aggregate profits. I summarize the key components shared by the most widely used designs into the following three core attributes:

Three Core Attributes The most commonly used type of RSs is **a limited number of suggested items in an ordered list**. The first attribute is **the selection of products** in the list of recommendations. The selected products in this list are more accessible to the target consumers than the other products outside the list. Surfing for a product incurs a search cost on the buyer. Search cost goes to consumers' utility function and is a function of the accessibility of the product. More accessible products save consumer search costs compared to less accessible products, and hence are more preferred by consumers, ceteris paribus. The differences in accessibility, between the listed items in RS and another product from the catalog, varies across companies and industries. So the impact of RS on consumer purchase decision, which depends on the difference in accessibility, also differ across companies and industries. The second attribute of RSs are **the prices of the recommended items**. Empirical studies show that recommendations not only increase sales but also give retailers flexibility to adjust their prices (Garfinkel et al., 2006). This means that since a consumer might have different willingness-to-pay for a listed product in RS than for the same product that is not listed, firm can charge consumers different prices for the listed product than the same product that is not listed. But setting prices for products in general is constrained by many factors, so such variation is only within a limited range. Practically, the profit optimization is a complicated problem: calculating the optimal prices requires not only solving analytically the profit maximization problems but also combining analytic solutions with business experience and avoiding the violation of related antitrust laws and regulations. The third attribute is **the listed order of a selected product** in the RS. Researchers and Internet marketers are interested in web page usability and report the location effects from their eye-tracking experiments (JoVE, 2016; TechWyse, 2012). They demonstrate that products at more salient positions such as left and top of a web page, on average receive more attentions and generate higher sales (TechWyse, 2012). What's more, the listed order of a product matters to consumers, in the sense that consumers think the order provides implicit information about the product's fitness to consumers. Particularly for experience goods like movies and books, consumers cannot predict their utilities from consuming a specific product merely based on the inference of quality according to the product specifications (Nelson, 1974; Garfinkel et al., 2006). Under this setting, ranking by the provider might be one of the reliable information sources that help consumers to make purchasing decisions. Consumers perceive products as ordered by priority. Furthermore, a product having a high priority gives them higher utility. In summary, selection, pricing, and ordering of products are supposed to be the three core attributes that influence consumers' purchase decisions and firm's design of RSs.

In the following sections, I am going to explore the welfare impact of online RSs when the RS provider can manipulate one or two of the three aforementioned core attributes in order to maximize profits. In Part I of this thesis, I empirically explore the effect from manipulating the listed product orders given the prices and selection of products unchanged, with the help of a randomized field experiment. Part I is an initial step to understand whether a real-world RS in use could possibly hurt consumers. Using the dataset from a specific application of RSs, I evaluate the conflict of interest between the firm and its consumers in terms of listed product orders, meanly whether the orders that are optimal to the firm are different from those optimal ones to the consumers. First of all, I estimate a exponential demand model of a specific RS application. The sales of each listed product in the RSs are predicted by their prices and the listed product orders in the list. The sales dataset I use for estimation is from a large-scale randomized experiment conducted in a newly launched RS of Video-on-Demand (VoD). The price and order elasticities in the exponential demand model are calibrated empirically by poisson regression models. With the unbiased estimates of elasticities from the randomized experiment, I am able to do counterfactual welfare analysis, comparing listed orders optimized for different objectives. The results show that the orders maximizing profit generate 8% less consumer surplus than the orders maximizing consumer surplus. Therefore, the proposed hypothesis that manipulated listed product orders may hurt consumer surplus is verified.

In the second part of the thesis, I evaluate the welfare impact when the firm can manipulate prices and selection of products such that profits are maximized. In real applications, it's hard to conduct an experiment to measure consumer demand given any combinations of products or arbitrary product prices; it's impractical for the firm to charge either an extremely high or a very low price just for experimental purposes. It is also impossible for the firm to recommend a set of products that are rarely purchased. Therefore, I choose to approach the research question by an analytical framework. Since in this part of analysis, I am not evaluating the conflict of interest that comes from the listed product order, the analytic model assumes the RSs include only one product at a time. In the conceptual framework, the firm is only able to manipulate prices and selection of products. Under this setting, I explore the influence of personalizing the RS on the purchase pattern of heterogeneous consumers in RS and the resulting welfares. Heterogeneous consumers refer to consumers with diverse tastes for products. Because in digital markets, a large variety of products are available and consumers usually have heterogeneous preferences, digital merchants are motivated to recommend products that match consumers' tastes in order to increase profits. As a result, personalization dominates the recommendation strategies of most online merchandise. There are two dimensions of such differentiation strategies that correspond to the manipulation of the two core attributes. The first one is differentiated pricing, also known as price discrimination, and the other one is the differentiated selection of products, which corresponds to "personalization" I use throughout this dissertation. Price discrimination means consumers are charged different prices for the same product. If every individual consumer is charged a unique price that equals their willingness-to-pay, then the firm exploits all surplus from consumers, which is called perfect (or first-degree) price discrimination. Various antitrust laws and regulations against price discrimination have been passed (J. Gifford and T. Kudrle, 2010). In contrast, there is no regulation on personalization strategies, even though, from my perspective, they generate comparable welfare results. My rationale of the resemblance between personalization and price discrimination is that if personalization takes in the form that each consumer is recommended a different product, firm is also able to charge the willingness-to-pay of each consumer and extracts all surpluses, the same welfare result as perfect price discrimination. What's more, similar to price discrimination, personalization might also make consumers experience unfairness when they get different products than their peers. Actually, several media and news reports have already revealed complaints from consumers about being treated unequally (Wilson, 2014; Dignan, 2012). Based on these understandings, if personalization and price discrimination resemble each other in several aspects, it is unfair that one is regulated but the other is not. The importance of this problem from the policy perspective makes it necessary to consolidate the understanding of the resemblance using in-depth theory and rigorous analysis. Specifically, I build an analytical model of heterogeneous consumer tastes and behaviors. Then I compare the two strategies, namely price discrimination versus personalization, in terms of the resultant profit, consumer surplus, and total welfare. I also investigate how levels of personalization impact the resulting welfare and the comparison of the personalization and price discrimination strategies.

In the third part of the thesis, I extend the one-period game in the second part into a multi-stage transaction between the firm and consumers because from practical lens, consumers purchase repeatedly and need an initial period of learning and forming opinions of the RS itself. Particularly in this part, I analyze the role of consumer's initial adoption in affecting firm's decision on pricing as well as increasing targetability in a two-period repeated game. Adding the first-period purchase in the analytic model allows the consumers' feedbacks from the first-period purchases to impact the second-period purchase decisions and thus the overall profits. Through this mechanism, the firm cannot arbitrarily appropriate the surplus, and is motivated to improve consumer surplus. The revised model is found to generate results that the firm as the RS provider lowers the prices of recommendations and improves the consumers' utilities from recommendations. Therefore, the setting mediates the conflicts of interest in RS between the firm and its consumers.

Chapter 2

Can Profit-Driven Recommender Systems Hurt Consumers?

2.1 Motivation

policymakers or lawyers are interested in learning the impact of digital tools like RSs on social welfare (Brynjolfsson et al., 2003; Hosanagar et al., 2014; Haubl and Trifts, 2000; Tyagi, 2004; Nijs et al., 2014; Wu et al., 2004). Before any regulations or acts are enacted on a new technology, such as Google search results or recommendations from Amazon, a comprehensive study needs to be done on its welfare impact. Specifically, the research question is, from a policymaker standpoint, whether a real-world RS can hurt consumers. The study to evaluate the welfare impact of RSs depends on the context in which RSs are applied, because RSs are applied in an explosive number of situations with diversified forms and contents. Even in its simplest form as a list of recommendations, it is likely that the number of items, the information provided for the items, and the way to arrange items vary across different types of products and different providers. Figure 2.1 presents three examples of recommendation lists that take different forms. Therefore a generalizable framework is needed to help the policymakers choose an analysis model and perform welfare evaluation in a systematic way.

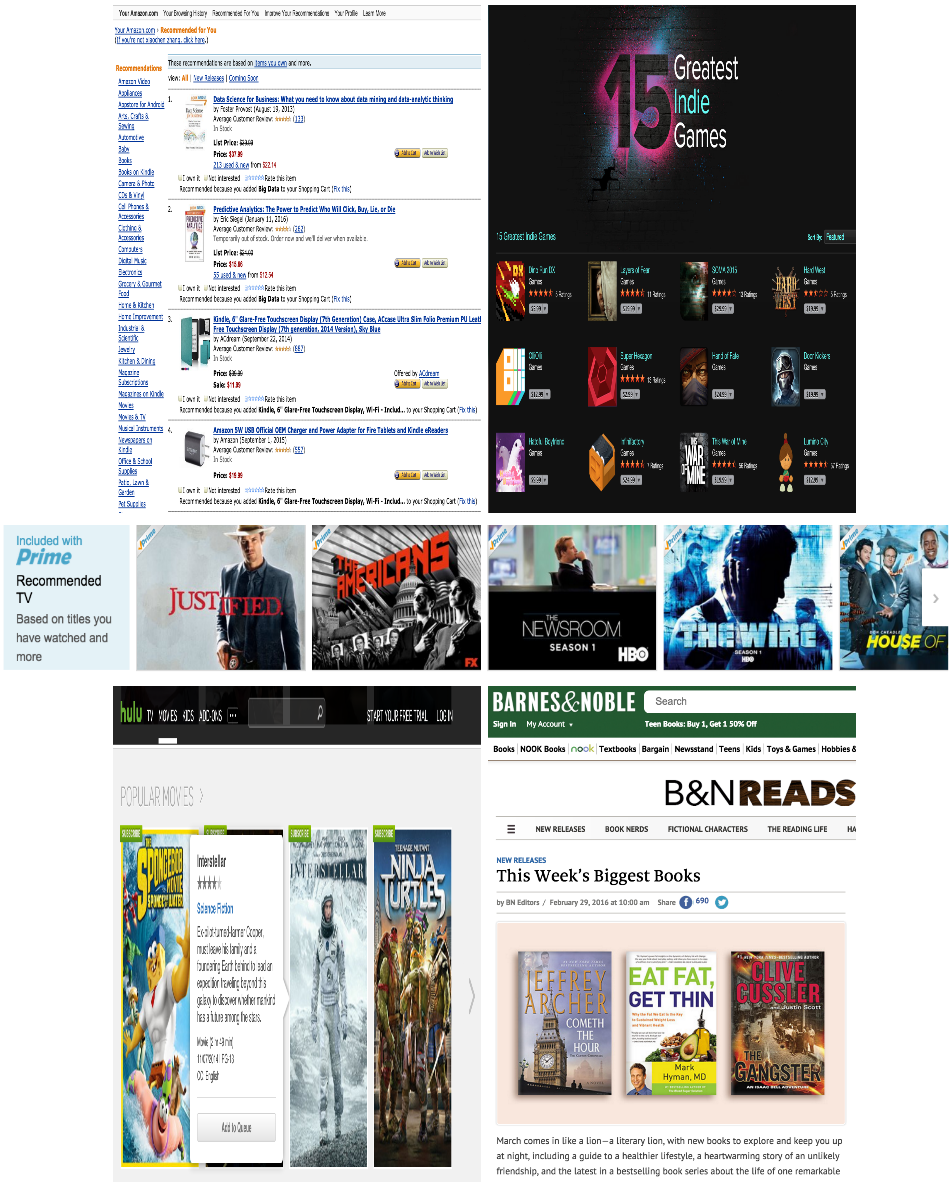

Figure 2.1: Examples of different forms of RS

2.2 Method

In this study, I propose a unified framework and exemplify the analysis process by implementing it in a typical RS. The methodology is generalizable to other application contexts of RS as well.

2.2.1 Overview: Empirical+Analytic Framework of Welfare Analysis

The framework combines empirical estimation using real experimental data and post-estimation counterfactual analysis with the aid of simulations.

The empirical estimation aims at constructing the demand model. The ultimate goal of the analysis is to evaluate the welfare impact through comparing, upon the adoption of RS, the difference between the optimal welfare function and the actual welfare. In most cases (Nicholson and Snyder, 2008), welfare measures are derived from product or service demand. Therefore, building an appropriate demand model for products or services with the application of RS is vital for exploring its welfare impact.

Constructing the demand model in my context is to empirically determine how sales are decided by the core attributes of RS and the product attributes. The resulting demand model is used to do the following counterfactual predictions on sales and welfare, which are functions of potential changes of the core attributes.

Down to the estimation procedures, I first find out what information of the RS is available to consumers and use econometric models to test which attributes are affecting consumers' purchase decision. For example, in certain recommendation sections of digital movies, consumers can view how movies are ordered in the list as well as each movie's price, year of release, director, genre, IMDB ratings, IMDB votes etc. A good choice of data analysis method is used to model how sales are predicted by different types of information perceived by consumers. Particularly, I perform several hypothesis tests on which set of information significantly explain the change in product sales. As a result, the estimated coefficients for significant properties are applied to construct the demand model in an appropriate form. Since exponential demand form, which assumes constant elasticities, is commonly used for digital products(Brynjolfsson et al., 2003), I adopt it to model the demand.

With the demand model, it can be predicted how demand changes when the firm changes the layout of RS, keeping product attributes as exogenous. When demand changes, firm's profit, consumer surplus, and total social welfare change as a result. Therefore, I am able to analyze how firm's different decisions impact the welfare. Decisions are modeled as solutions to specific optimization problems. I am interested in comparing the welfare results from decisions solving three separate optimization problems: maximizing profit, maximizing consumer surplus and maximizing total welfare. In the real world, firms aim at maximizing welfare, while the policymakers often focus on if the technology promotes social welfare without hurting consumer welfare too much. If the resultant consumer welfare, when the firm is maximizing profit welfare, is much lower than maximizing the consumer welfare, the two optimizations of maximizing profit and maximizing consumer surplus are conflicting with each other. If the firm and consumers actually have a conflict of interests, since the firm decides the core attributes of the actual RS, without any regulations, the system is trading consumer welfare off for firm's profit.

In the following sections, I implement the above framework for an application of RS in Video-on-Demand (VoD) system. Particularly, I would like to understand, when the firm's decision is merely choosing an ordering for the recommended movies, whether the consumers bear significant welfare loss as a result of firm's profit-driven decisions.

2.2.2 Theoretical Model

Demand Model

A monopoly firm offers a list of recommendations for its representative consumer i. To simplify there are two products, a and b in the list. Each product $j \in \{A, B\}$ is assigned a listed order s_j in the RSs. $s_j = 1$ means more salient order and 0 means an ordinary order. In such a RS, the demand of consumer i for the product j, denoted by d_i, is an exponential function of product listed order and the product price p_j.

$$d_j = p_j^{\beta_1+\beta_2 s_j} \ , \ \forall j \in \{A, B\} \tag{2.1}$$

, where β_1 and β_2 are elasticities that satisfy $\beta_1 < 0, \beta_2 > 0, \beta_1 + \beta_2 < -1$. Specifically, $\beta_1 < 0$ assumes the same negative baseline elasticity of demand for the two products. $\beta_2 > 0$ assumes that by putting the product at more salient order the consumer becomes inelastic. From the practical lens, this is consistent with observations that consumers usually prefer the top listed search items than the lower items.

Profit, Consumer Surplus, and Total Welfare

Assuming the cost information of both products are exogenously determined and denoted by c_A, c_B. After the firm chooses two prices p_A, p_B and the listed orders s_A, s_B, the profit (π), consumer surplus (CS), and total welfare (TW) from the RSs are presented by Eq.(2.3).

$$\begin{aligned}
\pi(s_A, s_B, p_A, p_B) &= d_A(p_A - c_A) + d_B(p_B - c_B) \\
&= p_A^{\beta_1+\beta_2 s_A}(p_A - c_A) + p_B^{\beta_1+\beta_2 s_B}(p_B - c_B) \\
CS(s_A, s_B, p_A, p_B) &= \int_{p_A}^{-\infty} d_A(p)dp + \int_{p_B}^{-\infty} d_B(p)dp \\
&= \frac{p_A^{1+\beta_1+\beta_2 s_A}}{1+\beta_1+\beta_2 s_A} + \frac{p_B^{1+\beta_1+\beta_2 s_B}}{1+\beta_1+\beta_2 s_B} \qquad (2.2) \\
TW(s_A, s_B, p_A, p_B) &= \pi(s_A, s_B, p_A, p_B) + CS(s_A, s_B, p_A, p_B) \qquad (2.3)
\end{aligned}$$

Firm could choose the order of the two products and set the prices of the two products in the RSs to maximize its total profit. There are two ways of ordering that firm could recommend products. $s = (s_A, s_B) = (1, 0)$, labeled by $A1B2$, or $(0, 1)$, labeled by $A2B1$.

Price

Writing down the lagrangian for solving the optimal prices and ordering simultaneously, it is easy to see that price optimization is independent of listed order optimization. Given a listed order s_j for product j, the optimal price of j is,

$$p_j^* = c_j \frac{\beta_1 + \beta_2 s_j}{1 + \beta_1 + \beta_2 s_j} \tag{2.4}$$

The price margin is

$$m_j = p_j - c_j = -\frac{1}{1 + \beta_1 + \beta_2 s_j} c_j \tag{2.5}$$

Optimal Ordering for the firm, Consumer Surplus, and Total Welfare

The following propositions for the optimal ordering for firm, consumers, and total welfare can be proved.

Proposition 1 Define $F = (\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})[\frac{\beta_1(1+\beta_1+\beta_2)}{(1+\beta_1)(\beta_1+\beta_2)}]^{\frac{\beta_1}{\beta_2}}$, if $c_A < c_B < F$, then $\pi(A1B2) \leq \pi(A2B1)$, and it's optimal to put product A in the salient order. If $F < c_A < c_B$, $A2B1$ generates more profit for the firm.

PROOF: See Appendix A.1

Proposition 2 Define $G = (\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})[\frac{\beta_1(1+\beta_1+\beta_2)}{(1+\beta_1)(\beta_1+\beta_2)}]^{\frac{1+\beta_1}{\beta_2}}$, if $c_A < c_B < G$, then $CS(A1B2) \leq CS(A2B1)$, and putting product A in the salient order generates higher consumer surplus. If $G < c_A < c_B$, $A2B1$ maximizes consumer surplus.

PROOF: See Appendix A.2

Proposition 3 Define $H = (\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})^{\frac{\beta_2+\beta_1}{\beta_2}}(\frac{\beta_1}{1+\beta_1})^{\frac{\beta_1}{\beta_2}}[\frac{(1+\beta_1+\beta_2)(1+2\beta_1)}{(1+\beta_1)(1+2\beta_1+2\beta_2)}]^{\frac{1}{\beta_2}}$, if $c_A < c_B < H$, then $TW(A1B2) \leq TW(A2B1)$, and putting product A in the salient order generates higher total welfare. If $H < c_A < c_B$, $A2B1$ maximizes total welfare.

PROOF: See Appendix A.3

Conflicts of Interests

Prop 1, 2, and 3 together show that when the values of F, G and H are on the two sides of the costs c_A, c_B, there exist conflicts between different optimization problems in preference of orderings. To figure out the relationship of different conflict condition, I need the following propositon 4.

Proposition 4 $G \leq H \leq F$

PROOF : See Appendix A.4

When $G < C_A < C_B < F$, I have $\pi(A1B2) < \pi(A2B1)$, and $CS(A1B2) > CS(A2B1)$. There is a conflict between firm and consumers: firm prefers low-cost product at salient order, while consumers prefer low-cost product at ordinary order.

When $H < C_A < C_B < F$, $\pi(A1B2) < \pi(A2B1)$, and $TW(A1B2) > TW(A2B1)$. Firm will choose the listed orders that produce lower total welfare.

Determinants of the conflicts

The essential component that leads to the conflict is the change of elasticity, β_2, caused by the difference in listed orders. The obvious evidence is when β_2 is 0, $F = G = H$, there is no region of conflicts for F, G, and H. Particularly, the following proposition 5 and 6 show how β_2 changes the conflict region size and the magnitude of conflicts.

Proposition 5 The sizes of conflict regions increase in β_2.

PROOF : See Appendix A.5

Proposition 6 The magnitude of conflicts increases in β_2. Specifically, when the firm and consumers prefer different assignments of listed orders, the loss in consumer surplus increases in β_2 when firm chooses the ordering that maximize profit.

PROOF : See Appendix A.6

With the idea of a conflict of interest between firm and consumers driven by the ordering of recommended products in theory, the following sections discuss how actual conflicts in welfare can be identified and measured empirically in a real life RSs.

2.2.3 Data and Experiment

VoD Movies

Some telecom companies offer the VoD service as part of their TV channel subscription packages, such as Xfinity by Comcast and Fios by Verizon. Households who subscribe to the service can select to watch any video listed from their service homepage on TV. There are two types of payment methods for the VoD service: pay-per-view payment, in which the customer pay a fee each time it requests to watch a video, and free-with-subscription payment wherein a customer pays the subscription fee upfront at the beginning of each billing period and can watch any video included in the subscription package. On the telecom company's side, it signs contracts with digital movie providers and, for each video, pays its suppliers a fixed cost or a fixed commission (such as 20% of the pay-per-view price) per streaming.

I have collaborated with our industrial partner, a European major telecom company who offers TV, internet, and phone services, to conduct a large-scale real experiments on their TV sector VoD system. This experiment was done before the analysis, initially to investigate the effect of price discounts. The details of the experiment setup are explained in Godinho de Matos et al. (2015b).

Recommendations offered by VoD Service

On the homepage of the TV VoD system provided by our collaborator, there is a relatively new list of recommendations, in which nine video titles are visible to the customers. For experimental purpose, during the experiment, those titles were chosen randomly from the popular items in previous sales and are arranged randomly. Yet those recommended video titles were perceived by the consumers as being carefully chosen and arranged to given them better consumer experiences. Consumers decided whether or not, and which movie to purchase based on the perceived movie characteristics, such as prices, genre, authors etc, and the implicit hints from the arrangement of the recommendations, such as the ranking of videos from left to right, and the mere fact that they had stund out from other videos and had been selected to be recommended. Movie characteristics, their positions in the recommendation section, and monthly sales of movies were collected during the 26 weeks of the experiments. In the analysis, the collected datasets are used for demand estimation. In the post-estimation counterfactual analysis, I plug in the exogenous movie characteristics into the demand model, and calculate the decisions on ordering of recommendations, with objectives of either maximizing firm's profit or maximizing consumer surplus. The details are explained in section 2.2.4.

Randomized Experiment

Motivation of Randomization: Endogeneity of Price and Ordering In the real world, prices are endogenous, because to maximize profit, firms set prices strategically in response to changes in consumers' demand. Likewise, the order of recommendations chosen by the firm is also affected by the sales of those recommended movies. For example, the firm might reserve the topmost positions for products with higher popularity (endogeneity of order), or with larger pricing margin (endogeneity of order caused by collinearity between order and endogenous prices). Therefore, the ordinary least squares estimation of price elasticity and ordering effect on consumer demand will be biased and erroneous. With the help of the randomized experiment in the

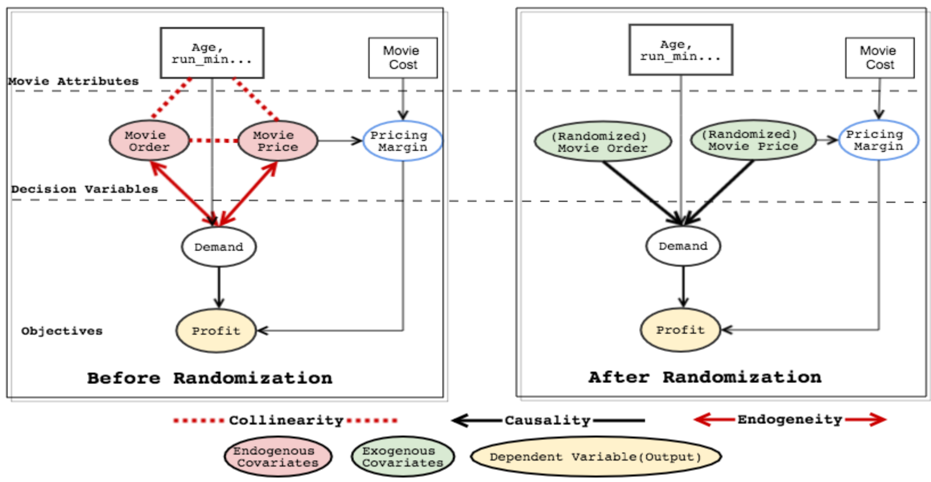

Figure 2.2: How randomization breaks the endogeneity of order and price effect

RS of VoD, I am able to break the price and order endogeneity with the randomized movie price as well as the randomized ordering of movies in each recommendation list. Figure 2.2 shows the relationship between movie price, order, demand, and the benefit of randomized design. The following two subsections describe the randomized design in details.

Households: Random Group Assignment The experiment was run for five consecutive test months from September 2013 to February 2014. Before the experiment, about 300,000 households (premium subscribers) were randomly partitioned into four control groups and four treatment groups, and each household remained in the same group throughout the experiment. In each test month, households within the same household group were recommended a list of nine movies with the same titles and in the same order.

Movies: Randomized Movie Order & Randomized Movie Price In each test month, a different set of 36 movies were randomly selected, which were all the best-selling titles from the most recent sales data. The selected movie titles were randomly partitioned into 4 groups, with 9 movies in each group.

The original prices of the movies are denoted by P_o. To break the price endogeneity, each movie also had a randomized price P_r. Each movie group is assigned to a different consumer group. Four movie groups with original prices P_o were recommended to the four control groups of the households, while the other four movie groups, which were duplicates of the first four movie group titles but with randomized prices P_d, were matched to the four treatment group of households.

To break the order endogeneity, the order of recommendation that each household saw was also random. However the random ordering was set according to a universal random rank of the 36[1] movies, which means if movie X and Y were both recommended to two household groups A and B, and X was ordered before Y to customers in group A, then X was also placed before Y to customers in group B.

2.2.4 Data Analysis

Estimate the Demand Model

To build up the exponential demand function for the VoD RSs, the first step is to find the significant factors that affect demand. The demand is described by how many times each movie was leased per months (*n_lease*). Since there are many zeros (not leased at all) in the data, I chose to use the Poisson regression model which is a standard econometric model for the count data with lots of zeros. I tested several movie attributes and analyzed which attributes affect demand significantly. The attributes included were movie prices, IMDB ratings(*IMDBRating*), IMDB votes(*IMDBVotes*), movie order(*order*), year of release (*age*), movie length(*runtime_mins*), etc.

[1] Actually there are 45 movies selected for each months. But only 36 movies, which partitioned to groups of nine movies, are visible to consumers on the homepage. There are additionally 9 popular movies selected. On the home page there are nine movies listed as recommendations. Consumer can see the hidden list of six movies by scrolling to the right. The movies in the hidden lists are randomly selected from the additional nine movies.

Since movie orders are discrete, *order* $\in \{1, 2, ..., 9\}$, I created a new variable *isOrderLowerThanR* as:

$$\textit{isOrderLowerThanR} := 1 \text{ if } \textit{order} < R,\ o.w.\ 0 \tag{2.6}$$

As previously mentioned in introduction section, items put into top orders are perceived by consumers as better. In equation 2.6, R describes how far down the recommendation list is the item a consumer will not perceive as more attractive.

I also created another variable, *TopQuartile*, to show the popularity of a movie:

$$\textit{Topquartile} := 1 \text{ if } \textit{IMDBVotes} > x, x = \min\{x : P(\textit{IMDBVotes} > x) = 0.25\},\ o.w.\ 0 \tag{2.7}$$

. The initial results of significance demonstrate that the influential factors are *price, isOrderLowerThanR consumerType* and *IMDBVotes*. Specifically, I have found that $R = 3$ fits the model better, which means consumers consider the two leftmost movies as better fitting to their tastes. The detailed results and discussions will be presented in section 2.3.1.

