Essays on consumer online search and digital content consumption

Shuo Zhang
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Part of the Advertising and Promotion Management Commons, Economics Commons, and the Marketing Commons

Recommended Citation
Zhang, Shuo, "Essays on consumer online search and digital content consumption" (2020). Arts & Sciences Electronic Theses and Dissertations. 2261.
https://openscholarship.wustl.edu/art_sci_etds/2261

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.
Essays on Consumer Online Search and Digital Content Consumption
by
Shuo Zhang

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2020
St. Louis, Missouri
© 2020, Shuo Zhang
# Table of Contents

List of Figures......................................................................................................................iv  
List of Tables .........................................................................................................................v  
Acknowledgements .............................................................................................................vi  
Abstract of the dissertation .................................................................................................viii  

Chapter 1: Online Shopping with Endogenous PC and Mobile Channel Choice .............1  
1.1 Introduction ..................................................................................................................1  
1.2 Literature Review .......................................................................................................5  
1.3 Data .............................................................................................................................7  
  1.3.1 Channel Choice .....................................................................................................10  
  1.3.2 Potential Explanations for the Conversion-Rate Difference ................................12  
1.4 Model ..........................................................................................................................14  
  1.4.1 Consumer Utility and Search .............................................................................15  
  1.4.2 Consumer Channel Choice .................................................................................18  
1.5 Model Estimation and Identification ...........................................................................21  
  1.5.1 Estimation Procedure ..........................................................................................21  
  1.5.2 Identification .......................................................................................................24  
1.6 Results .........................................................................................................................27  
1.7 Counterfactual .............................................................................................................33  
  1.7.1 The Optimal Pricing Policy on Two Channels ....................................................33  
  1.7.2 Optimal Retargeting Strategy for Sellers ............................................................38  
1.8 Conclusions and Limitations .......................................................................................40  
References ............................................................................................................................43  

Chapter 2: Time-inconsistent Preferences and Strategic Self-Control in Digital Content Consumption ........................................................................................................47  
2.1 Introduction ..................................................................................................................47  
2.2 Literature ....................................................................................................................51  
2.3 Data .............................................................................................................................54
2.3.1 Consumption Behaviors in Data .................................................................56
2.3.2 Alternative Explanations ........................................................................61
2.4 The model and its Estimation ......................................................................66
   2.4.1 An Analytical Model and Its Equilibrium ..............................................66
   2.4.2 An Econometric Model .........................................................................74
   2.4.3 Model Estimation ..................................................................................78
   2.4.4 Identification .......................................................................................82
2.5 Results .........................................................................................................83
   2.5.1 Estimation Results ................................................................................84
   2.5.2 Model Fit and Alternative Model Specifications ......................................89
   2.5.3 Counterfactuals ....................................................................................93
2.6 Conclusion ....................................................................................................98
References .........................................................................................................99
Appendices .......................................................................................................104
Appendix A. Additional Proofs for Chapter 2 ....................................................104
   A1. Proof of Proposition 1 in Chapter 2 ........................................................104
   A2. Proof of Proposition 2 in Chapter 2 ........................................................113
Appendix B. Additional Estimation Results for Chapter 2 ...............................115
   B1. Estimation results (time-consistent consumers) ......................................115
   B2. Estimation results (naive time-inconsistent consumers) ..........................116
List of Figures

Figure 1.1: Proportion of Consumers on Each Channel by Number of Products Searched .......... 9
Figure 1.2: Conversion Rate with Number of Products Searched ............................................. 13
Figure 1.3: Conversion Rate on Mobile and PC for Consumers Who Used Both Channels ...... 14
Figure 1.4: Model Fit by Comparing Actual and Model Simulated Data ................................. 32
Figure 1.5: The Estimated Demand Function on PC and Mobile .................................................. 35
Figure 2.1: Number of Chapters Read after Switching to Pay-per-Chapter .............................. 60
Figure 2.2: Dynamics of the Steady States .................................................................................. 73
Figure 2.3: Model Fit: Reading and Plan Choice ........................................................................... 89
Figure 2.4: Model Fit: Overpaying Ratio ....................................................................................... 91
# List of Tables

Table 1.1: Variable Description and Summary Statistics .......................................................... 9
Table 1.2: Channel Choice with Consumer Characteristics ....................................................... 11
Table 1.3: Results from Monte Carlo Simulation ......................................................................... 27
Table 1.4: Estimation Results ....................................................................................................... 28
Table 2.1 Reading Amount........................................................................................................... 56
Table 2.2: Probability of Switching Plans ................................................................................... 59
Table 2.3: Switching Probability after Overpay .......................................................................... 63
Table 2.4: Main Estimation Results ............................................................................................. 84
Table 2.5: Counterfactual Results ............................................................................................... 94
Acknowledgements

Over the past six years, I have received immense support and encouragement from a great number of individuals. I am especially indebted to my advisor, Tat Y. Chan, who not only have guided and encouraged me through the hardest days in my research and job-seeking activities, but also, with his dedication and enthusiasm, set an example for me to follow both as a research and as a teacher. I am grateful to each of my committee members. Additionally, George-Levi Gayle, Berliant Macus, Werner Ploberger and Seethu Seetharaman all have provided insightful and enlightening feedbacks to my research. Their valuable advises have helped me improve this dissertation dramatically.

Last but not the least, I would like to give special thanks to my parents who have always been supportive to me through this lengthy journey, whether they are nearby or far away.

Shuo Zhang

Washington University in St. Louis

May 2020
Dedicated to my beloved parents.
ABSTRACT OF THE DISSERTATION
Essays on Consumer Online Search and Digital Content Consumption

by

Shuo Zhang

Doctor of Philosophy in Economics

Washington University in St. Louis, 2020

Professor Tat Y. Chan, Co-Chair

Associate Professor George-Levi Gayle, Co-Chair

In my dissertation, I apply empirical quantitative methods to marketing research and investigate consumer online search and purchase patterns, as well as digital consumption behaviors and the potential implication for marketing managers. This dissertation consists of two chapters. Chapter 1 studies the consumer shopping channel choice when they search and shop for products online. Mobile phones have emerged as a major channel for online shopping as an alternative to PCs. Despite more consumers using mobile phones, the conversion rate on the mobile channel is lower than that on the PC channel. In this study, we propose a structural consumer search-and-purchase model that endogenizes the channel choice to explain the observed data pattern. Results suggest starting a search session using mobile phones is less costly, but intensive search is costlier. Consequently, mobile phones attract consumers who tend to have lower overall purchase interests and will search less. Based on the results, we use counterfactuals to explore how online retailers can customize their marketing strategies for consumers on the two channels. We find the optimal price on mobile is 2.7% lower than on PC. When sellers retarget non-purchasers by offering channel-specific coupons, the optimal coupon value is 6% higher for consumers on mobile than on PC. Sellers’ profit increase will be 5.1% higher when the retargeting coupons are channel specific.
Chapter 2 examines consumers’ time-inconsistent preferences in digital content consumption and their strategic self-control behaviors. We use a unique dataset obtained from a major digital book platform in China, where consumers can pay either by chapters or by monthly subscription. One-third of consumers consistently choose to pay by chapters, even though monthly subscription would significantly reduce the monetary cost. We propose a dynamic structural model that incorporates time-inconsistent preferences and strategic self-control behaviors to rationalize overpaying behavior. We first analytically demonstrate the existence of a unique equilibrium, and show how, under steady states, overpaying for reading may be optimal for consumers. We then estimate the model from the data. Results show that there is a large segment of consumers who are highly price-sensitive. They are also willing to overpay to curb future consumption. Our counterfactuals show that eliminating the pay-per-chapter plan would hurt consumer welfare and the platform’s profit. Eliminating the monthly subscription plan, however, would increase the platform’s profit but reduce consumer welfare. We introduce a novel nonlinear pricing plan with volume surcharge and illustrate how it can simultaneously improve the platform’s profit and consumer welfare.
Chapter 1

Online Shopping with Endogenous PC and Mobile Channel Choice
Co-authored with Zhengling Jiang and Hai Che

1.1 Introduction
In recent years, the online retail industry has seen a rapid increase in traffic from mobile devices compared to traditional PCs, including desktops and laptops. In the US, the average time adults spend using mobile devices to shop has surpassed that using PC since 2015.¹ Knowing the popularity of online shopping by smartphones, most major US retailers have been aggressively increasing their investment in both mobile application development and advertisement.²

Despite the more intensive usage of smartphones, consumers make fewer purchases from mobile devices than from PCs. A report from Business Insider Intelligence shows that although almost 60% of the time is allocated to the mobile device, only 15% of the total sales are generated from this channel.³ Such disproportionally low sales on mobile is consistent with the conversion-rate gap between the two channels. Based on data collected from over 1.9 billion shopping sessions in the US from 2015 Q4 to 2016 Q4, the conversion rate on PC is consistently much higher than

that on mobile (e.g., the average conversion rate is 4.14% on PC and 1.55% on mobile in 2016 Q4).\(^4\)

The systematic differences in browsing and purchase behaviors between PC and mobile channels offer online retailers an opportunity to differentiate and target consumers on the two channels. Traditional multi-channel retailers with online and offline channels have been engaging in channel-based price differentiation (Wolk and Ebling 2010, Cavallo 2017). With the emerging mobile channel, some companies have offered lower prices for mobile users. For example, anecdotal evidence shows Kayak and Orbitz quote lower hotel prices for mobile users than for PC users.\(^5\) Other companies do the opposite. Hannak et al. (2014) document that Home Depot provides more expensive products for mobile users than for desktop users. Many other companies do not engage in differential product offerings on the two channels. Clearly, what pricing strategy is more profitable depends on how consumers on the two channels differ from each other.

This paper has two main objectives. The first is to study how and, more importantly, why consumer search and purchase behaviors on PC and mobile channels differ. To achieve this goal, we develop a structural consumer search model with endogenous channel choice. The proposed model can explain how different types of consumers choose the shopping channel depending on the benefits and costs of using each channel. By modeling the consumer’s channel choice, our model rationalizes the intriguing data pattern of a higher usage rate but a significantly lower conversion rate on mobile. Estimation results from our model can help firms predict which segments of consumers would shop on mobile and PC channels, which enables us to achieve the

---


\(^5\) https://www.bostonglobe.com/business/2014/10/22/online-shopping-yields-different-prices-results-says-northeastern-study/ZbSVnoBxPjtA8STeWbpQ9H/story.html
second objective of the paper, which is to design channel-specific marketing strategies targeting consumers on the two channels. Without the structural model, whether and how prices should differ on the two channels is not ex-ante clear.

We estimate the proposed model using a unique clickstream dataset from both PC and mobile channels from Taobao, the largest online shopping platform in China. Consumers can use PCs or smartphones to browse and make purchases. The data set contains information on which channel consumers use to browse and purchase. We observe each consumer’s search activities (through browsing different product options) and purchase decisions. We also collect some additional information, such as consumer demographics and their smartphone attributes that may influence consumer channel choice.

Based on the data, we find (1) a higher proportion of consumer usage, (2) a smaller number of searches per customer, and (3) a lower conversion rate for the mobile channel than for the PC channel, consistent with the industry reports of the US market. Even after controlling for the difference in the number of searches on the two channels, the gap in the conversion rate remains unchanged. Estimation results show that, on average, the marginal search cost for an additional search is 1.55 (or US$0.23) higher on mobile than on PC. The average initial fixed search cost for starting a search session, however, is 1.66 (or US$0.25) higher on PC than on mobile. How does this difference influence consumers’ channel choice and conversion rate on each channel? When deciding which channel to shop, consumers consider the search-cost differences and choose the channel that maximizes the expected utility after search. Given the lower marginal search cost on PC, consumers who want to conduct more extensive search are more likely to choose PC over

---

mobile. Because consumers with higher overall valuation are willing to search more, they are more likely to self-select into using the PC channel. Consumers with a lower valuation of the category are more likely to conduct fewer searches and choose the mobile channel due to a lower initial fixed cost. This mechanism of consumer self-selection in our model therefore explains the observed conversion-rate gap between the two channels. We present evidence in the paper that several other alternative explanations, including the difference in transaction costs, cannot explain this difference.

The estimation results also show the heterogeneity in search costs and channel choices across different types of consumers. For example, consumers with more prior purchases and a longer registration history on the platform are associated with a lower fixed search cost on PCs, likely because these consumers were more accustomed to shopping from PCs before the mobile phones became popular. In terms of demographics, younger consumers and women are more likely to choose the mobile channel. Different types of smartphones influence the marginal search cost on mobile. We find that smartphones with a higher screen resolution (typically associated with a larger screen size) and better operating systems are associated with a lower marginal search cost, which increases the likelihood of using the mobile channel.

To guide how sellers can better target consumers on the two channels, we conduct counterfactual analyses. We first investigate the optimal strategy if sellers set different prices on PC versus mobile channels. Optimal prices can be different because consumers drawn to shopping on the two channels are systematically different. Our proposed model accounts for both channel choice and search activity. We find the optimal price on mobile is 2.7% lower than on PC, because consumers on the PC channel tend to have higher overall valuation due to the self-selection in channel choice. Next, we investigate the retargeting strategy by providing a coupon for consumers
who browsed but did not purchase. When sellers utilize the information of consumer channel choice, results suggest the optimal coupon value is about 6% higher for consumers on mobile than on PC. Although this analysis focuses on non-purchasers, the result is consistent with a lower optimal price on mobile suggested by the first counterfactual. Overall, sellers’ profit increase is 5.1% higher when the retargeting strategy is channel specific than when it does not differentiate channels. The counterfactual results illustrate the importance of considering consumers’ channel choice when planning marketing activities.

The rest of the paper is organized as follows. We discuss related literature in section 1.2 and present the data in section 1.3. We develop the model in section 1.4, followed by the estimation strategy and model identification in section 1.5. The estimation results are discussed in section 1.6. Section 1.7 presents the counterfactual regarding optimal channel-specific pricing and retargeting strategies. We conclude the paper and suggest future research in section 1.8.

### 1.2 Literature Review
Our paper is related to the multi-channel retailing literature. It has always been of interest for marketers to understand how to manage customers in a multi-channel environment. In the existing literature, researchers are primarily concerned about issues related to online shopping websites, physical stores, and catalogs (e.g., Neslin et al. 2006, Verhoef et al. 2007, Ansari et al. 2008, Neslin and Shankar 2009, Venkatesan et al. 2007, Wang and Goldfarb 2017, Forman et al. 2009). One of the questions of interest in this line of research is to understand the behavioral difference for consumers who use different channels. Hitt and Frei (2002) document the difference in consumer characteristics and behavior with PC and traditional banking. Degeratu et al. (2000) find that online and physical store environments can affect consumer choices in different ways. Our paper investigates the difference in behavioral patterns (e.g., the intensity of search, conversion rate, etc.)
for consumers who use smartphones or PCs to shop, which is a relatively new and increasingly important multi-channel context. Different from de Hann et al. (2018), who focus on the conversion rate for consumers who switch devices between mobile and PC, we explain the conversion-rate difference for consumers who choose either channel. By treating channel as an endogenous choice in our model, we can not only explain the observed behavioral difference on mobile and PC channels, but can also provide guidance on how sellers can offer *channel-specific* pricing and promotional strategies to increase profit.

This paper is also related to the growing literature about consumers using mobile devices. Existing research has studied how consumers respond to firms’ mobile marketing activities (Shankar and Balasubramanian 2009, Andrews et al. 2016), the impact of the mobile channel on consumer purchase (Einav et al. 2014, Wang et al. 2015, Xu et al. 2016) and news consumption (Xu et al. 2014), content generation and usage (Ghose and Han, 2011), and consumer search behaviors (Daurer et al. 2016). Using data from eBay, Einav et al. (2014) document that the mobile channel is more often used for strictly browsing, leading to a lower conversion rate than on PC. They also find the mobile channel is more often used for common products instead of idiosyncratic items that require more careful inspection. Ghose et al. (2012) find the search cost is higher on mobile than on PC, although local activities (distance) matter more. They do not explicitly model how consumers choose between the two channels. Different from the existing literature on the mobile channel, our paper studies the consumer channel choice using a structural model. Furthermore, our paper documents how channel choice differs across consumers with different demographics, purchase history, and mobile-device attributes.