As a result, an explanatory model for sales of an average movie in popularity from an average consumer is:

$$logE[n\_lease] = \beta_0 + \beta_1 \log price + \beta_2 \log price \cdot isOrderLowerThanR \tag{2.8}$$

Now with all the identified significant factors, I can plug in their estimated coefficients into an exponential form of demand model. For a specific group of consumers (considered with homogenous preferences, or a representative consumer), their aggregate demand is described by the following equation

$$\forall j \in \text{all movies}, d_j = A_j p_j^{\beta_1 + \beta_2 s_j} \tag{2.9}$$

, where d denotes the demand, p denotes price, s denotes whether the movie is put into the top orders. A captures the total market size and the movies' popularity. β_1 and β_2 in equation 2.9

and equation 2.8 are the same. β_1 denotes the price elasticity of the recommended movies except for the two leftmost movies. $\beta_1 + \beta_2$ is the price elasticity for the two leftmost movies in the list. Table 2.1 shows $\beta_1 = -0.443$, $\beta_2 = 0.08$.

Optimize Listed Orders of Simulated RS for Firm and Consumers

This section explains the counterfactual welfare analysis using the demand model constructed from previous empirical estimations.

In the real world, the firm designs the RS to maximize the total profit. To evaluate if consumers suffer the potential loss in surplus in RS, the real-world welfare is compared to the welfares of two counterfactual worlds: the world when the firm maximizes consumer surplus (CS) instead and the world when the total welfare-sum of profit and CS-is maximized.

Assuming the price of each item, and which set of products to put on the list are exogenously set, the firm's decision is only to choose the ordering of the selected items. Thus the conflict, if any, comes from the difference in optimal ordering as the solution to each maximization problem. Put the problem in the experimental context of VoD RS, given a selection of popular movies, which two movies should the firm put into the two leftmost positions, so that predicted profit, predicted CS, or the predicted total welfare is be maximized? Will CS be significantly lower than its maximum value, when the movies are sorted in a way that maximizes profit?

I compare the welfare results with different optimization targets using 1000 simulated sets of recommendations. Each set of recommendations is a set of 15 movies randomly selected from the 259 [2] distinct movie titles which I have used in the empirical estimation. In each simulation, with each optimization objective, I find two movies and put them into the two leftmost slots. After solving the 1000 optimal solutions of movie orders, I sum up the profits, consumer surplus and total welfare of all 1000 simulations. The following three sections describe the three optimization problems with equations.

[2] There are 45 movies/month *8 months = 360 movies used in the datasets. After deleting duplicates and invalid data points, 259 movies are left.

Table 2.1: Significant Covariates Result: movie order and price

	Dependent variable:							
	sales							
	NO order	R=2	R=3	R=4	R=5	R=6	R=7	EXACT order
log(price)	.427***	.438***	.442***	.440***	.438***	.424***	.425***	.391***
	(.128)	(.129)	(.129)	(.129)	(.129)	(.129)	(.129)	(.131)
log(price)*(is low-order movie)		.116***	.075**	.052*	.021	.005	.002	
		(.041)	(.033)	(.031)	(.027)	(.025)	(.024)	
log(price)*order								.006
								(.004)
intercept	.169	.004	.126	.204	.154	.171	.168	.047
	(1.014)	(1.017)	(1.013)	(1.013)	(1.013)	(1.014)	(1.014)	(1.016)
movie dummies?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
time dummies?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
consumer type dummies?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
consumer group dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,680	4,680	4,680	4,680	4,680	4,680	4,680	4,680
Log Likelihood	−3,428.283	−3,424.244	−3,425.671	−3,426.894	−3,427.960	−3,428.263	−3,428.279	−3,427.226
Akaike Inf. Crit.	7,446.567	7,440.487	7,443.342	7,445.788	7,447.921	7,448.526	7,448.559	7,446.451

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Profit Maximization The profit function derived from the demand function described by Eq.(2.9) is:

$$\pi(\{s_j\}_{j=1...15}) = \sum_{j \in \{1...15\}} d_j(p_j - c_j), d_j = A_j p_j^{\beta_1 + \beta_2 s_j} \tag{2.10}$$

where c_j denotes the cost of movie j and $s_j = 1$ if movie j is at the top two slots in the recommendation list. Parameters β_1, β_2, and A_j in the demand function can be estimated from Eq.(2.8). Combined with the price and cost information I have, sales and profits can be predicted for any ordering of a set of 15 movies selected from all the 259 movies. Put in another way, the sales and profits of the selected 15 movies are functions of their ordering (In fact, it only matters which movies are put into more salient position, i.e. ”low order”).

Therefore, with exogneous prices and costs, the profit maximization problem is described by Eq.(2.11):

$$\pi^* = \max_{s_j, j=1...15} \sum_{j \in \{1...15\}} A_j p_j^{\beta_1 + \beta_2 s_j}(p_j - c_j) \tag{2.11}$$
$$s.t. \sum_j s_j = 2, s_j \in \{0, 1\}$$

To solve the maximization problem, I rewrote equation (2.11) as:

$$\pi(\{s_j\}_{j=1...15}) = \sum_{j \in \{1...15\}} A_j p_j^{\beta_1}(p_j - c_j) + \sum_{k, s.t. s_k=1} A_k p_k^{\beta_1}(p_k - c_k)(p_k^{\beta_2} - 1) \tag{2.12}$$

So the movies to put into the top two slots down the recommendation list are those j’s that generate highest value of $A_j p_j^{\beta_1}(p_j - c_j)(p_j^{\beta_2} - 1)$

Consumer Welfare Maximization The most commonly used consumer welfare measure in public policy literature is consumer surplus (CS) [XXcitations of paperXX]. I computed the consumer surplus of the 15 recommended movies from each of the 1000 simulations in the following way

$$CS(\{s_j\}_{j=1...15}) = \sum_{j\in\{1...15\}} \int_{p_j}^{p_{max}} A_j \tau^{\beta_1+\beta_2 s_j} d\tau \tag{2.13}$$

where p_{max} is the upper bound of the movie prices. Since $\beta_1 + \beta_2 \in (0, 1)$, the original half-closed integral form of CS is divergent. Here by setting an upper bound for the integral, CS is computed as relative CS to a reference maximal CS.

$$\begin{aligned} CS^* = \max_{s_j, j=1...15} & \sum_{j\in\{1...15\}} \frac{A_j p_{max}^{1+\beta_1+\beta_2 s_j}}{1+\beta_1+\beta_2 s_j} - \frac{A_j p_j^{1+\beta_1+\beta_2 s_j}}{1+\beta_1+\beta_2 s_j} \\ & s.t. \sum_j s_j = 2, s_j \in \{0, 1\} \end{aligned} \tag{2.14}$$

To solve this optimization problem and search for best s_j's, rewrite equation (2.14):

$$\begin{aligned} CS^* = \max_{s_j, j=1...15} & \frac{\sum_{j\in\{1...15\}} A_j (p_{max}^{1+\beta_1} - p_j^{1+\beta_1})}{1+\beta_1} \\ & + \sum_{k, s.t. s_k=1} \frac{A_j p_{max}^{1+\beta_1+\beta_2}}{1+\beta_1+\beta_2} - \frac{A_j p_j^{1+\beta_1+\beta_2}}{1+\beta_1+\beta_2} - \frac{A_j p_{max}^{1+\beta_1}}{1+\beta_1} + \frac{A_j p_j^{1+\beta_1}}{1+\beta_1} \\ & s.t. \sum_j s_j = 2, s_j \in \{0, 1\} \end{aligned} \tag{2.15}$$

So the movies to put into the top two slots down the recommendation list are those j's that generate highest value of $\frac{A_j p_{max}^{1+\beta_1+\beta_2}}{1+\beta_1+\beta_2} - \frac{A_j p_j^{1+\beta_1+\beta_2}}{1+\beta_1+\beta_2} - \frac{A_j p_{max}^{1+\beta_1}}{1+\beta_1} + \frac{A_j p_j^{1+\beta_1}}{1+\beta_1}$

Measure Consumer/Total Surplus Loss of Simulated Profit-driven RS

Let $S^b = \{s_j\}_{j=1...15}$ denote the ordering of 15 items in simulation b. In each simulation b, after solving the optimal ordering for firm, S_f^b, and for consumer, S_c^b, the profits and CS in each cases were computed as well, i.e. $\pi(S_f^b)$, CS(S_f^b), and $\pi(S_c^b)$, CS(S_c^b). The total welfare, denoted by

W, was also measured as

$$
\begin{aligned}
W^b(S^b_f) &= \pi^b(S^b_f) + \text{CS}^b(S^b_f) \\
W^b(S^b_c) &= \pi^b(S^b_c) + \text{CS}^b(S^b_c)
\end{aligned}
\tag{2.16}
$$

Therefore if the firm chooses the order of recommendations to maximize profits, then the total profit is $\sum_{b=1}^{1000} \pi^b(S^b_f)$, the consumer surplus is $\sum_{b=1}^{1000} \text{CS}^b(S^b_f)$, the total welfare is $\sum_{b=1}^{1000} W^b(S^b_f)$. The best order for consumers on the other hand would generate the three welfare measures as $\sum_{b=1}^{1000} \pi^b(S^b_c)$, $\sum_{b=1}^{1000} \text{CS}^b(S^b_c)$, and $\sum_{b=1}^{1000} W^b(S^b_c)$. There is a conflict of interest between firm and consumer in how to order recommendations, if $\sum_{b=1}^{1000} \text{CS}^b(S^b_f) < \sum_{b=1}^{1000} \text{CS}^b(S^b_c)$. I measure this relative loss in CS as:

$$
RL = \frac{100\big(\sum_{b=1}^{1000} \text{CS}^b(S^b_c) - \sum_{b=1}^{1000} \text{CS}^b(S^b_f)\big)}{\sum_{b=1}^{1000} \text{CS}^b(S^b_c)}\% \tag{2.17}
$$

2.3 Empirical Results with Discussions

2.3.1 Initial Result of Significance & Estimated Demand Models

This section explains the initial result from the empirical estimation. Throwing all collected into the regresson model as predictor variables and using monthly movie sales as dependent variables, I am able to uncover which information sources really impact consumers' purchase decisions, i.e. how movie demand can be inferred by the various properties of RS.

General Form of the Demand Model: Price, Order and Fixed effects

I have found that movie sales are significantly predicted by movie prices and how movies are ordered in the RS (table 2.2,2.1,2.3, and 2.4). What's more, in table 2.2, model(1) and (2) include the following three types of fixed effects: the period of time that the movie is recommended to its target households during in the experiment (time dummy), whether the sales comes from premium or standard consumers[3] (consumer type dummy), and the size of household group that the movie is recommended to (consumer group dummy). Each of model(5), (6), (7) leaves out one of the above three fixed effects, and their resultant log-likelihoods are lower. In addition, comparing model(4) to model(3), after adding several movie essential attributes-year of release, IMDB ratings, and whether the movie is most voted (above 75 percentile) in IMDB, performance of the model increases. The signs of the coefficients reflect that an average consumer prefers the newer movie, and is more likely to buy it if the movie receives many votes and a high rating. In order to explain all differences between movies, I add the movie fixed effect instead of individual characteristics, as in model(2) which generates higher log-likelihood than model(3) and (4). These initial significant tests and regression results reveal that movie sales are time-variant, differentiated across consumer types, and correlated with prices and fixed movie characteristics, such as movie age, genre and reputations (table2.2). In the first step, I have identified the significant properties of RS. Now I want to test if the functional form of each predictor is appropriate and if there are any interactions among predictors. Table 2.1, 2.3, and 2.4 present different ways to include the influential factor movie order. Comparing different models in each table, I find that, $isLowerThanR$ generally performs better than excluding factor of the movie order or including the $order$ in exact term (discrete values from 1 to 9). $R = 3$ is the best of all $isLowerThanR$ model(column 2-7 in table 2.1, column 2-6 in table 2.3, and column 2-6 in table 2.4). Performance is measured in log-likelihood.

I also find there is an interaction between movie price and movie order, i.e. price elasticity varies with different levels of movie order. Leftmost movies generate higher sales because their sales are less elastic to prices than movies placed relatively right on the list, creating an opportunity for sellers to increase prices without lowering too much demand. The estimated price elasticity for the two leftmost movies is -0.443, and -0.357 for the other movies. In addition,

[3]Premium consumers pay higher membership fees than standard consumers, and shop regularly for our collaborator's products and services

Table 2.2: Initial result of significant tests on how RS attributes and movie attributes impact sales

	Dependent variable:						
	sales						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(price)	$-.430^{***}$	$-.430^{***}$	$-.057$	$-.138^{*}$	$-.457^{***}$	$-.420^{***}$	$-.554^{***}$
	(.128)	(.128)	(.068)	(.078)	(.121)	(.128)	(.111)
order	$-.021$						
	(.021)						
is low-order movie?		$.387^{**}$	$.175^{***}$	$.125^{*}$	$.317^{**}$	$.387^{**}$	.019
		(.193)	(.062)	(.068)	(.158)	(.193)	(.173)
years since- -first release				$-.017^{***}$ (.006)			
IMDB Rating				.026			
				(.025)			
is more voted?				-1.259^{***}			
				(.451)			
IMDB Rating *(is more voted?)				$.137^{**}$ (.063)			
intercept	.238	.083	-1.543^{***}	-1.370^{***}	$-.072$	.560	1.457
	(1.015)	(1.014)	(.388)	(.464)	(.966)	(1.013)	(.974)
movie dummies?	Yes	Yes	No	No	Yes	Yes	Yes
time dummies?	Yes	Yes	Yes	Yes	No	Yes	Yes
consumer type?	Yes	Yes	Yes	Yes	Yes	No	Yes
consumer group?	Yes	Yes	Yes	Yes	Yes	Yes	No
movie genre?	No	No	No	Yes	No	No	No
Observations	4,680	4,680	4,680	4,256	4,680	4,680	4,680
Log Likelihood	−3,427.782	−3,426.245	−3,814.510	−3,571.949	−3,566.871	−3,617.830	−3,447.640
Akaike Inf. Crit.	7,447.564	7,444.490	7,697.020	7,249.898	7,677.743	7,825.660	7,477.280

Note: *p<0.1; **p<0.05; ***p<0.01

The siginificant movie genres in model(4) are comedy(Positive***), drama(Positive***), kids(Positive*), suspense(Negative**).

Table 2.3: Significant Covariates Result: consumer type, movie order and price

	Dependent variable:							
	sales							
	NO order	R=2	R=3	R=4	R=5	R=6	R=7	EXACT order
log(price)	.288**	.298**	.303**	.301**	.301**	.286**	.288**	.250*
	(.134)	(.135)	(.135)	(.135)	(.135)	(.135)	(.135)	(.137)
is standard consumer?	1.672**	1.691**	1.673**	1.671**	1.666**	1.673**	1.671**	1.629**
	(.718)	(.718)	(.719)	(.718)	(.718)	(.718)	(.718)	(.719)
log(price)*(is low-order movie)		.122***	.074**	.052	.025	.004	.0004	
		(.042)	(.033)	(.032)	(.027)	(.025)	(.024)	
log(price)*order								.006
								(.004)
log(price)*(is standard consumer?)	.461***	.463***	.461***	.461***	.456***	.460***	.457***	.471***
	(.130)	(.130)	(.131)	(.131)	(.131)	(.131)	(.131)	(.131)
log(price)*(is low-order movie) *(is standard consumer?)		.023	.001	.0004	.012	.002	.008	
		(.030)	(.022)	(.020)	(.019)	(.017)	(.017)	
log(price)*order *(is standard consumer?)								.002
								(.002)
intercept	.267	.093	.224	.302	.256	.269	.269	.162
	(1.038)	(1.042)	(1.037)	(1.038)	(1.037)	(1.038)	(1.038)	(1.040)
movie dummies?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
time dummies?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
consumer type dummies?	No	No	No	No	No	No	No	No
consumer group dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,680	4,680	4,680	4,680	4,680	4,680	4,680	4,680
Log Likelihood	−3,422.051	−3,417.726	−3,419.441	−3,420.663	−3,421.502	−3,422.024	−3,421.923	−3,420.359
Akaike Inf. Crit.	7,436.102	7,431.451	7,434.881	7,437.326	7,439.003	7,440.049	7,439.846	7,436.719

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 2.4: Significant Covariates Result: movie popularity, movie order and price

	Dependent variable:						
	sales						
	NO order	R=2	R=3	R=4	R=5	R=6	EXACT order
log(price)	−.571***	−.576***	−.584***	−.582***	−.574***	−.563***	−.563***
	(.144)	(.144)	(.145)	(.145)	(.145)	(.145)	(.147)
log(price)*(is low-order movie)		.049	.039	.027	−.011	−.035	
		(.051)	(.039)	(.038)	(.031)	(.028)	
log(price)*order							−.002
							(.004)
log(price)*(more IMDB votes?)	.534**	.514**	.517**	.498*	.499*	.502*	.643**
	(.258)	(.258)	(.258)	(.258)	(.258)	(.258)	(.263)
log(price)*(is low-order movie)		.969	.149**	.160**	.151**	.136**	
*(more IMDB votes?)		(1.075)	(.069)	(.069)	(.061)	(.056)	
log(price)*order							−.011*
*(more IMDB votes?)							(.006)
intercept	.935	.955	.760	.774	.611	.635	.625
	(1.067)	(1.068)	(1.069)	(1.070)	(1.072)	(1.073)	(1.072)
movie dummies?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
time dummies?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
consumer type dummies?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
consumer group dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,592	4,592	4,592	4,592	4,592	4,592	4,592
Log Likelihood	−3,365.180	−3,360.242	−3,360.189	−3,360.302	−3,361.683	−3,362.165	−3,362.444
Akaike Inf. Crit.	7,308.360	7,302.483	7,302.379	7,302.604	7,305.366	7,306.330	7,306.887

Note: *p<0.1; **p<0.05; ***p<0.01

the interaction effect between prices and orders are different between consumer types (table 2.4), and varies with different levels of movie popularity (table 2.3). The details are presented in the following sections.

Heterogeneity of Consumers

The regression results show the heterogeneity across consumers in the experiment. As previously mentioned, I detect significant consumer type fixed effect in the regression, which indicates an average difference of demands between premium and standard consumers. Besides the difference in fixed effect, premium and standard consumers are also different in terms of how their demands are affect by prices, and the how order changed price elasticity.

Table 2.3 shows that movie orders have a significantly larger influence over premium consumer in $isLowerThanR$ models with $R \leq 3$ than standard consumers. When $R = 3$, the interaction between movie order and price (β_2) is $0.084 (p < 0.05)$ for the premium consumers and $0.084 - 0.028 = 0.066 (p < 0.1)$ for the standard consumers. Price elasticity of premium consumers, for movies not in the two leftmost orders, is -0.306 while that of the standard consumer is $-0.306 - 0.455 = -0.761$. The estimate of price elasticity, -0.443, from table 2.1 is the weighted average across all consumers including premium and standard consumers.

Different Interactions Effect of Movie Orders by Levels of Movie Popularity

259 distinct movies used in the Randomized experiments were picked from the most popular movies in previous three months before the experiments. They are comparable in terms of the popularity. However among them there are some relatively more popular movies. To capture that effect, in the estimation I also included a covariate called $top\_quartile$, defined earlier in equation 2.7, to measure the effect of movie popularity on the main estimates. The estimation result is reflected in table 2.4. I discovered that movie popularity affects the price elasticity and the magnitude of ordering effect on price elasticity. Still regression model with $R = 3$ gave the

best log-likelihood. For $R = 3$, movies that have the most IMDB Votes (more votes than 75% of the 259 movies) have very small price elasticity in absolute value, which is $-0.384 + 0.317 = -0.067$. For $R = 3$, $\beta_2 = 0.149$, which means, by putting the most popular movies in the top two slots down the recommendation list, the price elasticity increased by 0.149, and reaches a positive value of 0.082. It's very likely that the movies that the consumers interpreted price increase of most popular movies as a signal of increase in quality. In addition, the results of inelastic demands (price elasticity less than 1) revealed that most popular movies have stable demand, less affected by price change within a certain range.

Summary of Estimated Demand Models

Summarizing the empirical results, I identified the significance of movie, time, household group size, and consumer type fixed effects and, in the demand model, I used these estimates to calculate A_j for each movie. I also estimated the price elasticities and the interaction between movie price and order, which are represented by β_1, β_2 respectively for the demand model.

Since I have found a heterogenity in consumer demand, can design the RS that fits two types of demand model-general demand model, and consumer-type demand model.

1. General demand model, expressed in equation 2.9, which is estimated from the following regression model:

$$log(E(n\_lease|X)) = A + \beta_1 \cdot logprice + \beta_2 \cdot logprice \cdot (order < R) + M + T + C \tag{2.18}$$

2. Two-type consumer demand model, expressed by

$$\forall j \in \text{all movies}, d_j = A_{j,prem} \cdot p_j^{\beta_{1,prem}+\beta_{2,prem}(s_j)} + A_{j,stand} \cdot p_j^{\beta_{1,stand}+\beta_{2,stand}(s_j)} \tag{2.19}$$

, estimated from:

$$
\begin{aligned}
log(E(n\_lease|X)) = {} & A_{standard\_consumers} + A_{premium\_consumers} \\
& + \beta_{1,standard\_consumers} \cdot logprice + \beta_{1,premium\_consumers} \cdot logprice \\
& + \beta_{2,standard\_consumers} \cdot logprice \cdot (order < R) \\
& + \beta_{2,premium\_consumers} \cdot logprice \cdot (order < R) + M + T + C
\end{aligned}
\tag{2.20}
$$

In equation 2.20 and 2.18, M denotes movie fixed effect, T denotes the time fixed effect, and C denotes the fixed effect of household group. The estimates from the regression models are shown in table 2.3.

2.3.2 Welfare Results

Comparisons of Welfares from Different Optimizations

With the general and consumer-type demand model, I am able to search the orderings for the 1000 simulated 15-movie sets and find the optimal RS with objectives of maximizing profit, CS and total profit. Thre are three types of RS that correspond to different types of demand models:

1. general: assuming general demand model
2. ct: Assuming two types of consumers- premium and standard(half and half in the population). I choose ordering of items to optimize profit/cs/social welfare for each group individually, and the total profit/cs/social welfare is the sum of the individually optimized ones. This is a personalized version of RS.
3. ct.noPersi: Assuming two types of consumers: premium and standard(half and half in the population), choose the same ordering of items to optimize the sum of the two groups' welfare. This is a non-personalized version of RS.

The results show that the CS, when ordering of RS is chosen to maximize profit, is less than the maximal CS. Using equation (2.17) to calculate the relative loss, I find that in this specific RS of VoD movies, the profit maximization could hurt consumers' potential welfare by a magnitude of about 8%.

Consumer relative loss of welfare is one way to measure the confict of interest between the firm and the consumers. Profit gain of switching objective from maximizing CS to maximizing profit is another measure of the conflict. It's also the motivation for the firm to trade off CS for the profit. Additionally, it's also important to compare the total welfare from maximing profit to the maximal total welfare. Therefore I define the following three measure of conflict:

$$\begin{aligned} CS\ Loss &= 1 - \frac{CS_{\ maximizing\ \pi}}{maximized\ CS} \\ Profit\ Gain &= \frac{maximized\ \pi}{\pi_{\ maximizing\ CS}} - 1 \\ Welfare\ Loss &= 1 - \frac{total\ welfare_{\ maximizing\ \pi}}{total\ welfare_{maximizing\ CS}} \end{aligned} \tag{2.21}$$

The calculation using the above equations generates the gain and loss in welfare:

- Gain in Profit: $0.07170098 \approx 7\%$ (compared to CS-driven RS)
- Loss in Consumer Surplus: $0.084426 \approx 8\%$
- Loss in Total Welfare: $0.02073565 \approx 2\%$

The above results were calculated for the ct type of RS assuming two-type-consumer. However these results are robust The above three results provide evidence to the existence of conflicts

Table 2.5: CS, profit and total welfare loss table

RS-type	Demand model	Personalized?	CS loss	Profit loss	Welfare loss
General	General	No	7.14%	8.41%	2.06%
ct	Consumer-type	Y	7.17%	8.44%	2.07%
ct.noPersi	Consumer-type	N	7.08%	8.41%	2.06%

between the firm and the consumers. Welfare results from the simulations demonstrate that the firm is able to gain larger profits by sorting the recommendations differently than the way that maximizes total CS. As a result, the profit-driven RS generates lower CS and total welfare. Consumers endure a loss in CS and total welfare is not at its optimal level.

Another way to present the conflict of interest is to graphically compare the three types of welfare measure (profit, CS, and total welfare) of the optimal RSs with different objectives (Fig.2.3, 2.4, and 2.5).

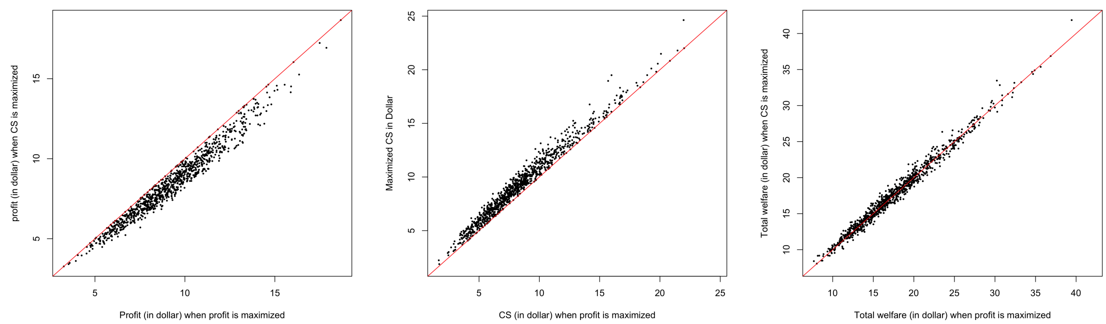

Figure 2.3: **General Model**: Two-way plots of three welfare measures (profit, CS and total welfare) comparing maximizing CS versus maximizing profit All numbers are based on sales from 15-movie RS monthly sale

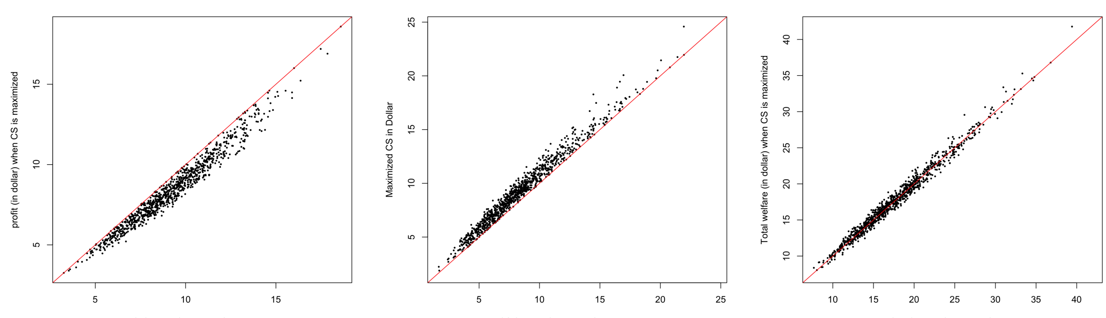

Figure 2.4: **ct model**:Two-way plots of three welfare measures (profit, CS and total welfare) comparing maximizing CS versus maximizing profit. All numbers are based on sales from 15-movie RS monthly predicted sale.

Fig.2.3, Fig.2.4, and Fig.2.5. In all nine plots, each dot represents one of the 1000 simulations, an ordered list of 15-movie set ,and in each plot,the horizontal and vertical axis are the same type of welfare measures, profit (left), CS (middle), or total welfare (right). The red line is the diagonal line.

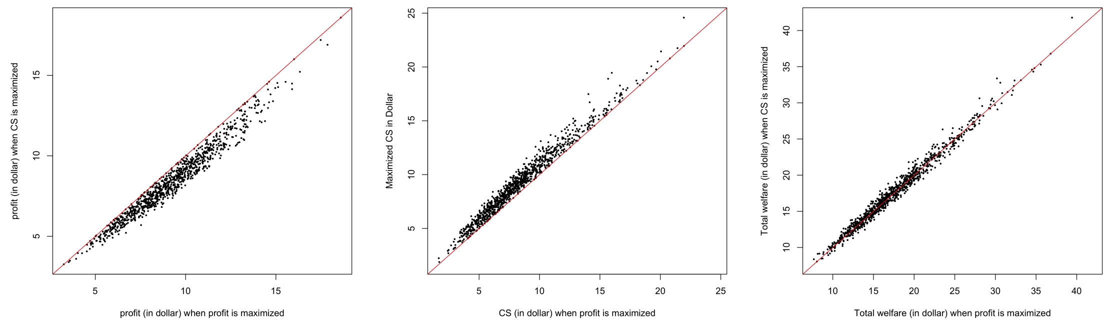

Figure 2.5: **ct-noPersi model**: Two-way plots of three welfare measures (profit, CS and total welfare) comparing maximizing CS versus maximizing profit. All numbers are based on sales from 15-movie RS monthly sale

Fig.2.4 illustrates the welfare results from RS that assumes two-type-consumer demand model and personalizes the recommendations according to consumer types(ct type). Horizontal axis are profit, CS and total welfare as a result of the firm's profit-maximization strategy. In the left and middle subplots of Fig.2.4, vertical axis indicates the value of profits and CS generate from the ordering of each RS that maximizes CS. Since left subplot has most points below the diagonal line, meaning that the profit from maximizing CS is less than the optimal profit. In the middle subplot, most points scattered above the diagonal line, so the CS that maximizes profit generates lower than the highest value of CS. Last, in the right subplot of total welfare, vertical axis corresponds to the maximal acheivable total welfare in each simulated RS. As points are scattered around the 45 degree line, it's not straightforward to see the relationship of welfare values under different optimization objectives. Refering to results from computation in equation 2.3.2, the profit-maximization orderings give less than optimal total welfare on average.

These three subplots in Fig.2.4 visualize the result from equation 2.21, providing an graphical evidence to the conflicts of interests. The plots under the general and the ct.noPersi types of RSs show the same relationships (Fig.2.3 and 2.5).

Comparisons of Welfare with Other Popular RSs

Apart from the RS that maximizes profit, CS, and total welfare, I have also computed other four popular types of RS: ranking movies by their past sales (monthly sale from August, 2013) as ”mostsold”, IMDB rating as ”highestrating”, and IMDB votes as ”mostvoted”, as well as ranking randomly, as ”random” RS in Fig.2.6. Each subplot in Fig.2.6 corresponds to one of three welfare measures, and each bar corresponds to one of the seven RSs of the ct type, i.e. personalized recommendation for premium and standard consumers.

Comparison of Profits

- Profit-driven RS generates significantly higher profit than any other type of RS.
- The second highest level of profit is obtained from RS that maximize total welfare.
- CS-driven RS is at least as profitable as those other popular types of RS, such as Rating-driven, Votes-driven and previous sales-driven RS.

Comparison of CS

- CS-driven RS generates highest level of CS, but CS from Total welfare-driven RS is not significantly lower than that from CS-driven RS.
- Profit-driven RS generates as much CS as other popular types of RS (previous sales-driven, rating driven, votes-driven, random?).

Comparison of Total Welfare

- Total welfare-driven, CS-driven and profit-driven RSs yield significantly higher total social welfare than other popular types of RS.

I have computed the optimizations with other two types of RS, and the results remain robust across different types of RS. From Fig.2.6,I have found that the RSs that maximize profit, CS and total welfare outperform other popular types of RS in all three welfare measures. The RS

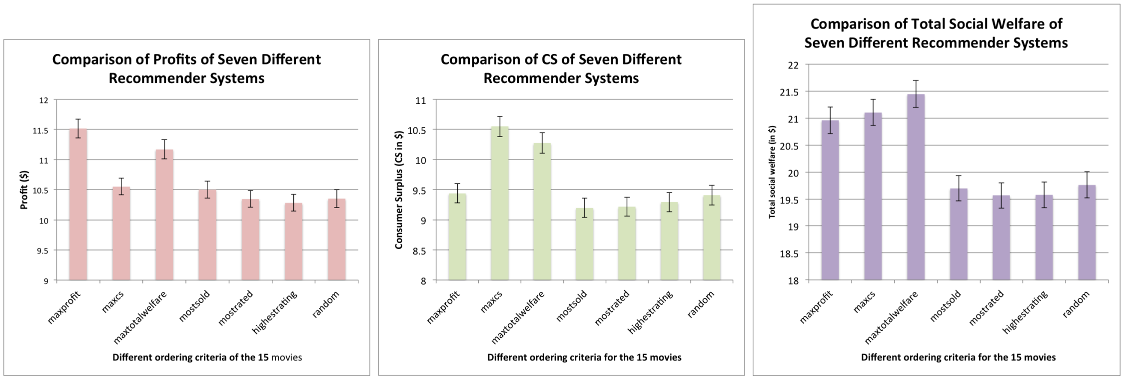

Figure 2.6: Bar plots with 95% confidence Interval of three welfare measures (profit, CS and total welfare) comparing 6 RSs

that maximizes profit(or CS) reaches its highest level of profit (or CS) with a cost of lowering the CS (or profit). As the total-welfare-driven RS maximizes the total welfare while still keeping a sufficiently high level of the other two welfare measures, it is probably the favorite option from the policymaker's perspective.

2.4 Conclusions and Limitations

In summary, I explore whether RS can potentially hurt consumers in this part of analysis.

First I have empirically estimated an exponential consumer demand model using datasets from the large-scale randomized experiment on Recommender System (RS) of Video-on-Demand (VoD) movies. Applying the estimated demand functions to 1000 simulations of RSs, I was able to calculate the loss in consumer surplus and total welfare if the firm designs the RS to maximize profit. The results show that the firm is motivated to choose an ordering strategy of recommendations that generates 7% more profits than the strategy to maximize consumer surplus. As a result, consumer surplus is traded off by 8% and the total welfare is 2% less than the maximum level of the total welfare. In addition, I have compared the RSs that maximize the three welfare measures to four other common RSs: sorting by IMDb votes, IMDb ratings, and previous sales as well as sorting randomly. The three types of optimal RSs are demonstrated to outperform the other four common types of sorting strategies, and among them, the RS that maximizes total welfare has the best overall performance in the three welfare measures.

The potential conflict I have identified in the paper is helpful to guide future public policy making in E-commerce. If recommender systems are kept unregulated, the welfare conflict can increase to a level that significantly hurts consumers. Although the result of welfare loss from empirical data is only a small portion of total consumer surplus, it's enough to raise attention and concern. Certain market mechanisms or policy interventions need to be introduced to mediate this loss in welfare. In the analysis, I have discovered that if a RS is constructed to maximize total welfare instead of profit, both profit and CS will still be kept at relatively high levels, and the conflict can be attenuated. As such, the total-welfare-driven RS is a possible compromise solution between the firm and policymakers.

This study contributes to the literature by being among the first to empirically investigate the welfare conflicts that arise in RSs. As an initial step in welfare analysis of the RS, it attracts the attention of regulators to evaluate the welfare impact of RSs with more comprehensive analysis. However, the study has some limitations. I have analyzed the case when only the ordering of items can be changed, in the context of Video-On-Demand movies, but in reality RS can be applied in various forms and contexts. Besides the ordering, firms can choose which items to put into the RSs. In some scenarios, even the prices of recommendations can be changed, such as offering coupons to some targeted consumers. Despite all these limitations, the framework that combines empirical estimations and counterfactual analysis can be applied in a much wider application context of recommender systems.

For most of the study I consider homogeneous consumers and the only personalization is differentiating two types of consumers in estimating demand function and optimizing separately for the premium and standard consumers. In future analysis, the analysis should be based on consumers with more heterogeneity in tastes. The second part of this thesis introduces differentiated consumer tastes into the world of RS. I analytically model the targeted price discrimination and personalized selection of products which are omitted in the first part of analysis.

Chapter 3

Does Personalization Resemble Price Discrimination?

3.1 Introduction

Today, personalized recommendations have become ubiquitously adopted on the web (ListrakNews, 2014). In terms of online shopping, personalized RSs bring in considerable amount of additional revenue for retailers by increasing consumer engagement and improving the purchase experience. For example, Amazon's pillar of business, its recommendation engine, boosted sales by more than 29% in 2015 for the online retailing industry titan (Mangalindan, 2012a).

Personalization is the key to success for RSs. Previous research and studies have uncovered the myth of why personalization increases consumers' probability of purchase. To mention two widely accepted reasons, one observation is that personalizing the product recommendations can raise the conversion rate of an exposure to actual purchase. A case in point are weekly emails that recommend a feature product each time based on past purchases. According to a study released by personalization technology company Magnet and Retail TouchPoints, 41% of consumers who received a highly relevant digital advertisement or email said they spent slightly or significantly more with that retailer. The second observation is that personalization reduces consumers' searching behaviors and encourages more of them to accept recommendations. The reduction in search cost is highly valued if the recommended products are better fitted to consumers' tastes.

Personalized marketing can be categorized into personalized product and personalized pricing. Personalized products are offered to consumers of differentiated preferences over product categories, while personalized pricing is based on different consumers' willingness to pay for the same product.

3.1.1 Personalized Product

The underpinnings of online personalization is, on the demand side, consumers have diversified preferences over product brands and features. The theory of differentiated preferences of consumers are explained by classical economic literature (Dixit and Stiglitz, 1977). On the supply side, the online market provide an explosively large number of product options for consumers to choose from. Brynjolfsson et al. 2003 verified that online bookstores like Barnes & Nobles have larger catalogs than their brick-and-mortar correspondents. Furthermore, recent research interests (Hinz and Eckert, 2010; Matt et al., 2013; Fleder and Hosanagar, 2009; Hosanagar et al., 2014; Oestreicher-Singer and Sundararajan, 2010; Brynjolfsson et al., 2006) have been focused on the second-order effects of online RSs - the long tail effects. That is to say, the RSs designed to match consumers' diverse preferences have changed consumers' shopping behavior, even resulting in more evenly distributed preferences.

3.1.2 Personalized Pricing

Besides personalized product items, another highly adopted personalized marketing strategy reported by (ListrakNews, 2014) is the personalized price discount (Godinho de Matos et al., 2015a). Personalized price discount uses the price discrimination strategy in microeconomic theory, where sellers charge different buyers different prices for the same product. In markets where consumers have differentiated tastes and elastic demand, each consumer has their own willingness to pay (WTP) for a specific product depending on how it matches their taste. If retailers have accurately gauged such information, they can price discriminate across consumers. When each consumer is charged their WTP, all consumer surplus is transformed to revenues, which is called first-degree price discrimination. There are two other types of price discrimination: second-degree price discrimination by which sellers charge different unit prices for different quantities sold, such as wholesale pricing and reward programs, and third-degree price discrimination which means selling to different consumer groups at different prices, such as student discounts and membership prices.

Even though price discrimination makes the sellers wealthy, it is accused of taking advantage of consumers by exploiting their surplus and treating them unequally. The discussion of the morality of price discrimination dates back to the early 20th century. In the 1930's people strongly protested against any form of price discrimination, so 1935 came out the introduction of Robinson-Patman Act, also called the Anti-Price Discrimination Act. Recently, price discrimination has become more desirable in an economic sense as researchers have established that price discrimination strategies can increase the efficiency of the economy (Odlyzko, 2003; Phlips, 1983; Varian, 1996). Nevertheless, many consumers still resent price discrimination for charging unfair prices compared to their peers. In the last ten years the media has been continuously reporting complaints from consumers and suggesting ways to avoid being price discriminated and shop for lowest prices. (CBSNews, 2017a,b, 2016; Mashable, 2014; Economist, 2003). One study with surveys conducted from the Internet identified various instances of price discrimination and price steering among the top e-commerce websites (Hannak et al., 2014). Now the debate is efficiency v.s. equality and there is a continuing tension between the two opposite sides.

As personalized pricing faces increasing challenges, companies are forced to implement use of price discrimination in a parsimonious manner or more implicitly. They mainly price discriminate downward, i.e. in the form of price discount rather than markup, which is more receptive to consumers. In fact, the current commonly adopted price discrimination strategies are most of the time legal, but consumers can penalize the sellers by lowering the reputation or loyalty to the firm and no longer shop there. Studies by Li and Jain 2015 found that without the consideration of the consumers' fairness concern companies might incorrectly implement price discrimination, resulting in loss of both profits and consumer surplus.

3.1.3 Connecting Personalization to Price Discrimination

The two aforementioned types of personalization are based on different criteria. The first criterion is the heterogeneous consumer preference for different products, and the second is the heterogeneous consumers' WTPs for the same product. Furthermore, since heterogeneous consumer preferences for products lead to consumers' differentiated WTPs, it's natural to ask the question: can companies offer personalized products and charge at every one's WTP? Since each person gets a different product, the firm is not literally charging different prices for the same product and thus not implementing a price discrimination strategy.

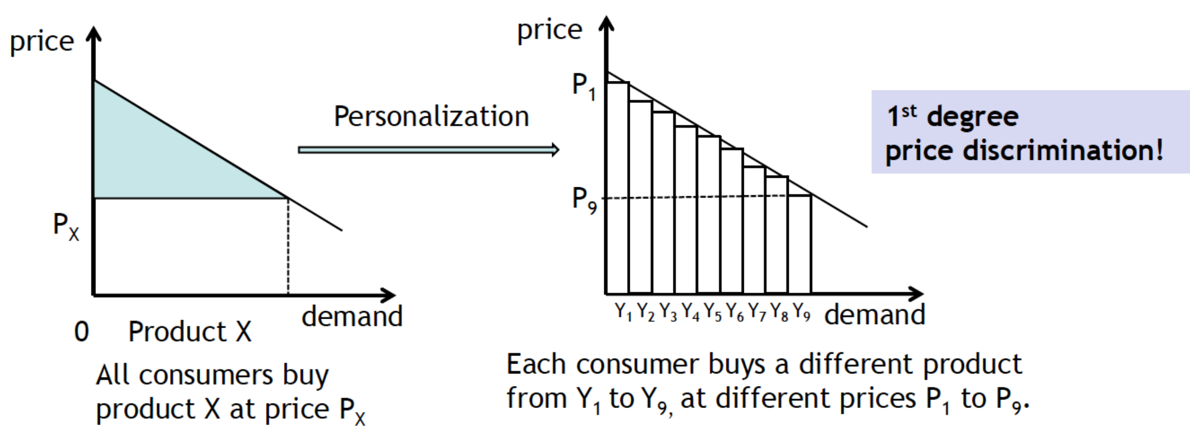

Figure 3.1: A simple example to illustrate the resemblance of personalization to price discrimination

Let's look at Figure 3.1, the left figure shows the aggregate demand curve of a product X. On the x axis, each point corresponds to consumers with a specific level of WTP for product X. If the seller recommends the same product X at price P_X for all these consumers, it collects revenue represented by the rectangle below the P_X price line. If the seller recommends each consumer on the x-axis a different but similar product, $Y_1, Y_2, \dots Y_9$, and charges each consumer their WTP for the product, the firm is able to collect revenue as the area under the demand curve. This is also the revenue size of first-degree price discrimination, when a firm recommends only product X but is allowed to charge each consumer a different price for the same product. The idea is essentially, by recommending different personalized products, a firm can transform consumer surplus into revenue. In such a way, the company might not only avoid being criticized for unfair prices but also appraised for good consumer experience, such as saving search effort and highly matching the expectation.

There are requirements for implementing this strategy: abundant differentiated product items, heterogeneous consumer preferences, and accurate identification of each consumer's WTP for each product. In the electronic marketplace, these conditions are highly probably to be satisfied. I have already explained the heterogeneous consumer preferences and high level of differentiation of the online marketplace compared to brick-and-mortar stores. In terms of the WTP estimation, there is a large body of literature with discussions of various approaches (Bishop and Heberlein, 1990; Weaver and Luloff, 1992; Alberini, 1995; Blend and Van Ravenswaay, 1999; Lusk et al., 2001; Lusk, 2003; Wang et al., 2007; Miller et al., 2011) to quantify the WTPs using survey data, observational purchase data, randomized experiments, or any combinations of them. Miller et al. 2011 is a good review of the state-of-the-art methods for obtaining WTP in marketing and compared their performances. Furthermore, alternatively to estimating WTPs of consumers for products directly, many business practitioners choose to match consumers with their favorite products based on user profile and transaction data using state-of-the-art machine learning algorithms (Davis et al., 1995; Anderson et al., 1999; Glaser et al., 2006).

3.1.4 Previous Works

As a trending topic of information systems and marketing, personalized recommendation has been explored both analytically and empirically in the literature. There is an extensive body of empirical works on personalized RSs. Some researchers focused exclusively on utilizing various information sources, such as user reviews, ratings, and tags, to improve RSs' performance. Others discussed metrics of performances of RSs including accuracy, serendipity, novelty, and utility (Zhang et al., 2012; Iaquinta et al., 2008b; Adamopoulos and Tuzhilin, 2014; McNee et al., 2006b). Analytically, researchers studied the impact of personalization on product exposure and visibility (Li et al., 2014), and a firm's strategic choice of personalization accuracy and consumer search costs to optimize profits (Choudhary and Zhang, 2014).

3.1.5 Summary of Results

This part of analysis contributes to the analytical literature of personalized RSs from a new angle. I connect personalized product recommendations to personalized price discounts through a consumer choice model between the catalog and RSs. In the proposed model, monopolistic vendors charge higher-than-catalog prices for recommended products that better match consumers' preferences and save consumers' effort incurred in searching the catalog. The model shows initial results consistent with previous literature (J. Gifford and T. Kudrle, 2010; Bakos, 1997): reduction in consumer search cost explains why personalization is both profitable to retailers and beneficial for consumers. Innovatively, after introducing a new concept of the personalization level, I show that perfect personalization, i.e. each consumer is recommended a different product and personalization level is infinite, is equivalent to first-degree price discrimination in terms of resulting profit, consumer surplus, and total welfare. A finite level of personalization is analogous to third-degree price discrimination. Higher personalization level always generates higher profits for the vendor. When personalization level is moderate, consumer surplus increases with personalization level, but when personalization level is higher than a certain threshold, consumer surplus is negatively correlated with the personalization level.

The structure of this part of analysis is summarized as follows: section 3.2 describes the model settings and section 3.2.2 explains how the optimal pricing and product selection strategy of the RS can be derived. Section 3.3.1 presents the welfare results from a baseline RS model comparing before and after the introduction of the RS. Section 3.3.2 presents how the welfare results change with increasing personalization levels. Section 3.3.3 compares the price discrimination to personalized product strategy. Section 3.3.4 and Section 3.3.5 present welfare analysis for the effect of two potential policies. Section 3.4 concludes the paper, discusses the managerial and policy implications, and points out future research directions.

3.2 Model

3.2.1 Model Setting

Consumers and Products

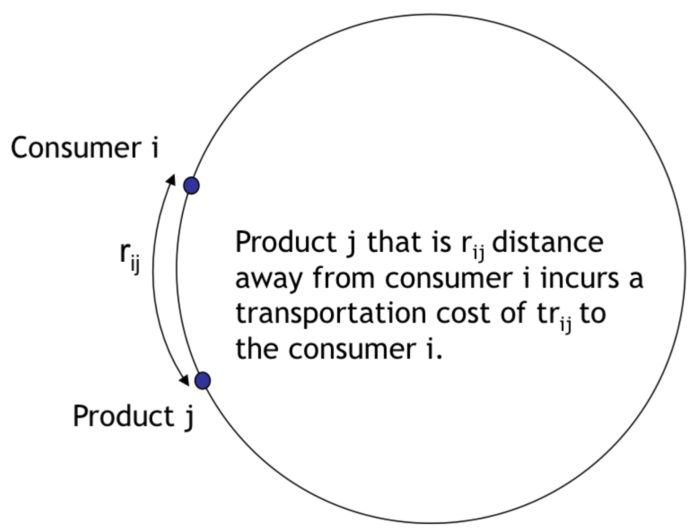

Figure 3.2: The circular preference, product characteristic space, and misfit cost

This part of analytical work focuses on a monopolistic vendor who offers a large number of products of horizontal differentiation. Taking the definition from Salop's model of horizontal differentiation (Salop, 1979), products are modeled as evenly dispersed along a circle of unit circumference. This circle represents the product characteristic space. Each point on the circle maps to a vector of product characteristics. The circle is also the consumer preference space. Each point corresponds to a certain product that perfectly matches the preference of the consumer at the same location. Consumers are uniformly distributed around the circle. In this paper, I assume that firm has perfect information of consumers: it knows the exact location of each consumer. A consumer incurs a misfit cost (**transportation cost**), in another word, a loss in utility, as a function of the distance from the purchased product, denoted by $t{\cdot}r$, where r measures the distance, and t is the cost per unit distance (unit transportation cost). The transportation cost is illustrated by Figure 3.2. All consumers have the same baseline evaluation for all products offered by the firm, denoted by V. In order to make a purchase, each consumer can search through the catalog, where all products are visible, or accepts the recommendations offered by the firm.

Searching

By searching, each consumer is able to find the perfect fit (zero transportation cost), but pays an average **search cost** of s (s denotes how exhaustive the search process is). To simplify the model, prices of all products in the catalog are the same, denoted by p_0. All consumers purchase. Consumer i's utility from purchasing product j is $U_{ij} = V - tr_{ij} - p_0 - s$, and $U_{ij} = V - p_0 - s \geq U_0$, where U_0 is the utility from the outside option. Zero or proportional production cost is assumed, thus maximizing the profit is equivalent to maximizing the revenue.

Recommender System with Personalization Level L

Without loss of generality, in this paper I model a type of RSs that sends one recommendation to each customer per period and I focus on the single-period interaction. The practical example is the monthly or weekly email that on-line retailers sent to the members on its subscription list with

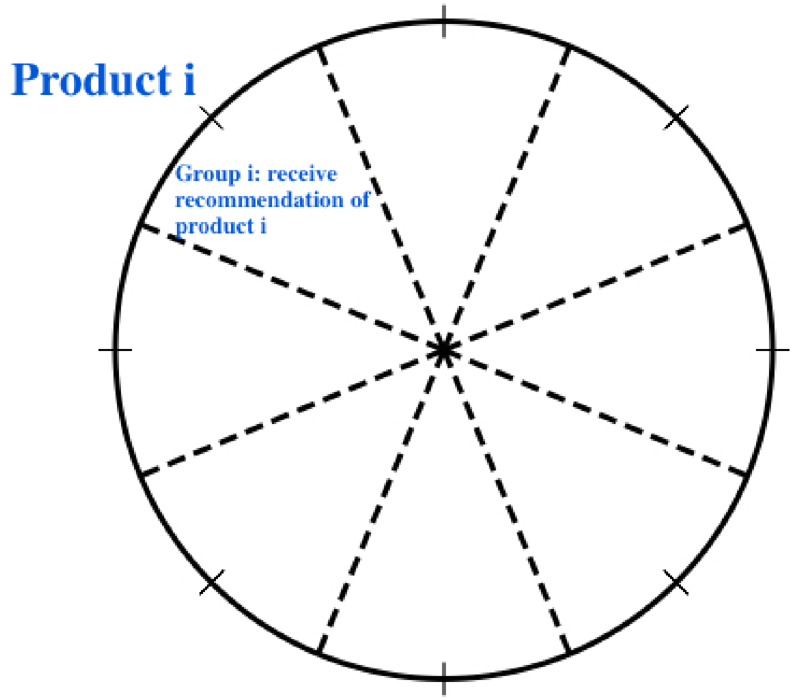

Figure 3.3: Personalized recommendation strategy when $L = 8$

a featured product. To design such a RS, the vendor usually first chooses how many different types of products in total it offers to consumers, or put another way, how it segments consumers such that consumers from each segment receive a recommendation that is personalized to them. I define the number of different product types offered as the level of personalization, denoted by L. In the real world, the choice of L might be endogenous, and firm strategically chooses this level to maximize profit, or the choice is exogenous and constrained by the performance of the vendor's personalization technology[1]. Since the consumers and products are uniformly distributed, it's natural to say that each segment of consumers is an arc on the circle of equal distance. Selected products are located at the center of each consumer segment and equidistant ($\frac{1}{L}$) on the circle. After choosing L, the vendor chooses the price of each recommended product $\{p_l\}_{1\leq l\leq L}$, shown in Figure 3.3. Due to symmetry, all products will be charged the same price

[1]Improving the personalization technology usually has costs, which I haven't included in this paper. But future research can investigate the strategic choice of L as a trade-off between improved revenue and increased technology cost.

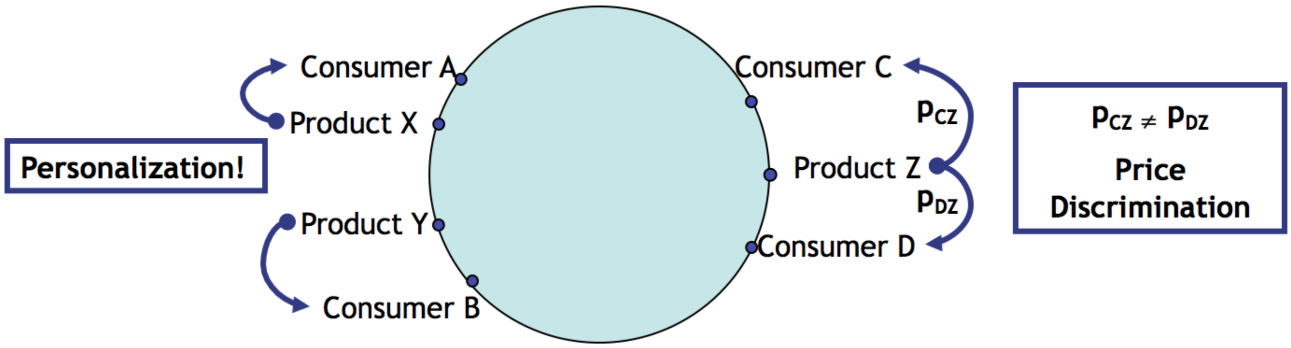

Figure 3.4: Personalization and price discrimination in RS

$(p_1 = p_2 = .. = p_L)$.[2]

Recommendation Strategies

I am interested in whether personalization and price discrimination are comparable in terms of the welfare impact. To make it clear, by price discrimination I mean the online vendor recommends same products to consumers but charges different prices, and by personalization I mean it only personalizes recommendations to consumers by products types. The difference of price discrimination and personalization in RS are demonstrated by Figure 3.4.

Specifically, I formally define four types of recommendation strategies:

1. **noPD-noPers**: (baseline strategy) The vendor recommends one product to all consumers at the same price.

2. **noPD-Pers**: The vendor charges only one price for the same recommended product, but two consumers are possible to be recommended different products.

3. **PD-noPers**: The vendor recommends the same product to all its consumers, but is possible to charge different prices for the same recommended product.

4. **PD-Persi**: The vendor is not constrained to charge the same prices for the same recom-

[2]To avoid different prices between catalog and RSs, if the firm recommends j to i. it sets catalog price of j be p_l. Even though implemented for all consumers, this practice will not influence the decision of consumers who don't receive j, because as later I show $p_l > p_0$, raising the price will make them even less likely to buy j.

mended product. Neither is it constrained to recommend same product.

I solve for the profit maximization problem in each of the above strategies and compare the resulting profit, consumer surplus, and total welfare. Solving strategy 1 (baseline strategy) helps us understand the role of search cost and product differentiation in making the RS profitable. Strategies 2 (personalization only) and 3 (price discrimination only) will be the focuses to compare personalization to price discrimination when the two strategies are implemented separately. Strategy 4 adds price discrimination to strategy 2 and adds personalization to strategy 3. In the end, I will prove that by price discrimination firm only makes profit by ”exploiting” consumers, while personalization can benefit both consumers (at moderate exploring level) and the firm. When personalization is perfect, strategy 4 is exactly the same as strategy 3.

3.2.2 Optimal Pricing and Product Selection

In order to select the optimal prices and products for the RS, it's necessary to analyze how each consumer reacts to the price and location of the recommended product. Since the transaction between the vendor and its consumers is a two-stage sequential game, the vendor's optimal strategy is solved by backward induction.

Stage 2: Consumers Accept Recommendation or Search and Buy from the Catalog?

Figure 3.5 compares the consumers' decisions before and after the introduction of the RS. Without the RS, only by searching consumers can make a purchase. With the RS, each consumer faces two choices: accepts the recommendation (pays the recommendation price p_r and probably a utility loss tr), or ignores the recommendation and searches and purchases from the catalog instead. I assume here from the catalog purchase, consumers pay the uniform search cost s for the perfectly fit products and catalog price p_0. On the other hand, by accepting and purchasing from the RS, it's likely that not all consumers get perfectly fit products. In addition, consumers might be charged the prices higher than the catalog price p_0 for the recommended items.