Finally, the paper is related to the literature of consumer search. Because information gathering is costly (i.e., requiring time and effort), consumers cannot review all possible options
when making a purchase. Recent empirical studies have estimated consumer search models to describe how consumers make search and purchase decisions (e.g., Kim et al. 2010, Koulayev 2014, Honka 2014, Chen and Yao 2016, Kim et al. 2016, Honka and Chintagunta 2016). Understanding consumer search is important for firms when making marketing decisions, such as pricing (e.g., Hong and Shum 2006, Wildenbeest 2011, Zhang et al. 2018). Most of the existing literature considers consumer search behavior on one channel, which is likely driven by the availability of browsing data only from one channel (e.g., Chen and Yao [2016] and Ursu [2018] study consumer search behaviors using online browsing data). Honka (2014) considers different channels by allowing the search cost to differ when obtaining an insurance quote through the insurer website, online quote service, or call center. In this paper, we obtain consumers’ browsing and purchase data as well as which channel, PC or mobile, consumers use. Our search model endogenizes consumers’ channel choice, which allows us to study the optimal channel-specific pricing and promotional strategies. A recent working paper by Jiang et al. (2019) uses a consumer search model to explore the effectiveness of retargeting strategies. We also study how to improve the effectiveness of retargeting strategies in one of the counterfactuals; however, our focus is on channel-specific strategies.

1.3 Data
Our dataset comes from Taobao, which is the largest online shopping platform in China and is owned by Alibaba. Taobao has both mobile and PC channels for consumers to browse and make purchases. The product offerings and their attributes, including prices, are the same on the two channels. From the dataset, we observe detailed individual-level browsing history and purchase decisions and, more importantly, through which channel, mobile or PC, a browsing activity happens. The dataset also contains additional consumer characteristics including demographic
information, smartphone attributes (even for those who did not use the mobile channel to make purchases in our data), and prior shopping history on the platform. We collect data for consumers who had browsed the fishing pole category. We observe search and purchases of 133,896 unique consumers during the data-observation period from October 15, 2014, to November 15, 2014. Among those consumers, 51% had browsed at least one product option from the mobile channel and 49% from the PC channel. Moreover, only 6% of them had used both PC and mobile channels during the one-month data-observation period. Most purchasers (99.2%) bought only one product during the sample period. Thus, we assume consumers have a unit demand in the model.

The data show the browsing and purchase patterns are very different on mobile versus PC. First, the conversion rate, defined as the percentage of consumers who made a purchase out of those who browsed, is significantly lower on mobile (9.93%) than on PC (13.59%). Second, the search intensity, defined as the number of unique products browsed, is higher on PC than on mobile: 58% of consumers browse one product on PC, compared to 65% on mobile, and 28% of consumers browse at least three products on PC, compared to 20% on mobile. Figure 1.1 graphically compares the proportion of consumers shopping on the two channels conditional on the number of searches. More consumers choose the mobile channel if they only search one option; however, for those who search three options or more, the proportion who shop via PC is significantly larger.
Figure 1.1 Proportion of Consumers on Each Channel by Number of Products Searched

Table 1.1 reports the average and the standard deviation of prices and number of searched options. We observe consumer demographics, gender and age, for 65% of the sample. We also collect consumers’ smartphone-device information including the model, screen size, and the phone’s operating system, for 82% of the sample. The rest of Table 1.1 reports the variable descriptions and summary statistics for consumer demographics and mobile-device characteristics.

Table 1.1 Variable Description and Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>Unit price for fishing poles</td>
<td>263.7</td>
<td>63.69</td>
</tr>
<tr>
<td>Search times</td>
<td>Number of products browsed by consumers</td>
<td>1.89</td>
<td>1.27</td>
</tr>
<tr>
<td>Buyer rating</td>
<td>Based on buyer’s prior purchase history</td>
<td>3.8</td>
<td>1.96</td>
</tr>
<tr>
<td>Buyer rating missing</td>
<td>Indicator variable; equals 1 if buyer rating is missing</td>
<td>0.005</td>
<td>–</td>
</tr>
<tr>
<td>Buyer spending</td>
<td>Buyer total spending in ¥ before data observation period</td>
<td>183.2</td>
<td>575.81</td>
</tr>
<tr>
<td></td>
<td>Description</td>
<td>Value 1</td>
<td>Value 2</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Buyer history</td>
<td>Number of days passed since the buyer registered on the website</td>
<td>1099</td>
<td>831.47</td>
</tr>
<tr>
<td>Screen resolution (length)</td>
<td>Smartphone screen resolution in pixels (width)</td>
<td>1184</td>
<td>392.86</td>
</tr>
<tr>
<td>Screen resolution (width)</td>
<td>Smartphone screen resolution in pixels (height)</td>
<td>782.3</td>
<td>299.42</td>
</tr>
<tr>
<td>IOS</td>
<td>Indicator variable; equals 1 for IOS operating system</td>
<td>0.34</td>
<td>–</td>
</tr>
<tr>
<td>Android</td>
<td>Indicator variable; equals 1 for Android operating system</td>
<td>0.15</td>
<td>–</td>
</tr>
<tr>
<td>Mobile browsing</td>
<td>Total number of products browsed on a smartphone before data observation period</td>
<td>173.9</td>
<td>295.90</td>
</tr>
<tr>
<td>Male</td>
<td>Indicator variable; equals 1 for male</td>
<td>0.56</td>
<td>–</td>
</tr>
<tr>
<td>Age</td>
<td>Buyer’s age</td>
<td>30.6</td>
<td>8.47</td>
</tr>
<tr>
<td>Male missing</td>
<td>Indicator variable; equals 1 if gender information is missing</td>
<td>0.09</td>
<td>–</td>
</tr>
<tr>
<td>Age missing</td>
<td>Indicator variable; equals 1 if age information is missing</td>
<td>0.13</td>
<td>–</td>
</tr>
<tr>
<td>Mobile missing</td>
<td>Indicator variable; equals 1 if there is no smartphone information</td>
<td>0.34</td>
<td>–</td>
</tr>
</tbody>
</table>

1.3.1 Channel Choice
The prices for fishing poles did not change over time during our sample observation period. Other product attributes are also identical on mobile and PC, and thus do not affect the channel choice. Consumer characteristics, on the other hand, may affect the choice. We use a reduced-form regression to test how consumers who choose to use PC or mobile are systematically different. Using channel choice as the dependent variable, which equals 1 if the consumer chooses PC, and 0 if he chooses mobile, we run a probit regression to study how the channel choice correlates with
various observed consumer characteristics (described in Table 1.1). Results are reported in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-1.12</td>
<td>0.08</td>
<td>***</td>
</tr>
<tr>
<td>Buyer rating</td>
<td>0.12</td>
<td>4.06E-03</td>
<td>***</td>
</tr>
<tr>
<td>Buyer rating missing</td>
<td>-0.32</td>
<td>0.03</td>
<td>***</td>
</tr>
<tr>
<td>Buyer spending</td>
<td>8.38E-05</td>
<td>2.11E-05</td>
<td>***</td>
</tr>
<tr>
<td>Buyer history</td>
<td>7.89E-05</td>
<td>6.66E-06</td>
<td>***</td>
</tr>
<tr>
<td>Screen resolution</td>
<td>-1.12E-07</td>
<td>6.52E-09</td>
<td>***</td>
</tr>
<tr>
<td>IOS</td>
<td>-0.03</td>
<td>0.01</td>
<td>**</td>
</tr>
<tr>
<td>Android</td>
<td>-0.05</td>
<td>0.01</td>
<td>***</td>
</tr>
<tr>
<td>Mobile browsing</td>
<td>-2.03E-03</td>
<td>2.84E-05</td>
<td>***</td>
</tr>
<tr>
<td>Mobile missing</td>
<td>1.82E-03</td>
<td>1.22E-02</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.40</td>
<td>0.08</td>
<td>***</td>
</tr>
<tr>
<td>Age</td>
<td>0.01</td>
<td>5.97E-04</td>
<td>***</td>
</tr>
<tr>
<td>Gender missing</td>
<td>-0.04</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Age missing</td>
<td>0.23</td>
<td>0.08</td>
<td>**</td>
</tr>
</tbody>
</table>

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

Considerable heterogeneity exists among consumers who choose PC or mobile. We find that younger consumers and consumers who use mobile phones with higher screen resolution and

---

7 We multiply the screen resolution in pixels in length and width, and use the demeaned value to represent screen resolution in the model estimation.
more advanced operating systems\textsuperscript{8} are more likely to use the mobile channel. In addition, consumers with a higher buyer rating (based on a higher number of prior purchases) and higher prior spending are more likely to use the PC channel, both of which positively correlate with the consumer’s past experience on Taobao. These consumers are likely more familiar with the PC channel than the mobile channel because Taobao only introduced the mobile channel in 2008.\textsuperscript{9} The reduced-form evidence suggests the observed consumer characteristics significantly correlate with their channel choice. We incorporate these characteristics in the structural model to account for consumer heterogeneity.

1.3.2 Potential Explanations for the Conversion-Rate Difference
The underlying mechanism driving the observed data pattern in our model is that consumers endogenously choose which channel to browse. We identify and estimate both a marginal search cost (for an addition search) as well as an initial fixed cost (for starting a search session) for the two channels. The channel choice depends on the level of overall valuation as well as the cost to search on the two channels. Before describing the full model, we discuss in this subsection several possible explanations for the lower conversion rate on mobile compared to PC to help justify our model setup. Note we assume consumers have a choice between using mobile or PC. The CNNIC (the Chinese administrative agency responsible for Internet affairs) reports that among Internet users, the smartphone penetration is 85.8\%, and desktop and laptop penetrations are 70.8\% and 43.2\% during 2014.\textsuperscript{10} Therefore, it is reasonable to assume that consumers have access to both types of devices.

\textsuperscript{8} Apple and Android operating systems were considered advanced in China during 2014, when many other smartphones used operating systems developed by local manufacturers.

\textsuperscript{9} Source: https://yq.aliyun.com/articles/583335.

The first potential explanation is that the lower conversion rate on mobile is driven by a higher marginal search cost. With a higher marginal search cost, consumers browse fewer options and are less likely to find a good match and make a purchase on mobile. To test the hypothesis that the difference in the marginal search costs is the only cause for the conversion-rate gap, we compare the conversion rates for consumers who browsed the same number of products. Figure 1.2 shows the conversion rate on PC is still consistently higher than that on mobile among consumers who browse the same number of products. Therefore, although the marginal search-cost difference between the two channels can lead to an overall conversion-rate gap, it cannot explain the gap after controlling for the number of products browsed.

![Figure 1.2 Conversion Rate with Number of Products Searched](image)

The second potential explanation is the difference in transaction cost for completing a purchase on mobile versus PC. For example, consumers may have difficulty typing in the shipping address or the payment information when using a smartphone without a keyboard. In that case, consumers might be more likely to abandon the shopping session on mobile without purchase. To test this explanation, we focus on a small group of consumers (6%) who use both channels to browse the products. If the transaction cost is higher on mobile, we would expect a higher
conversion rate on PC among these consumers as well. Figure 2.3 shows that among the consumers who browse both channels, the conversion rates on the two channels are almost the same (12.9% on PC, 12.2% on mobile). The interpretation for the equal conversion rates is that either the transaction cost is the same on both channels or the transaction cost is trivial, so it does not play an important role in determining where to purchase. In reality, once a debit or credit card is linked to the account, consumers on Taobao only need to type in a six-digit password for payment using mobile devices. Therefore, the time and effort required for payment on mobile is not distinctively higher than that on PC.

![Figure 1.3 Conversion Rate on Mobile and PC for Consumers Who Used Both Channels](image)

**1.4 Model**

We propose a consumer search-and-purchase model that incorporates endogenous channel choice. Before starting the search, consumers first choose through which channel (mobile or PC) to browse the products. We assume consumers can only choose one channel, due to the empirical observation that only 6% of consumers ever switch devices in our data. We exclude this small group of consumers in our empirical analysis to keep the model tractable.

Conditional on the channel choice, consumers then decide how many product options to search. In the literature, both simultaneous and sequential search models have been applied to
study consumer search behavior. We do not observe the order of search from data. This data limitation makes estimating a sequential search very difficult. Prior empirical studies (e.g., De Los Santos et al. 2012 and Honka 2013) have tested the two search models and found evidence to support the simultaneous-search model. Therefore, we follow these studies by assuming consumers conduct simultaneous search. We note that if the data on the search order are available, our proposed framework of channel choice can be easily carried through to scenarios where consumers search sequentially.

Finally, given the channel choice and the number of product options to search, consumers will search on the retail platform. After the search, they will decide whether to purchase from the searched options and, if they do, which option they should buy.

1.4.1 Consumer Utility and Search
We first describe consumers’ search and purchase decisions after they have selected a channel to browse. Suppose there are $I$ consumers and $J$ products. The utility of product $j$ for consumer $i$ is specified as

$$u_{ij} = a_i - \lambda \cdot P_j + e_{ij},$$

(1)

where $a_i$ is consumer $i$’s valuation for the product category. We allow $a_i$ to be heterogeneous across consumers with a normal distribution $a_i \sim N(\mu_a, \sigma_a^2)$. $P_j$ is the price of product $j$ and $e_{ij}$ is the individual match value. We assume $e_{ij}$ follows i.i.d. extreme-value type-I distribution across consumers and products. If the consumer decides not to purchase any product after search, he chooses the outside option denoted by $e_{i0}$. The outside option $e_{i0}$ represents consumer $i$’s valuation of purchasing from other websites or purchasing other products. We assume consumers
know their own outside option before conducting the search activities. $e_{t0}$ is assumed to follow i.i.d. extreme-value type-I distribution across consumers.

Denote channel choice as $s_t \in \{1,0\}$, where $s_t = 1$ if consumer $i$ chooses the PC channel, and $s_t = 0$ if choosing mobile. We first describe how consumers decide the number of product options to search, conditional on choosing channel $s$. Before the search, consumer $i$ knows his initial utility level $a_i$. We assume the consumer knows the overall distribution of $p_j$ and $e_{ij}$, but he has no information on $p_j$ and $e_{ij}$ for a specific retailer $j$, which are only revealed if he clicks into the product detail page. Therefore, the expected $u_{ij}$ conditional on purchase for all product options are the same to the consumer before the search, but the overall expected utility from search will be different due to individuals having different levels of $a_i$ and thus different purchase probabilities. The justification of this assumption is that many small sellers are on Taobao, and none of them belong to well-known branded manufacturers. Consumers are unlikely to have a priori information on the quality of any specific seller. Furthermore, each seller sells multiple brands and models of fishing poles; without searching for detailed information on the product page, consumers are unlikely to know anything about the price or other product and service attributes.

Under this assumption, our simultaneous search model focuses on how many product options the consumer chooses to search, denoted by $b_i$. Consumer $i$ incurs a marginal search cost $c_i^s$ for each product he browses. We allow the marginal search cost to vary across the two channels and individuals. Furthermore, as is common in the search literature, our data do not include consumers who do not search at all. Thus, we require that consumers search at least once in the model. A consumer chooses $b_i$ to maximize the expected utility taking account of the search cost.
Following Chade and Smith (2010), the consumer maximizes the following indirect utility by choosing the number of searches:

\[
IU_i(b) = E \left[ \max_{j \in C_{ib}} \{u_{ij}\} \right] - b_i \cdot c_i^S,
\]  

(2)

where \( C_{ib} \) is the set of the searched options (the outside option \( e_{i0} \) is always an element in \( C_{ib} \)).

The probability that consumer \( i \) chooses to search \( b_i \) times is

\[
P_{ib|a,s} = P\{IU_i(b) \geq IU_i(b') | a_i, s_i\}.
\]  

(3)

After the search, consumers make their purchase decisions by comparing the realized utilities among the choice set (knowing the price and individual match value) and the outside option. Consumer \( i' \)'s conditional purchase probability for product \( k \) is

\[
P_{ik|e,b,a,s} = P\{u_{ik} > u_{ik'}, \forall k' \in C_{ib} | e_{ij}, P_j, a_i, s_i\}.
\]  

(4)

In other words, the consumer will choose option \( k \) if the realized utility is larger than any other options \( k' \) in the choice set.