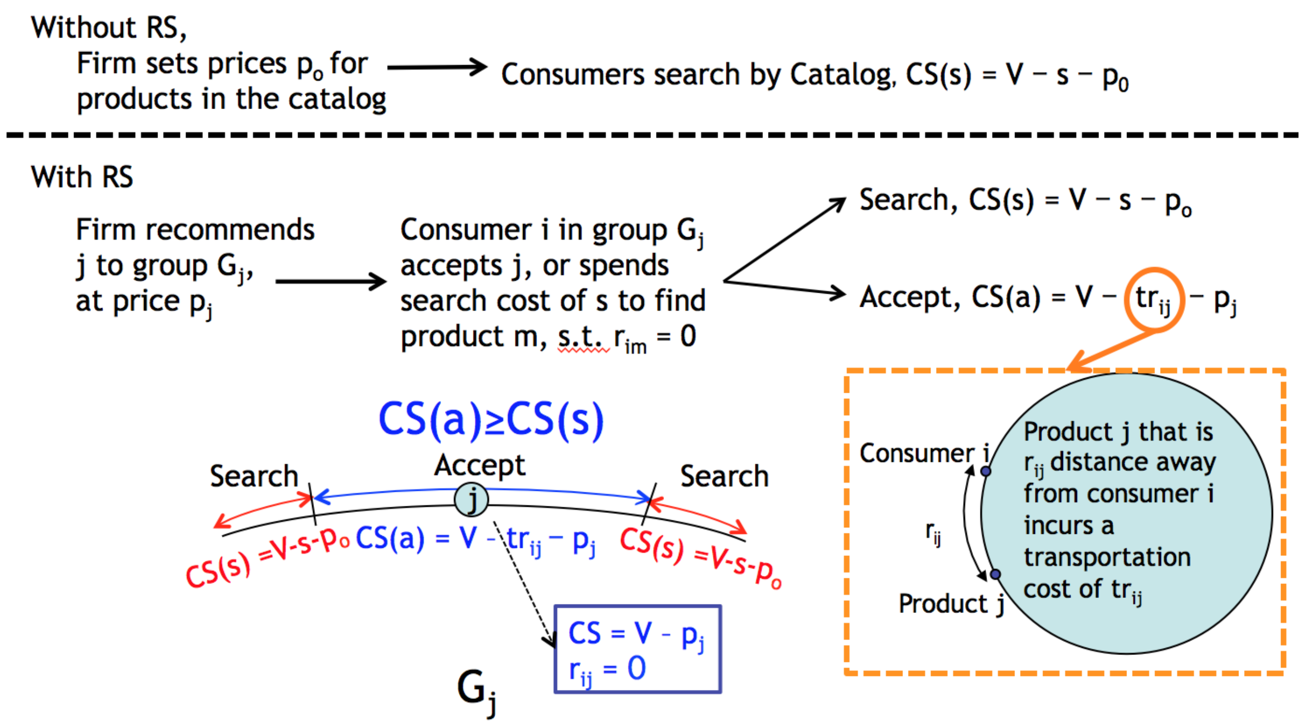

Figure 3.5: Dynamics between the firm and the consumers in RS

When the price discriminations are not allowed, after the firm selects one recommended product with one uniform price for each group of consumers, the consumers decide to accept or search and purchase. I want to predict consumers' decision given every combination of price and product location of RS (whichever generates no lower profit than the catalog, i.e. $p_l \geq p_0$) under the utility maximization framework. Suppose a consumer is recommended a product with attributes (p_l, r). The consumer surplus from searching the catalog is $CS_0 = V - p_0 - s$ and from purchasing recommended products is $CS_r = V - p_l - tr$. The marginal consumer that is indifferent between accepting recommendation and searching the catalog is at:

$$r = r_m := \frac{s - (p_l - p_0)}{t} \tag{3.1}$$

Recall that in each consumer group, the distances between consumers and the product recommended to them range from 0 to $\frac{1}{2L}$. So the position of the marginal consumers cannot be over $\frac{1}{2L}$ distance away from their recommended product. Actually the firm won't charge a price such that $r_m > \frac{1}{2L}$. Because under $r_m > \frac{1}{2L}$, or $p_l < p_0 + s - \frac{t}{2L}$, the farthest consumers from the recommendation in each group has surplus higher than the surplus from searching and purchasing from the catalog.

$$CS_r = V - p_l - \frac{t}{2L} > V - p_0 - s = CS_0$$

The firm is motivated to increase the price until the farthest consumers are indifferent between accepting and searching the catalog for purchase. So if $r_m = \frac{1}{2L}$, the marginal consumer are r_m distance away from recommendations. If $r_m \geq \frac{1}{2L}$, the farthest consumers from the recommendation in each group are marginal consumers.

Therefore, the position of the marginal consumers is $min(r_m, \frac{1}{2L})$. By symmetry, the sizes of acceptance regions of recommendations in each consumer group are the same, which are $2min(r_m, \frac{1}{2L}) = min(\frac{2(s-(p_l-p_0))}{t}, \frac{1}{L})$.

When the vendor price discriminates, p_l can be different across consumers within the same group of consumers.

Stage 1: Profit Maximization

In the current setting, without the RS, the vendor earns p_0 from each consumer on the circle. So collectively the profit from catalog purchase is $p_0 \times 1 = p_0$. With RS, the firm is switching consumers from the catalog purchases to purchases in RS. So the maximization of total profit for the firm is equivalent to maximization of the total increase in profits from RS than the catalog profits without RS. As explained previously, the optimal prices satisfy $p_l \geq p_0$, so the net increase in profit is collected from consumers who accepts recommendations under p_l. I represent the change in profit from a specifically selected product j by $\Delta\pi_j$, along with the change in consumer surplus and total welfare by ΔCS_j and ΔTW_j.

If price discrimination is allowed (strategies 3 & 4), it's straightforward to prove that it's optimal for the vendor to charge each consumer's willingness to pay for switching from searching and purchasing from the catalog to accepting the recommended products, as long as this willingness to pay is larger than the catalog price p_0. So each consumer becomes a marginal consumer as long as $p_l \geq p_0$, i.e.

$$\forall i \in \{i : r_i < \frac{s}{t}\}, CS_0 = CS_r, \iff p_l = p_0 + s - tr_i \tag{3.2}$$

Therefore, when $\frac{s}{t} < \frac{1}{2L}$, or $L < \frac{t}{2s}$ the size of the acceptance region of each segment is $\frac{s}{t}$. In total, $1 - \frac{2sL}{t}$ consumers search the catalog. Comparing strategies 3 & 4, as personalization level L increases, more consumers accept the recommendations ($\frac{2sL}{t}$ increases) and the price discrimination profit increases. When $L \geq \frac{t}{2s}$, all consumers accept recommendations and are charged at their willingness to pay to switch from searching and buying from the catalog. Increasing L only makes firm to charge higher prices. Consumer surplus don't change with L.

If price discrimination is not allowed (strategies 1 & 2),

Given certain p_l,

$$\Delta\pi_j(p_l) = 2min(r_m, \frac{1}{2L})(p_l - p_0) = 2min(\frac{s - (p_l - p_0)}{t}, \frac{1}{2L})(p_l - p_0) \tag{3.3}$$

$$\begin{aligned} \Delta CS_j(p_l) &= 2\int_0^{min(r_m, \frac{1}{2L})} (V - p_l - tr) - (V - p_0 - s)dr \\ &= 2\int_0^{min(r_m, \frac{1}{2L})} (p_0 + s - p_l - tr)dr \end{aligned} \tag{3.4}$$

$$\Delta TW_j(p_l) = \Delta\pi_j(p_l) + \Delta CS_j(p_l) = 2\int_0^{min(r_m, \frac{1}{2L})} (s - tr)dr \tag{3.5}$$

When $p_l \geq p_0 + s - \frac{t}{2L}, r_m \leq \frac{1}{2L}$, maximizing the profit function from equation (3.3) gives the optimal price:

$$p_l^* | (p_l \geq p_0 + s - \frac{t}{2L}) = p_0 + \frac{s}{2}, r_m^* = \frac{s}{2t} \tag{3.6}$$

For prices lower than $p_0 + \frac{s}{2}$, profits increases in price, and for prices higher than $p_0 + \frac{s}{2}$, profits decreases with price. So if $p_l \geq p_0 + s - \frac{t}{2L}$ includes $p_0 + \frac{s}{2}$, the optimal price is $p_0 + \frac{s}{2}$. If $p_0 + \frac{s}{2} \leq p_0 + s - \frac{t}{2L}$, or $L \geq \frac{t}{s}$, the optimal price in this region is $p_0 + s - \frac{t}{2L}$.

When $p_l \leq p_0 + s - \frac{t}{2L}, r_m \geq \frac{1}{2L}$, equation (3.3) gives total profits: $\Delta\pi_j(p_l | p_l \leq p_0 + s - \frac{t}{2L}) = p_l - p_0$, so the optimal price is

$$p_l^*|(p \leq p_0 + s - \frac{t}{2L}) = p_0 + s - \frac{t}{2L} \tag{3.7}$$

, which confirm the previous argument that firm will not charge a price higher than $p_0 + s - \frac{t}{2L}$ for any recommended item.

From the above rationale and mathematical representations, the optimal price as a result of profit maximization depend on the personalization level. As L assumed exogenously determined, the optimal choices of L for the firm won't be discussed here, but in the end it can be shown how welfare changes with varying levels of L. The following sections discuss and visualize how optimal price changes in levels of personalization. Particularly, I need to discuss cases when personalization is at low levels ($L < \frac{t}{s}$) and at high levels ($L > \frac{t}{s}$) respectively.

Notice that since personalization level L is an integer larger than 1, when $\frac{t}{s}$ is less than 1, i.e. $t < s$ there will be only high level case, since $L \geq 1 > \frac{t}{s}$ is satisfied naturally.

Low Level of Personalization, $L < \frac{t}{s}$: When $L < \frac{t}{s}$, I have $\frac{s}{2t} < \frac{1}{2L}$. The unconstrained optimal prices from (3.6) satisfy this constraint. So when personalization level is relatively low at $L < \frac{t}{s}$, it's optimal for the firm to set each recommended product at $p^*_{l,lowlevel} = p_0 + \frac{s}{2t}$. The optimal pricing strategy when the firm chooses low level of personalization is illustrated by Figure 3.6 (left) using a concrete example at $L = 4, \frac{s}{t} = 0.2$. The acceptance regions are plotted for the selected products for the two neighboring segments of consumers.

High Level of Personalization, $L > \frac{t}{s}$: When $L > \frac{t}{s}$, I have $\frac{s}{2t} > \frac{1}{2L}$. It means if the firm sets $p_l^* = p_0 + \frac{s}{2t}$ for each recommended product, it is not able to collect profits from consumers

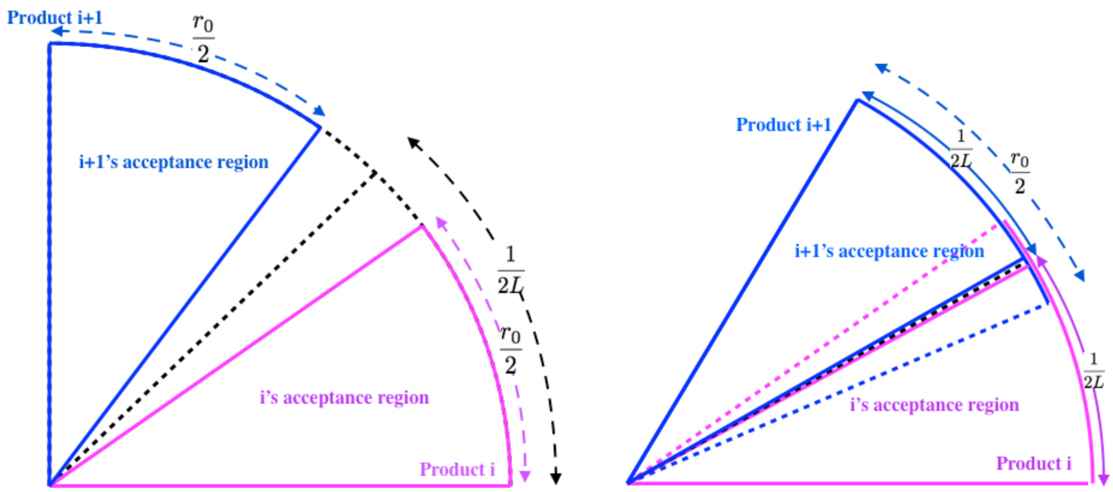

Figure 3.6: Region of acceptance of recommendation between two neighboring products

who are $(\frac{1}{2L}, \frac{s}{2t})$ distance away from the product, since only consumers who are less than $\frac{1}{2L}$ away from product j will be recommended j. In order to find the optimal prices for high-level recommended products, $p^*_{l,highlevel}$, I can graphically view the increase in profit as a function of the price, as plotted in Figure 3.7. The dashed and solid lines describe the unconstrained and constrained profit function respectively. The optimal price of high level of recommendation, denoted by $p^*_{l,highlevel}$ is

$$p^*_{l,highlevel} = p_0 + s - \frac{t}{2L}; r^*_{highlevel} = \frac{1}{2L} \tag{3.8}$$

Note that in this scenario, all consumers in the market will accept recommendations. As an example, Figure 3.6(right) plots the regions of acceptance for recommendations to two neighboring consumer segments at $L = 6, \frac{s}{t} = 0.2$. Obviously, in such case when all consumers accept recommendations, additional personalization will no long make consumers benefit since all search costs have been saved. But the optimal price increases in L. So as the total profit.

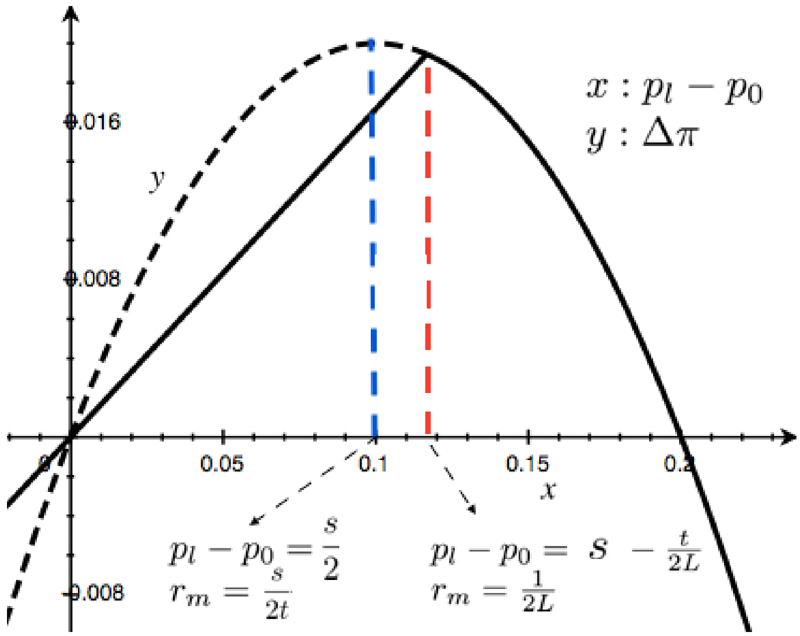

Figure 3.7: Optimal pricing in the case of high-level personalization: $L = 6, \frac{s}{t} = 0.2, t = 1$.

3.3 Results and Discussions

3.3.1 Initial Result from Baseline Model: Recommender System and Search Costs

After solving the optimal price and product selection for baseline strategy, welfare impact of the introduction of the RS can be evaluated. Baseline strategy 1 is when price discrimination is not allowed and when $L = 1$ as the condition for no personalization. So substituting $L = 1$ and (3.6) into (3.4) and (3.5), I get,

$$\Delta\pi = min(\frac{s^2}{2t}, s - \frac{t}{2}), \Delta CS = min(\frac{s^2}{4t}, \frac{t}{4}), \Delta TW = min(\frac{3s^2}{4t}, s - \frac{t}{4}) \tag{3.9}$$

When $t < s$, $L > \frac{t}{s}$ is satisfied as long as $L \geq 1$. We have only high levels of personalization and consumers all accept even if the recommendations are not personalized, i.e. $L = 1$. With optimal price being $p^* = p_0 + s - \frac{t}{2L}$ and marginal consumers being farthest consumers who are

$\frac{1}{2L}$ distance away from recommended product, equation (3.9) becomes

$$\Delta\pi = s - \frac{t}{2}, \Delta CS = \frac{t}{4}, \Delta TW = s - \frac{t}{4} \tag{3.10}$$

When $t \geq s$ and $L = 1$, $L \leq \frac{t}{s}$. L satisfies the condition of low levels of personalization. The optimal price of this case is $p^* = p_0 + \frac{s}{2}$, and $r_m \leq \frac{1}{2L}$. Then equation (3.9) becomes

$$\Delta\pi = \frac{s^2}{2t}, \Delta CS = \frac{s^2}{4t}, \Delta TW = \frac{3s^2}{4t} \tag{3.11}$$

Equations (3.9) presents the overall welfare changes caused by the introduction of RS: profit, consumers surplus and total welfare all increase. Decomposition of the three types of resulting welfare reflects why such improvement occurs. I find that welfares increase in search cost s. With the cost of searching perfectly fit product, s, unchanged, profit and total welfare decreases with unit transportation cost t, and CS increases with t at low levels and decreases with t at high levels. How welfare changes with s and t are illustrated in Fig. 3.8.

Therefore, the initial results from the baseline strategy reveal that the reduction of search cost caused by introducing the RS leads to the improvement of profit, consumer surplus, and social efficiency. The larger the search cost is the larger such improvements are. The explanation is that the focus RS works because it helps consumer save search costs, and so search cost reduction is a proxy for the efficiency of the system. This systems have larger impact on a market that needs more searching efforts.

Unit transportation cost, represented by t, is also a measure of the level of product differentiation in the market. Fig. 3.8 shows that price decreases monotonically with t when search cost s is fixed. The explanation of this trend is that if products are more differentiated to consumers, consumers are more sensitive to the level of misfit of the product, and hence the vendor has to charge lower price since consumers are willingness to pay less. On the other hand, consumers prefer more product differentiation level when t is lower than search cost s and less differentiation level when $t > s$. It's optimal for consumers to have $t = s$, when the product differentiation level and search cost match.

Both search cost s for a perfectly fit product and the product differentiation level t are features of the market, and the welfare impact of recommender systems highly depend on these market parameters.

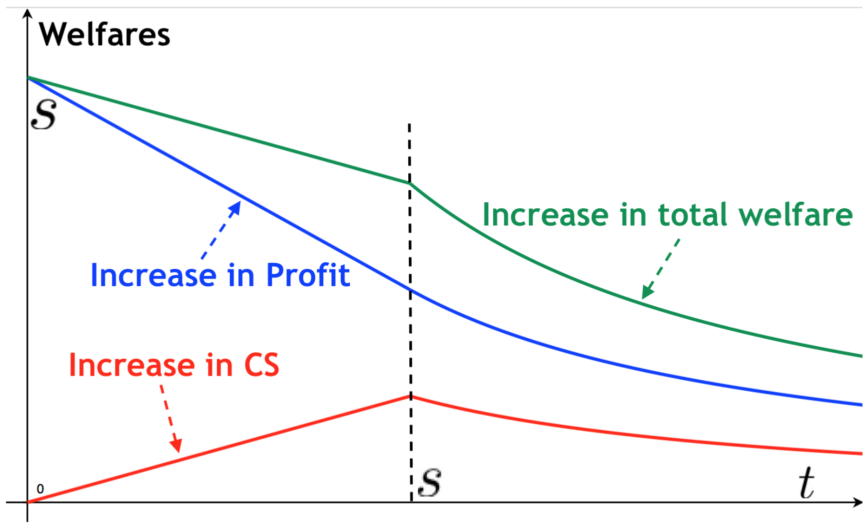

Figure 3.8: How welfare improvements change with search cost s and unit transportation cost t of the nonpersonalized RS $L = 1$

3.3.2 How Welfare Changes in Personalization Levels

Table 3.1: Summary of welfare result: how increases in profit, CS and total welfare change with the level of personalization L, when price discrimination is NOT allowed.

L	Δ Profit	Δ CS	Δ Total Welfare
L=0: No RS	0	0	0
$0 < L \leq \frac{t}{s}$: Low$L$	$\frac{s^2L}{2t}$	$\frac{s^2L}{4t}$	$\frac{3s^2L}{4t}$
$L \geq \frac{t}{s}$: high$L$	$s - \frac{t}{2L}$	$\frac{t}{4L}$	$s - \frac{t}{4L}$
$L \to +\infty$: perfect personalization	s	0	s

Table 3.1 summarizes the profit, consumer surplus, and total welfare from the vendor's profit maximization for strategies 1 & 2. Table 3.2 summarizes the results after adding price discrimination. The two tables together reflect how the three welfare measures change as the level of

Table 3.2: Summary of welfare result: how increases in profit, CS and total welfare change with the level of personalization L when perfect price discrimination is allowed

L	Δ Profit	Δ CS	Δ Total Welfare
L=1: No personalization	$s - \frac{t}{4}$	0	$s - \frac{t}{4}$
$2 \leq L \leq \frac{t}{2s}$	$\frac{s^2 L}{t}$	0	$\frac{s^2 L}{t}$
$L \geq \frac{t}{2s}$	$s - \frac{t}{4L}$	0	$s - \frac{t}{4L}$
$L \to +\infty$: perfect personalization	s	0	s

personalization L goes up. To compare the trends of three welfares graphically, Figure 3.9 is

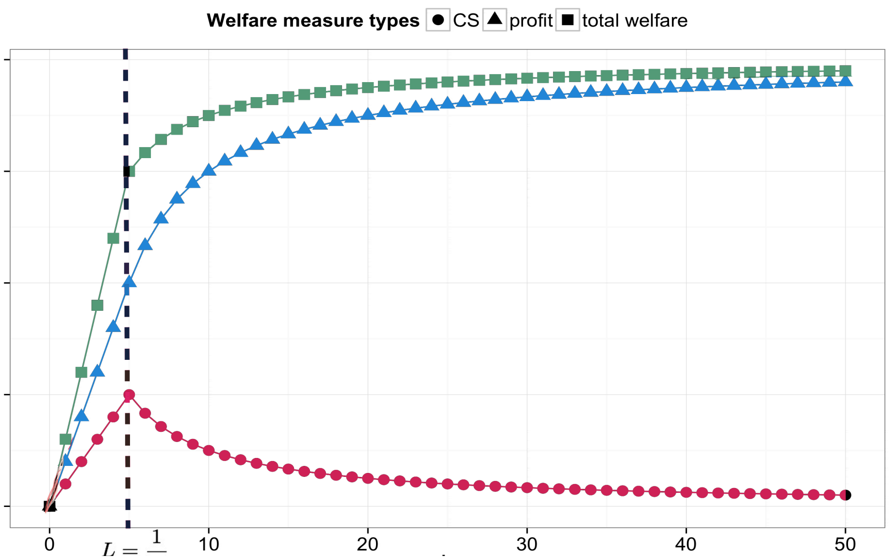

Figure 3.9: Comparing how three welfare measures change in L ($\frac{s}{t} = 0.2, t = 1$)

generated using $s = 0.2, t = 1, \frac{s}{t} = 0.2$ without price discrimination (strategies 1 & 2).

I have discovered that profit and total welfare always increase with personalization, whether price discrimination is allowed or not. With price discrimination, change in consumer surplus won't improve from higher level of personalization. Without price discrimination, at lower level of personalization, consumer surplus first increases with L, but after the level is higher than a threshold, drops with higher L. The optimal personalization level is achieved at $L = \frac{t}{s}$

This non-monotonic trend of consumer surplus is the result of two profit-maximizing behavior of the vendor that change the consumer surplus in two opposite directions. The first one is firms' "**exploring**" behavior: it tries to persuade more consumers to accept recommendations through reducing their search costs. This behavior occurs when the personalization level is relatively low and when there still exist consumers who prefer searching the catalog. The firm increases the size of the pie of profit by which consumer surplus increases in the meanwhile, because more consumers accepting recommendations means more search costs are saved. The second profit- maximizing behavior is to increase the thickness of the profit pie, the price margin. When personalization level is so high that all consumers accept the recommendation, the main channel of maximizing profit is to transforming consumer surplus into profit, which is the "**exploiting**" behavior. Therefore, the personalization level is moderate, i.e. $L = \frac{t}{s}$, the consumer surplus is at the maximal level, when the vendor is just able to persuade all consumers to accept recommendations.

The optimal level of $L = \frac{t}{s}$ for consumers increases in degree of product differentiation and decreases in search cost. When searching costs less effort, and products are highly differentiated (or searching is easy), firm needs to personalize at a higher level, i.e. segments consumers more granularly to persuade all consumers to accept recommendations, while in contrast exhaustive search and less product differentiation make consumers accept recommendations easily and it costs the vendor less personalization to attract consumers to higher-priced recommendations.

In contrast, as depicted in Figure 3.9, profit and total welfare improvement increase monotonically when the level of personalization L goes up. Specifically, when $L \leq \frac{t}{s}$, profit and total welfare improve in L linearly, and beyond $\frac{t}{s}$, the two measures converge to $t\frac{s}{t} = s$ with rate $1/L$.

3.3.3 Personalized Products vs Price Discrimination

Let's look at Table 3.1 and Table 3.2. Row 2 in Table 3.2 shows the result of price discrimination without personalization in RS (strategy 2: PD-noPers), and row 3-5 in Table 3.1 demonstrate the result of personalized recommendation when price discrimination is not allowed (strategy 3: noPD-Pers). Comparisons of table results of strategy 2 to strategy 1 shows that personalization-only and price discrimination-only generates comparable profits, and personalization-only generates even higher profits than price discrimination-only. In addition, personalization always increases consumers surplus (even though when L goes to infinity, such increase dies down to zero).

Rows 3-5 in Table 3.2 demonstrates the resulting welfare from strategy 4 in which both dimensions of personalization are allowed. The change in resulting welfare from strategy 3 to strategy 4 is the result of personalization on top of price discrimination. Obviously, personalization enables the vendor to make even more profit from price discrimination and benefits the social welfare. Comparison of strategy 2 to strategy 4 shows the welfare effect of price discrimination on top of personalization. I have found that price discrimination is just transferring the surplus from consumers to the firm.

However, I discover when personalization level goes to infinity, i.e. perfect personalization, strategy 3 acts exactly the same way as strategy 4 in terms of the three welfare impacts. Perfect personalization charges the willing-to-pay on each consumer and leaves zero surplus for consumers.

3.3.4 Possible Pricing Policy 1-Effect of Price Cap

In the mechanism that firm can exploit consumer surplus by increasing the personalization level, the essential part is that firm can charge higher price with better fitting recommendations with smaller distances to the consumers. To restraint such a exploitive behavior of the firm in high levels of personalization, one potential policy is to put a price cap on the RS. This cap should be applied to any recommendation strategy of the firm, i.e. independent of L. Since in the low levels of recommendations, firm charges price $p_0 + \frac{s}{2}$ and under this price, consumers benefit from higher personalization, the most intuitive and simple price cap is $p_l \leq p_0 + \frac{s}{2}$.

With price cap $p_0 + \frac{s}{2}$, under low levels of personalization $L \leq \frac{t}{s}$, the firm charges $p_0 + \frac{s}{2}$ as without the cap, and all three welfares increase with higher L. Under high levels of personalization $L \geq \frac{t}{s}$, as shown by Fig.3.7, profit increases with price when $p_l \leq p_0 + \frac{s}{2} \leq p_0 + s - \frac{t}{2L}$, so it's optimal for the firm to charge $p_l = p_0 + \frac{s}{2}$ under the price cap. So regardless of levels of personalization, the optimal price for the firm is $p^* = p_0 + \frac{s}{2}$. The resulting welfares are summarized in Table 3.3.

Table 3.3: Summary of welfare result: how increases in profit, CS and total welfare change with the level of personalization L when perfect price discrimination is allowed

L	Δ Profit	Δ CS	Δ Total Welfare
$1 \leq L \leq \frac{t}{s}$	$\frac{s^2 L}{2t}$	$\frac{s^2 L}{4t}$	$\frac{3s^2 L}{4t}$
$L \geq \frac{t}{s}$	$\frac{s}{2}$	$\frac{s}{2} - \frac{t}{4L}$	$s - \frac{t}{4L}$
$L \to +\infty$: perfect personalization	$\frac{s}{2}$	$\frac{s}{2}$	s

In this case higher personalization will no longer lower consumer surplus. CS increases with L instead. In the personalization level case, the firm and its consumers evenly splits the amount of total welfare improvement that equals to total search costs $s \times 1 = s$. However, under this price cap, the firm is not motivated to increase the level of personalization beyond $\frac{t}{s}$ since under the cap, the profits are unchanged with higher L.

3.3.5 Possible Policy 2-Effect of a Competitor

In the case of high levels of personalization, the reason why firm can charge consumers the exploiting price is that consumers have only two options: either searching and buying from the catalog or accepting and purchasing the recommended item. Under exploiting price, consumers still accept because the RS option generates higher surplus than searching the catalog. The monopoly RS provider itself offers low catalog price such that no outside option can make consumers buy nothing from the monopoly firm and switch to other companies. In another word, consumers are "locked in" the company. What if now consumers have a chance to leave the old company and switch to another competitor company that implements a RS that provides surplus higher than the catalog option of the old company? Could the competition motivate the company to lower the price below the exploiting price, especially when the personalization level is high?

If the competition works in providing incentives, is the effect enough to restrict firm's exploitive behavior in equilibrium?

Suppose there are two companies A and B, forming a duopoly in the market, with shares σ_A and σ_B respectively. $\sigma_A + \sigma_B = 1$. Both firms have consumers and catalog products that are uniformly distributed on the circle with price p_0 and search cost s for perfectly fit products. The two companies are assumed to implement RS with the same selection of products, but charging different prices. Company A charges p_A in RS and company B charges p_B in RS. In addition, by switching to another company consumers pay switching costs c. Explained in Klemperer (1987), the classic switching cost paper, the switching cost could be the learning cost, transaction costs, or the artificial switching cost such as the reward programs that encourage repeated purchases.

First we assume company A has already introduced the RS with price p_A under personalization L. Company A focused on its own consumers and was not aware of the competition from company B. Then we can find company B's optimal price and how B's pricing strategy motivates company A to adjust its pricing strategy.

Suppose company A charges the price p_A for RS as we have described without competition, and gets profit π_A as the total profits collected from catalog purchase and RS purchases in each consumer group.

$$\forall L \in [1, \frac{t}{s}], p_A = p_0 + \frac{s}{2}, \pi_A = \frac{p_0}{L} + \frac{s^2}{2t}$$

$$\forall L \in [\frac{t}{s}, +\infty), p_A = p_0 + s - \frac{t}{2L}, \pi_A = \frac{p_0 + s - \frac{t}{2L}}{L}$$

Now company B charges a price p_B to maximize its profits. Company B's profits are composed of three parts: profits from B's catalog purchases, profits from B's RS purchases, and profits from A's consumers. Each consumer i of company A has distance r_{iA} from the recommended product in A's RS and r_{iB} from the recommended product in B's RS. The utilities of different options, using U_s for catalog purchase, U_A for RS purchase of A's products, and U_B for RS purchase of

B's products are:

$$U_s = V - p_0 - s$$

$$U_A = V - p_A - t \times r_{iA}$$

$$U_B = V - p_B - c - t \times r_{iB}$$

Since we assume the company B has the same product selection for RS as A does, i.e. $r_{iA} = r_{iB} \equiv r_i$, the choices of company A's consumers depend on the consumers' distances from recommendations, r_i.

- searching: $r_i \geq max\{\frac{p_0+s-p_A}{t}, \frac{p_0+s-p_B-c}{t}\}$
- accept RS of A: $p_A \leq p_B + c$ and $r_i \leq \frac{p_0+s-p_A}{t}$
- switch to RS of B: $p_A > p_B + c$ and $r_i \leq \frac{p_0+s-p_B-c}{t}$

So some consumers of company search and buy from the catalog of A, and the others all accept the recommended products provided by the same company, either A or B. Those consumers who don't search accept and buy from A if $p_A \leq p_B + c$, and switch to B otherwise. Consumers of B have three options similar to those of A. The condition for B's consumers to switch to A's RS is $p_B \geq p_A + c$.

Combing the choices of consumers of both A and B, the condition for the case when consumers of neither companies switch is:

$$p_A - c \leq p_B \leq p_A + c, p_B - c \leq p_A \leq p_B + c \tag{3.12}$$

Assuming now company B charge the same price for the recommended product, now let's discuss whether firm B is motivated to steal company A's consumers by cutting prices of RS, or put in another way, whether the price p_B with $p_B \leq p_A - c$ is more profitable than $p_B = p_A$ [3].

[3]It's easy to see that a price p_B with $p_B \geq p_A + c$ will only make B worse off because under such price some consumers of B switch to A's RS.

The profits of B from each group of consumers are:

$$\pi_B(p_B) = \sigma_B p_B min(\frac{p_0+s-p_B}{t}, \frac{1}{2L}) + \sigma_B p_0 max(\frac{1}{2L} - \frac{p_0+s-p_B}{t}, 0) + \sigma_A p_B min(\frac{p_0+s-p_B-c}{t}, \frac{1}{2L}) \quad (3.13)$$

- If $\frac{p_0+s-p_B}{t} \leq \frac{1}{2L}$, $\pi_B(p_B) = \frac{1}{t}[-p_B^2 + p_B(p_0+s+\sigma_B \times p_0 - \sigma_A \times c) + \sigma_A(\frac{t}{2L}p_0 - p_0^2 - p_0 \times s)]$
- if $\frac{p_0+s-p_B-c}{t} \leq \frac{1}{2L} \leq \frac{p_0+s-p_B}{t}$, $\pi_B(p_B) = \frac{1}{t}[-\sigma_A p_B^2 + p_B(\frac{\sigma_B}{2L} + \sigma_A(p_0+s-c)]$
- if $\frac{p_0+s-p_B-c}{t} \geq \frac{1}{2L}$, $\pi_B(p_B) = \frac{1}{t}(\frac{p_B}{2L})$

$$p_B^* = \begin{cases} p_0 + \frac{s-\sigma_A(p_0+c)}{2} & \text{if} \quad \frac{p_0+s-p_B}{t} \leq \frac{1}{2L} \\ \frac{p_0+s-c}{2} + \frac{\sigma_B}{4L\sigma_A} & \text{if} \quad \frac{p_0+s-p_B-c}{t} \leq \frac{1}{2L} \leq \frac{p_0+s-p_B}{t} \\ p_0 + s - \frac{t}{2L} - c & \text{if} \quad \frac{p_0+s-p_B-c}{t} \geq \frac{1}{2L} \end{cases} \quad (3.14)$$

Since two firms have the same settings, we can assume that they have equal shares in the market, i.e. $\sigma_A = \sigma_B = \frac{1}{2}$. Then we can derive the condition under which firm B is motivated to lower its price below $p_A - c$ to steal consumers from A, assuming company A sets price according to equation (3.13).

The objective of introducing competition is to mediate the firm's exploitive behavior in high level of personalization, here the focus is on the case $L \geq \frac{t}{s}$.

When company B doesn't cut price, the price, $p_{B,nCut}$ and profit $\pi_{B,nCut}$ will be the same as company A.

When company B steals consumers from A by cutting price, the low price that B offers satisfies $p_{B,cut} \leq p_A - c = p_0 + s - \frac{t}{2L} - c$. This leads to the case when all consumers of both A and B adopt RS of B, since:

$$\frac{p_0+s-c-p_{B,cut}}{t} \geq \frac{p_0+s-c-(p_0+s-\frac{t}{2L}-c)}{t} \geq \frac{1}{2L}$$
$$\frac{p_0+s-p_{B,cut}}{t} \geq \frac{p_0+s-c-p_{B,cut}}{t} \geq \frac{1}{2L}$$

So equation (3.13) generates $\pi_{B,cut}(p_{B,cut}) = \frac{1}{L}p_{B,cut}$. The optimal price is at $p_{B,cut} =$

$p_0 + s - c - \frac{1}{2L}$. The maximally achievable profit from cutting price is $\pi_{B,cut} = \frac{1}{2}\frac{p_0+s-c-\frac{t}{2L}}{L}$. Comparing the profit of vendor B when B cuts prices to the profit without cutting price, we have

$$\begin{aligned}\pi_{B,cut} - \pi_{B,nCut} &= \frac{p_0 + s - c - \frac{t}{2L}}{L} - \frac{1}{2}(\frac{p_0 + s - \frac{t}{2L}}{L}) \\ &= \frac{1}{2L}[p_0 + s - \frac{t}{2L} - 2c]\end{aligned}$$

Therefore, vendor B is motivated to cut price to compete aggressively with A when

$$p_0 + s - 2c - \frac{t}{2L} \geq 0 \tag{3.15}$$

Specifically from equation (3.15), we learn how switching cost c influences whether a competitor would like to divert from the monopoly price and aggressively attract the consumers of another company by cutting prices.

When switching cost satisfies $c \leq \frac{p_0}{2} + \frac{s}{4}$, B will be motivated to cut price. When $c > \frac{p_0}{2} + \frac{s}{2}$, B will not cut price. Within the range $[\frac{p_0}{2} + \frac{s}{4}, \frac{p_0}{2} + \frac{s}{2}]$, when personalization level is relatively higher, $L \geq \frac{t}{c(p_0+s-2c)}$, B cuts price, otherwise B continues charging the monopoly price. We can see that when the switching cost between companies in the market is within some range, both companies are motivated to compete against each other aggressively. Competition between RSs of different companies leads to vendors lowering prices of RS, especially at the high levels of personalizations.

Under condition in equation 3.15, company B is willing to cut price. As a response, A will respond to B's such behavior by not doing anything and losing all its consumers, or cutting the price of A's RS based on B's adjusted price for its RS. By symmetry, the two companies stop at the same prices of their RS, when cutting prices further will make them worse off. We can derive the equilibrium prices.

Since the equilibrium prices, denoted by p_e, can remain only if both companies have no incentives to cut prices, cutting prices generate lower prices than without cutting prices. So we have $\pi_{i,cut} < \pi_{i,nCut}, i \in \{A, B\}$. Since $\pi^*_{i,cut} = \frac{p_e - c}{L}$, $\pi_{i,nCut}(p_e) = \frac{p_e}{2L}$, the condition is:

$$p_e \leq 2c \tag{3.16}$$

As starting from the monopoly price $p_0 + \frac{t}{2L} + s$, the firm is cutting price by c each time, the final price at equilibrium will be:

$$p_e = p_0 + s - \frac{t}{2L} - N_e \times c \tag{3.17}$$

, where $N_e = min_n\{p_0 + s - \frac{t}{2L} \leq (n+2)c\}$. In Equation (3.16), N_e represents the number of rounds of cutting prices after which, for the first time, the two vendors charge prices $p_{i,cut} \leq 2c, i \in \{A, B\}$ that satisfy the condition for equilibrium.

Therefore, by introducing the RS for a competitor, the monopolistic RS provider has motivation to lower the prices. Encouraging competition might be a potential policy for the regulators to restrict surplus exploitive behavior of the RS provider.

3.4 Conclusions, Implications, and Future Works

This research presents an analytical framework to evaluate the welfare impact of a monopolistic personalized online recommender system in a market with horizontally differentiated products and heterogeneous consumers. The analysis shows that through perfectly identifying the preference of consumers and reducing the search cost, the personalized recommender systems can serve as an alternative business tool to price discrimination, converting the willingness-to-pay of each consumer into profits. Such a personalized tool not only increases profit by a magnitude no smaller than price discrimination does, but also improves consumer surplus through reducing search costs.

The unique contribution of the authors is to evaluate personalization from an innovative perspective of how it resembles price discrimination. The proposed model is the first attempt to formally separate the two dimensions of personalization, i.e. prices and non-price product characteristics, and build the correspondence between them by a novel concept of personalization level. Personalization level is defined as the number of segments by which the firm personalizes recommendation strategies. Profit is maximized when personalization is “perfect”, i.e. each segment has a single consumer, and each one (segment) is recommended a different product. Consumer surplus will be zero in this case, resembling the effect of first-degree price discrimination. Finite levels of personalization correspond to third-degree price discrimination.

The managerial implication of this research is to demonstrate to industry practitioners that a personalized recommender system is an alternative to price discrimination in the form of price markup since consumers pay the markup of prices for reduced search costs. This type of recommender systems are more effective and profitable in industries where consumer search is exhaustive and difficult, or where the degree of product differentiation is not too high. Strategically, online vendors should develop technologies to segment and personalize product recommendations to a level that maximize the expected profits.

The authors’ work has important policy implication for regulators. It has proven how personalized recommendation is able to benefit both a monopoly firm and its consumers by reducing search cost. However, when personalization level is high enough, there is no space for search cost reduction and the firm exhibits too much exploitive behavior such that consumer surplus becomes negatively correlated with personalization level. Therefore, policymakers should be aware of that even though reduction of search costs are beneficial to consumers, the existence of a fair amount of searching behavior in the market is healthy for consumers. To restrict a firm’s exploitive behavior when there is no reduction in search cost, the essential part is to restrict the firm from charging high exploiting prices. As demonstrated in the discussion, policymakers can make higher personalization always benefit both the monopoly firm and consumers through the implementation of a price cap. Another potential policy analyzed is the introduction of a competitor offering RS that uses the same selection of products for the competitor's RS. The introduced competitor has been found to be willing to cut price to aggressively attract the consumers of the incumbent vendor if the switching cost for the incumbent company's consumers to switch to the competitor's RS is within some range of values. Such competition motivates the incumbent company to lower the price of its RS to fight back. At equilibrium, the prices that two companies charge are lower than the monopoly price without competition. In fact, the switching cost determines whether the competitor competes aggressively, and if the competitor competes by cutting price, the equilibrium prices are also functions of the switching cost. Policymakers can implement regulations, such as enabling the transfer of points between reward programs of different companies and advocating formation of alliance of companies, in order to encourage the easiness of switching between RS providers so as to get a desirable welfare result.

Several assumptions of the analytical model this work has been built on are tractable and generalizable, but assumptions of single period, perfect information, and zero technology cost are too strong to generalize. Future work can be done to expand the model to multiple periods and include uncertainties of both firm and consumers. Such a revised model incorporates the effect of a recommender system on helping consumers reduce uncertainties. Consumers build trust in a recommender system through experiences of previous interactions. This is not captured yet in the current single-period model with complete information. It's expected that the online vendor and consumers' interest will be more aligned with regard to the level of personalization. In addition, it's an interesting question to figure out by including the cost of improving segmentation technology, how a firm balances between raising personalization level to create more revenue and reducing such technology costs.

Chapter 4

Uncertainties and Consumer Learning in a Multi-Period Recommender System

4.1 Introduction

In a one-shot game like the model in Chapter 3, the shortsighted seller, as the designer of the RS, chooses prices and recommendations that maximize the one-period profit. Yet in the real world, it is unlikely that the seller only has one-period interactions with all of its consumers. Online consumers shop repeatedly. In general, consumers have repeated needs for a considerable variety of products. As found by the famous marketing consultant Jack Trout, an American family on average buys the same 150 items repeatedly which occupy 85% of the household spending (Schneider and Hall, 2011). Compared to brick-and-mortar stores, online shops have less geographic constraints, and consumers have more brand and platform choices. Online consumers might meet their online shopping needs at different stores and switch between stores in different periods of time (Sharp, 2013), which makes the online market more competitive from the sellers' side. Therefore, how to make consumers come back is the key to success for e-vendors. A large body of literature investigates the factors that influence consumers' online repeat-purchase intention (Chiu et al., 2014; Abdul-Muhmin, 2010; Gefen et al., 2003; Hoffman et al., 1999).

Nowadays major online companies are forward-looking. As an effective marketing tool, online RSs are deliberately utilized to boost both consumers' initial adoptions and their later repeated purchases. To examine the motives of consumers' purchase behavior, it is natural to model them as dynamic decision makers and predict their actions in multiple periods.

Previously in Chapter 3, the model describes one-period transactions between a monopoly firm and its consumers. Each consumer is assumed to know the exact location of the recommended product. Instead, in Chapter 4, I release this assumption and model consumers as being uncertain about the true location. Only through purchasing are they able to learn the true location of the recommended product.

The type of scenario I model in Chapter 4 can be illustrated by a synthetic example:

Suppose a monopoly online retailer has a large consumer base and all consumers used the catalog for searching and purchasing before. Now the vendor wants to introduce a new RS to its online platform. No one has used the system before and cannot observe the location of the product being recommended. But consumers have a guess on the location. Whether or not they want to buy depends on their guess. Consumers who purchase from the RS learn the true location and update their guess of the RS.

The purpose of the multi-period model is to predict a firm's recommendation strategy with consideration of the dynamic choices and learning behavior of consumers. I am interested to evaluate how a firm's exploitive behavior, as the result of Chapter 3, changes by its intention to retain consumers in the initial interactions. If having the power to influence consumers' perception of the product location, how the online merchant revises its recommendation strategies? For policymakers and regulators, this study explains how RSs might be used as a profit exploiting tool of e-vendors and highlights the key factors of how online RSs can be made healthy.

4.2 Model

To simplify the assumptions, I model the multi-period transactions between the monopoly firm and its consumers by a two-period game. In this game, one firm sets price and recommendations upfront before the first period, and prices are the same across periods. Without any purchase, consumers have a uniform prediction of the recommendation quality (closeness in preference space). Consumers who purchased previously are able to completely learn the quality. Firm is maximizing the sum of the two-period profits.

4.2.1 The Game

The two-period game can be described by four stages. In stage 1 and stage 2, the firm chooses the prices and the locations of the products to be recommended for two periods. Since firm doesn't know the exact location of each consumer, or firm can only choose a few product for recommendations, it chooses the number of different products, L, for recommendations to L groups of consumers. This setting is the same as in Chapter 3. Even though in Chapter 4, there are two periods, the prices and locations of the product recommended to each consumer are the same for period one and period two. Fig.4.1 illustrates the two-players multi-stage sequential game. After describing the four stages individually, I solve the optimal price by backward induction.

Stage 1

In stage 1, the e-vendor selects L unique products, each of which is to be recommended to each of the L consumer groups. Due to symmetry, the vendor selects products that are located equidistantly along the circular preference space. The personalization level L is exogenously determined. One explanation of exogenous L is limited targeting technology, which means the vendor is only able to specify the range of each consumer's location, i.e., which group on the preference space in Fig.3.3 each consumer belongs to. So the vendor splits the circle into L arcs of equal length and treats consumers situated on the same arc the same. The other explanation for exogenous L is that the vendor has a limited number of products available to recommend (L).

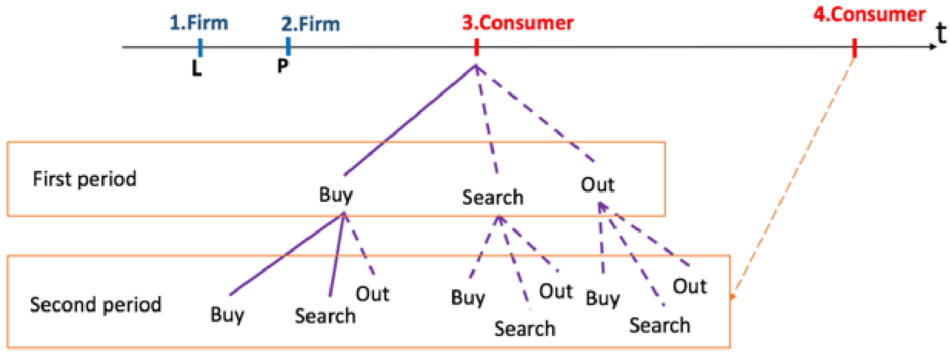

Figure 4.1: Timeline of decisions for a RS game of two-period transaction

Stage 2

In stage 2, firm chooses prices, denoted by p_{it}, where $t \in \{1, 2\}$ denotes time period. $p_{i1} = p_{i2} = p_i$. By symmetry and no price discrimination condition, $p_i = p$.

Stage 3

In the first period, none of the consumers have previous experiences with the RSs. Consumers have an initial guess on the recommendation quality. For horizontally differentiated products, qualities are proxied by the transportation cost from the product to a consumer, as the product of distance in preference space and the unit transportation cost. Specifically, I use r_{0ij} to denote consumer i's initial predicted distance of the recommended product j, and $r_{0ij} = r_0, \forall i, j$. Each consumer has three choices: accept and buy the recommended product, search and buy the catalog, or buy nothing. As shown in Chapter 3, consumers have a positive utility from searching the catalog, so they will always buy something. Whether they accept or search depends on which option gives a higher predicted consumer utility.

Under price $p > p_0 + s - tr_0$, predicted utility from accepting the recommended product is always lower than searching the catalog, so none of the consumers are motivated to take an initial step to experience the RSs. Then in the later periods, there is, no chance of learning and consumers always hold their intitial guess of r_0, so they don't accept as well. The e-vendor earns zero profit from recommendations if charging $p > p_0 + s - tr_0$.

If firm charges price $p \le p_0 + s - tr_0$, all consumers accept recommendations in the first periods, and pay p. Whether those consumers continue buying from recommendations in the second period or not, the vendor earns profit p from recommendations to whoever accepts in a specific period.

However if $p \le p_0 + s - tr_0 \le p_0, i.e.\ r_0 \ge \frac{s}{t}$, firm could have earn more profit $p_0 - p$ from encouraging those consumers who accept to search and buy from the catalog instead. So when $r_0 \ge \frac{s}{t}$, firm encourages consumers to search and charges recommendation price $p \ge p_0 + s - r_0$.

When $r_0 \le \frac{s}{t}$, the vendor charges $p \le p_0 + s - r_0$, and all consumers accept in the first period. Specifically, in the first period,

$$\forall i, \text{Consumer } i \begin{cases} \text{accepts recommendation} \quad \text{if } U_{accept} \ge U_{search} \quad V - tr_0 - p \ge V - s - p_0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad p \le p_0 + s - tr_0 \\ \hline \text{searches the catalog} \qquad \text{if } U_{accept} < U_{search} \qquad\qquad p > p_0 + s - tr_0 \end{cases}$$

Stage 4

If $r_0 > \frac{s}{t}$, and the vendor charges $p > p_0 + s - r_0$, then the game stops at stage 3 and firm has no incentives to introduce the RSs. Therefore it is more meaningful to talk about the scenario under the constraint $r_0 \le \frac{s}{t}$.

If $r_0 \le \frac{s}{t}, p \le p_0 + s - r_0$, in period one, all consumers buy directly from the catalog. At the beginning of stage 4, consumers completely learn the true location of the recommended product from the first-period purchasing experience. That is to say, consumers update their knowledge about the product distance by $r_i = r_0 + (r_i - r_0)$. Consumers not only learn the specific location of the product, but also learn the quality of the RS (how close). In period two, consumers choose to "accept" or "search", whichever generates higher surplus based on the true location of the recommended product.

$$
\text{So}, \forall i, \text{Consumer } i \begin{cases} \text{accept recommendation} & \text{if } U_{accept} \geq U_{search} \quad V - tr_i - p \geq V - s - p_0, \\ & p \leq p_0 + s - tr_i \\ & r_i \leq \frac{p_0+s-p}{t} \ \& \ r_i \leq \frac{1}{2L} \\ \hline \text{search the catalog} & \text{if } U_{accept} < U_{search} \quad p > p_0 + s - tr_i \\ & r_i > \frac{p_0+s-p}{t} \ \& \ r_i \leq \frac{1}{2L} \end{cases}
$$

, where r_i denotes the distace of the recommended product from consumer i.

It's straightforward that consumers who overestimate the distance ($r_i \leq r_0$) in the first period will buy directly in the second period for sure. Only consumers $\{i : r_i \in [r_0, \frac{1}{2L}]\}$ need to be discussed. If $r_0 \geq \frac{1}{2L}$, this case is trivial, since all consumers overestimate in the first period and so if they accept in the first period, they will also accept in the second period. In the more interesting case of $r_0 < \frac{1}{2L}$, viewing price p, the marginal consumer who is indifferent between "accept" directly and "search" is $r_m = \frac{p_0+s-p}{t} \geq r_0$. The farthest consumer from the recommended product is $\frac{1}{2L}$ distance away. So the position of the "marginal" consumer relative to the farthest consumer, as shown in Fig. 4.2, determines how many consumers accept recommendations. If $r_m < \frac{1}{2L}$, then consumers with $r_i \in (r_m, \frac{1}{2L}]$ search and buy from the catalog, while the other consumers accept and purchase from the RS. If $r_m \geq \frac{1}{2L}$, consumers all accept in the second period.

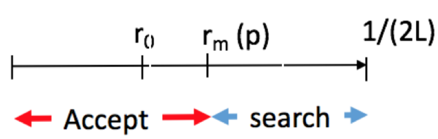

Figure 4.2: Consumer choice when the marginal consumer is at $r_m \in [r_0, \frac{1}{2L}]$

4.3 Analysis and Results

4.3.1 Optimal Price

In previous description of the game, I discussed the condition under which the firm is motivated to encourage consumers to accept recommendations, i.e. $r_0 \leq \frac{s}{t}$. When this condition is satisfied, the e-vendor will charge a price that makes all consumers accept in the first period, i.e. $p_0 \leq p \leq p_0 + s - tr_0$. Under this constraint, optimal price p that firm chooses at stage 2 maximizes the sum of the two-period profits.

So, when $r_0 \leq \frac{s}{t}$,

$$\begin{aligned} p|p_0 \leq p \leq p_0 + s - tr_0 &= argmax_p \sum_{period=1,2} \pi_{period} \\ &= argmax_p f(p; L) \\ where\ f(p; L) &= p + 2L \int_0^{\frac{1}{2L}} 1(search_r)p_0 + 1(accept_r)pdr \end{aligned} \tag{4.1}$$

Let $f_0(L) = 2p_0, \Delta p = p - p_0$, then

$$f_r(\Delta p; L) = f(p, L) - f_0(L) = \Delta p + 2L \int_0^{\frac{1}{2L}} 1(accept_r)\Delta p dr \tag{4.2}$$

It's easy to see that,

$$p|r_0 \leq \frac{s}{t} = p_0 + argmax_p f_r(\Delta p|L, \Delta p \leq s - tr_0) \tag{4.3}$$

In other words, the optimal price maximize the profit difference between with and without RSs.

The "accept" condition in stage 4 can also be represented by Δp in:

$$1(accept_r|r \leq \frac{1}{2L}) = \begin{cases} 1 & \text{if } r \leq \frac{s-\Delta p}{t} \\ 0 & \text{if } r > \frac{s-\Delta p}{t} \end{cases} \tag{4.4}$$

Substituting Eq.(4.4) into Eq.(4.2),

$$f_r(\Delta p|L, 0 \leq \Delta p \leq s - tr_0) = \begin{cases} \Delta p(1 + 2L(\frac{s-\Delta p}{t})) & \text{if} \quad r_0 < \frac{1}{2L} \ \& \ \Delta p \geq \max(0, s - \frac{t}{2L}) \\ 2\Delta p & \text{if} \quad r_0 < \frac{1}{2L} < \frac{s}{t} \ \& \ \Delta p < s - \frac{t}{2L} \\ 2\Delta p & \text{if} \quad r_0 \geq \frac{1}{2L} \end{cases} \tag{4.5}$$

Therefore, the optimal price is a function of consumers' initial guess on the quality of recommendations r_0, and the level of personalization L. Eq.(4.6) and Fig.4.3 presents the mathematical solution to profit maximization and the corresponding visualizations respectively.