Note that various factors, including the ranking of product options (e.g., Ursu 2018), may affect the final outcome. We do not observe those factors from data. The impact of these factors on the purchase decision is captured by \( e_{ij} \), which is unknown to the consumer when he decides the optimal \( b_i \). Conditional on \( b_i \), these factors may affect which product options the consumer will search, as well as the order of the search. Our model is agnostic about what options are searched and how they are searched. Importantly, these unobserved factors do not affect our main focus on consumer channel choice.
1.4.2 Consumer Channel Choice
Before starting the search process, consumers choose whether to use a smartphone or a PC to shop. We introduce a fixed search cost in addition to the marginal search cost for both channels. Different from the marginal search cost, which depends on how many products a consumer browses, the fixed search cost is a one-time upfront cost to start a search session. The fixed cost can come from the time and effort required to use a PC or a smartphone to initialize the search process, whereas the marginal search cost is associated with the time and effort required to gather information from the product page. Prior literature (Ghose et al. 2012) and the data pattern of a higher number of searches on PCs suggests the marginal search cost on mobile should be higher than that on PCs, likely because of the smaller screen and lack of keyboard on a smartphone. On the other hand, we expect the PC channel to have a higher fixed cost than the mobile channel, because the portability of a smartphone allows consumers to access it from anywhere.\textsuperscript{11}

We allow individual heterogeneity in both the fixed and marginal search costs given the consumer’s demographic information, mobile-device features, and past usage patterns. For example, younger consumers may be more proficient in using their smartphones for online shopping. In addition, smartphones with larger screen sizes or advanced operating systems could make the search process more effortless and thus are associated with a lower marginal search cost. Because consumers choose one of the channels to search, for model identification, the fixed cost of the mobile channel is normalized to 0. We specify the fixed cost of the PC channel as

\[ f_{c_i} = \mu_{fc} + \beta Z_i + v_{ifc}, \tag{5} \]

\textsuperscript{11} We assume consumers have access to both channels. If a consumer cannot access a channel (e.g., cannot use PC to shop while in transit), the model interprets such cases as the consumers having a very high fixed search cost to start a shopping session on PC.
where $\mu_{fc}$ is a constant term, $Z_i$ is a list of relevant consumer characteristics and device attributes, and $\nu_{ife}$ captures the unobservable heterogeneity and is assumed to follow a standard normal distribution. We do not impose the fixed cost on PC to be higher or lower than that on mobile. The estimated parameters determine the sign and magnitude of the fixed cost on PC for different consumers.

Consumers pay a marginal search cost for an additional search. The marginal search cost for consumer $i$ on the PC channel ($s_i = 1$) is

$$c^1_i = \exp(\mu_c + \sigma_c \nu_{ic}),$$

where $\nu_{ic}$ follows a standard normal distribution. The marginal search cost is guaranteed to be positive in this specification (e.g., Hortaçsu and Syverson 2004).

The marginal search for consumer $i$ on the mobile channel ($s_i = 0$) can be systematically different from his marginal search cost on PC. We specify the marginal search cost as

$$c^0_i = c^1_i + sc_0 + \gamma X_i,$$

where $sc_0$ represents the average difference in marginal search cost between mobile and PC, $X_i$ is a list of consumer $i$'s smartphone characteristics and his past mobile shopping experience that may affect his marginal search cost on mobile, and $\gamma$ captures the heterogeneity in marginal search cost with observed characteristics $X_i$. We do not impose the difference in marginal search cost between mobile and PC, $sc_0 + \gamma X_i$, to be negative or positive. The estimated parameters determine the marginal search cost for different consumers.
We assume that before the search, consumer $i$ is aware of the distribution for prices and individual match values. He knows his level of interest in the product category $a_i$ and his outside option $e_{i0}$. He also knows his marginal and fixed search costs for both channels. Based on the information, the consumer forms expectations on the utility for each channel. Let $F^s_{ib}$ be the cumulative distribution function of the expected maximum utility among $b$ products searched by consumer $i$ on channel $s$, and $f^s_{ib}$ is the corresponding pdf function. The calculation of $F^s_{ib}$ is shown in detail in the next section. The consumer’s expected utility for channel $s$ is

$$ECU^s_i = \max_b[F^s_{ib}(e_{i0}) \cdot e_{i0} + \int_{e_{i0}}^{+\infty} f^s_{ib}(u)\, du - f c_i \cdot s_i - b_i \cdot c^s_i].$$  \hspace{1cm} (8)$$

When the maximum utility from the $b$ browsed products is lower than the outside option, the consumer chooses the outside option. Otherwise, he will choose the maximum of the searched options. The consumer chooses the channel that offers a higher expected utility. The channel choice probability thus is

$$P_{is|a} = P(ECU^s_i \geq ECU^{s'}_i | a_i), s' \in \{0,1\}. \hspace{1cm} (9)$$

To summarize, the channel choice depends on their overall valuation, outside option value, and the fixed and marginal search costs on the two channels. The proposed model is able to capture the difference in channel choices among consumers with different observed characteristics by incorporating heterogeneous fixed and marginal search costs. Moreover, it provides a mechanism of how consumers with different product valuation and search costs tend to select certain channel. This endogenous channel choice is key to understanding the observed conversion rate and search patterns between the two channels.
1.5 Model Estimation and Identification

In this section, we lay out detailed model-estimation procedures, present results from a Monte Carlo simulation study, and discuss the model identification.

1.5.1 Estimation Procedure

The likelihood function comprises the three parts of consumer decisions: choosing a channel (channel choice probability $P_{i,s|a}$), searching $b$ product options (optimal search-time probability, $P_{i,b|a,s}$), and purchase decisions (purchase probability $P_{k|e,b,a,s}$). The likelihood function integrates over the distribution of the outside option $e_{i0}$, the individual shock for fixed search cost $v_{ifc}$ and marginal search cost $v_{ic}$, and the valuation of the product category $a_i$:

$$ LL = \sum_{i=1}^I \log \left( \int \int \prod_{s=0}^1 \prod_{b=1}^{b_i} P_{i,k|e,b,a,s} P_{i,b|a,s} P_{i,s|a} dF(e) dF(v_{fc}) dH(v_c) dG(\alpha) \right). $$

(10)

The probability functions in the equation do not have a closed-form solution. We use simulated maximum likelihood to estimate the model by drawing from the corresponding distributions for numerical integration. More specifically, we draw the following variables $Q$ times. Consumer $i$’s match value for product $j$ $e_{ij}^q$ and the outside option $e_{i0}^q$ are drawn independently from extreme-value type-I distribution. The error terms for fixed search cost and marginal search cost, $v_{ic}^q$ and $v_{ifc}^q$, are drawn i.i.d. from a standard normal distribution. Consumers’ utility constant term is parameterized as $a_i = \mu_\alpha + \sigma_\alpha \cdot e_{ia}^q$, where $e_{ia}^q$ is drawn from a standard normal distribution.

We assume consumers know the distribution of prices prior to search, but the actual values are only realized after they browse the product detail pages and pay the corresponding search cost.
Before the main model estimation, we first estimate the price distribution, which determines the benefit from an additional price search. Following prior literature on price-search models (e.g., Hong and Shum 2006, Moraga-González and Wildenbeest 2008, Honka 2014), we assume prices follow an extreme-value type-I distribution and estimate the price-distribution parameters. We use the estimated price-distribution parameters in the model estimation.

Consumers form expectations of the benefit they receive under a specific number of searches. We evaluate the distribution of the benefit consumers receive from drawing the price and individual match value \( b \) times. To calculate the distribution of the expected benefit from search given one set of parameters, we draw from the price and individual match-value distributions \( b \) times, and calculate the expected maximum value as \( V_b = \max \{-\lambda p_1 + e_1, \ldots, -\lambda p_b + e_b\} \). The process is repeated \( Q \) times. We get a \( Q \)-length vector of \( V_b \) for \( b \) number of searches, which represents the distribution of the expected benefit from searching \( b \) times.

To calculate channel choice probability (equation 8), we evaluate the expected utility from choosing channel \( s \) (equation 7). For consumer \( i \), the expected utility from searching \( b_i \) times on channel \( s_i \) is

\[
\overline{ECU}_i^s = \max_b [u_{ib}^s] - f c_i \cdot s_i,
\]

where \( u_{ib}^s \) is the maximum utility from the searched products and the outside option minus the corresponding marginal search cost. To calculate \( u_{ib}^s \) through simulation, we draw \( Q \) times from the distributions for overall product valuation, outside option, and marginal search cost. We calculate the utility with each set of random draws, and \( u_{ib}^s \) is evaluated as the average from the \( Q \) values:
\[
    u_{ib}^{s} = \frac{1}{Q} \sum_{q} \{ I(a_{i}^{q} + V_{b}^{q} > e_{i0}^{q}) \cdot (a_{i}^{q} + V_{b}^{q}) + I(a_{i}^{q} + V_{b}^{q} < e_{i0}^{q}) \cdot e_{i0}^{q} \} - b_{i} \cdot c_{i}^{s,q} \}.
\]

We draw the fixed-search-cost random-error term \( Q \) times to calculate \( f_c \) as specified in equation 5. The expected utility for channel \( s \) \( ECU_{i}^{s} \) is the maximum of \( u_{ib}^{s} \) by selecting the optimal number of searches \( b_{i} \) minus the corresponding fixed search cost.

Consumers choose the channel that gives them higher expected utility \( ECU_{i}^{s} \), \( s \in (0,1) \). The channel-choice probability calculated from the simulations is not a smooth function. Following prior literature (McFadden 1989, Honka 2014), we apply a kernel-smoothing method where the choice probability is represented by a scaled multivariate logistic CDF. The probability of consumer \( i \) choosing channel \( s_{i} \) is

\[
    P_{is} = \frac{1}{1 + \exp (\omega_{1} \cdot (ECU_{i}^{s} - ECU_{i}^{1-s}))},
\]

where \( \omega_{1} \) is a scaling parameter.

Next, we evaluate the probability of searching \( b_{i} \) times. Consumers choose the number of searches by maximizing the expected utility (equation 3). Applying the kernel-smoothing method, the probability of consumer \( i \) choosing to search \( b_{i} \) times conditional on choosing channel \( s_{i} \) is

\[
    P_{ib|s} = \frac{1}{1 + \exp (\omega_{2} \cdot (IU_{i,b} - \max(IU_{i,-b}))},
\]

where \( \omega_{2} \) is a scaling parameter, and \(-b\) denotes search times other than \( b \).

Finally, we evaluate the purchase probability for consumers after they have chosen a channel and have selected the number of products to browse. The prices and individual match
values are realized for options in the consumers’ consideration set $C_{i,b}$ (the $b_i$ products consumer $i$ browses). The probability that consumer $i$ chooses option $k$ from the consideration set $C_{i,b}$ on channel $s_i$ is

$$
P_{ik|c_{i,b}s} = \frac{1}{1 + \exp (-\omega_3 \cdot (u_{ik} - \max(u_{ik}')))} ,
$$

where $k'$ denotes choices other than option $k$, including the outside option $k = 0$ when consumers do not make a purchase, and $\omega_3$ is a scaling parameter.

Combining the three sets of probabilities together, we obtain the overall probability of observing consumer $i$ choosing channel $s_i$, searching $b_i$ times, and choosing option $k$. We evaluate this probability through simulation by drawing the error terms for overall product valuation $\alpha_i$, fixed and marginal search costs $\nu_{ifc}, \nu_{ic}$, individual match value for each product searched $e_{ij}$, and outside option $e_{i0}$ $Q$ times. The overall likelihood considers channel-choice probability $P_{is}^q$, number-of-searches probability $P_{ib|s}^q$, and purchase probability $P_{ik|c_{i,b}s}^q$:

$$
p_{ij} = \frac{1}{Q} \sum_{q} P_{is}^q P_{ib|s}^q P_{ik|c_{i,b}s}^q .
$$

1.5.2 Identification
We discuss the identification of the model parameters. The parameters can be divided into three categories: the marginal search-cost parameters $\{\mu_c, \sigma_c, sc_0, \gamma\}$, the fixed search-cost parameters $\{\mu_{fc}, \beta\}$, and the utility parameters $\{\mu_\alpha, \sigma_\alpha, \lambda\}$.

For the marginal search-cost parameters, we identify the constant term and the standard deviation of the error terms from the distribution of search times on both PC and mobile channels. $sc_0$ captures the average difference in marginal search cost on PC and mobile. It is identified from
the difference in the mean of the number of searches for consumers on the PC and mobile channels. The systematic difference in the number of searches for consumers with different mobile attributes identifies the observed heterogeneity in marginal search cost across consumers on the mobile channel.

The identification of the fixed search cost on the PC channel comes from consumers’ channel choice for browsing. Recall that the fixed search cost on the mobile channel is normalized to 0. The constant in the fixed cost \( \mu_{fc} \) is identified from the proportion of the consumers who choose the PC channel, after accounting for the difference in marginal search cost. If the fixed cost on the PC channel is higher, more consumers will choose the mobile channel. The systematic difference in channel choice among consumers with different demographics, user behaviors, and device features identifies the observed heterogeneity in fixed cost across consumers.

The mean of the product-category valuation \( \mu_\alpha \) is identified from the overall level of the conversion rate after search, and price sensitivity \( \lambda \) is identified from the purchase data. The variation of the overall product valuation among consumers, \( \sigma_\alpha \), leads to the systematic difference in consumers who select a certain channel for browsing. Consumers with a higher level of overall product valuation may systematically choose a channel given its search-cost structure. For example, when the average fixed cost on PC is higher than on mobile and the marginal search cost is lower, consumers with a high value of \( a_t \) will be more likely to choose PC. In general, if \( \sigma_\alpha \) is greater, the average utility difference of consumers who use PC will be greater than for those who
use mobile, which will lead to a larger difference in conversion rates across the two channels. Thus, the value of $\sigma_\omega$ is identified by the systematic conversion-rate gap observed in our data.  

We run a Monte Carlo study to test the model identification. We simulate data for 10,000 consumers and set the maximum number of searches at 5. The simulation procedure is as follows. We draw the error terms of marginal search cost $\nu_{lc}$ and fixed search cost $\nu_{lfc}$ i.i.d. from a standard normal distribution. The outside option $e_{i0}$ is drawn from extreme-value type-I distribution. The expected channel utility and the optimal search times are calculated as in equations 7 and 8. With the chosen channel $s$ and search times $b$, consumers sample $b$ products. After search, consumers see $b$ prices (drawn i.i.d. from the price distribution) and the match values for each product $e_{ij}$ (drawn i.i.d. from extreme-value type-I distribution). Consumer $i$ makes purchase decisions depending on the realized utility.

In the estimation, we set all scaling factors ($\omega_1, \omega_2, \omega_3$) in the kernel-smoothing logit functions to be 20. The number of simulations $Q$ is 50. Results from the Monte Carlo study are reported in Table 1.3. Column (1) shows the true value of the parameters, and columns (2) and (3) show the estimated value and standard error. Thus, the proposed estimation procedure can successfully recover the true parameters.

---

12 When $\sigma_\omega = 0$, the systematic conversion-rate gap between PC and mobile will no longer exist.
<table>
<thead>
<tr>
<th>Variable</th>
<th>True Value (1)</th>
<th>Estimated Value (2)</th>
<th>Standard Error (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Utility parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_a$: Mean of valuation</td>
<td>-45.0</td>
<td>-43.47</td>
<td>0.770</td>
</tr>
<tr>
<td>$\sigma_a$: Std. dev. of valuation</td>
<td>110.0</td>
<td>125.71</td>
<td>13.48</td>
</tr>
<tr>
<td>$\lambda$: Price coefficient</td>
<td>-1.5</td>
<td>-1.45</td>
<td>0.005</td>
</tr>
<tr>
<td><strong>Search-cost parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_c$: Mean of marginal search cost</td>
<td>4.0</td>
<td>4.09</td>
<td>0.006</td>
</tr>
<tr>
<td>$\sigma_c$: Std. dev. of marginal search cost</td>
<td>0.4</td>
<td>0.42</td>
<td>0.01</td>
</tr>
<tr>
<td>$\Delta c_0$: Difference in marginal search cost on PC from mobile</td>
<td>10.0</td>
<td>9.34</td>
<td>0.372</td>
</tr>
<tr>
<td>$\mu_f c$: Fixed search cost on PC (normalized to 0 on mobile)</td>
<td>0.3</td>
<td>0.27</td>
<td>0.007</td>
</tr>
</tbody>
</table>

### 1.6 Results

We apply and estimate the proposed model using the Taobao data. In this section, we report and discuss the model-estimation results. In particular, we highlight how the model of channel choice can explain the lower conversion rate on mobile compared to PC. We show the estimated model can reproduce the conversion rates and the number of searches very well across both channels.

The estimation results are shown in Table 1.4. The estimated parameters are presented in four panels. Starting from the first panel, the price coefficient is negative at -5.16 for ¥1 (or -$0.77 for US$1). We transform the utility parameters into dollar value by dividing the estimated parameters.
by the price coefficient. The mean valuation for the product category $\mu_{\alpha}$ is ¥148 (or US$21.7) and
the standard deviation across consumers $\sigma_{\alpha}$ is ¥53 (or US$7.9).