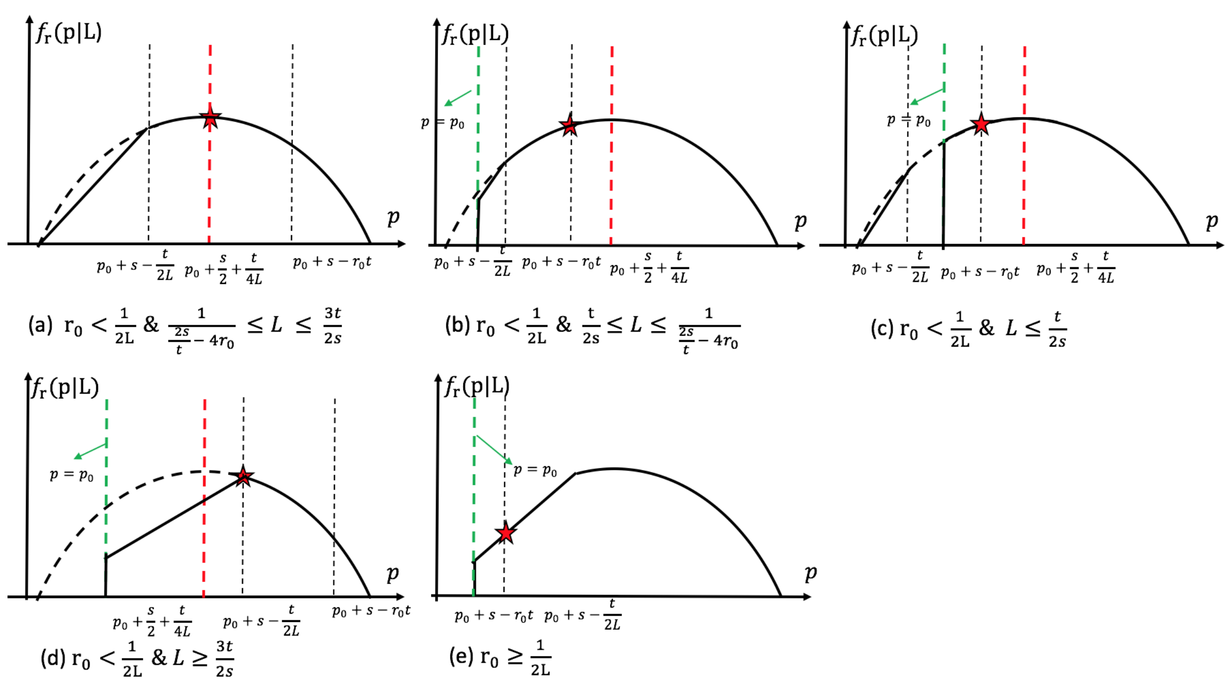

Figure 4.3: How optimal prices change with r_0 and L

$$p^*(r_0,L)=\begin{cases} p_0+s-\frac{t}{2L} & \text{if} \quad r_0<\frac{1}{2L}\ \&\ p_0+\frac{s}{2}+\frac{t}{4L}\le p_0+s-\frac{t}{2L} \iff L\le\frac{1}{2r_0}\ \&\ L\ge\frac{3t}{2s} \\ p_0+s-tr_0 & \text{if} \quad r_0<\frac{1}{2L}\ \&\ p_0+\frac{s}{2}+\frac{t}{4L}\ge p_0+s-tr_0 \iff L\le\frac{1}{2r_0}\ \&\ L\le\frac{1}{\frac{2s}{t}-4r_0} \\ p_0+\frac{s}{2}+\frac{t}{4L} & \text{if} \quad r_0<\frac{1}{2L}\ \&\ p_0+s-\frac{t}{2L}\le p_0+\frac{s}{2}+\frac{t}{4L}\le p_0+s-tr_0 \\ & \qquad \iff L\le\frac{1}{2r_0}\ \ \&\ \frac{1}{\frac{2s}{t}-4r_0}\le L\le\frac{3t}{2s} \\ p_0+s-tr_0 & \text{if} \quad r_0\ge\frac{1}{2L} \iff L>\frac{1}{2r_0} \end{cases} \tag{4.6}$$

The above Eq.(4.6) shows the optimal prices when r_0, L have mass in specific area of the parameter space. There are three optimal prices in total:

$$\begin{aligned} p_1 &= p_0+s-tr_0 \\ p_2 &= p_0+\frac{s}{2}+\frac{t}{4L} \\ p_3 &= p_0+s-\frac{t}{2L} \end{aligned} \tag{4.7}$$

p_1 : **Price Upper Bound**

p_1 is the highest price that can make all consumers accept in the first period, which is the price that makes consumers who accurately predict the quality of recommendations marginal consumers (with zero surplus). In the second period, only consumers who overestimated in the first period accept and purchase the recommended item. Naturally, when consumers all overestimated the distances, which means $r_0 \ge \frac{1}{2L}$, demand is fixed and profit is a linear function of price $f_r(\Delta p, p \le p_1) = \Delta p$ that monotonically increases in p. So it's optimal to choose p_1.

p_2 : **Optimal Price of the Unconstrained Problem**

Similar to the model in Chapter 3, firm's profit maximization is composed of two parts: the first part is to increase the margin obtained from each consumer who accepts $(p-p_0)$. The second part is to increase the number of second period demand from recommendations ($r_m = \frac{s-(p-p_0)}{t}$). The two conflicting factors balance at p_2 in theory. Considering the demand of each recommended product is at most $\frac{1}{L}$ and the farthest consumers are at most $\frac{1}{2L}$, if the distance of a marginal consumer under p_2 is within $[0, \frac{1}{2L}]$, then it's optimal for the firm to charge p_2.

p_3 : **Exploiting Price**

On the other hand, if under p_2, the distance of marginal consumers r_m is above maximum demand, then the actual demand is $\frac{1}{L}$ in each group. In this case, the firm can do better by increasing the price a little bit above p_2 without decreasing the demand. So p_2 is not optimal for the firm. In fact, firm can continue increasing the price until the the farthest consumers become marginal consumers (with zero surplus, $r_m = \frac{1}{2L}$). p_3 is corresponding price that leads to such r_m.

One thing to notice is that the two-period price also has a lower bound p_0 to ensure non-negative change in profit compared with the case without the RS.

The aforementioned cases of optimal prices are combined into Eq.(4.6) and visualized by Fig.4.3.

4.3.2 Consumers' Uncertainty r_0 and Firm's Uncertainty L

As previously shown, optimal prices depend on the pair of values of (r_0, L). The various optimal prices and sizes of "accept" or "search" consumers' sizes spanning over parameter space of (r_0, L) are summarized in table 4.1. To better visualize the region of (r_0, L) for different optimal

Table 4.1: Different optimal prices determined by (r_0, L)

Area	r_0	L	Optimal price	2^{nd} Period:"accept"	2^{nd} Period:"search"
II	$(0, \frac{s}{3t})$	$[1, \frac{1}{\frac{2s}{t}-4r_0})$	p_1	$(0, r_0)$	$(r_0, \frac{1}{2L})$
III		$(\frac{1}{\frac{2s}{t}-4r_0}, \frac{3t}{2s})$	p_2	$(0, \frac{s}{2t} - \frac{1}{4L})$	$(\frac{s}{2t} - \frac{1}{4L}, \frac{1}{2L})$
I		$(\frac{3t}{2s}, \frac{1}{2r_0})$	p_3	$(0, \frac{1}{2L})$	$\emptyset$
IV		$(\frac{1}{2r_0}, +\infty)$	p_1	$(0, \frac{1}{2L})$	$\emptyset$
II	$(\frac{s}{3t}, \frac{s}{t})$	$[1, \frac{1}{2r_0})$	p_1	$(0, r_0)$	$(r_0, \frac{1}{2L})$
IV		$(\frac{1}{2r_0}, +\infty)$		$(0, \frac{1}{2L})$	$\emptyset$

prices p_1, p_2, p_3, the parameter space is plotted in Fig.4.4. Each colored area represents one condition with different combination of (r_0, L). Specifically, when both r_0 and L are large (area IV

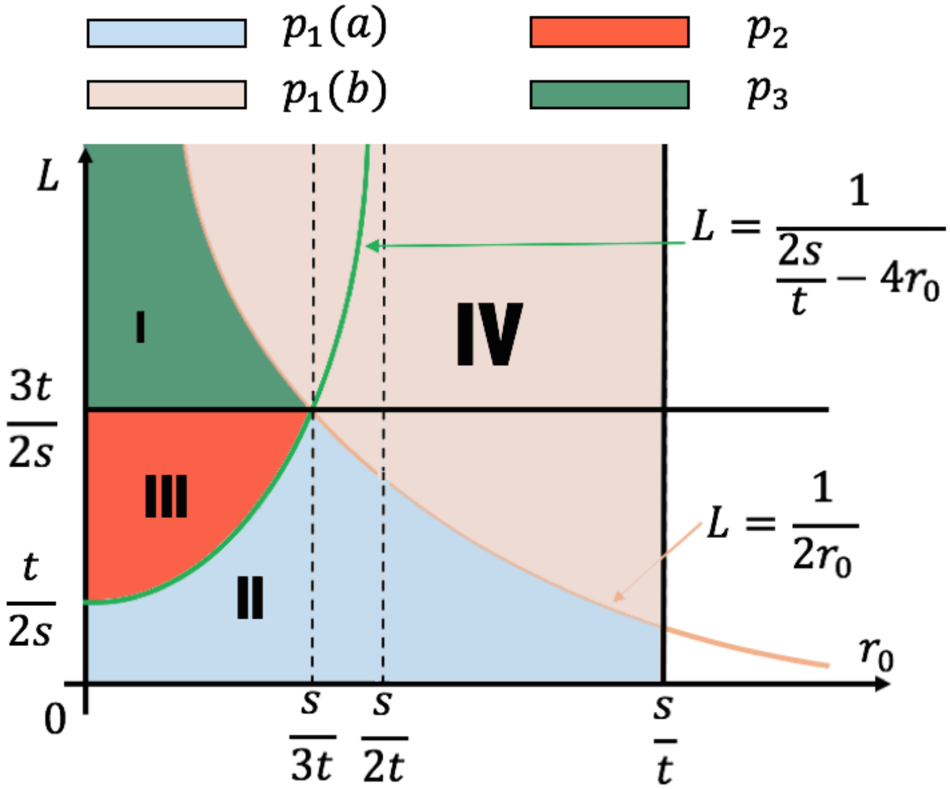

Figure 4.4: Different conditions of optimal price, as functions of (r_0, L)

of Fig.4.4), which means a large number of consumers overestimate the distances, if consumers accept in the first period, most of them also accept in the second period. Then firm is most likely to charge the highest price below which consumers are willing to adopt RS in the first period (p_1). In this case, the optimal price $p_1 = p_0 + s - tr_0$, and thus the maximal obtainable profit, is negatively correlated with r_0. On the other hand, larger r_0 means larger consumer surplus.

If r_0 is small, and L is also very small, most consumers underestimate the distances, so in the second period few consumers accept and the profits collected in the second period will be a small portion compared to the first-period profit. In this case firm will be short-sighted and charge the highest price as possible to collect more margins from the first period, and the optimal price is also p_1. As L increases, the farthest consumers in each group will be closer to the recommended product. The second period profits will be more important to the firm, so it feels more motivated to lower the price to increase demand in the second period. In the range of medium levels of L (area III of Fig.4.4), firm charges the optimal price p_2 as the balance between higher margin in two periods and larger demand in the second period. When L is so large that each group becomes so small and there is no room for expanding demand, firm will charge the highest price that makes all consumers accept in both periods (area I of Fig.4.4, p_2).

Recall that r_0 is the consumers' initial guess on the product quality, indicating their uncertainties about the product, while L is determined by how many groups of consumers can firm segment referring to its current knowledge of consumers' locations. When the above mathematical results are combined with the physical meaning of r_0 and L, I have the following result of how two-sided uncertainties interact and alltogether determine optimal price and the resulting welfares.

- High targetability (L) of the vendor and high initial evaluation (r_0) of the consumers will make the vendor shortsighted and charge the highest price at which consumers adopt recommendations in the first period.
- Higher targetability is always beneficial to both the firm and the total welfare.
- Low initial consumer evaluation of the RSs can give the consumer more power to motivate firm to lower price and then increase consumer surplus (shown more in detail in the next section).
- The hint of low intial consumer evaluation can limit firm's exploitive behavior.

4.3.3 Welfare

Previously it is demonstrated that different values of (r_0, L) lead to different optimal prices. This section analyzes the welfare results from not only different optimal prices, but also different acceptance conditions. It's shown from Table 4.1, even under the same price p_1, area II and area IV have different acceptance conditions, with partial acceptance ($(0, r_0)$ in each group, in total

$\frac{r_0}{2L}$ of consumers) and all consumers' acceptance respectively.

Particularly, the three welfare measures, profit, consumer surplus, and total welfare are presented under area I, II, III, and IV in Table 4.2. Because the objective is to analyze how welfares change before and after adding the RS, the welfare differences are used as the measures for, i.e. change in profit $\Delta\pi$, change in consumer surplus ΔCS, and change in total welfare ΔTW. From Table 4.2, it can be calculated how different welfare meansures change with r_0 and L, and

Table 4.2: Resulting welfare depending on (r_0, L)

Area	Optimal price	$\Delta\pi$	ΔCS	ΔTW
I	p_3	$2s - \frac{t}{L}$	$\frac{t}{2L}$	$2s - \frac{t}{2L}$
II	p_1	$-2tLr_0^2 + (2Ls - t)r_0 + s$	$tLr_0^2 + tr_0 - \frac{t}{4}\frac{1}{L}$	$-tLr_0^2 + 2sLr_0 - \frac{t}{4L} + s$
III	p_2	$\frac{s^2}{2t}L + \frac{t}{8}\frac{1}{L} + \frac{s}{2}$	$\frac{s^2}{4t}L - \frac{7t}{16}\frac{1}{L} + \frac{s}{4}$	$\frac{3s^2}{4t}L - \frac{5t}{16}\frac{1}{L} + \frac{3}{4}s$
IV	p_1	$2(s - tr_0)$	$2tr_0 - \frac{t}{2L}$	$2s - \frac{t}{2L}$

if there exist any conflicts between the vendor and consumers in preferences of r_0 and L. As a result, I generate Table 4.3.

Table 4.3: How welfares change with (r_0, L)

Welfare measure	(r_0, L)			
	Area I	Area II	Area III	Area IV
$\Delta\pi$	(=, ↗)	(↘, ↗)	(=, ↘↗)	(↘,=)
ΔCS	(=,↘)	(↗, ↗)	(=,↗)	(↗, ↗)
ΔTW	(=, ↗)	(↗,↗)	(=,↗)	(=, ↗)
Notes: = means "not depend on" ↘ means "decreases with", ↗ means "increases with".				

Profit Looking at row 1 of Table 4.3, for most of the time, profit decreases with r_0 and increases with L. The only exception is area III, in which profit is a bell-shaped function of L. However, when $L \geq \frac{s}{2t}$, as satisfied by area III values, profit still increases in L. Physical meaning of this result is that the vendor would be better off if its consumers have a good evaluation of its quality at the very beginning (small r_0), regardless of whether that guess is accurate or not, and if it is more accurate in the inference of consumers' locations (large L). This result is consistent with firm's commonly observed behavior of running advertising campaign to improve reputations and user experience, as well as its efforts to improve targeting technology.

Consumer Surplus Looking at row 2 of Table 4.3, consumer surplus decreases with L in area I and increases with L in area II, III, and IV. This means when consumers have a high initial guess on the recommendation quality, increasing targetability of the firm makes consumers worse off. This is because in area I & III , r_0, consumers' initial evaluation of recommendation in the first period, is so high and consumers enter so easily that price's upper bound from encouraging first-period entry into the RS almost plays no effects. In the region of area I & III, consumers and firm behave in a way similar to that of one-period Chapter 3 model: at low levels of personalization level L (area III), increased targetability makes more consumers "accept" and therefore increases profits and consumer surplus at the same time. But after the personalization level surpasses a threshold (area I), which is $\frac{s}{t}$ in Chapter 3 and $\frac{3t}{2s}$ here, all consumers "accept" and the vendor starts to "exploit" consumers and shifts consumer surplus to profits through increasing targetability L.

In area IV, higher targetability benefits consumers. This area is defined by the condition: $L > \frac{1}{r_0}$, which, with simple transformation, is equivalent to $r_0 > \frac{1}{2L}$. r_0 is the a consumer's initial inference on the distance of the recommendation, and $\frac{1}{2L}$ is the true distance of the farthest consumers from a recommended product. The incentive of the vendor to encourage all consumers' first-period entry, combined with the condition $r_0 > \frac{1}{2L}$, plays an effect now: to encourage all consumer to adopt the RS in the first period, the vendor will charge a price that makes a "virtual" consumer with distance r_0 accepts. Since actual distances of all consumers are less than r_0, all of them get positive surplus in both period, as

$$\begin{aligned}\Delta CS_i &= CS_{accept} - CS_{search} \\ &= (V - tr_i - p) - (V - s - p_0) \\ &= s + p_0 - tr_i - p \\ &\geq s + p_0 - t\frac{1}{2L} - p \\ &> s + p_0 - tr_0 - p \geq 0\end{aligned}$$

So all consumers overestimate the distances of the recommendations, once they accept in the first period, they accept in the second period as well. Firm charges the highest price that ensures first-period entries. With r_0 fixed, if the vendor increases L within the range $(\frac{1}{2r_0}, +\infty)$, say from L_1 to $L_2(> L_1)$, the willingness-to-pay of the farthest consumers increase by $(-\frac{t}{2L_2}) - (-\frac{t}{2L_1})$ under the same price p_1. Since the price is bounded above by the first-period price to ensure entries, the vendor cannot charge higher price to exploit such extra willingness-to-pay. This is different from the Chapter 3 model, where higher L comes with higher price in the exploiting part of the firm. In the component of individual consumer surplus equation $\Delta CS_i = s + p_0 - tr_i - p$, as L increases, only r_i shrinks, with p fixed ($p = p_1$), so the total consumers surplus increases monotonically with targetability.

In area II, consumers surplus also increases with higher targetability, but because of a different reason. In area II, $r_0 \leq \frac{1}{2L}$, so in the second period there are some consumers who underestimate the true distances of the recommended products. Which consumers accept in the second period, among those who are $(r_0, \frac{1}{2L})$ distances away from the recommendations, depends on the price p. Under area II condition, the total profit increases monotonically in p, as illustrated by Fig.4.3 (b,c). So the vendor will charge p_1 and give up all consumers at $(r_0, \frac{1}{2L})$ distances away from recommendations in the second period. In each group in the second period, only consumers at $(0, r_0)$ accept. Therefore, higher targetability means more groups of consumers and thus more consumers surplus in total.

From analysis of area II and area IV, I have demonstrated the role of r_0 in lowering price and increasing consumer surplus. In fact row 2 of Table 4.3 shows that in area II & IV, consumer surplus increases in r_0. This indicates that when consumers question more about the quality of recommendation, they have more power to let the firm charge low price in order to ensure first-period adoption of the RS.

Total Welfare Looking at row 3 of Table 4.3, total welfare, which is the sum of profit and consumer surplus, grows monotonically with the targetability. The one-period model in Chapter 3 and the two-period model are consistent on the relationship between total welfare and the level of personalization L. For most of the time the total welfare is not correlated with consumers' initial evaluation of the recommendations, with the exception of area II. Area II features optimal price p_1 by which firm encourages first-period all consumers' entries and second-period entries of only consumers who overestimates in the first period. So $2r_0L$ describes the second-period size of "accept" consumers. As demand of using recommendations increases, total welfare grows.

4.3.4 Compared to Chapter 3 after Extending to Two Periods

The objective of the two-period model is to analyze firm's behavior in the RSs when it is promoting adoptions by its current consumer base. The two-period model adds one period of intial adoption to the one-period model in Chapter 3 and assumes consumers' uncertainties of the recommendation quality.

Remember in Chapter 3, firm's behavior is splitted into two parts, first of which is the exploring behavior at low levels of personalization and both firm and consumers benefit from more personalization as adoption increases. The second one is the exploitive behavior under high levels of personalization. All consumers adopt recommendations and firm starts to increase price through increasing the level of personalization. The result of the new two-period model limits such exploitive behavior from the following perspectives:

Limit Exploiting Price As shown in Chapter 3, the exploiting price charged by the firm in the one-period model is

$$p = p_0 + s - \frac{t}{2L} \tag{4.8}$$

In the two-period model, this is exactly p_3 and firm only chooses p_3 in area I defined by $r_0 \leq \frac{1}{2L}$ and $L \geq \frac{3t}{2s}$. When $r_0 \leq \frac{1}{2L}$, i.e. consumers have good initial evaluation of the recommendations, the uncertainty doesn't affect firm's decisions. In comparison, when r_0 is large, the vendor has to choose lower-than-exploiting price in order to ensure all consumers adopt recommendations in the first period, either p_1 or p_2.

The two-period model still exhibits patterns of exploring and exploiting, as shown in the previous section. The turning point from exploring to exploitive behavior is to have a level of personalization more than $\frac{t}{s}$, which is lower than the necessary condition for the two-period model, $\frac{3t}{2s}$. Another condition, if combined with $L \geq \frac{3t}{2s}$ constitute the sufficient condition for exploitive behavior, is $L \leq \frac{1}{2r_0}$. r_0 is assumed only in the two-period model, and the exploiting condition is harder for the firm to satisfy as r_0 becomes larger. Reflected in the Fig.4.4 is the shrinkage of green area I (exploiting area) as r_0 gets higher.

Increase Consumer Surplus Chapter 3 also shows that consumer surplus first increases and then decreases with L. The turning point is also $\frac{t}{s}$. For the two-period model, the turning point between area I and area III is $\frac{3t}{2s}$, which is larger than $\frac{t}{s}$. Even under large targetability, if a consumer has large uncertainty and predicts a bad quality of recommendation initially (r_0), then exploiting condition can be avoided. The introduction of r_0 allows the consumer surplus to be remediated by large r_0.

In a nutshell, by adding the first period when consumers have uniform guess on the distance, firm needs to set a relatively lower price to persuade them to buy directly. This constraint puts an upper bound on the price, so firm cannot exploit all consumers surplus. To some extent I have achieved the goal of limiting the firm's exploitive behavior.

4.3.5 Discussion of L, Search Costs s, and Product Differentiation t

As the personalization level L is discrete and no less than 1, the parameter space plotted in Fig.4.4 might not be all valid. The parameter space that is valid under $L \geq 1$ should be above the line $L = 1$. The relative position of $L = 1$ in the space determines whether each one of the four areas discussed above is valid or not. Such relative position is a function of search cost s, and the product differentiation level/unit transportation cost t.

In fact, the relative position of $L = 1$ to the two critical values $\frac{t}{2s}$ and $\frac{3t}{2s}$ is a function of $\frac{t}{s}$. The valid areas with different conditions of $\frac{t}{s}$ is the following:

$$\begin{Bmatrix} \text{Area I \& IV} & \text{if} & \frac{t}{s} < \frac{2}{3} \\ \text{Area I, II, III \& IV} & \text{if} & \frac{t}{s} \geq \frac{2}{3} \end{Bmatrix} \tag{4.9}$$

The validity conditions of four areas represented by Eq.(4.9) can be put into the parameter space as of Fig.4.4. The result is Fig.4.5.

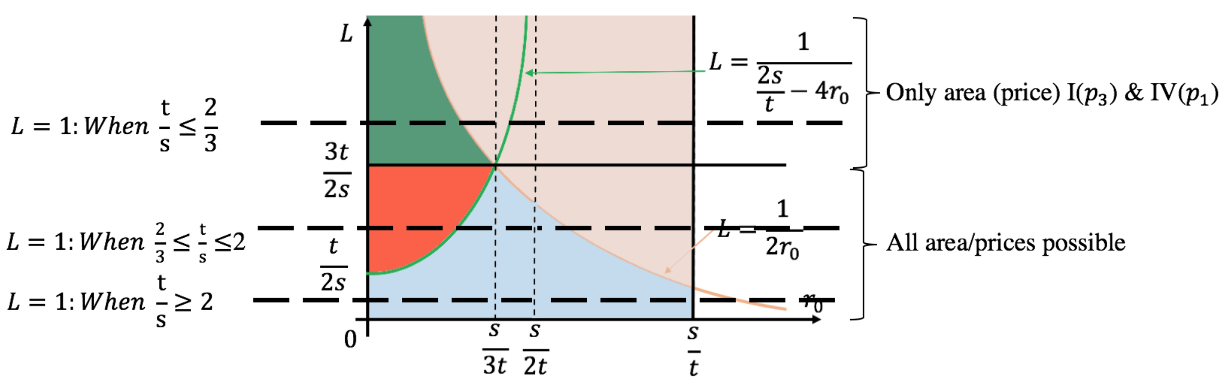

Figure 4.5: The relative position of line $L = 1$

4.3.6 Discussion of r_0

As the essential part of the two-period model, the value of r_0 as consumers' initial evaluation of the recommendation quality determines how much the concern of first-period adoption can limit the second period exploitive behavior of the e-vendor. It is therefore important to discuss the valid region of r_0.

When an existing consumer of the vendor is new to the RS, it is usually assumed that the consumer has some prior information about the RS. Such information might be the exact true location of the recommended product, which is the case of Chapter 3, or it might be consumers' prior belief on how the recommended product location is distributed. What is more relevant information in my model is the firm's targetability, and if the consumer gets to know such information of L correctly, the location information can be estimated/ derived according to the value of L.

If a consumer is able to obtain accurate information of L, r_0 can be assumed to follow some specific distribution supported by $[0, \frac{1}{2L}]$. If no other information is available, consumers can select uniform distribution on $[0, \frac{1}{2L}]$, and if product information is partially observed, consumers might assume normal or triangle distribution centered at the point on the circle where the observed product characteristics have the most occurrences.

Consumers' prior distribution is assumed to be homogeneous in the model. However, for most of the time in reality, it is heterogenous. For example, even if the consumers' prior distribution belongs to the same family such as normal distribution, if firm reveal different partial characteristics of the product to consumers, or the sources of information obtained by the consumers are different, the means of prior distribution are heterogenous across consumers.

Strategic Choices of r_0 on Both Sides

As the initial guess of product location r_0 goes into both the profit function of the vendor and the consumer surplus functions of the consumers, it is relevant to discuss how firm and consumers might strategically choose or influence r_0 to improve their own welfare.

First, on the vendor's side, profit decreases in r_0. If it learns how consumers form opinions about the recommendation quality, such as their prior distribution of r_0, the vendor might strategically personalize the information to be revealed to each consumers that makes them choose a distribution of r_0 with low mean. In practice, if the firm knows a specific consumer prefers old romantic movies, the firm might reveal the year of release information and hide the genre information of a old horror movie it decides to put into the RS. The objective of firm's strategic decision is to increase r_0 as much as possible.

On the other side, for consumers, consumer surplus increases in r_0. r_0 increases consumers surplus by influencing the vendor's pricing decision. In order to ensure all adoptions in the first period, the highest price to charge decreases in r_0. So in order to achieve the same purpose, consumers might strategically give wrong hints that they have a bad evaluation of the recommendations and threaten the vendor to not adopt it in the first period.

To summarize, the strategies of both the firm and consumers to increase welfare through influencing r_0 are based on their current information of the other players. If both firm and consumers make strategic decisions, they are actually competing for the amount and accuracy of information they acquire.

A Simple Deterministic and Homogenous r_0

As previously explained, consumers' prior information on r_0 can be a random variable with some specific distribution, either homogenous or heterogenous across individual consumers on the circle. To begin with, I analyze a simple case of deterministic and homogenous initial evaluation of r_0:

$$r_0 = \frac{1}{4L} \tag{4.10}$$

As shown in the previous model, in the first period, all consumers accept under price $p_0 + s - p \geq tr_0$, i.e. $p \leq p_0 + s - tr_0 = p_0 + s - \frac{t}{4L}$. and in the second period, marginal consumers on the unbounded Hotelling line:

$$r_m = \frac{p_0 + s - p}{t}$$

As the actual groups are on bounded Hotelling arc, the actual marginal consumer who accepts recommendations is at $r_i = min(\frac{p_0+s-p}{t}, \frac{1}{2L})$ distance away from the product.

As concluded previously, the consumer will always search or purchase recommendation, firm can always earn not-lower-than search price from each consumer. So firm is maximizing the profit difference between the RS and the catalog. The two-period profit difference is:

$$\Delta\pi(p|p \leq p_0 + s - tr_0) = \Delta\pi_1(p) + \Delta\pi_2(p) = (p - p_0)(1 + 2L * min(r_m, \frac{1}{2L}))$$

Specifically, when $p \leq p_0 + s - \frac{t}{2L}$, $r_m \geq \frac{1}{2L}$. When $p_0 + s - \frac{t}{2L} \leq p \leq p_0 + s - \frac{t}{4L}$, $\frac{1}{4L} \leq r_m \leq \frac{1}{2L}$.

So the two-period profit as a function of price,

$$\Delta\pi(p|p \leq p_0 + s - \frac{t}{2L}) = 2(p - p_0)$$
$$\Delta\pi(p| \leq p_0 + s - \frac{t}{2L} \leq p \leq p_0 + s - \frac{t}{4L}) = -\frac{2L}{t}[(p - p_0)^2 - (\frac{t}{2L} + s)(p - p_0)]$$

The optimal prices:

$$p^*|p \leq p_0 + s - \frac{t}{2L} = p_0 + s - \frac{t}{2L}$$

$$p^*| \leq p_0 + s - \frac{t}{2L} \leq p \leq p_0 + s - \frac{t}{4L} = \begin{cases} p_0 + s - \frac{t}{4L} & \text{if } L \leq \frac{t}{s} \\ p_0 + \frac{s}{2} + \frac{t}{4L} & \text{if } \frac{t}{s} \leq L \leq \frac{3t}{2s} \\ p_0 + s - \frac{t}{2L} & \text{if } L \geq \frac{3t}{2s} \end{cases}$$

Combining two cases of $r_m \geq \frac{1}{2L}$ and $r_m \leq \frac{1}{2L}$, I plot the profit function and mark optimal prices for each scenarios (different levels of personalization) as red stars in Fig.4.6. Plugging the

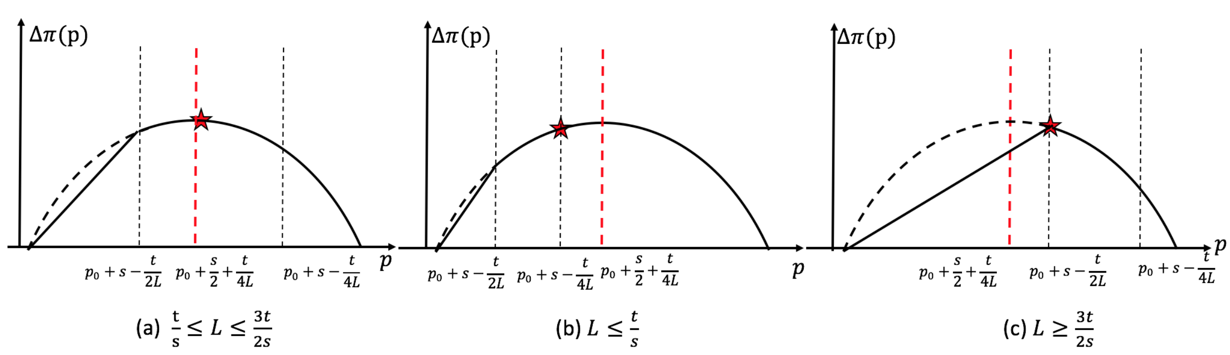

Figure 4.6: Optimal prices with different level of personalization

optimal prices, the resulting optimal profit as a function of personalization level is as follows.

$$\Delta\pi^*(L) = \begin{cases} \frac{3s}{2} - \frac{3t}{8L} & \text{if } L \leq \frac{t}{s} \\ \frac{s^2}{2t}L + \frac{t}{8}\frac{1}{L} + \frac{s}{2} & \text{if } \frac{t}{s} \leq L \leq \frac{3t}{2s} \\ 2s - \frac{t}{L} & \text{if } L \geq \frac{3t}{2s} \end{cases} \tag{4.11}$$

It's easy to find that the optimal profit increases monotonically in L as shown in Eq.(4.11). Therefore in this case of $r_0 = \frac{1}{4L}$, if L is endogeneous, firm will choose infinite level of personalization.

4.3.7 Discussion of L: the Return to Investment of Improving Targetability

In previous sections and Chapter 3, I have shown that firm always prefers higher level of personalization L as well as higher targetability, and all previous discussion of this Chapter 4 is around how low evaluation of recommended product in consumers' initial adoption is able to mediate the exploiting-through-better-targeting behavior of the firm. However, in real life, improvement of targeting technology has a cost, and if the benefit from more targetability cannot afford to cover such cost, the vendor might not be willing to do so.

Using the simple case of $r_0 = \frac{1}{4L}$ as consumers' prior quality information, Fig.4.7 plots the profit function against increasing level of personalization L when $s = 1, t = 6$, according to Eq. (4.11) . From Fig.4.7, it is easy to observe that the diminishing marginal returns to

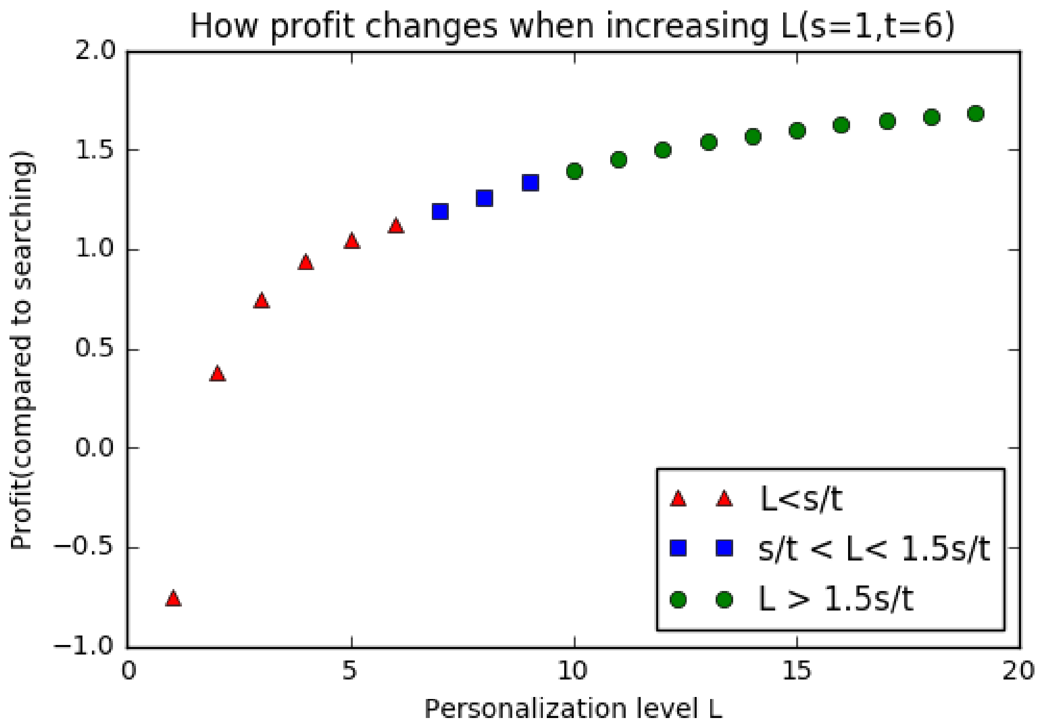

Figure 4.7: How profit increases with the level of personalization

increasing L. In contrast, the technology cost of improving targetability usually has a increasing marginal cost of increasing L. Several researchers assumed convex cost function of investing in improving targetability of consumers. For example, Chen and Iyer (2002) assume $c(x) = \frac{1}{2}kx^2$, where x denotes addressability in width. Chen et al. (2001) apply $c(I) = cI^2$, where I denotes targetability, defined in the paper as the accuracy of correctly identifying the consumer stype. To optimize profit, firm should choose the level of personalization at which marginal benefit equals the marginal cost of increasing targetability.

4.4 Concluding Remarks and Future Work

In this part of the thesis, I extend the one-period model developed in Chapter 3 to a two-period model by adding one adoption period before the one-period model. In the two-period model, I assume consumers don't observe the true location of the recommended product before any purchase, but can learn completely after one period purchase. This setting motivates the vendor to lower the price to ensure first-period adoption of consumers. By solving the extended model, I have demonstrated that the vendor's optimal price in two periods is lower than the optimal price in a one-shot game, and the firm exhibits less exploitive behavior.

Particularly, the profit increases with targetability and decreases with consumers' initial evaluation of the distances from the recommended products. Targetability is measured in the level of personalization, while the distance between a consumer and the recommended product is a proxy for the quality of horizontally differentiated products. In contrast, consumer surplus exhibits a non-monotonic trend when targetability increases, but increases monotonically as the initial evaluation of product quality is lower.

Furthermore, the potentials are discussed for both firm and consumers to strategically influence the knowledge of each other on observed targetability and initial inference of quality respectively in order to increase their own welfares. As the firm and the consumers have conflict of interest on how the two factors should change, two players can compete for the accuracy of information they have and change the two factors in the directions they prefer.

In addition, I have also pointed out that the firm's decision of improving targetability not only depends on the benefits from higher personalization, but also relies on the cost of technology improvements.

This part of the study is an initial step in exploring a monopolistic vendor's long term behavior and resulting welfare impact in personalized RSs. This work contributes to the literature by analyzing the conflicts of interest between the vendor and its consumers in such a multi-period transaction model along with two important factors, which are the firm's targetability and consumers' initial evaluation of the recommender systems. In the future, the topics I mentioned, such as the two-sided strategic choices as well as the benefit-cost analysis of the targeting technology, should be formally analyzed in a similar framework.

Chapter 5

Conclusions and Implications

In the first part of my thesis work, the research question is whether the profit-driven firm will choose a recommendation mechanism that hurts or is suboptimal to its consumers. I explore this question empirically with a concrete recommender system created by our industry collaborator for their Video-on-Demand (VoD) system. Empirically, in order to evaluate the counterfactual profit-maximizing recommender system, I first chose the commonly adopted exponential demand function and calibrated the price and slot elasticities using the sales dataset of the online VoD recommender system. Particularly, I implemented a Poisson regression model, which is typically used for count data, regressing sales on video features. Thanks to the large-scale (300,000 users) randomized experiment, I was able to get consistent estimates for elasticities. The average price elasticity is about -0.4 and the average elasticity of demand, by moving movies from the right to left (considered superior) slots, decreases by 0.08. The heterogeneous fixed effect is also estimated for each movie. Next, in order to explore the profit-maximizing recommendation mechanism, 1000 simulations were conducted. In each iteration, 15 movies were randomly sampled and used for recommendation. Plugging the exogenous prices, fixed effects, as well as price and slot elasticities into the demand functions, the firm's optimal slot assignments of the selected 15 movies were calculated. As a comparison, the mechanisms to maximize consumer surplus and total welfare, along with other popular schemes like ranking by previous sales, IMDb ratings or IMDb votes, were also calculated. In the end, three aggregate welfare measures, i.e. profits, consumer surplus and total welfare, were evaluated for each 15-movie set, for each simulation and for seven mechanisms of recommendations. As a result, the profit-maximizing recommender system was found to generate 8% less than the consumer surplus-maximizing recommender system.

From this study, I conclude that in the real world, there exists a conflict of interest between the firm and the consumers in recommender systems, and this identified significant conflict only arises from the different preference in assignment of listed orders. Practically, since there are more design variables besides listed orders, this evaluation of the conflict, i.e. the potential relative loss of consumer welfare, is an merely underestimates of the actual consumer loss. This study indicates a potential need for institutional or governmental regulatory interference in order to protect consumers. As shown in the result, there might exist an assignment of listed orders that maximize total welfare as the sum of the firm and consumers' welfares, this type of optimization is potentially preferred by policymakers and regulators.

What's more, in the first part what I quantified is a recommender systems for a representative consumer, and in the econometrics model I used to estimate demand function, I merely differentiate consumers by two types: whether they are premium or standard consumers. This is not enough for a world teeming with personalizations. Encouraged by this idea, I investigated the topic of personalization in recommender systems.

In the second part of my research, I explored the role of personalization in online recommender systems on improving firm's profit, consumer surplus, and total social welfare, especially how the effect of personalization resembles that of price discrimination strategy. I proposed an analytical framework to model a monopoly firm offering horizontally differentiated products in the catalog and deploying recommendations with a specific level of personalization. By evaluating the profit, consumer surplus, and total welfare after introducing the recommender system, I was able to show that the profit-driven recommender system improves all three welfare measures, mainly because it reduces the consumer search cost. While the magnitude of improvement of profits and total welfare increases monotonically with the level of personalization, it is important to point out that the increase in consumer surplus first gets higher and then shrinks when the personalization level grows further. In extreme cases, when the firm offers each consumer a different recommendation, i.e. perfect personalization, it approximates the effect of price discrimination strategy, wherein the company captures all the increase of surplus from consumers. This result motivates policymakers and regulators to rethink the effect of personalized recommendations. The welfare impact of different recommendation strategies largely varies across different application situations and depends largely on the format of recommender systems, the search cost of the navigation system of the websites, and the level of differentiation of the available products. It's important to conduct a comprehensive study of recommender systems that carefully calibrates the specific factors I have talked about in this thesis. What's more, researcher and executors should treat the negative aspect of personalization as seriously as price discrimination.

In the third part of the dissertation, motivated by consumer's repeated purchase behavior and uncertainties of the seller on consumer preferences and consumers' uncertainty about recommendation quality, I analyzed welfare properties of recommender systems in a framework similar to the second part but with the addition of an initial period for consumer learning about the recommendation quality. Extension to the two-period model motivates the monopoly e-vendor to lower price in order to ensure consumers' entries in the initial period, which increases consumer surplus of both periods. The firm's exploitive behavior through high levels of personalization in the one-period model is restricted by such an incentive from the initial period.

The third part highlights two important factors that have conflicting welfare effects: targetability and consumers' initial guess. Even through increasing targetability, the monopolistic vendor is able to provide each consumer with the recommendation that is so close to them that consumers have high willingness-to-pay. However, the vendor cannot take all the surplus by charging the price as much as consumers' willingness-to-pay, because in the initial adoption period, inexperienced consumers don't know the recommendation is of high quality and therefore the adoption price should match their initial guess on the quality.

The results have significant implications for policymakers. In part one, I have suggested regulations of the listed orders or other designs of recommender systems than the profit-driven design. The third study suggests a capital way to protect consumers' welfare that arises from a healthy market mechanism, which is a long-term collaboration between the firm and its consumers. As such, encouraging feedback mechanisms and buyer and sellers' loyalty that promote long term transaction could give policymakers less barriers and achieve more efficient welfare effects.

Chapter 6

Future Works

The work of this thesis is an initial investigation into the potentially existing conflict of interest between the e-vendor and the targeted consumers of its recommender systems. Following this thesis work, there are two major directions to develop:

6.1 Empirical Investigation of Personalized Recommender Systems

The future application of the first part's empirical work can be focused on applying the empirical and analytic framework to highly personalized recommender systems that implement state-of-art algorithms such as matrix completions (Candes and Recht, 2012). The organic results of recommendations by the state-of-art algorithms are analogous to the recommender systems in the randomized experiment of the study. Based on the empirically estimated demand function for organic results, the next step is to simulate how the firm can use the demand function for each individual consumer to optimize listed orders with the objective of profit maximization. The profit-driven assignment is then compared to the other assignment discussed in the current study so that the existence of conflicts of interest can be identified and quantified, if any.

Furthermore, another layer on top of the above proposed work is to have two parameters as decision variables, which are the targetability and addressability. Targetability is the accuracy of recommendations for each consumers, and the addressability describes how many consumers the vendor can approach and recommend. It would be interesting to see how welfare properties of the personalized recommender systems vary by the depth and width of personalizations.

6.2 Extensions of the Analytic Model

In the second part during the discussion of two factors, consumers' initial quality inference for the recommender systems and the provider's targetability of consumer locations, two possible topics have been mentioned.

The first one is to model both the firm and its consumers as strategic players. That means firm and consumers can influence consumers' initial guess of the product locations. The consumers' initial quality inference can be a function of the firm's decision and investment, while the firm's knowledge of the consumers' inference can also be strategically affect by consumers.

In addition, the future work on investigating a firm's decision to improve targetability can be focused on developing an analytic model including a convex technology cost on targetability improvement similar to those developed by Chen et al. (2001) and Chen and Iyer (2002).

References

Alhassan G Abdul-Muhmin. Repeat purchase intentions in online shopping: The role of satisfaction, attitude, and online retailers' performance. *Journal of International Consumer Marketing*, 23(1):5–20, 2010. 4.1

Panagiotis Adamopoulos and Alexander Tuzhilin. On unexpectedness in recommender systems: Or how to better expect the unexpected. *ACM Trans. Intell. Syst. Technol.*, 5(4):54:1–54:32, December 2014. ISSN 2157-6904. doi: 10.1145/2559952. URL http://doi.acm.org/10.1145/2559952. 1.1, 3.1.4

Anna Alberini. Efficiency vs bias of willingness-to-pay estimates: bivariate and interval-data models. *Journal of environmental economics and management*, 29(2):169–180, 1995. 3.1.3

M.W. Anderson, D. Hargreaves, S.M. Hoellig, D.S. Johnson, and B.D. Mischel. Method and system for gathering and analyzing consumer purchasing information based on product and consumer clustering relationships, October 26 1999. URL https://www.google.com/patents/US5974396. US Patent 5,974,396. 3.1.3

J. Yannis Bakos. Reducing buyer search costs: Implications for electronic marketplaces. *Management Science*, 43(12):1676–1692, 1997. doi: 10.1287/mnsc.43.12.1676. URL http://dx.doi.org/10.1287/mnsc.43.12.1676. 1.1, 3.1.5

Richard C Bishop and Thomas A Heberlein. The contingent valuation method. *Economic valuation of natural resources: Issues, theory, and applications*, pages 81–104, 1990. 3.1.3

Jeffrey R Blend and Eileen O Van Ravenswaay. Measuring consumer demand for ecolabeled apples. *American Journal of Agricultural Economics*, 81(5):1072–1077, 1999. 3.1.3

Erik Brynjolfsson, Yu (Jeffrey) Hu, and Michael D. Smith. Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Manage. Sci.*, 49(11):1580–1596, November 2003. ISSN 0025-1909. doi: 10.1287/mnsc.49.11.1580.20580. URL `http://dx.doi.org/10.1287/mnsc.49.11.1580.20580`. 1.1, 2.1, 2.2.1, 3.1.1

Erik Brynjolfsson, Yu Jeffrey Hu, and Michael D. Smith. From niches to riches: The anatomy of the long tail. *Sloan Management Review*, 47(2):67–71, July 2006. 1.1, 3.1.1

Emmanuel Candes and Benjamin Recht. Exact matrix completion via convex optimization. *Communications of the ACM*, 55(6):111–119, 2012. 6.1

CBSNews. Cbs news goes undercover to reveal gender price discrimination. 2016. URL `http://www.cbsnews.com/news/price-discrimination-gender-gap-cbs-news-undercover-dry-cleaners/`. 3.1.2

CBSNews. Can shopping online make you a victim of price discrimination? 2017a. URL `http://www.cbsnews.com/news/shopping-online-could-make-you-a-victim-of-price-discrimination/`. 3.1.2

CBSNews. Browser history and location can change what you pay online. 2017b. URL `http://www.cbsnews.com/videos/browser-history-and-location-can-change-what-you-pay-online/`. 3.1.2

Yuxin Chen and Ganesh Iyer. Research note consumer addressability and customized pricing. *Marketing Science*, 21(2):197–208, 2002. 4.3.7, 6.2

Yuxin Chen, Chakravarthi Narasimhan, and Z John Zhang. Individual marketing with imperfect targetability. *Marketing Science*, 20(1):23–41, 2001. 4.3.7, 6.2

Chao-Min Chiu, Eric TG Wang, Yu-Hui Fang, and Hsin-Yi Huang. Understanding customers' repeat purchase intentions in b2c e-commerce: the roles of utilitarian value, hedonic value and perceived risk. *Information Systems Journal*, 24(1):85–114, 2014. 4.1

Vidyanand Choudhary and Zhe (James) Zhang. Recommender systems and consumer product search. 2014. 3.1.4

Scott Davis, Eitan Gerstner, and Michael Hagerty. Money back guarantees in retailing: matching products to consumer tastes. *Journal of Retailing*, 71(1):7 – 22, 1995. ISSN 0022-4359. doi: http://dx.doi.org/10.1016/0022-4359(95)90010-1. URL `http://www.sciencedirect.com/science/article/pii/0022435995900101`. 3.1.3

Larry Dignan. "mac users pay more than pc users, says orbitz", 2012. URL `http://www.cnet.com/news/mac-users-pay-more-than-pc-users-says-orbitz/`. 1.2

Avinash K Dixit and Joseph E Stiglitz. Monopolistic competition and optimum product diversity. *The American Economic Review*, 67(3):297–308, 1977. 3.1.1

Economist. They're watching you. 2003. URL `http://www.economist.com/node/2137947`. 3.1.2

Daniel Fleder and Kartik Hosanagar. Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management science*, 55(5):697–712, 2009. 1.1, 3.1.1

Robert Garfinkel, Ram D. Gopal, Bhavik Pathak, and Fang Yin. Empirical analysis of the business value of recommender systems. November 2006. URL `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=958770`. 1.1, 1.2

David Gefen, Elena Karahanna, and Detmar W Straub. Trust and tam in online shopping: an integrated model. *MIS quarterly*, 27(1):51–90, 2003. 4.1

W.T. Glaser, T.B. Westergren, J.P. Stearns, and J.M. Kraft. Consumer item matching method and system, February 21 2006. URL `https://www.google.com/patents/US7003515`. US Patent 7,003,515. 3.1.3

Miguel Godinho de Matos, Pedro Ferreira, and Rodrigo Belo. Price discounts and peer effects in information goods: Results from an in-vivo organic randomized experiment. *Available at SSRN 2551800*, 2015a. 3.1.2

Miguel Godinho de Matos, Pedro Ferreira, and Rodrigo Belo. Price discounts and peer effects in information goods: Results from a randomized experiment. 2015b. URL `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2551800`. 2.2.3

Carlos A. Gomez-Uribe and Neil Hunt. The netflix recommender system: Algorithms, business value, and innovation. *ACM Trans. Manage. Inf. Syst.*, 6(4):13:1–13:19, December 2015. ISSN 2158-656X. doi: 10.1145/2843948. URL `http://doi.acm.org/10.1145/2843948`. 1.1

Aniko Hannak, Gary Soeller, David Lazer, Alan Mislove, and Christo Wilson. Measuring price discrimination and steering on e-commerce web sites. In *Proceedings of the 2014 conference on internet measurement conference*, pages 305–318. ACM, 2014. 3.1.2

Gerald Haubl and Valerie Trifts. Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing Science*, 19(1):4–21, January 2000. ISSN 1526-548X. doi: 10.1287/mksc.19.1.4.15178. URL `http://dx.doi.org/10.1287/mksc.19.1.4.15178`. 2.1

Juniorprofessor Dr Oliver Hinz and Dipl-Kfm Jochen Eckert. The impact of search and recommendation systems on sales in electronic commerce. *Business & Information Systems Engineering*, 2(2):67–77, 2010. 3.1.1

Donna L Hoffman, Thomas P Novak, and Marcos Peralta. Building consumer trust online. *Communications of the ACM*, 42(4):80–85, 1999. 4.1

Kartik Hosanagar, Daniel M. Fleder, Dokyun Lee, and Andreas Buja. Will the global village fracture into tribes? recommender systems and their effects on consumer fragmentation. *Manage. Sci.*, 60(4):805–823, April 2014. ISSN 0025-1909. doi: 10.1287/mnsc.2013.1808. URL `http://dx.doi.org/10.1287/mnsc.2013.1808`. 2.1, 3.1.1

L. Iaquinta, M. d. Gemmis, P. Lops, G. Semeraro, M. Filannino, and P. Molino. Introducing serendipity in a content-based recommender system. In *Hybrid Intelligent Systems, 2008. HIS '08. Eighth International Conference on*, pages 168–173, Sept 2008a. doi: 10.1109/HIS.2008.25. 1.1

Leo Iaquinta, Marco De Gemmis, Pasquale Lops, Giovanni Semeraro, Michele Filannino, and Piero Molino. Introducing serendipity in a content-based recommender system. In *Hybrid Intelligent Systems, 2008. HIS'08. Eighth International Conference on*, pages 168–173. IEEE, 2008b. 3.1.4

Daniel J. Gifford and Robert T. Kudrle. The law and economics of price discrimination in modern economies: Time for reconciliation? *UC Davis Law Review*, 43(4):1235, 4 2010. 1.2, 3.1.5

JoVE. Eye tracking in cognitive experiments. `http://www.jove.com/science-education/5421/eye-tracking-in-cognitive-experiments`, 2016. 1.2

Paul Klemperer. The competitiveness of markets with switching costs. *The RAND Journal of Economics*, pages 138–150, 1987. 3.3.5

Krista Jiaoyang Li and Sanjay Jain. Behavior-based pricing: An analysis of the impact of peer-induced fairness. *Management Science*, 3 2015. URL `http://dx.doi.org/`. 3.1.2

Lusi Li, Jianqing Chen, and Srinivasan Raghunathan. Recommender system rethink: Implications for an electronic marketplace with competing manufacturers. *WISE 2014*, 2014. 3.1.4

Greg Linden, Brent Smith, and Jeremy York. Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1):76–80, January 2003. ISSN 1089-

7801. doi: 10.1109/MIC.2003.1167344. URL http://dx.doi.org/10.1109/MIC.2003.1167344. 1.1

ListrakNews. Survey reveals 80% of email readers find it useful when emails feature recommended products based on past purchases. 2014. URL https://www.listrak.com/about/news-events/Press-Release-Survey-Reveals-80-Percent-of-Email-Readers-Find-It-Useful-
3.1, 3.1.2

Jayson L Lusk. Effects of cheap talk on consumer willingness-to-pay for golden rice. *American Journal of Agricultural Economics*, 85(4):840–856, 2003. 3.1.3

Jayson L Lusk, M Scott Daniel, Darrell R Mark, and Christine L Lusk. Alternative calibration and auction institutions for predicting consumer willingness to pay for nongenetically modified corn chips. *Journal of Agricultural and Resource Economics*, pages 40–57, 2001. 3.1.3

JP Mangalindan. Amazon's recommendation secret. 2012a. URL http://fortune.com/2012/07/30/amazons-recommendation-secret/. 3.1

JP Mangalindan. "amazon's recommendation secret", 2012b. URL http://fortune.com/2012/07/30/amazons-recommendation-secret/. 1.1

Mashable. Study will make you question prices on travel websites. 2014. URL http://mashable.com/2014/10/27/online-price-discrimination/#I6nxZEf8Esqb. 3.1.2

Christian Matt, Thomas Hess, and Christian Weiß. The differences between recommender technologies in their impact on sales diversity. 2013. 3.1.1

Sean M. McNee, John Riedl, and Joseph A. Konstan. Being accurate is not enough: How accuracy metrics have hurt recommender systems. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '06, pages 1097–1101, New York, NY, USA, 2006a. ACM. ISBN 1-59593-298-4. doi: 10.1145/1125451.1125659. URL http://doi.acm.

org/10.1145/1125451.1125659. 1.1

Sean M McNee, John Riedl, and Joseph A Konstan. Being accurate is not enough: how accuracy metrics have hurt recommender systems. In *CHI'06 extended abstracts on Human factors in computing systems*, pages 1097–1101. ACM, 2006b. 3.1.4

Klaus M Miller, Reto Hofstetter, Harley Krohmer, and Z John Zhang. How should consumers' willingness to pay be measured? an empirical comparison of state-of-the-art approaches. *Journal of Marketing Research*, 48(1):172–184, 2011. 3.1.3

Phillip Nelson. Advertising as information. *Journal of Political Economy*, 82(4):729–754, 1974. ISSN 00223808, 1537534X. URL http://www.jstor.org/stable/1837143. 1.2

Walter Nicholson and Christopher Snyder. *Microeconomic Theory: Basic Principles and Extensions*. 2005 Thomson South-Western, 10 edition, 2008. ISBN 10: 0-324-42162-1. 2.2.1

Vincent R. Nijs, Kanishka Misra, and Karsten Hansen. Outsourcing retail pricing to a category captain: The role of information firewalls. *Marketing Science*, 33(1):66–81, 2014. doi: 10.1287/mksc.2013.0822. URL http://dx.doi.org/10.1287/mksc.2013.0822. 2.1

Andrew Odlyzko. Privacy, economics, and price discrimination on the internet. In *Proceedings of the 5th international conference on Electronic commerce*, pages 355–366. ACM, 2003. 3.1.2

Gal Oestreicher-Singer and Arun Sundararajan. Recommendation networks and the long tail of electronic commerce. *Available at SSRN 1324064*, 2010. 3.1.1

Louis Phlips. *The economics of price discrimination*. Cambridge University Press, 1983. 3.1.2

Rufus Pollock. Is google the next microsoft: competition, welfare and regulation in online search. *Review of Network Economics*, 9(4), 2010. 1.1

Steven C. Salop. Monopolistic competition with outside goods. *The Bell Journal of Economics*, 10(1):141–156, 1979. ISSN 0361915X. URL http://www.jstor.org/stable/3003323. 3.2.1

J Ben Schafer, Joseph Konstan, and John Riedl. Recommender systems in e-commerce. In *Proceedings of the 1st ACM conference on Electronic commerce*, pages 158–166. ACM, 1999. 1.1