Table 1.4 Estimation Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimated Value</th>
<th>Standard Error</th>
<th>p-value$^{13}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Utility parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_{\alpha}$: Mean of valuation</td>
<td>763</td>
<td>10.55</td>
<td>***</td>
</tr>
<tr>
<td>$\sigma_{\alpha}$: Std. dev. of valuation</td>
<td>272</td>
<td>5.20</td>
<td>***</td>
</tr>
<tr>
<td>$\lambda$: Price coefficient</td>
<td>-5.16</td>
<td>0.01</td>
<td>***</td>
</tr>
<tr>
<td><strong>Search-cost parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_{c}$: Mean of marginal search cost</td>
<td>5.09</td>
<td>1.35E-03</td>
<td>***</td>
</tr>
<tr>
<td>$\sigma_{c}$: Std. dev. of marginal search cost</td>
<td>0.82</td>
<td>3.91E-04</td>
<td>***</td>
</tr>
<tr>
<td>$sc_{0}$: Difference in marginal search cost on PC from mobile</td>
<td>8.02</td>
<td>0.04</td>
<td>***</td>
</tr>
<tr>
<td>$\mu_{fc}$: Fixed search cost on PC (normalized to 0 on mobile)</td>
<td>8.57</td>
<td>0.05</td>
<td>***</td>
</tr>
<tr>
<td><strong>Fixed-cost heterogeneity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buyer rating</td>
<td>-0.03</td>
<td>2.43E-03</td>
<td>***</td>
</tr>
<tr>
<td>Buyer rating missing</td>
<td>0.01</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Buyer spending</td>
<td>-1.10E-04</td>
<td>1.94E-05</td>
<td>***</td>
</tr>
<tr>
<td>Buyer history</td>
<td>-9.47E-05</td>
<td>5.95E-05</td>
<td>*</td>
</tr>
<tr>
<td>Male</td>
<td>-0.06</td>
<td>0.01</td>
<td>***</td>
</tr>
<tr>
<td>Gender missing</td>
<td>-7.85E-06</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.02</td>
<td>6.46E-04</td>
<td>***</td>
</tr>
<tr>
<td>Age missing</td>
<td>3.51E-05</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td><strong>Marginal-cost heterogeneity (Mobile)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screen resolution</td>
<td>-1.32E-02</td>
<td>4.87E-04</td>
<td>***</td>
</tr>
<tr>
<td>IOS</td>
<td>-1.95E-04</td>
<td>1.08E-04</td>
<td>*</td>
</tr>
<tr>
<td>Android</td>
<td>-3.49E-05</td>
<td>1.44E-05</td>
<td>**</td>
</tr>
</tbody>
</table>

$^{13}$ *: p<0.1; **: p<0.05; ***: p<0.01.
The second panel shows the search-cost parameter estimates. Note the marginal search cost is assumed to follow a log-normal distribution. We calculate the mean marginal search cost for using PCs as \( \exp(5.09+0.822/2)=227.3 \). Divided by the price coefficient, the mean cost is ¥44.05 (or US$6.61). Using the same procedure, the mean marginal search cost for using mobiles is ¥45.60 (or US$6.84). Thus, the average marginal search cost is 3.5% (¥1.55 or US$0.23) higher on mobile than on PC. The difference is statistically significant but not very large in magnitude. The marginal search cost determines the number of searches. This result is consistent with the data pattern showing the average number of searches on mobile is lower than on PC.

The fixed search cost on mobile is normalized to 0. Dividing the fixed search-cost parameter by the price coefficient, the cost for using PCs is higher by ¥1.66 (or US$0.25). Compared to the average difference in marginal search cost ¥1.55 ($0.23), the one-time fixed cost is higher (by about 6.8%) than the difference in marginal search cost between the two channels. Therefore, an average consumer who searches only one time would prefer using the mobile channel to PC because of the lower fixed search cost. When the optimal search times increases, PC becomes increasingly appealing to consumers because of its lower marginal search cost. The results are consistent with the data pattern showing that a larger proportion of consumers who search less tend to shop on mobile phones, whereas those who search more are more likely to choose PCs.

In addition to explaining the difference in the number of searches for consumers on the two channels, the marginal and fixed search-cost difference contributes to the observed gap in conversion rates between the two channels. When deciding which channel to search, consumers consider the search-cost differences and choose the channel that maximizes the expected utility.
after search. For consumers with higher overall valuation for the product category, the probability of making a purchase after search is high. Consumers who are likely to buy have a higher expected number of searches, because one additional search can have a higher marginal benefit in terms of a lower price and/or a higher individual match value. With a higher expected number of searches, these consumers are more likely to choose the PC channel with a lower marginal search cost. Therefore, PC is more likely to attract consumers with a higher overall valuation, who are expected to have a higher number of searches. Such self-selection of consumers leads to a higher conversion rate on the PC channel.

The third panel reports the observed heterogeneity of fixed search cost across consumers. Because Taobao started with the website optimized for PC and only introduced the mobile interface later, long-time consumers may have started shopping on Taobao before the introduction of the mobile-shopping option, and therefore become used to the PC shopping channel. We include measures that positively correlate with long-time usage history on the platform. Results support our hypothesis. Consumers with a higher buyer rating, more purchases in the past, and a longer buyer history on the platform are associated with a lower fixed search cost on the PC channel, which leads to a higher likelihood of using the PC channel compared to other consumers. This finding is also consistent with the probit regression results (Table 2).

In addition to length of usage history, consumer demographics may also play a role in explaining the choice of PC or mobile. We find age is negatively correlated with the fixed search cost for the PC channel. In other words, older consumers are more likely to have a lower fixed search cost on PC and are therefore more likely to use the PC channel for shopping. Male consumers have a lower fixed search cost for PC, which means they are more likely than women to use the PC channel for shopping. These estimates are again consistent with the reduced form in
Table 2.

In the fourth panel, we explore how the marginal search cost varies with different types of mobile devices. Because the marginal search cost is influenced by the effort in gathering information from an additional search, such a process should be less costly if gathering information on some mobile devices is easier. For example, consumers may find shopping using smartphones with a higher screen resolution (typically associated with a larger screen size) and a more robust operating system is easier. We find the parameter estimates for screen resolution, IOS, and Android operating systems are all negative and statistically significant. For smartphones with higher screen resolution and better operating systems, the marginal search cost becomes lower on the mobile channel. The results are consistent with reduced-form analysis showing that consumers with the more advanced smartphones are more likely to choose the mobile channel. Our results suggest that as the smartphone technology continues to improve, the marginal search cost on the mobile channel will decrease, leading to a higher number of consumers using the mobile channel for shopping.

Lastly, we examine the model fit by simulating consumer actions (channel choice, number of searches, and purchase decision) with the model estimates, and compare simulation results with the actual data. We run the simulation 100 times and take the average. We compare the conversion rate by search times on mobile (Figure 1.4A) and PC (Figure 1.4B), and the proportion of consumers who search one to five times on mobile (Figure 1.4C) and PC (Figure 1.4D). Both the conversion rates and the search times match well between simulated and actual data on both channels. The proposed model can predict the key empirical patterns. First, the conversion rate is higher with a higher number of searches on both channels. Second, the conversion rate is higher on PC than on mobile for the same number of searches. Third, consumers with more intensive
searches (who search at least three times) are more likely to choose the PC channel, which matches well between the simulated and actual data.

Figure 1.4 Model Fit by Comparing Actual and Model Simulated Data

To summarize, our results suggest the self-selection of consumers can explain the gap in conversion rates between the two channels. The PC channel has a higher fixed search cost and a lower marginal search cost, and it attracts consumers with higher valuation toward the product category who are more likely to make a purchase. The mobile channel has the advantage of a lower fixed search cost, because of the channel’s great portability and ease of access anywhere. It attracts consumers who may not find searching on PC to be worthwhile. Therefore, the pool of consumers the two channels attract can be systematically different before the start of any search activity.
1.7 Counterfactual
Consumers who choose PC and mobile channels are systematically different. Taking the different pools of consumers into account, we study how sellers can improve profits by utilizing channel-specific pricing and promotion strategies. Whether sellers are better off charging a lower price or offering a larger promotion deal on mobile is not obvious. On the one hand, consumers have a smaller consideration set (lower search intensity) on mobile, which reduces price competition and allows sellers to set a higher price. On the other hand, the conversion rate is lower on mobile, which suggests consumers are less inclined to make a purchase and sellers could be better off lowering prices. The proposed structural model accounts for both effects. With the estimated model, we can provide a more complete picture for sellers about consumer preferences using channel-choice information in addition to the search and purchase activities.

1.7.1 The Optimal Pricing Policy on Two Channels
In the first counterfactual, we study how sellers can utilize the information revealed by the consumer channel choice by offering different prices across channels. In practice, sellers can offer mobile-only prices for consumers using their smartphones to shop. With different prices on mobile and PC, in equilibrium, consumers will consider the price distribution on both channels and select channels accordingly. Therefore, channel-specific prices will also lead to changes in the pool of customers on both channels. Using our estimated model, we calculate the new equilibrium situation where sellers set different prices on PC and mobile and consumers have rational expectations of the price distribution, which influences their channel-choice decisions.

To find the optimal channel-specific prices, we need to estimate the marginal cost of sellers and the consumer demand function. This approach allows us to find the equilibrium condition, in which sellers set prices accounting for the customer base on both channels and, additionally,
consumers choose a channel considering the channel-specific price distribution. Our dataset contains more than 100 different products. Recovering the marginal cost for each one is computationally infeasible. We focus on the top 10 products, which account for more than 60% of the total sales during the data period. The prices of these products range from ¥117 to ¥208 (US$17.6 to $31.2).

To estimate the marginal cost of each seller, we assume the observed prices are the equilibrium prices when sellers can only choose the same price level for both channels. We first estimate the consumer demand function. The demand of product $j$ with price $p_j$ in channel $s$ is

\[
D^s(p_j) = \pi^s \cdot \left( \sum_{b=1}^{5} \pi_b^s \cdot \frac{b}{N} \cdot P[U_b^s(p_j) > \max(U_b^s(p_{-j}), 0)] \right),
\]

where $\pi^s$ is the proportion of consumers who choose channel $s$. $\pi_b^s$ is the proportion of consumers searching $b$ products on channel $s$. They search product $j$ with probability $\frac{b}{N}$, where $N$ is the number of all available products. $U_b^s(p_j)$ denotes the utility of product $j$ for consumers who search $b$ times on channel $s$ minus the outside option value. A consumer chooses to purchase $j$ if and only if the utility is higher than the utility of all other products browsed $U_b^s(p_{-j}^b)$, and is larger than 0 (i.e., buying product $j$ is more appealing than leaving without a purchase).

To evaluate $D^s(p_j)$, we draw the error terms in the model and simulate consumer search and purchase decisions using the model parameters. With the simulation results, we estimate $\pi^s$ and $\pi_b^s$ by the corresponding average values. For consumers who search $b$ times on channel $s$, we evaluate the probability that product $j$ offers the highest utility. For each product $j$, we obtain price
draws for the other \( b - 1 \) products, as well as their individual match values. With 1,000 sets of draws, we approximate the probability \( P[U_b^s(p_j) > \max(U_b^s(p_{-j}), 0)] \) by its corresponding sample average.

We calculate the demand function by changing \( p_j \) from 0 to ¥1000, which covers all observed prices in our dataset. Figure 1.5 plots the demand functions for PC (black dashed line) and mobile (grey solid line). The demand on PC is higher than on mobile at any given price, due to the self-selection by which consumers on PC are likely to have higher valuation of the product category than those on mobile. On both channels, price elasticity of demand is larger for moderate prices. When price is very low, demand is bounded above by the probability of the product being browsed. When price is very high, demand converges to 0 because the utility is likely to be lower than that of the other products or the outside option.

Figure 1.5 The Estimated Demand Function on PC and Mobile

---

\(^{14}\) Note the individual match value \( e_{ij} \) is only realized after consumer \( i \) searches product \( j \). This is different from search models (e.g., Honka 2014) where consumers know all the individual match values prior to search. In our model setting, the distribution of \( e_{ij} \) is not subject to selection. Similarly, prices are also realized after search. Therefore, we can take unconditional draws from the price distribution and extreme-value type-I distribution in the model simulation.
With the demand function, we then infer the marginal cost for product $j$ assuming that, given the prices of other sellers, the observed price maximizes the seller profit when the seller sets a single price for both channels. The marginal cost for seller $j$ $mc_j$ satisfies the condition

$$\hat{p}_j = \arg\max_{p_j} R(p_j, mc_j) = \arg\max_{p_j} \sum_s (p_j - mc_j)D^s(p_j),$$

where $R(p_j, mc_j)$ is the profit function for product $j$ with price $p_j$ and marginal cost $mc_j$. With the profit-maximizing assumption that observed price $\hat{p}_j$ maximizes the seller’s profit, we estimate the marginal costs for the top 10 sellers.\(^{15}\)

Instead of a single price on both channels, sellers can charge channel-specific prices to maximize profits. Seller $j$ chooses prices $p^0_j$ on mobile and $p^1_j$ on PC to maximize his expected profit as a function of the two prices and marginal cost:

$$\text{Max}_{p^0_j, p^1_j} R(p^0_j, p^1_j, mc_j) = \sum_s (p^s_j - mc_j)D^s(p_j),$$

where $D^s(p_j)$ is the channel-specific demand function under the new counterfactual prices. Consumers form rational expectations of the new price distributions on both channels, which will affect consumers’ channel choice. For example, if prices on mobile are lower than on PC, more consumers will choose the mobile channel, which will further influence the seller’s optimal prices on both channels. To find the equilibrium, we iterate between sellers choosing channel-specific prices given consumer channel choice, and consumers choosing a channel given channel-specific

\(^{15}\) We assume the remaining sellers keep their original uniform pricing on both channels in the counterfactual exercise.
prices. The process converges when the changes in channel-specific prices are less than 0.1 between iterations.

We find that when sellers charge channel-specific prices, the optimal price on mobile is lower than that on PC. Across the top 10 sellers, the average optimal price on mobile is ¥163.94, which is lower than the original uniform price at ¥165.74, whereas the optimal price on PC is ¥168.43, which is higher than the original price. On average, the price on mobile is lower by ¥4.49 (95% confidence interval: ¥3.73 – ¥5.11) or 2.7%. For the top 10 sellers, the optimal price on mobile is always lower than that on PC, with the magnitude of difference ranging from 1% to 4% across the sellers. With prices becoming lower on mobile under channel-specific prices, the conversion rate on mobile increases from 12.51% to 12.85% (or 2.7% in relative terms). We see the opposite story on PC where the prices become higher under channel-specific prices, and the overall conversion rate decreases from 16.35% to 16.10% (or 1.5% in relative terms). The overall pattern of a higher conversion rate on PC than on mobile continues, although the gap becomes slightly smaller. With channel-specific prices, sellers are able to make higher profits than under uniform pricing on both channels. Overall, the average profit increases by 0.55% (95% confidence interval: 0.06% – 0.70%) for the top 10 sellers.

To summarize, we find the optimal prices on mobile is 2.7% lower than that on PC. The proposed model considers the differences in search pattern and conversion rate on PC and mobile and shows that prices should be lower on the mobile channel, because of the lower valuation for consumers shopping on the channel. Ignoring consumer self-selection between the two channels can lead to incorrect channel-specific prices.
1.7.2 Optimal Retargeting Strategy for Sellers

In the second counterfactual, we investigate a retargeting strategy by offering coupons to consumers who have browsed but have not purchased. Similar to the first counterfactual, we consider the different pools of consumers on mobile and PC. However, consumers who abandon the search without purchase are systematically different from the total consumer population targeted in the first counterfactual (see Jiang et al. 2019). We focus on how sellers can use the channel choice information to offer optimal channel-specific coupons to attract consumers who have browsed without purchase. Such a retargeting strategy can be economically impactful because as many as 85% of consumers browse without making a purchase.

We calculate the optimal coupon values offered to retargeted consumers on mobile and PC. To focus on how the channel choice provides valuable information for sellers, we assume sellers know which channel consumers chose but not which products they have browsed. Seller $j$ chooses the coupon value $x$ on each channel to maximize the expected profit $r_j(x)$:

$$Max_x r_j(x) = \sum_s (p_j - mc_j - x)B^s_j(x)l^s,$$

where $p_j - mc_j - x$ represents the profit for seller $j$ considering the marginal cost (estimated in the first counterfactual) and coupon value $x$. $B^s_j(x)$ denotes the purchase probability on channel $s$ for seller $j$ when he offers a coupon value $x$ (the estimation procedure is described later). $l^s$ represents the number of consumers who browsed without purchase.\(^{16}\)

\(^{16}\) We assume consumers do not anticipate the retargeting coupon (i.e., they will not choose to search and abandon in order to get a retargeting coupon). Therefore, the percentage of non-purchasers $l^s$ does not change when the coupon value $x$ varies.
We calculate the purchase probability \( B_j^s(x) \) using simulation. We assume that when sending the coupons, sellers also provide detailed product information including price. Therefore, retargeted consumers do not need to search for the information and pay the search costs again. Using the estimated model, we simulate consumer channel-choice, search, and purchase decisions by drawing \( Q = 50 \) times from the error-term distributions and price distribution for each consumer. Let \( I_s^s q \) be the number of consumers who do not make a purchase on channel \( s \) at simulation \( q \), and \( a_{t,s,q} \) is the overall category valuation for these non-purchasers, whose outside option value is \( e_{0t,s,q} \) and the individual match value toward seller \( j \) is \( e_{j, t,s,q} \).

The purchase probability for seller \( j \) on channel \( s \) when the seller offers a coupon value \( x \) is

\[
B_j^s(x) = \frac{1}{Q} \sum_q \frac{1}{I_s^s q} \left[ a_{t,s,q} - \lambda \cdot (p_j - x) + e_{j,t,s,q} > e_{0t,s,q} \right],
\]

where the numerator calculates the number of non-purchasers who will make a purchase after receiving coupon \( x \). Dividing it by the total number of non-purchasers, we get the purchase probability for the retargeting coupon \( x \). \( B_j^s(x) \) represents the expected purchase probability when seller \( j \) sends a coupon worth value \( x \) to retarget consumers on channel \( s \).

With estimated \( B_j^s(x) \), we calculate the optimal coupon value \( x \) for seller \( j \) on PC and mobile given its original price and marginal cost. Similar to the first counterfactual, we focus on the top 10 sellers. We find the optimal retargeting coupon value is higher for consumers on mobile than on PC. The optimal coupon value for mobile consumers is ¥5.11 (about 3% of the original
price) and ¥4.81 for PC consumers. The difference in the coupon values is about ¥0.3 (or 6%) between the two channels with a 95% confidence interval from ¥0.0076 to ¥0.6057.

With the retargeting coupon, sellers can improve profits by 9.97% on the mobile channel, and by 10.05% on PC. The overall expected profit increases by 10.01%. We compare it with a scenario where sellers do not know the consumers’ channel choice. Sellers can only set one retargeting coupon value for all non-purchasers, regardless of their chosen channel. The optimal coupon value in this case is ¥4.92. The overall expected profit is lower by 5.1% than the profit under channel-specific coupons. This finding demonstrates how online sellers can utilize the channel-choice information for a more effective promotional strategy such as sending out retargeting coupons.

To summarize, we find sellers’ profit increase is higher when they offer channel-specific retargeting coupons than when channel choice is not considered. The optimal coupon value is higher for consumers on mobile than on PC. The result is consistent with that in the first counterfactual, which suggests a lower optimal price on mobile than on PC. Both results are driven by the difference in what types of consumers will self-select to browse on mobile phones or PCs.

1.8 Conclusions and Limitations
In this paper, we develop a model of consumer channel choice in addition to search and purchase. The proposed model can explain an intriguing phenomenon whereby, although more consumers use mobile phones to shop, the conversion rate is significantly lower than that on PCs. We find the PC channel has a lower marginal search cost but a higher fixed search cost than the mobile channel. Consumers with higher product valuation are more likely to use the PC channel because they have a higher search intensity and will benefit from the lower marginal search cost. Consumers with
lower product valuation, on the other hand, are more likely to choose the mobile channel because of its lower fixed cost to start a shopping session.

The estimated model allows us to study channel-specific marketing strategies for sellers. We find the optimal price on mobile is 2.7% lower than on PC. For non-purchasers, the optimal retargeting coupon value is 6% higher for consumers on mobile than on PC. Overall, sellers’ profit will increase if their marketing strategies are channel specific. Both counterfactual analyses demonstrate how the proposed model can provide sellers with important managerial insights. Ignoring consumer self-selection between the two channels can lead to incorrect channel-specific marketing strategies.

The contributions of this paper are two-fold. From a methodological prospective, we propose a flexible framework that incorporates endogenous consumer channel choice in addition to the search and purchase decisions. The proposed model can capture the observed search activities and purchase decisions on both channels. From a managerial perspective, our results offer guidance to sellers on the optimal channel-specific marketing strategies. We consider channel-specific prices and retargeting coupons and show how they should be different on the two channels.

Like all research, our study has limitations. First, the optimal channel-specific prices and marketing strategies are from counterfactual analysis using one product category on Taobao. We call for future research to further test these recommendations with actual field experiments. They should also be tested with a broader range of product categories and in different countries for generalizability. Second, our proposed model makes several strong assumptions on consumer search and purchase behaviors. In particular, we assume consumers use simultaneous search
strategy. Future research with richer datasets should further explore consumers’ channel choice in other scenarios, such as when consumers use sequential search and when they have prior knowledge on the differentiated quality of sellers. Results on how consumers who choose to shop on the two channels are systematically different will help test the robustness of our findings.
References


Chapter 2

Time-inconsistent Preferences and Strategic Self-Control in Digital Content Consumption

Co-authored with Tat Y. Chen, Xueming Luo and Xiaoyi Wang

2.1 Introduction
Consumers widely embrace digital content, such as social media, videos, music, and books. According to a report from eMarketer,\(^{17}\) adults in the United States on average spent 6 hours and 19 minutes per day on digital media in 2018, overtaking the time spent on traditional media for the first time. Online video has been the main driver of this phenomenon. Netflix, for example, gained more than 700 million subscribers in just four years from 2013 to 2017. Demand for other content has also significantly increased. The rapid growth in the consumption of digital content has raised concerns about the mental and physical well-being of consumers. Research on consumer digital content consumption decisions (e.g., Zhang et al. 2012; Boumosleh and Jaalouk, 2017) has shown that when consumers indulge themselves inside the fantasy world created by video games, online videos, and web fictions, they can lose self-control and spend more time and money than they originally intended to. Other consequences include various mental and physical problems. Such over-consumption behavior has been categorized as a clinical disorder in China since 2008.\(^{18}\) Being aware of the negative implications of over-consumption may induce consumers to conduct self-control strategies in advance. Wertenbroch (1998), for example, documents that smokers may

\(^{17}\) Source: https://www.emarketer.com/content/us-time-spent-with-media-2018.

\(^{18}\) Source: https://www.theguardian.com/news/blog/2008/nov/11/china-internet
purchase small cigarette packets, to curb future smoking. This strategy may also be used to control over-consumption in digital content.

The goal of our study is to examine consumers’ digital content consumption behaviors and their self-control strategies. We focus on a specific type of digital content—web fiction. Web fiction is available primarily on the Internet, usually released serially by chapters daily. Reading web fiction is prevalent among Asian consumers: according to a report, more than 400 million unique readers in China pay for some form of web fiction. A successful web fiction writer can earn more than 16 million dollars.\textsuperscript{19} The market size in China reached 18 billion RMB\textsuperscript{20} in 2018.\textsuperscript{21} We use a unique dataset obtained from a major digital book platform in China that provides the content to subscribers. We observe several unique data features that are related to our research goal. First, individual consumers, on average, read about 500 chapters per month, with one-fourth reading 1,500 chapters in average. Assuming one chapter takes five minutes to read,\textsuperscript{22} reading time per month is 42 hours for an average consumer and 125 hours for the top one-fourth, indicating individuals spend as much time on web fictions as on other digital media.\textsuperscript{23} Second, web fiction is different from literary fiction. It enables readers to immerse themselves in “fantasy pleasure” and can cause withdrawal symptoms when they try to stop reading.\textsuperscript{24} Further, as payment from the online platform to authors is linked to how many chapters an individual reads, authors have an inherent incentive to include plot twists to keep readers following their books every day. These

\textsuperscript{19} Source: https://www.nytimes.com/2016/11/01/world/asia/china-online-literature-zhang-wei.html
\textsuperscript{20} One RMB is about 0.15 US dollar.
\textsuperscript{22} A chapter contains about 1,000 Chinese characters
\textsuperscript{23} An average Chinese adult spent nearly 4 hours a day (120 hours per month) on digital media. (Source: https://www.emarketer.com/content/china-time-spent-with-media-2019)
\textsuperscript{24} Source: http://www.cnki.com.cn/Article/CJFDTotal-DDWT201701036.htm
features may cause consumers to read more than they intend to regularly. Finally, we observe an interesting overpaying phenomenon among consumers: to read web fictions, an individual can choose either a pay-per-chapter plan (0.1 RMB for each chapter) or a monthly subscription plan (12 RMB each month). Anyone who reads more than 120 chapters under the former plan or fewer than 120 chapters under the latter plan will overpay for the reading. We find that, while only 6% of consumers overpay by choosing the monthly subscription, one-third of consumers consistently overpay by choosing the pay-per-chapter plan. Among consumers who overpay, 60% switch plans later if they are under the monthly subscription, but only 20% switch under the pay-per-chapter plan. Furthermore, most switchers from the monthly subscription (including those who do not overpay) continue to read a lot after switching, implying that the monetary cost would have been significantly lower had they not switched. We argue that these overpaying behaviors are consistent with the idea that consumers use strategic self-control measures to curb over-consumption, which can have negative long-term consequences.

To formalize the idea, we propose a dynamic structural model that allows consumers to have different consumption preferences during the reading plan and consumption choice stages. It characterizes a type of “time-inconsistent” preferences that can lead to over-consumption behavior. Such time-inconsistent preferences apply to not only digital content consumption behaviors but also other types of behaviors such as gambling, use of drugs, alcohol, and tobacco. Furthermore, the model allows for habit formation, as reading more today may increase the reading preference (leading to reading even more) in the future. Anticipating the potential downsides of the time-inconsistent preferences and habit formation, rational consumers will impose strategic self-controls. One of the strategies is to choose a reading plan that will help curb their future reading, even though doing so could incur a higher monetary cost. This decision implies that a
forward-looking consumer, when choosing a plan, plays a game against the myopic self during the consumption stage.

We first use an analytical model to illustrate this argument. We show that, given a set of model parameters and state of reading preference, a unique equilibrium exists. We also show that, under reasonable assumptions, the equilibrium will converge globally to one of the three steady states: (1) consumers choose the pay-per-chapter plan, and stay at a low reading preference; (2) consumers choose the pay-per-chapter plan and remain at a medium state; and (3) consumers pay monthly subscription and stay at a high state. Depending on the utility parameters, we show that overpaying by choosing the pay-per-chapter plan can be optimal. We further show that the standard recursive method that solves the optimal policy function in the dynamic programming literature can be used in our model, even though the agent exhibits time-inconsistent preferences.

We then construct an econometric model for the empirical analysis by incorporating unobservable and individual heterogeneities. Estimation results show that, out of the three latent segments of consumers, the first segment is more price-sensitive and has a higher non-monetary cost of over-consumption. Interestingly, despite being more price-sensitive, this segment actually overpays for the consumption due to the self-control reason. In contrast, segment 2 is less likely to overpay, and about half of the segment chooses the monthly subscription plan. Segment 3 is small in size; it has the highest reading preference and is most likely to select the monthly subscription plan.

The findings of this study can have substantive implications not only for public policymakers but also for producers and distributors of digital content. We use counterfactuals to study the impacts of pricing plans on consumer welfare and the platform’s profit. We find that if
the platform were to eliminate the pay-per-chapter plan that helps curb consumption, not only would it hurt consumer welfare but also its profit, which would drop by 76%. On the other hand, if the platform were to eliminate the monthly subscription that encourages more consumption, its profit would increase by 46%, whereas consumer welfare only decreases by 4.5%. Finally, we find that introducing a new nonlinear pricing plan with volume surcharge can simultaneously improve the platform’s profit by 46% (almost the same as eliminating the monthly subscription) and the consumer welfare by 2.5%. It also dominates another nonlinear pricing plan with volume discount on both dimensions. Overall, our results show that in contrast to the common belief that firms should offer pricing plans that encourages consumption (e.g., monthly subscription only), offering a plan that helps curb consumption (e.g., pay-by-chapter and volume surcharge) can increase the firm profit when consumers have the motivation of self-control. It highlights the necessity of considering consumers’ self-control behaviors when marketing managers design the pricing structure for digital content distribution.

The rest of the paper is organized as follows: Section 2.2 reviews the related literature. We describe the data and present some empirical data patterns in section 2.3 to motivate our modeling approach. Section 2.4 describes our structural model and estimation strategy. Section 2.5 presents the estimation and counterfactual results. Finally, section 2.6 concludes.

2.2 Literature
This study is related to the literature on time inconsistency, a concept first formally introduced by Strotz (1956). Because preferences evolve over time, the optimal choice today may not be the best in the future. Since Strotz (1956), numerous experimental and empirical studies have shown different forms of time inconsistency (see Loewenstein and O'donoghue, 2002 for a summary). An example of time inconsistency is hyperbolic discounting, that is, the discounting rate is much
higher for future outcomes. Hyperbolic discounting has found support in numerous studies using experiment or field data (e.g., Thaler, 1981; Benzion et al., 1989; Redelmeier and Heller, 1993; Chapman and Elstein, 1995; Chapman, 1996; Pender, 1996). Previous literature has also shown other forms of time inconsistency, for example, the “sign effect,” whereby consumers value future loss more than gains (Mischel et al., 1969; Yates and Watts, 1975; Loewenstein, 1987; Benzion et al., 1989; MacKeigan et al., 1993; Redelmeier and Heller, 1993), and the “magnitude effect,” whereby consumers discount small numbers more than large numbers (Thaler, 1981; Ainslie and Haendel, 1983; Kirby and Loewenstein, 1987; Benzion et al., 1989; Green et al., 1994a, 1994b; Kirby and Marakovic, 1995; Kirby, 1997). We model how individuals may ignore the consumption cost during the consumption stage, although they are fully rational when choosing the price plan. This is consistent with the general definition of time inconsistency in Strotz (1956) but is different from a typical hyperbolic discounting model setting.

Given the ubiquitous evidence of time-inconsistent behaviors, theoretical and empirical researchers have further studied to what extent consumers use self-control strategies to solve this problem, mostly in the form of hyperbolic discounting. Laibson (1997), for example, constructs a theory to show how dynamically inconsistent preferences could incentivize consumers to constrain their future choice. O'Donoghue and Rabin (1999) also use theoretical models to study how individuals self-control when the cost and reward of consumer decisions do not realize at the same time. Other studies also examine how consumers can achieve self-control by restricting the opportunity for additional purchases (Rachlin, 1995), or reducing temptation through substitution (Hoch and Loewenstein, 1991). The concept of time inconsistency has also been adopted in empirical studies to explain various types of consumer behaviors that seem to be inconsistent with classical economic theories. Wertenbroch (1998), for example, uses both experiment and field data
to show how consumers strategically ration the purchased quantity to restrict excessive consumption. DellaVigna and Malmendier (2006) find from data that individuals overpay for a gym membership. They argue that this overpaying is a strategy to increase future gym use. We study a similar behavior; however, we show that consumers overpay to curb over-consumption that will result in future utility loss. Gruber and Köszegi (2001) develop a utility function with time inconsistency and addictive consumption behaviors. Due to the technical barrier of estimating such a dynamic model, they calibrate model parameters to show that, due to smokers’ time-inconsistent preferences, the optimal cigarette tax should be higher than under the time-consistent assumption. Our study makes both methodological and substantive contributions to this stream of literature. On the methodological side, Caplin and Leahy (2006) have shown that the standard recursive iteration method cannot be applied to time inconsistency models when more than three time periods exist. They further suggest that equilibrium may not exist. We show that a unique equilibrium exists in our model setting and can be computed using the recursive method. Future researchers could use our method to study other types of time-inconsistent consumer behaviors in a wide range of markets, including other digital contents and traditional product categories (e.g., tobacco, alcohol, and drugs). For the substantive contribution, we use counterfactuals to study the impacts of marketing actions on firms’ profit and consumer welfare. The implications will be entirely different without considering the consumer strategic self-control induced by time-inconsistent preferences.