Joan Schneider and Julie Hall. Why most product launches fail? *Harvard Business Review*, 4 2011. URL `https://hbr.org/2011/04/why-most-product-launches-fail`. 4.1

Byron Sharp. Two type of repeat purchase market, with different loyalty patterns. 1 2013. URL `https://byronsharp.wordpress.com/2013/01/30/two-type-of-repeat-purchase-market-with-different-loyalty-patterns/`. 4.1

Bette Ann Stead and Jackie Gilbert. Ethical issues in electronic commerce. *Journal of Business Ethics*, 34(2):75–85, 2001. 1.1

TechWyse. Eye tracking study – web design for conversion. `https://www.techwyse.com/blog/website-conversion/eye-tracking-study-web-design-for-conversion/`, 2012. 1.2

Rajeev K. Tyagi. Technological advances, transaction costs, and consumer welfare. *Marketing Science*, 23(3):335–344, 2004. doi: 10.1287/mksc.1030.0038. URL `http://dx.doi.org/10.1287/mksc.1030.0038`. 2.1

Hal R Varian. Differential pricing and efficiency. *First Monday*, 1(2), 1996. 3.1.2

Tuo Wang, Ramaswamy Venkatesh, and Rabikar Chatterjee. Reservation price as a range: An incentive-compatible measurement approach. *Journal of Marketing Research*, 44(2):200–213, 2007. 3.1.3

Robert D Weaver and AE Luloff. Willingness-to-pay. *Agribusiness*, 8(2):131–142, 1992. 3.1.3

Christo Wilson. ”if you use a mac or an android, e-commerce sites may be charging you more”, 2014. URL `https://www.`

washingtonpost.com/posteverything/wp/2014/11/03/
if-you-use-a-mac-or-an-android-e-commerce-sites-may-be-charging-you-more
1.2

Dazhong Wu, Gautam Ray, Xianjun Geng, and Andrew Whinston. Implications of reduced search cost and free riding in e-commerce. *Marketing Science*, 23(2):255–262, 2004. doi: 10.1287/mksc.1040.0047. URL http://dx.doi.org/10.1287/mksc.1040.0047. 2.1

Yuan Cao Zhang, Diarmuid Ó Séaghdha, Daniele Quercia, and Tamas Jambor. Auralist: Introducing serendipity into music recommendation. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*, WSDM '12, pages 13–22, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-0747-5. doi: 10.1145/2124295.2124300. URL http://doi.acm.org/10.1145/2124295.2124300. 1.1, 3.1.4

Appendices

Appendix A

Technical Proofs

A.1 Proof of Proposition 1

$$
\begin{aligned}
\pi(A1B2) &= m_A p_A^{\beta_1+\beta_2} + m_B p_B^{\beta_1} \\
&= -\frac{(\beta_1+\beta_2)^{\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{1+\beta_1+\beta_2}} c_A^{1+\beta_1+\beta_2} - \frac{\beta_1^{\beta_1}}{(1+\beta_1)^{1+\beta_1}} c_B^{1+\beta_1} \\
CS(A1B2) &= -\frac{p_A^{1+\beta_1+\beta_2}}{1+\beta_1+\beta_2} - \frac{p_B^{1+\beta_1}}{1+\beta_1} \\
&= -\frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}} c_A^{1+\beta_1+\beta_2} - \frac{\beta_1^{1+\beta_1}}{(1+\beta_1)^{2+\beta_1}} c_B^{1+\beta_1}
\end{aligned}
$$

Substituting the optimal price function into the welfare functions, the profit and consumers surplus from ordering $A2B1$ are:

$$\begin{aligned}
\pi(A2B1) &= m_A p_A^{\beta_1} + m_B p_B^{\beta_1+\beta_2} \\
&= -\frac{\beta_1^{\beta_1}}{(1+\beta_1)^{1+\beta_1}} c_A^{1+\beta_1} - \frac{(\beta_1+\beta_2)^{\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{1+\beta_1+\beta_2}} c_B^{1+\beta_1+\beta_2} \\
CS(A2B1) &= -\frac{p_A^{1+\beta_1}}{1+\beta_1} - \frac{p_B^{1+\beta_1+\beta_2}}{1+\beta_1+\beta_2} \\
&= -\frac{\beta_1^{1+\beta_1}}{(1+\beta_1)^{2+\beta_1}} c_A^{1+\beta_1} - \frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}} c_B^{1+\beta_1+\beta_2}
\end{aligned}$$

The optimal ordering to maximize price can be found by calculating the profit difference under different orderings.

$$\begin{aligned}
\pi(A1B2) - \pi(A2B1) =& [-\frac{(\beta_1+\beta_2)^{\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{1+\beta_1+\beta_2}} c_A^{1+\beta_1+\beta_2} - \frac{\beta_1^{\beta_1}}{(1+\beta_1)^{1+\beta_1}} c_B^{1+\beta_1}] \\
& - [-\frac{\beta_1^{\beta_1}}{(1+\beta_1)^{1+\beta_1}} c_A^{1+\beta_1} - \frac{(\beta_1+\beta_2)^{\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{1+\beta_1+\beta_2}} c_B^{1+\beta_1+\beta_2}] \\
=& [-\frac{(\beta_1+\beta_2)^{\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{1+\beta_1+\beta_2}} c_A^{1+\beta_1+\beta_2} + \frac{\beta_1^{\beta_1}}{(1+\beta_1)^{1+\beta_1}} c_A^{1+\beta_1}] \\
& - [-\frac{(\beta_1+\beta_2)^{\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{1+\beta_1+\beta_2}} c_B^{1+\beta_1+\beta_2} + \frac{\beta_1^{\beta_1}}{(1+\beta_1)^{1+\beta_1}} c_B^{1+\beta_1}]
\end{aligned}$$

Define an intermediate function f, such that

$$f(c) = -\frac{(\beta_1+\beta_2)^{\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{1+\beta_1+\beta_2}} c^{1+\beta_1+\beta_2} + \frac{\beta_1^{\beta_1}}{(1+\beta_1)^{1+\beta_1}} c^{1+\beta_1} \tag{A.1}$$

Using $f(c)$ to represent the profit difference under two orderings, $\pi(A1B2) - \pi(A2B1) = f(C_A) - f(C_B)$. Without loss of generality, assume $c_A \le c_B$,

$$\begin{aligned}
f'(c) =& -(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2})^{\beta_1+\beta_2} c^{\beta_1+\beta_2} + (\frac{\beta_1}{1+\beta_1})^{\beta_1} c^{\beta_1} \\
=& -(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2})^{\beta_1+\beta_2} c^{\beta_1} [c^{\beta_2} - (\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})^{\beta_1+\beta_2} (\frac{\beta_1}{1+\beta_1})^{\beta_1}]
\end{aligned}$$

Since $F = (\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})[\frac{\beta_1(1+\beta_1+\beta_2)}{(1+\beta_1)(\beta_1+\beta_2)}]^{\frac{\beta_1}{\beta_2}}$

$$f'(c) = -(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2})^{\beta_1+\beta_2} c^{\beta_1}[c^{\beta_2} - F^{\beta_2}]$$

If $c < F$, then $f'(c) > 0$, if $C_A < C_B$, then $f(C_A) < f(C_B), \pi(A1B2) < \pi(A2B1)$.. So **for profit** firm will put high cost movie in the first place.

A.2 Proof of Proposition 2

$$\begin{aligned}
CS(A1B2) - CS(A2B1) =& [-\frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}} c_A^{1+\beta_1+\beta_2} - \frac{\beta_1^{1+\beta_1}}{(1+\beta_1)^{2+\beta_1}} c_B^{1+\beta_1}] \\
& - [-\frac{\beta_1^{1+\beta_1}}{(1+\beta_1)^{2+\beta_1}} c_A^{1+\beta_1} - \frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}} c_B^{1+\beta_1+\beta_2}] \\
=& [-\frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}} c_A^{1+\beta_1+\beta_2} + \frac{\beta_1^{1+\beta_1}}{(1+\beta_1)^{2+\beta_1}} c_A^{1+\beta_1}] \\
& - [-\frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}} c_B^{1+\beta_1+\beta_2} + \frac{\beta_1^{1+\beta_1}}{(1+\beta_1)^{2+\beta_1}} c_B^{1+\beta_1}]
\end{aligned}$$

Let

$$g(c) = -\frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}} c^{1+\beta_1+\beta_2} + \frac{\beta_1^{1+\beta_1}}{(1+\beta_1)^{2+\beta_1}} c^{1+\beta_1}$$

Then $g(C_A) - g(C_B) = CS(A1B2) - CS(A2B1)$

$$\begin{aligned}
g'(c) =& -(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2})^{1+\beta_1+\beta_2} c^{\beta_1+\beta_2} + (\frac{\beta_1}{1+\beta_1})^{1+\beta_1} c^{\beta_1} \\
=& -(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2})^{1+\beta_1+\beta_2} c^{\beta_1}[c^{\beta_2} - (\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})^{1+\beta_1+\beta_2}(\frac{\beta_1}{1+\beta_1})^{1+\beta_1}]
\end{aligned}$$

Let

$$G^{\beta_2} =(\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})^{1+\beta_1+\beta_2}(\frac{\beta_1}{1+\beta_1})^{1+\beta_1}$$
$$G =(\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})[\frac{\beta_1(1+\beta_1+\beta_2)}{(1+\beta_1)(\beta_1+\beta_2)}]^{\frac{1+\beta_1}{\beta_2}}$$

So if $c < G, g'(c) > 0$, then I have if $C_A < C_B, g(C_A) - g(C_B) = CS(A1B2) - CS(A2B1) < 0$,

which means it's more beneficial for **the consumer** if firm recommends high cost in the first place.

A.3 Proof of Proposition 3

Total social welfare, denoted by TS, is $TS = CS + \pi$.

$$\begin{aligned} TS(A1B2) - TS(A2B1) =& [CS(A1B2) - CS(A2B1)] + [\pi(A1B2) - \pi(A2B1)] \\ =& [g(C_A) - g(C_B)] + [f(C_A) - f(C_B)] \\ =& [g(C_A) + f(C_A)] - [g(C_B) + f(C_B)] \end{aligned}$$

Let $h(c) = g(c) + f(c)$, then

$$
\begin{aligned}
TS(A1B2) - TS(A2B1) =& h(C_A) - h(C_B) \\
h'(c) =& f'(c) + g'(c) \\
=& -(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2})^{\beta_1+\beta_2} c^{\beta_1+\beta_2} + (\frac{\beta_1}{1+\beta_1})^{\beta_1} c^{\beta_1} \\
& -(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2})^{1+\beta_1+\beta_2} c^{\beta_1+\beta_2} + (\frac{\beta_1}{1+\beta_1})^{1+\beta_1} c^{\beta_1} \\
=& -(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2})^{\beta_1+\beta_2} c^{\beta_1} [c^{\beta_2} - (\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2})^{\beta_1+\beta_2} (\frac{\beta_1}{1+\beta_1})^{\beta_1} \\
& + \frac{\beta_1+\beta_2}{1+\beta_1+\beta_2} c^{\beta_2} - (\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})^{\beta_1+\beta_2} (\frac{\beta_1}{1+\beta_1})^{1+\beta_1}] \\
=& -(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2})^{\beta_1+\beta_2} c^{\beta_1} (1 + \frac{\beta_1+\beta_2}{1+\beta_1+\beta_2}) [c^{\beta_2} \\
& -(\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})^{\beta_1+\beta_2} (\frac{\beta_1}{1+\beta_1})^{\beta_1} \frac{(1+\beta_1+\beta_2)(1+2\beta_1)}{(1+\beta_1)(1+2\beta_1+2\beta_2)}]
\end{aligned}
$$

Let $H^{\beta_2} = (\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})^{\beta_1+\beta_2}(\frac{\beta_1}{1+\beta_1})^{\beta_1}\frac{(1+\beta_1+\beta_2)(1+2\beta_1)}{(1+\beta_1)(1+2\beta_1+2\beta_2)}$ So if $c < H, h'(c) > 0$, I then have if $C_A < C_B, h(C_A) < h(C_B), TW(A1B2) < TW(A2B1)$, which means to maximize total social welfare, firm should recommend high cost movie in the first place in this situation.

A.4 Proof of Proposition 4

As $F^{\beta_2} \cdot \frac{\beta_1(1+\beta_1+\beta_2)}{(1+\beta_1)(\beta_1+\beta_2)} = F^{\beta_2} \cdot \frac{\beta_1^2+\beta_1+\beta_1\beta_2}{\beta_1^2+\beta_1+\beta_2+\beta_1\beta_2} = G^{\beta_2}, 0 < \frac{\beta_1^2+\beta_1+\beta_1\beta_2}{\beta_1^2+\beta_1+\beta_2+\beta_1\beta_2} < 1, \beta_2 > 0$, so $F > G$.

$$
\begin{aligned}
F^{\beta_2} =& (\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})^{\beta_1+\beta_2}(\frac{\beta_1}{1+\beta_1})^{\beta_1} \\
G^{\beta_2} =& (\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})^{1+\beta_1+\beta_2}(\frac{\beta_1}{1+\beta_1})^{1+\beta_1} = (\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})^{\beta_1+\beta_2}(\frac{\beta_1}{1+\beta_1})^{\beta_1}\frac{(1+\beta_1+\beta_2)(\beta_1)}{(1+\beta_1)(\beta_1+\beta_2)} \\
H^{\beta_2} =& (\frac{1+\beta_1+\beta_2}{\beta_1+\beta_2})^{\beta_1+\beta_2}(\frac{\beta_1}{1+\beta_1})^{\beta_1}\frac{(1+\beta_1+\beta_2)(1+2\beta_1)}{(1+\beta_1)(1+2\beta_1+2\beta_2)} \\
\frac{H^{\beta_2}}{F^{\beta_2}} =& \frac{(1+\beta_1+\beta_2)(1+2\beta_1)}{(1+\beta_1)(1+2\beta_1+2\beta_2)} = \frac{(1+\beta_1)(1+2\beta_1)+\beta_2(1+2\beta_1)}{(1+\beta_1)(1+2\beta_1)+2\beta_2(1+\beta_1)} < 1 \\
\therefore F >& H \\
\frac{G^{\beta_2}}{H^{\beta_2}} =& \frac{\beta_1(1+2\beta_1+2\beta_2)}{(\beta_1+\beta_2)(1+2\beta_1)} = \frac{\beta_1+2\beta_1^2+2\beta_1\beta_2}{\beta_1+2\beta_1^2+2\beta_1\beta_2+\beta_2} < 1 \\
\therefore G <& H, G < H < F
\end{aligned}
$$

Proof is done.

A.5 Proof for proposition 5

A.5.1 $F - G$ increases in β_2

Since I have

$$
F = \frac{1+\beta_1+\beta_2}{\beta_1+\beta_2}[\frac{\beta_1(1+\beta_1+\beta_2)}{(1+\beta_1)(\beta_1+\beta_2)}]^{\frac{\beta_1}{\beta_2}}
$$

, taking the derivative after taking the log, with respect to β_2, I have

$$
\begin{aligned}
\frac{\partial log(F)}{\partial \beta_2} &= -\frac{\beta_1}{\beta_2^2}log(\frac{\beta_1(1+\beta_1+\beta_2)}{(1+\beta_1)(\beta_1+\beta_2)}) - \frac{1}{\beta_2(1+\beta_1+\beta_2)} \\
&= \frac{\beta_1}{\beta_2^2}[log(1+\frac{\beta_2}{\beta_1(1+\beta_1+\beta_2)}) - \frac{\beta_2}{\beta_1(1+\beta_1+\beta_2)}]
\end{aligned}
$$

Let $x = \frac{\beta_2}{\beta_1(1+\beta_1+\beta_2)}$, I have,

$$\frac{\partial log(F)}{\partial \beta_2} = -\frac{\beta_1}{\beta_2^2}(log(1+x) - x)$$

Since

$$\frac{d[log(1+x)-x]}{dx} = \frac{1}{1+x} - 1 < 0, \ \forall x > 0$$

,

$$log(1+x) - x < log(1+0) - 0 = 0$$

$$\begin{aligned}
&\because \beta_2 > 0, \beta_1 < 0, 1 + \beta_1 + \beta_2 < 0 \\
&\therefore x > 0 \\
&\therefore log(1+x) - x = log(1 + \frac{\beta_2}{\beta_1(1+\beta_1+\beta_2)}) - \frac{\beta_2}{\beta_1(1+\beta_1+\beta_2)} < 0 \\
&\because \frac{\beta_1}{\beta_2^2} < 0, \frac{\partial log(F)}{\partial \beta_2} > 0
\end{aligned}$$

Therefore, I have proved that F increase in β_2.

$$\begin{aligned}
&\because G = F[\frac{\beta_1(1+\beta_1+\beta_2)}{(1+\beta_1)(\beta_1+\beta_2)}]^{\frac{1}{\beta_2}} \\
\therefore \frac{\partial log(G)}{\partial \beta_2} &= \frac{\partial log(G)}{\partial \beta_2} + \frac{1}{\beta_2^2} log[\frac{\beta_2}{\beta_1(1+\beta_1+\beta_2)} + 1] - \frac{1}{\beta_2(\beta_1+\beta_2)(1+\beta_1+\beta_2)} \\
&= (\frac{\beta_1+1}{\beta_2^2})[log(1 + \frac{\beta_2}{\beta_1(1+\beta_1+\beta_2)}) - \frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)}] \\
&= (\frac{\beta_1+1}{\beta_2^2})[-\frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)} - log(1 - \frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)})]
\end{aligned}$$

Define $x = -\frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)}$, then I have

$$\frac{\partial log(G)}{\partial \beta_2} = (\frac{\beta_1+1}{\beta_2^2})[x - log(1+x)]$$

$$\because \beta_2 > 0, 1+\beta_1 < 0, \beta_1+\beta_2 < 0, \therefore x < 0$$

$$1 + x = \frac{\beta_1(1+\beta_1+\beta_2)}{(\beta_1+\beta_2)(1+\beta_1)} > 0, \therefore -1 < x < 0$$

$$\frac{d[x - log(1+x)]}{dx} = 1 - \frac{1}{1+x} < 0,\ \forall -1 < x < 0$$

$$x - log(1+x) > 0 - log(1+0) = 0$$

$$\because \beta_1 + 1 < 0, \beta_2 > 0, \therefore \frac{\partial log(G)}{\partial \beta_2} < 0$$

Therefore I have proved that G decreases in β_2.

A.5.2 $F - H$ increases in β_2

Since I have shown $F > H$. The size of conflict region between total welfare and profit is $F - H$,

$$F - H = F[1 - \frac{H}{F}]$$

If defining M, s.t.

$$\frac{H}{F} = [\frac{(1+2\beta_1)(1+\beta_1+\beta_2)}{(1+2\beta_1+2\beta_2)(1+\beta_1)}]^{\frac{1}{\beta_2}}$$

Since it is already showed that F increases β_2, if $\frac{H}{F}$ decreases in β_2, then the size of region $F-H$ also increases in β_2.

$$\begin{aligned}
log(\frac{H}{F}) &= \frac{1}{\beta_2}log[\frac{(1+2\beta_1)(1+\beta_1+\beta_2)}{(1+2\beta_1+2\beta_2)(1+\beta_1)}] \\
\frac{\partial log(\frac{H}{F})}{\partial \beta_2} &= -\frac{1}{\beta_2^2}log[\frac{(1+2\beta_1)(1+\beta_1+\beta_2)}{(1+2\beta_1+2\beta_2)(1+\beta_1)}] \\
&+ \frac{1}{\beta_2}\frac{(1+2\beta_1+2\beta_2)(1+\beta_1)}{(1+2\beta_1)(1+\beta_1+\beta_2)}\frac{\partial}{\partial \beta_2}[\frac{(1+2\beta_1)(1+\beta_1+\beta_2)}{(1+2\beta_1+2\beta_2)(1+\beta_1)}] \\
&= -\frac{1}{\beta_2^2}log[\frac{(1+2\beta_1)(1+\beta_1+\beta_2)}{(1+2\beta_1+2\beta_2)(1+\beta_1)}] - \frac{1}{\beta_2}\frac{1}{(1+2\beta_1+2\beta_2)(1+\beta_1+\beta_2)} \\
&= -\frac{1}{\beta_2^2}\{\frac{\beta_2}{(1+2\beta_1+2\beta_2)(1+\beta_1+\beta_2)} - log[1+\frac{\beta_2}{(1+2\beta_1)(1+\beta_1+\beta_2)}]\}
\end{aligned}$$

let $x = \frac{\beta_2}{(1+2\beta_1)(1+\beta_1+\beta_2)}$, then

$$\begin{aligned}
&\because\ \beta_2 > 0, \beta_1+\beta_2 < -1 \\
&\therefore\ \frac{\beta_2}{(1+2\beta_1)(1+\beta_1+\beta_2)} > 0 \\
&\because\ \forall x > 0, x - log(1+x) > 0 \\
&\therefore\ -log[1+\frac{\beta_2}{(1+2\beta_1)(1+\beta_1+\beta_2)}] > -\frac{\beta_2}{(1+2\beta_1)(1+\beta_1+\beta_2)} \\
&\because\ \beta_2 > 0 \\
&\therefore\ \frac{\partial log(\frac{H}{F})}{\partial \beta_2} < -\frac{1}{\beta_2^2}\{\frac{\beta_2}{(1+2\beta_1+2\beta_2)(1+\beta_1+\beta_2)} - \frac{\beta_2}{(1+2\beta_1)(1+\beta_1+\beta_2)}\} \\
&\because\ (1+2\beta_1+2\beta_2)(1+\beta_1+\beta_2) < (1+2\beta_1)(1+\beta_1+\beta_2) \\
&\therefore\ \frac{\beta_2}{(1+2\beta_1+2\beta_2)(1+\beta_1+\beta_2)} - \frac{\beta_2}{(1+2\beta_1)(1+\beta_1+\beta_2)} > 0 \\
&\therefore\ -\frac{1}{\beta_2^2}\{\frac{\beta_2}{(1+2\beta_1+2\beta_2)(1+\beta_1+\beta_2)} - \frac{\beta_2}{(1+2\beta_1)(1+\beta_1+\beta_2)}\} < 0 \\
&\therefore\ \frac{\partial log(\frac{H}{F})}{\partial \beta_2} < 0
\end{aligned}$$

So I conclude that $\frac{H}{F}$ decreases in β_2, so $F - H$ increase in β_2, the size of conflict region is larger for larger order effects.

A.6 Proof for proposition 6

A.6.1 $CS(A2B1) - CS(A1B0)$ increases in β_2 when $G < c_A < c_B < F$

When $G < c_B < c_A < F$, the firm adopts A1B0 to maximize profit, but consumers endure potential loss in consumer surplus (CS). The magnitude of CS loss is: $\Delta CS = CS(A2B1) - CS(A1B2)$ Let

$$K = -\frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}}[c_B^{1+\beta_1+\beta_2} - c_A^{1+\beta_1+\beta_2}] \tag{A.2}$$

Taking derivatives wrt β_2, I have:

$$\begin{aligned}
\frac{\partial \Delta CS}{\partial \beta_2} =& \frac{\partial K}{\partial \beta_2} \\
=& -\frac{\partial}{\partial \beta_2}[\frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}}](c_B^{1+\beta_1+\beta_2} - c_A^{1+\beta_1+\beta_2}) \\
& - \frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}}[log(c_B)c_B^{1+\beta_1+\beta_2} - log(c_A)c_A^{1+\beta_1+\beta_2}]
\end{aligned} \tag{A.3}$$

Let $T = \frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}}$, so

$$\begin{aligned}
\frac{\partial log(-T)}{\partial \beta_2} &= log(-(\beta_1+\beta_2)) + \frac{1+\beta_1+\beta_2}{\beta_1+\beta_2} - log(-(\beta_1+\beta_2)) - \frac{2+\beta_1+\beta_2}{1+\beta_1+\beta_2} \\
&= log(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2}) + \frac{1}{(1+\beta_1+\beta_2)(\beta_1+\beta_2)} \\
\because \frac{\partial log(-T)}{\partial \beta_2} &= \frac{1}{T}\frac{\partial T}{\partial \beta_2} \\
\therefore \frac{\partial T}{\partial \beta_2} &= T\frac{\partial log(-T)}{\partial \beta_2} \\
&= \frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_2+\beta_2}}[log(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2}) + \frac{1}{(1+\beta_1+\beta_2)(\beta_1+\beta_2)}]
\end{aligned} \tag{A.4}$$

Plugging equation A.4 into equation A.3, I have:

$$\frac{\partial \Delta CS}{\partial \beta_2} = -\frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_2+\beta_2}}[log(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2}) + \frac{1}{(1+\beta_1+\beta_2)(\beta_1+\beta_2)}](c_B^{1+\beta_1+\beta_2} - c_A^{1+\beta_1+\beta_2}$$
$$- \frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}}[log(c_B)c_B^{1+\beta_1+\beta_2} - log(c_A)c_A^{1+\beta_1+\beta_2}]$$

Let $S(c) = [log(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2}) + \frac{1}{(1+\beta_1+\beta_2)(\beta_1+\beta_2)} + log(c)]c^{1+\beta_1+\beta_2}$, then I have:

$$\begin{aligned}\frac{\partial \Delta CS}{\partial \beta_2} &= -\frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_1+\beta_2}}\\ &= s(c_B) - s(c_A)\\ s'(c) &= [log(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2}) + \frac{1}{(1+\beta_1+\beta_2)(\beta_1+\beta_2)} + log(c) + \frac{1}{1+\beta_1+\beta_2}](1+\beta_1+\beta_2)c^{\beta_1+\beta_2}\\ &= [log(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2}) + \frac{1}{\beta_1+\beta_2} + log(c)](1+\beta_1+\beta_2)c^{\beta_1+\beta_2} \qquad \text{(A.5)}\end{aligned}$$

Since the conflict region is $[G, F]$, i.e. $G < c_B < c_A$, here I have

$$G = \frac{1+\beta_1+\beta_2}{\beta_1+\beta_2}[\frac{\beta_1(1+\beta_1+\beta_2)}{(1+\beta_1)(\beta_1+\beta_2)}]^{\frac{1+\beta_1}{\beta_2}} < c \qquad \text{(A.6)}$$

Taking the log on the two side of equation A.6

$$\begin{aligned} log(c) &> -log(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2}) + \frac{1+\beta_1}{\beta_2}log(1 - \frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)})\\ log(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2}) + log(c) + \frac{1}{\beta_1+\beta_2} &> \frac{1+\beta_1}{\beta_2}log(1 - \frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)}) + \frac{1}{\beta_1+\beta_2}\\ &> \frac{1+\beta_1}{\beta_2}[log(1 - \frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)}) + \frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)}]\end{aligned}$$

Since $log(1-x) + x < 0, \forall x > 0$, and $\frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)} > 0$, therefore

$$log(1 - \frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)}) + \frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)} < 0$$

Since $\frac{1+\beta_1}{\beta_2} < 0$,

$$\frac{1+\beta_1}{\beta_2}[log(1 - \frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)}) + \frac{\beta_2}{(1+\beta_1)(\beta_1+\beta_2)}] > 0$$

So

$$log(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2}) + log(c) + \frac{1}{\beta_1+\beta_2} > 0 \tag{A.7}$$

Plugging in equation A.7 into equation A.5,

$$\begin{aligned}
&\because\ \beta_1+\beta_2 < \beta_1+\beta_2+1 < 0\\
&\therefore\ \frac{(\beta_1+\beta_2)^{1+\beta_1+\beta_2}}{(1+\beta_1+\beta_2)^{2+\beta_2+\beta_2}} < 0, log(\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2}) > 0, \frac{1}{(1+\beta_1+\beta_2)(\beta_1+\beta_2)} > 0\\
&\quad (\frac{\beta_1+\beta_2}{1+\beta_1+\beta_2})^{1+\beta_1+\beta_2} > 0\\
&\because\ c_B < c_A\\
&\therefore\ c_B^{1+\beta_1+\beta_2} - c_A^{1+\beta_1+\beta_2} > 0,\ c_B^{\beta_1+\beta_2} - c_A^{\beta_1+\beta_2} > 0\\
&\therefore\ \frac{\partial \Delta CS}{\partial \beta_2} < 0
\end{aligned}$$