The reason behind time-inconsistent preferences can be related to additive consumption. The theory of rational addiction in Becker and Murphy (1988) is based on the assumption that consumers can evaluate the monetary and non-monetary benefits and costs from consumption. This theory has been adopted in later empirical research studying different consumption behaviors,
including cigarettes (Chaloupka, 1991; Becker et al., 1994), alcohol (Baltagi and Griffin, 2002), drugs (Grossman and Chaloupka, 1998; Liu et al., 1999; Olekalns and Bardsley, 1996), and gambling (Mobilia, 1993). Arcidiacono et al. (2007) study how forward-looking consumers make decisions for consuming alcohol and tobacco. Researchers have also investigated addiction in social media consumption, Internet browsing, and mobile apps usage (e.g., Young, 1998; Pelling and White, 2009; Wan, 2009; Kuss and Griffiths, 2011; Kwon et al. 2016). Our model differs from the rational addiction literature, as we allow consumers to be myopic during the consumption stage and ignorant of the cost of over-consumption. This is consistent with the medical research (e.g., Nestler, 2013) which argues that, because consumption is intrinsically rewarding, consumers may not correctly evaluate the short- and long-term costs during consumption. A recent study on addictive usage of smartphones by Boumosleh and Jaalouk (2017) also finds users usually do not consider health consequences during usage. However, following Becker and Murphy (1988), we allow consumers to form expectations about their future consumption when choosing the price plan.

2.3 Data
Our data come from a major digital book platform in China. The data sample includes the reading activity from 11,346 unique consumers, randomly selected among the existing subscribers of the platform, for six months from January to June 2017. The platform offers a rich collection of web fictions. Unlike online books provided by Amazon Kindle, web fictions are mostly written by amateurs. They are primarily published online, updated daily by chapters. The length of each book

---

25 In marketing, researchers have documented addictive cigarette consumption and the consequences of a cigarette tax on demand (Chen, Sun, and Singh, 2009; Wang et al., 2015). Gordon and Sun (2015) use a dynamic model of rational addiction to study the impacts of a permanent price shift induced by a new cigarette tax on the demand for cigarettes.
varies depending on whether it attracts readership. A complete one could easily exceed a thousand chapters. Each chapter is usually about 1,000 to 5,000 Chinese characters, requiring a few minutes of reading time. Distinct from literary fictions, which typically refer to fictions with literary merit, web fictions are classified by genres such as fantasy, romance, and science. Readers follow web fictions for entertainment, a riveting story, and escape from reality. Given that each chapter is short, consumers often read multiple books each day. This is different from the “binge-watch” behavior, wherein consumers spend a short period on intensively watching a whole season of TV series and then stop.

The platform offers two pricing plans for readers: pay-per-chapter and monthly subscription. The prices are 0.1 RMB for each chapter under the former plan, and 12 RMB under the latter.26 Under the pay-per-chapter plan, readers have to swipe a bar on their mobile phones to confirm the payment before reading each chapter. Under the monthly subscription, readers receive a text reminder for the payment due a few days before the subscription expires. If readers stop the subscription, they are automatically switched to the pay-per-chapter plan.

Each period is a month in our analysis. The number of chapters an individual reads in a month represents his/her consumption level. Because readers start and end monthly subscriptions at different times, we make an assumption when constructing the dataset: If a reader starts a subscription before the 15th of a month, we assume he/she starts the subscription at the beginning

---

26 These prices are equivalent to US $0.015 for each chapter and $1.80 for each month.
of this month; otherwise, his/her subscription starts at the beginning of next month. We tried different data-construction methods and found robust results.\textsuperscript{27}

2.3.1 Consumption Behaviors in Data
We use the data to examine consumers’ reading behaviors. We first present an overview of the reading amount, then look at whether there is a habit formation that is well-known in the marketing literature. Habit formation is characterized as past consumption enhancing the current consumption preference, thus creating dynamics in reading behaviors. Next, we investigate possible time-inconsistent preferences exhibited from the consumption and plan choice stages. Finally, we examine the evidence that consumers use self-control strategies. These are the two important components in our dynamic structural model.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
Reader group & Monthly reading (chapter) & Percentage of Population \\
\hline
Overall & 506.4 & 100.00\% \\
Pay-per-Chapter & 392.19 & 74.51\% \\
Monthly Subscription & 827.06 & 25.49\% \\
Overpay for Monthly Subscription & 45.92 & 6.60\% \\
Overpay for Paid-per--by-Chapter & 836.03 & 34.16\% \\
\hline
\end{tabular}
\caption{Reading Amount}
\end{table}

\textsuperscript{27} For example, we change the date to the 7th or 23rd in each month, and find the data patterns that we present below remain similar. In the data, 84\% of consumers change plans in the first or last week of a month; therefore, using different cut-off dates does not affect the results.
Table 2.1 presents a glimpse of the consumption level. The first row shows that an average consumer reads about 500 chapters per month. Assuming consumers spend five minutes reading each chapter, this amount represents spending 2,500 minutes per month reading web fiction, or 83 minutes per day. Furthermore, 25% of consumers in the data read more than 700 chapters in a month, implying they spend about 60 hours a month, or almost two hours each day. If the average workday is eight hours, the above statistics suggest that in a month, the reading time is equivalent to 5.2 working days for an average consumer, and 7.5 working days for the top 25%. These suggest that the non-monetary costs (e.g., time cost and other adverse consequences) of reading could be very significant. A report from eMarketer shows that US adults, on average, spent 379 minutes per day on digital media in 2018. Broken down by format, consumers spend 51 minutes per day on video games\(^{28}\) and 135 minutes on social media.\(^{29}\) The comparison suggests reading web fictions is as time-consuming as other digital contents.

We test whether increasing exposure from past consumption will increase the current consumption in our data. Following the model in Becker and Murphy (1994), we specify the habit stock as an accumulation of past consumption under depreciation. Denoting the state of individual \(i\) in month \(t\) as \(h_{it}\), the depreciation rate as \(\delta\), and the consumption as \(c_{it}\), the habit stock evolves as the following:

\[
h_{i,t+1} = (1 - \delta) h_{it} + c_{it}. \tag{1}
\]

To test the relationship between the habit stock and consumption, we run an ordinary least squares regression with the reading amount (from second to the last month) as the dependent

---

\(^{28}\) Source: https://www.limelight.com/resources/white-paper/state-of-online-gaming-2018/#spend

\(^{29}\) Source: https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide
variable and the habit stock in the previous month as an independent variable. Because the habit stock cannot be directly observed in the data, we calculate it in equation (1), restricting the value of $\delta$ with a lower bound at 0.36, as suggested by Becker and Murphy (1990). We also assume every consumer starts at $h_{t1} = 0$ in the first month. To control for the heterogeneity in reading preferences across consumers, we include individual fixed effects in the regression.

Regression results show that the coefficients for $h_{t\ell}$ range from 0.05 to 0.005, and for $\delta$ between 0.36 and 1. They are all significant at the 0.001 significance level. The results suggest that past consumption is positively correlated with future consumption.

To show evidence of time-inconsistent preferences is less straightforward because we do not observe consumers’ utility during the plan choice and consumption stages. Our strategy is to show inconsistencies between the pricing plan consumers choose and their reading amount, as indirect support of the assumption. The second and third rows of Table 2.1 show that the majority (three-fourth) of consumers choose the pay-per-chapter plan, and as expected, the number of chapters these consumers read is lower than those read by consumers who choose the monthly subscription plan. What is surprising is that the average number of chapters read by the consumers is 392, significantly higher than the 120 chapters over which the optimal plan choice should be a monthly subscription. Furthermore, the last two rows of Table 2.1 show that the majority of consumers who overpay choose the pay-per-chapter plan. Their average reading amount is 836 chapters, similar to consumers who choose the monthly subscription plan. All these numbers suggest that the reading amount of the majority of consumers who choose the pay-per-chapter plan is not consistent with their choice, indicating possible time-inconsistent preferences in the two stages.
Finally, we look for supporting evidence for the strategic self-control assumption. Assuming consumers know that during consumption, they will not sufficiently consider the costs, they will have an incentive to take actions during the plan choice stage to curb future consumption. The most striking data pattern is the proportion of consumers who overpay under the pay-per-chapter and monthly subscription plan, as shown in Table 2.1. Over one-third of consumers overpay by reading too much under the pay-per-chapter plan, whereas only 6.6% consumers overpay by reading too little under the monthly subscription plan. For the former consumers, the reading amount is 836 chapters, far higher than the 120 chapters beyond which they should choose the monthly subscription. The asymmetric overpay ratios of consumers under the two pricing plans are consistent with the assumption that consumers strategically choose the pay-per-chapter plan to curb their future consumption.

Table 2.2 Probability of Switching Plans

<table>
<thead>
<tr>
<th>Switching from</th>
<th>Overall</th>
<th>Overpay</th>
<th>No overpay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pay-per-Chapter</td>
<td>17.55%</td>
<td>22.65%</td>
<td>12.45%</td>
</tr>
<tr>
<td>Monthly Subscription</td>
<td>46.43%</td>
<td>58.44%</td>
<td>34.41%</td>
</tr>
</tbody>
</table>

Table 2.2 offers further evidence in support of the strategic self-control assumption. The first column reports the probabilities that consumers switch away from the plan they chose last month. The proportion of consumers who switch away from monthly subscription is far larger than those who switch away from pay-per-chapter. The next column shows that for consumers who overpay for pay-per-chapter (by reading too much), only 22.7% switch to the monthly subscription, whereas the switching probability for those who overpay for the monthly subscription (by reading too little) is 58.4%. The high switch probability for monthly subscription suggests
consumers pay attention to the monetary cost and adjust the pricing plan accordingly. The low switch probability for pay-per-chapter, on the other hand, implies consumers are willing to incur a higher monetary cost to curb future consumption.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always subs:</td>
<td>Consumers who always purchase monthly subscription during the data time period</td>
</tr>
<tr>
<td>Always chpt:</td>
<td>Consumers who always choose pay-per-chapter during the data time period</td>
</tr>
<tr>
<td>Subs in xy:</td>
<td>Consumers who choose monthly subscription from month x to y and then quit</td>
</tr>
</tbody>
</table>
For self-control to work, one necessary assumption is that consumers are responsive to monetary cost during the consumption stage. We find from data that after a consumer switches to pay-per-chapter, the average reading amount drops from 794 to 395 chapters per month. Figure 2.1 breaks down reading amount in the six months of the sample period by consumers who always choose the monthly subscription, always choose pay-per-chapter, and subscribe in months 1 and 2, 2 and 3, and 3 and 4, before they switch to pay-per-chapter. The figure shows that the consumption level of consumers who always choose monthly subscription increases over time (from 1,063 to 1,171 chapters). The consumption level of consumers who always choose pay-per-chapter is steady between 200 and 300 chapters a month. The consumption level of consumers who switch from the monthly subscription to pay-per-chapter drops significantly after the switch. These data patterns suggest switching to pay-per-chapter helps curb consumption. Interestingly, we find these consumers still overpay for their reading after they switch.

2.3.2 Alternative Explanations
In this subsection, we examine whether several alternative explanations adopted from the past literature can explain the data patterns presented above. The first one is the rational addiction theory developed by Becker and Murphy (1988). The habit formation specification that we presented above is based on their work. However, they do not consider time-inconsistent preferences. Consequently, the consumption is optimal, and consumers will choose the pricing plan consistent with their consumption level. In this case, the asymmetrical pattern of overpaying we show in Table 2.1 should not exist. Therefore, we conclude that the rational addiction theory cannot explain why overpaying predominantly originates from pay-per-chapter consumers. Later, in the model estimation, we further show that without time inconsistency, the model fails to predict
the asymmetric overpaying behavior where a large proportion of consumers stay with their pay-per-chapter plan and continue overpaying.

Another alternative explanation is that because consumers first make the plan choice and then go through the consumption process, their reading preferences can experience random shocks in the second stage. This explanation still cannot explain the asymmetric overpay patterns. We use a simulation exercise for illustration. For each individual, we assume the reading amount follows an individual-specific normal distribution. Before making the pricing plan decision, the individual knows the distribution of his/her reading amount in the next period but not the exact amount and makes the plan choice that minimizes the expected cost based on the information. We run the simulation that draws from the empirical distribution of the reading amount for each individual. We find that the average overpay ratio for consumers who choose pay-per-chapter is about 11.7% compared with a significantly higher 17.9% for those who choose the monthly subscription. Overpaying for the monthly subscription is higher because the plan puts an upper bound on the monetary cost if a positive shock occurs in reading preferences. The result contradicts the empirical data pattern. The data pattern also rejects the risk aversion or option value explanation for consumers who choose not to switch to the monthly subscription. This is because the likelihood that a consumer reads fewer than 120 chapters per month is very small, but the probability that he/she reads a large amount is very substantial. A consumer should choose the monthly subscription if he/she is averse to overpaying. Likewise, choosing monthly subscription will have a higher option value (in the case he/she reads a lot in a month) than the pay-per-chapter plan.

Consumer learning is another explanation. It argues that overpaying behavior can be a result of people learning about their true preferences over time. To test this explanation, we repeat the simulation described above, allowing each individual to update his/her belief of the mean reading
amount in each month based on the reading amount in the previous month. With the learning, the proportion of consumers overpaying is reduced under the two pricing plans; however, the overpaying ratio under monthly subscription is still higher than that under the pay-per-chapter plan.

Table 2.3 Switching Probability after Overpay

<table>
<thead>
<tr>
<th>Consumer group</th>
<th>Overpay for 1 months</th>
<th>Overpay for 2 months</th>
<th>Overpay for 3 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overpay for Pay-per-Chapter</td>
<td>22.65%</td>
<td>16.03%</td>
<td>4.95%</td>
</tr>
<tr>
<td>Overpay for Monthly Subscription</td>
<td>58.44%</td>
<td>52.64%</td>
<td>27.64%</td>
</tr>
</tbody>
</table>

Another way to test the learning story is to see how consumers adjust their pricing plans after overpaying. We calculate the average switch probabilities for consumers who overpay for one, two, and three months. The results are reported in Table 2.3. Even though consumers who choose pay-per-chapter predominantly overpay, their switching probabilities are consistently lower than those who choose the monthly subscription. Furthermore, the switch probabilities actually decline for those who overpay for a longer period. This clearly contradicts the learning story, which predicts consumers will adjust their choices accordingly over time.

As another alternative explanation, consumer inattention suggests consumers are not aware they are overpaying. Table 2.2 shows the switching probability is high (58.4%) when consumers overpay under monthly subscription for the previous month. Only consumers who overpay under the pay-by-chapter plan do not switch. Consumer inattention cannot explain this asymmetric switching pattern. Furthermore, in our empirical context, consumers under pay-per-chapter have
to agree to pay 0.1 RMB each time they read a new chapter. Because they are constantly reminded of the payment, inattention does not seem to be the main reason behind the massive scale of overpaying for such a pricing plan.

Transaction cost or consumer inertia as an alternative explanation for the asymmetric overpay patterns suggests that consumers find switching from pay-per-chapter to monthly subscription to be too costly. We do not have direct evidence to rule out such an explanation. In our empirical context, however, consumers are already registered users of the platform. Because they do not have to provide any personal information or change their payment method, consumers can switch with just one click. Given the average overpaying amount is more than four times the cost of a monthly subscription, transaction costs are unlikely to be the reason for not switching.

One may argue that the costs are not about registration but are related to the effort of making payment. However, under the pay-per-chapter plan, consumers need to swipe the bar on their smartphones to pay for each chapter they read. Such effort is costlier than making one upfront payment for a monthly subscription. Again, the transaction cost explanation does not seem to hold in our empirical context.

The asymmetric overpay pattern may be explained by the non-monetary benefit of pay-per-chapter. For example, once a consumer pays for a chapter, he/she can always re-read the chapter, whereas that option is not available for the monthly subscription once he/she stops the subscription. This is not true in our context. Once the consumer pays for one chapter, he/she can only read it within 30 days. This restriction makes access to any chapter under pay-per-chapter identical to the monthly subscription. Another potential non-monetary benefit of pay-per-chapter for consumers lies in its option value, in the sense that consumers can always switch from pay-per-chapter to
monthly subscription but cannot freely switch back to pay-per-chapter within one subscription cycle (30 days). There can be a psychological benefit due to payment flexibility. We find from data that consumers who pay by chapters read on average 800 chapters per month, and more than 80% of them never read below 120 chapters. Given the consistently high reading amount, consumers need to perceive an exceedingly large psychological benefit of the pay-per-chapter plan for them not to switch.

Another alternative explanation, over-confidence, suggests that consumers overestimate their self-control ability in the future, and thus, underestimate the future reading amount. The assumption behind this is that they are naive time-inconsistent consumers who are not aware of how they will make choices during the consumption stage. Although this explanation cannot be completely ruled out, we find that it is not likely to be the main reason consumers overpay for their reading. The majority of those who overpay with pay-per-chapter continue with their plan for more than four months. It is unreasonable to assume that, after such a long time, consumers still believe that they can keep their reading amount below 120 chapters, when the actual average reading amount is consistently far higher than that. We should observe consumers to eventually switch to the monthly subscription. The same argument is made in DellaVigna and Malmendier (2006). As a more formal test, we estimate a model with naive time-inconsistent consumers. Estimation results suggest that consumers would need to underestimate future reading by an unreasonably large amount. Details are in the next section.

The last potential explanation is that consumers do not care about the small price of 0.1 RMB under pay-per-chapter and end up overpaying. In this case, we should not find consumers reduce their consumption after they switch from monthly subscription to pay-per-chapter. In our data, however, the average reading amount after consumers switch to pay-per-chapter drops from
794 to 395 chapters per month, suggesting their consumption is responsive to the price of each additional chapter they read.

To conclude, although the alternative explanations we list above may explain some of the data patterns, they are inconsistent with either the asymmetric overpaying and switching patterns or the change in consumption behaviors in the data. We acknowledge that time-inconsistent preference and strategic self-control may not be the only mechanism driving the data observations; other potential behavioral explanations that we have not examined may still exist.

2.4 The model and its Estimation
In this section, we first use an analytical model to demonstrate under what conditions consumers will overpay as an optimal behavior. We then develop an econometric model with stochastic components and estimate the model from data. We discuss the estimation method and the model identification.

2.4.1 An Analytical Model and Its Equilibrium
In the dynamic programming literature, the recursive method has proved to be a useful tool to solve dynamic problems. When agents have time-inconsistent preferences, however, equilibrium may not be computed using the solution concept (e.g., Peleg and Yaari, 1973; Caplin and Leathy, 2006), because value-function iteration may not converge, and thus the optimal policy function does not exist. Our model features both habit formation and time-inconsistent preferences under rational agent behavior. We show a unique equilibrium exists in the model. We further show the existence of a unique steady-state equilibrium. We characterize how the equilibrium varies depending on model parameters and the state of consumption stock.
In the model, a consumer chooses either pay-per-chapter \((s = 0)\) or monthly subscription \((s = 1)\) at the beginning of each period, then decides the level of consumption conditional on the chosen plan. The price for each chapter is \(p_c\) under pay-per-chapter, and \(p_s\) for the monthly subscription. The consumption utility is influenced by the habit stock, \(h\). The utility during the consumption process is different from the utility when the consumer makes the pricing-plan choice.

Starting with the utility function during consumption, we specify a quadratic utility function as follows:

\[
u_c(c, s, h) = (\alpha_c + \alpha_{ch} \cdot h) \cdot c - \alpha_{cc} \cdot c^2 - \mu \cdot p_c \cdot 1\{s = 0\} \cdot c ,\]

where \(\alpha_{ch}\) captures how the habit stock \(h\) may change the marginal utility of consumption, and \(\mu\) is the price coefficient representing the marginal disutility of the monetary cost during the consumption stage. It only occurs if the consumer chooses pay-per-chapter; otherwise, the price of reading a chapter is zero. Given \(s\) and \(h\), the consumer chooses the optimal \(c^*(s, h)\). It is straightforward to derive that

\[
c^*(s, h) = \left(\frac{\alpha_{cc} + \alpha_{ch} \cdot h - \mu \cdot p_s \cdot 1\{s = 0\}}{2 \cdot \alpha_{cc}}\right) \text{ if } c^* \geq 0, \text{ and } 0 \text{ otherwise.}
\]

Suppose \(\mu\) is positive. The optimal \(c^*\) when \(s=0\) is lower than that when \(s=1\).

Time-inconsistency in our model comes from consumers who, when choosing the pricing plan, consider the non-monetary costs of consumption (e.g., time, negative health impact from excessive smartphone usage) that they do not fully take into account in \(u_c(c, s, h)\). A difference in the monetary cost between the consumption stage and plan-choice stage can also exist (e.g., the consumer may care less about price once he/she has indulged in reading). We use \(\gamma\) to represent
the total difference in the marginal cost for reading each chapter. The utility function when the consumer chooses the pricing plan is specified as follows:

$$u_p(c, s, h) = u_c(c, s, h) - \gamma \cdot c. \tag{4}$$

We assume the consumer is a “sophisticated” type, as proposed by Strotz (1956). That is, the consumer is aware that during consumption her choice is $c^*(s, h)$ in equation (3) without the cost $\gamma \cdot c$. Furthermore, she is forward-looking as she considers how her current consumption can affect her future habit and thus the consumption. Formally, the consumer chooses a pricing plan by solving the following value function:

$$V(h) = \max_s u_p(c, s, h) - \mu \cdot p_s \cdot 1\{s = 1\} + \beta V(h'), \tag{5}$$

where

$$h' = (1 - \delta)h + c \tag{6}$$

Equation (6) illustrates how the habit stock in the next period will evolve following the current consumption $c$.

Because the utility function is continuous under either pricing plan, value function $V$ is continuous in state space $h$. Assuming $h$ is a compact set bounded above, for any parameter set, the contraction mapping theorem will hold so that the dynamic programming problem is guaranteed to have a unique fixed point for $V$. Therefore, for any $h$, a unique equilibrium exists in our model. This finding is different from other studies of time inconsistency in preferences, because we assume the consumer is myopic during the consumption stage. The optimal
consumption thus can be solved as in equation (3) and, as a result, the optimal plan choice in equation (5) is reduced to a standard dynamic programming problem.

The unique optimal policy function that solves the dynamic problem in equation (5) is a function of $h$ and model parameters. First, define the following variables:

$$A = \frac{\alpha_{Ch}^2}{4\alpha_{cc}}; \; e = (1 - \delta) + \frac{\alpha_{ch}}{2\alpha_{cc}};$$

$$B_0 = \frac{(a_c - \mu p_c - \gamma) \alpha_{ch}}{2\alpha_{cc}}; \; C_0 = \frac{(a_c - \mu p_c)(a_c - \mu p_c - 2\gamma)}{4\alpha_{cc}}; \; f_0 = \frac{a_c - \mu p_c}{2\alpha_{cc}};$$

$$B_1 = \frac{(a_c - \gamma) \alpha_{ch}}{2\alpha_{cc}}; \; C_1 = \frac{a_c(a_c - 2\gamma)}{4\alpha_{cc}} - \mu p_s; \; f_1 = \frac{a_c}{2\alpha_{cc}}.$$

Let

$$\begin{align*}
a_0 &= \frac{A}{1 - \beta e^2}, \\
b_0 &= \frac{B_0 + 2\beta e f_0 a_0}{1 - \beta e}, \\
c_0 &= \frac{C_0 + \beta a_0 f_0^2 + \beta b_0 f_0}{1 - \beta}.
\end{align*}$$

and

$$\begin{align*}
a_1 &= \frac{A}{1 - \beta e^2}, \\
b_1 &= \frac{B_1 + 2\beta e f_1 a_1}{1 - \beta e}, \\
c_1 &= \frac{C_1 + \beta a_1 f_1^2 + \beta b_1 f_1}{1 - \beta}.
\end{align*}$$

Finally, define the “cut-offs” as

$$h_1^{ss} = \frac{a_c - \mu p_c}{2\delta \alpha_{cc} - \alpha_{ch}}, \; h_2^{ss} = \frac{a_c}{2\delta \alpha_{cc} - \alpha_{ch}}, \; \text{and}$$
\[ h_1^* = \frac{(c_0 - c_1)}{(b_1 - b_0)}; \]
\[ h_2^* = \frac{(1-\beta)(c_0 - c_1) + \beta(b_1 - b_0)f_1}{(1-\beta)e(b_1 - b_0)}, \]
\[ h_2^{**} = \frac{(1-\beta)(c_1 - c_0) + \beta(b_1 - b_0)f_0}{(1-\beta)e(b_1 - b_0)}. \]

We use the standard recursive method to solve for the value function and the optimal policy function, and obtain the following proposition:

**Proposition 1:** Assuming parameters \( \alpha_c, \alpha_{ch}, \) and \( \lambda \) are nonnegative, the following unique equilibrium exists:

1A. If \( h_1^{ss} \leq h_1^* < h_2^{ss} \),
\[ s^*(h) = \begin{cases} 1 & \text{for } h > h_1^* \\ 0 & \text{for } h \leq h_1^* \end{cases} \]

1B. If \( h_2^{ss} \leq h_1^* \),
\[ s^*(h) = \begin{cases} 1 & \text{for } h > h_2^* \\ 0 & \text{for } h \leq h_2^* \end{cases} \]

1C. If \( h_1^* < h_1^{ss} \),
\[ s^*(h) = \begin{cases} 1 & \text{for } h > h_2^{**} \\ 0 & \text{for } h \leq h_2^{**} \end{cases} \]

**Proof:** See Appendix A.1.

Proposition 1 establishes that a unique equilibrium always exists. Depending on the parameter condition listed in 1A, 1B, and 1C, the cutoff points \( h_1^*, h_2^*, \) and \( h_2^{**} \) will vary, and thus the optimal plan choice will be different. Following the equilibrium plan choice, the consumption

---

\(^{30}\) We require steady states to be nonnegative in the proof. This requirement is not necessary for the general setup.
level follows equation (3). Because now we have an analytical solution to our model, given an initial value $h$, we can predict how it will evolve and where the steady state is. Based on Proposition 1, we can derive how the plan choice and consumption level will converge to the steady-state equilibrium as follows:

**Lemma:** Based on Proposition 1, with any initial habit stock $h_0$, the following steady-state equilibria exist:

i) if the condition in 1A is satisfied,

$$
\lim_{t \to \infty} h_t = h_1^{ss} \cdot \{h_0 < h_1^*\} + h_2^{ss} \cdot \{h_0 \geq h_1^*\};
$$

ii) if the condition in 1B is satisfied,

$$
\lim_{t \to \infty} h_t = h_1^{ss}
$$

iii) if the condition in 1C is satisfied,

$$
\lim_{t \to \infty} h_t = h_2^{ss}
$$

The lemma essentially guarantees all steady states are globally convergent steady states. Figure 2.2 provides a graphic illustration. In each panel of the figure, the x-axis denotes the current-period habit stock, and the y-axis the next-period habit stock. The black solid line is the policy function which returns a unique value of next-period given current habit stock level, with arrows indicating how it evolves over time. The steady states are located at the intersection of the path and the 45-degree line, on which the state in the next period is the same as the current period. Under the condition in 1A in the top panel, if the habit stock starts below $h_1^*$, consumers will always choose pay-per-chapter, and $h$ will converge to the “low” steady state $h_1^{ss}$; otherwise, they will choose the monthly subscription, and $h$ will converge to the “high” steady state $h_2^{ss}$. In the second
panel, when the condition in 1B is satisfied, \( h \) will converge to the steady state \( h_{1}^{ss} \), regardless of where the initial consumption stock state is. If the initial state \( h_{0} \) is very high (> \( h_{2}^{*} \)), consumers first choose monthly subscription, then switch to pay-per-chapter when \( h \) drops below \( h_{2}^{*} \). The gap on \( h_{2}^{*} \) indicates the change in consumption level when they switch plans. The last panel in the figure demonstrates when the condition in 1C is satisfied. Regardless of where the initial consumption stock state is, consumers will eventually choose monthly subscription and converge to the high steady state \( h_{2}^{ss} \).
Figure 2.2 Dynamics of the Steady States

When no time-inconsistency exists (either \( \gamma = 0 \) or \( \gamma \cdot c \) also affects the utility function in the consumption stage), overpaying can never be the optimal choice. For any consumption amount above \( p_s / p_c \) chapters, consumers should always choose the monthly subscription. Under time-inconsistent preferences, however, consumers may overpay by choosing pay-per-chapter to prevent excessive consumption. The equilibrium characterized in proposition 1 provides a clear illustration on how time-inconsistent preferences and strategic self-control together can explain the unique data features we observe in section 2.3:

1. Consumers overpay by consistently choosing pay-per-chapter: In the first two panels of Figure 2.2, consumers choose the plan at the steady-state equilibrium even though \( h_{t+1}^{ss} \) can be at a level that is costlier for pay-per-chapter.
2. Consumers switch to pay-per-chapter but still overpay: The second panel indicates consumers in a high state will switch from subscription to pay-per-chapter and decrease their consumption toward \( h_{t+1}^{ss} \). During the process, consumers may overpay for pay-per-chapter. They may still overpay at the steady state \( h_{t+1}^{ss} \).
3. Asymmetric overpay patterns: We have shown consumers have incentive to overpay under
pay-per-chapter. It is easy to show that when consumers choose the monthly subscription (i.e., $s^*(h) = 1$ in Proposition 1), their reading amount will never be below $p_s/p_c$ chapters a month.

2.4.2 An Econometric Model

In the analytical model, the plan choice and consumption level are deterministic. In reality, however, observed data will not be perfectly aligned with model predictions due to unobserved factors. To explain fluctuations in the plan choice and consumption across and within individuals, we construct an econometric model so that such fluctuations can be estimated from data. The model is similar to the analytical model, but it incorporates stochastic components in the utility functions. Furthermore, we allow heterogeneous model parameters across consumers to capture the fact that some consumers’ reading preferences can be systematically different from the others. And we also allow the non-monetary cost of reading web fiction to change given different level of habit stock $h$.\(^{31}\)

For individual $i$ in period $t$, the utility function during the consumption stage that corresponds to equation (2) is modified as follows:

$$u_{it}(c_{it}, s_{it}, h_{it}) = (\alpha_{it} + \omega_{it} + \alpha_{i,eh} \cdot h_{it}) \cdot c_{it} - \alpha_{i,ce} \cdot c_{it}^2 - \mu_i \cdot p_c \cdot 1{s_{it} = 0} \cdot c_{it} (7)$$

In this function, $\omega_{it}$ represents the unobserved factors that may affect the marginal utility of consumption in each period. The individual-specific model parameters capture the heterogeneity across consumers.

The consumption level that corresponds to equation (3) therefore is

\(^{31}\) Since such change of model setting do not affect the consumption stage decisions when compared to the baseline model, it is straightforward to see all essential model properties and proof are unaffected even if we apply such change into the analytical model we discussed in the last section.
Note the consumption level cannot be negative; therefore, \( c^*_it(s_{it}, h_{it}) \) in equation (8) is bounded below by 0.

For the marginal cost of reading each chapter, we allow \( \gamma \) (see equation 4) to be heterogeneous among consumers. In addition, we allow the cost to change as the habit state increases. That is, \( \gamma_{it} = \gamma_{i0} + \gamma_{i1} \cdot h_{it} \). This setting implies that an individual’s time-inconsistency in terms of plan and consumption choices can be dynamically evolving.

We assume that when making the plan choice, the consumer only knows the distribution of \( \omega_{it} \) and not the exact value. The consumer will choose a plan that maximizes the expected value function. Two additional stochastic terms, \( e^0_{it} \) and \( e^1_{it} \), will affect the utility of choosing pay-per-chapter and monthly subscription, respectively. Corresponding to equations (5) and (6), the consumer’s dynamic problem of plan choice is specified as follows:

\[
V_{it}(h_{it}, e^0_{it}, e^1_{it}) = \max_{s_{it}=0,1} E_\omega u_{it}(c^*_it(s_{it}, h_{it}), s_{it}, h_{it}) - \mu_t \cdot p_s \cdot 1\{s_{it} = 1\}
\]

\[
- (\gamma_{i0} + \gamma_{i1} \cdot h_{it}) \cdot c^*(s_{it}, h_{it})
\]

\[
+ e^0_{it} \cdot 1\{s_{it} = 0\} + e^1_{it} \cdot 1\{s_{it} = 1\}
\]

\[
+ \beta \cdot E_\omega E_e V_{i,t+1}(h_{i,t+1}, e^0_{i,t+1}, e^1_{i,t+1}),
\]

where

\[
h_{i,t+1} = (1 - \delta)h_{it} + c^*(s_{it}, h_{it}).
\]
The expectation operator $E_\omega$ in equation (9) integrates over $\omega_{lt}$ and another $E_e$ integrates over $e_{lt,t+1}^0$ and $e_{lt,t+1}^1$ in the next period. Given $\omega_{lt}$, the expected value function in the third line of equation (9) can be specified as

$$E_e V_{i,t+1}(h_{i,t+1}, e_{i,t+1}^0, e_{i,t+1}^1) = E_e \max \{V_{i,t+1}^0(h_{i,t+1}, e_{i,t+1}^0), V_{i,t+1}^1(h_{i,t+1}, e_{i,t+1}^1)\},$$

(11)

where $V_{i,t+1}^0$ and $V_{i,t+1}^1$ represent the value function conditional on choosing pay-per-chapter and monthly subscription, respectively. In the empirical application, we assume $\omega_{lt}$ is distributed as $N(0, \sigma_\omega^2)$, where $\sigma_\omega^2$ is the variance, and is independent and identically distributed across consumers and periods. Furthermore, $e_{lt}^0$ and $e_{lt}^1$ are extreme-value type I distributed with zero location parameter and a scale parameter $\tau$.

In the value function, the state variable is $h_{lt}$. Suppose the state space is a closed interval on $\mathbb{R}^1$ denoted by $[0, H]$. We discretized the state space into $N$ grid points, and assume $h_{lt}$ is constant within an interval $\left[\frac{(k-1)H}{N}, \frac{kH}{N}\right]$, where $k \in \{1, ..., N\}$. Based on the distribution assumption of $\omega_{lt}$ and our model setting, we can derive that given the plan choice $s_{lt}$ and current state $h_{lt}$, the conditional distribution of $h_{i,t+1}$ follows a truncated normal distribution $g$ with support on $[0, H]$, with mean $\mu(h_{lt}, s_{lt}) = (1 - \delta)h_{lt} + \frac{\alpha_c + \alpha_{ch} h_{lt} - \mu_p c^{-1}[s_{lt} = 0]}{2 \cdot \alpha_{cc}}$ and variance equal to $\frac{\sigma_\omega^2}{4 \alpha_{cc}}$. The distribution satisfies the Markov property of memorylessness. Let $\theta$ be the collection of model parameters. The distribution function of $h_{i,t+1}$, unconditional on $s_{lt}$, is

$$f(h_{i,t+1}|h_{lt}, \theta) \sim p(s_{lt} = 0|h_{lt}, \theta) g\left(\mu(h_{lt}, 0|\theta), \frac{\sigma_\omega^2}{4 \alpha_{cc}}\right) +$$

$$p(s_{lt} = 1|h_{lt}, \theta) g\left(\mu(h_{lt}, 1|\theta), \frac{\sigma_\omega^2}{4 \alpha_{cc}}\right),$$

76
where \( p \) is the probability of choosing a pricing plan. The probability of \( h \) in interval \( m \) falling into another interval \( n \) in the next period therefore is

\[
p^{1}_{mn}(\theta) = \frac{n^H}{N} \int_{\frac{(n-1)H}{N}}^{\frac{nH}{N}} f'(h| h \in m, \theta) dh'.
\]

Denote the transition matrix

\[
M^1(\theta) = \begin{bmatrix}
p^{1}_{11}(\theta) & \cdots & p^{1}_{1N}(\theta) \\
\vdots & \ddots & \vdots \\
p^{1}_{N1}(\theta) & \cdots & p^{1}_{NN}(\theta)
\end{bmatrix}.
\]

Similarly, denote the transition probability from state \( m \) to state \( n \) after \( s \) periods as \( p^{s}_{mn}(\theta) \) and similarly the transition matrix as \( M^s(\theta) \). We have the following proposition:

**Proposition 2:** For any model parameters \( \theta \), there exists a unique stationary distribution (or limiting distribution) equilibrium for the state. That is, \( \lim_{s \to \infty} p^{s}_{mn}(\theta) = p_n(\theta) \) exists for any \( m, n = 1, 2, ..., N \). Let \( p(\theta) = (p_1(\theta), p_2(\theta), ..., p_N(\theta)) \), where \( p_n(\theta) \geq 0 \), \( \sum_{n=1}^{N} p_n(\theta) = 1 \), under the stationary distribution equilibrium, we have \( p(\theta)M^1(\theta) = p(\theta) \). This implies

\[
\lim_{s \to \infty} M^s(\theta) = \begin{bmatrix}
p_1(\theta) & p_2(\theta) & \cdots & p_N(\theta) \\
\vdots & \ddots & \vdots \\
p_1(\theta) & p_2(\theta) & \cdots & p_N(\theta)
\end{bmatrix}.
\]

**Proof:** See Appendix A.2.

Proposition 2 guarantees that, for any set of model parameters, a unique stationary distribution for \( h \) exists. Therefore, regardless of the initial distribution, a unique distribution of \( h \)
will exist under a sufficiently large number of iterations. This property helps us solve the initial value problem in the model estimation.

### 2.4.3 Model Estimation

Based on the assumption that $e_{it}^0$ and $e_{it}^1$ are extreme-value type I distributed with zero location parameter and a scale parameter $\tau$, we can rewrite

$$E_e V_{i,t+1}(h_{i,t+1}, e_{i,t+1}^0, e_{i,t+1}^1) = \tau \cdot r + \tau \cdot \ln \left( \sum_{s=\{0,1\}} \exp \left( \frac{\bar{V}_i^s(h_{i,t+1})}{\tau} \right) \right), \quad (12)$$

where $r$ is the Euler constant, and

$$\bar{V}_i^s(h_{i,t+1}) = E_\omega u_{it}(c_{i,t+1}^*(s_{i,t+1} = s, h_{i,t+1}), s_{i,t+1} = s, h_{i,t+1}) \nonumber$$

$$-(\gamma_{i0} + \gamma_{i1} \cdot h_{it}) \cdot c_{i,t+1}^*(s_{i,t+1} = s, h_{i,t+1}) + \beta \cdot E_\omega \left( \tau \cdot r + \tau \cdot \ln \left( \sum_{s'_{i,t+2}} \exp \left( \frac{\bar{V}_i^{s'}(h_{i,t+2})}{\tau} \right) \right) \right). \quad (13)$$

Substitute equation (12) into equation (9), we can rewrite the dynamic plan-choice problem as

$$V_{it}(h_{it}, e_{it}^0, e_{it}^1) = \max \{\bar{V}_i^0(h_{it}) + e_{it}^0, \bar{V}_i^1(h_{it}) + e_{it}^1\}, \quad (14)$$

where

$$\bar{V}_i^s(h_{it}) = E_\omega u_{it}(c_{it}^*(s_{it} = s, h_{it}), s_{it} = s, h_{it}) - (\gamma_{i0} + \gamma_{i1} \cdot h_{it}) \cdot c^*(s_{it} = s, h_{it})$$

$$+ \beta \cdot E_\omega \left( \tau \cdot r + \tau \cdot \ln \left( \sum_{s'_{i,t+1}} \exp \left( \frac{\bar{V}_i^{s'}(h_{i,t+1})}{\tau} \right) \right) \right).$$

78
Proposition 1 shows that for the analytical model, a unique equilibrium exists; therefore, the value function \( \bar{V}_i(h_{it}) \) can be solved through the iteration method. For the econometric model, however, the existence of the unique equilibrium is difficult to prove. For each \( h_{it} \) in the state space, we use different initial values for \( V_i(h_{it}) \) in the model estimation, and find the iteration method always converges to the same value. Therefore, we assume \( \bar{V}_i(h_{it}) \) will converge to the unique equilibrium in practice even with the stochastic components.

Given the distribution assumption for \( e_{it}^0 \) and \( e_{it}^1 \), Rust (1987) shows the plan-choice probability has the following analytical expression:

\[
P_{it}^s(h_{it}) = \frac{\exp(\bar{V}_i^s(h_{it})/\tau)}{\exp(\bar{V}_i^0(h_{it})/\tau) + \exp(\bar{V}_i^1(h_{it})/\tau)}
\]

Let \( c_{it} \) be the observed consumption level that is bounded below by zero. Conditional on the plan choice \( s_{it} \) and the assumption that \( \omega_{it} \) is distributed as \( N(0, \sigma_\omega^2) \), the likelihood of observing \( c_{it} \) can be derived from equation (8) as follows:

\[
P_{it}^{c|s}(c_{it}|s_{it}, h_{it}) = \begin{cases} 
1 - \Phi(\alpha_{ic} + \alpha_{i, ch} \cdot h_{it} - \mu_i \cdot p_c \cdot 1\{s_{it} = 0\}), & \text{if } c_{it} = 0 \\
\phi \left( \frac{2 \cdot c_{it} \cdot \alpha_{i, cc} - \alpha_{ic} + \alpha_{i, ch} \cdot h_{it} - \mu_i \cdot p_c \cdot 1\{s_{it} = 0\}}{2 \cdot a_{i, cc}} \right), & \text{if } c_{it} > 0
\end{cases}
\]

where \( \phi \) and \( \Phi \) are the probability density function and cumulative density function of the normal distribution with mean zero and variance \( \sigma_\omega^2 \).

Combining the likelihoods in (15) and (16), the full likelihood function for observation \((c_{it}, s_{it})\), conditional on the habit stock \( h_{it} \) and individual-specific model parameters \( \theta_i \), is
To evaluate the likelihood function, however, we need to solve an initial value problem, because the consumption stock state when the sample period starts, \( h_{t0} \), is unobserved in the data. Given individual-specific model parameters, \( h_{t0} \) can be systematically different across consumers. Ignoring this problem can lead to biased model estimates. To deal with the problem, recall that Proposition 2 guarantees the existence of a stationary distribution of \( h \). We assume at the beginning of period 1 the state variable \( h_{t0} \) of each individual comes from the stationary distribution. We simulate the stationary distribution as a function of the individual-specific model parameters, and draw \( h_{t0} \) multiple times from this distribution. We then calculate the likelihood function conditional on the simulated \( h_{t0} \), and finally take the average of the likelihoods across draws to obtain the simulated likelihood. Given a trial value \( \theta \), the detailed estimation procedure is as follows:

1. We first numerically solve the value function \( V_{it}(h_{it}, e_{it}^0, e_{it}^1) \) defined in equation (14). We set the state space for habit stock \( h_{it} \) as \([0, 2000]\).\(^{32}\) We discretize the state space into grids with length of 10, and linearly interpolate the value function within the range.

2. Next, we calculate the stationary distribution of the state variable \( h \). With the numerical solution of the value function, we can calculate the value of the plan-choice probability \( P^s(h) \) with equation (15) for any given \( h_{it} \). Define interval \([10(m-1), 10m]\) as state \( m \) for \( h \). Starting from any value of \( h_m \) with state \( m \), the probability for \( h \) to transfer from state \( m \) to state \( n \), \( p_{mn}(\theta) \), can be calculated by

\[
L(c_{it}, s_{it} | h_{it}, \theta_i) = P^1_{it}(h_{it}, \theta_i)^{s_{it}=1} \cdot P^0_{it}(h_{it}, \theta_i)^{s_{it}=0} \cdot P^{\epsilon|s}_{it}(c_{it} | s_{it}, h_{it}, \theta_i). \tag{17}
\]

\(^{32}\) We choose the maximum state space as 2,000 because it covers most of the data range (95%). We have also tried other values as large as 10,000 and the estimation results are robust.
\[ p_{mn}(\theta) = \sum_{s \in \{0,1\}} P^s(h_m) \int_{0(n-1)}^{10n} g\left( \mu(h_n, s|\theta), \frac{\sigma_w^2}{4\alpha_{i,c}} \right) dh_n. \]

Define the transfer matrix as

\[ M^1(\theta) = \begin{bmatrix} p_{11}(\theta) & \cdots & p_{1,200}(\theta) \\ \vdots & \ddots & \vdots \\ p_{200,1}(\theta) & \cdots & p_{200,200}(\theta) \end{bmatrix}. \]

Proposition 2 shows a unique nonnegative vector \( P(\theta) = (p_1(\theta), p_2(\theta), \ldots, p_{200}(\theta)) \) exists such that

\begin{enumerate}
\item \( \sum_{n=1}^{200} p_n(\theta) = 1 \)
\item \( \lim_{n \to \infty} (M(\theta))^n \begin{bmatrix} P(\theta) \\ \vdots \\ P(\theta) \end{bmatrix} = \begin{bmatrix} P(\theta) \\ \vdots \\ P(\theta) \end{bmatrix} \)
\end{enumerate}

To compute the limit, we calculate \((M(\theta))^n\) recursively until \( |M(\theta)^{n+1} - M(\theta)^n| < 0.01 \), and obtain the stationary distribution \( P(\theta) \).

3. Finally, we take 50 draws of the initial value \( \{h_{i0}^n\} \) for each individual. With an initial \( h_{i0}^n, h_{it} \) can be computed from the monthly reading amount. This approach enables us to calculate the simulated maximum likelihood function as

\[ L(\bar{c}_{it}, \bar{s}_{it}, h_{it}) = \prod_{t} \frac{1}{50} \sum_{n=1}^{50} \prod_{t} P^{\bar{s}_{it}}(h_{it}^n)P^{1-\bar{s}_{it}}(h_{it}^n)P_{ict}(\bar{c}_{it}, \bar{s}_{it}, h_{it}^n). \]

During the parameter search, we use the gradient optimization method (BFGS). After it converges, we then switch to the Nelder-Mead numerical optimization. We repeat the procedure until the increase in the log likelihood becomes trivial (<1e-2). From different start points, we find such an algorithm can effectively reach the same global optimum without falling into some local optimum when we only use one method.
2.4.4 Identification

For ease of discussion, we ignore the heterogeneity in model parameters across consumers. The parameters in the model include all parameters in the utility functions, and the variance for \( \omega \)'s and the scale parameter for \( e \)'s, that is, \( \{\alpha_c, \alpha_{cc}, \alpha_{ch}, \gamma_0, \gamma_1, \mu, \tau, \sigma_\omega\} \). Two additional parameters, \( \{\beta, \delta\} \), represent the discounting factor and the depreciation rate of the consumption stock state. Following the previous literature, we fix \( \beta = 0.98 \) as the monthly discounting factor. This value is equivalent to the daily discounting factor 0.998, as suggested in DellaVigna and Malmendier (2006). For the depreciation rate \( \delta \), we find from practice that it is difficult to be separately identified from \( \alpha_{ch} \), the parameter that captures the effect of the consumption stock state on the marginal utility of consumption. During the estimation, we vary \( \delta \) from 0.1 to 0.9, and choose the one (\( \delta = 0.8 \)) that maximizes the likelihood function. We test the robustness by varying the value of the two parameters, and find the main results remain unchanged.

Parameters in the consumption utility function, including \( \{\alpha_c, \alpha_{cc}, \alpha_{ch}, \sigma_\omega\} \), are identified from the data on monthly reading amounts. For those who choose monthly subscription, equation (8) shows that multiplying a constant on \( \alpha_c, \alpha_{cc} \), and \( \alpha_{ch} \) will not change the consumption level. Therefore, we normalize \( \alpha_{cc} = 0.5 \). The optimal consumption level thus is \( c_{lt} = \max\{\alpha_c + \alpha_{ch} h_{lt} + \omega_{lt}, 0\} \). Given the consumption states across consumers, we can identify parameters \( \alpha_c, \alpha_{ch}, \) and the variance \( \sigma_\omega \).

The price coefficient \( \mu \) is identified from the difference in the reading amount for the same individual after she switches pricing plans. Suppose the consumer switches from monthly subscription to pay-per-chapter and the reading amount significantly drops. This finding would suggest she has a higher value of \( \mu \). Given the parameters in the consumption utility, the disutility parameters from excessive consumption, \( (\gamma_0, \gamma_1) \), are identified by the proportion of consumers
who overpay for pay-per-chapter, relative to the proportion of consumers who overpay under monthly subscription across different level of habit stock. In other words, the identification comes from the asymmetric overpay pattern across consumers with different consumption pattern history. The larger the proportion of consumers who read more than 120 chapters per month and do not choose monthly subscription, the larger \( \gamma_0 \) is among consumers. And the larger overpay amount for consumers who have higher habit stock after controlling for the effect of \( \gamma_0 \), the larger \( \gamma_1 \) is among consumers.

Finally, the scale parameter \( \tau \) is identified from the plan choice and the corresponding reading amount. Suppose, given all other model parameters, the expected value from one pricing plan is higher than the other. If the probability that consumers choose the high-valued plan is not much higher than the low-valued plan, this scenario implies a large \( \tau \). Note that \( \tau \) cannot be identified if we only observe the plan choice. We need the reading amount together with the plan choice to pin down the parameter.

### 2.5 Results

In this section, we first discuss the estimation results of the proposed model. We will then discuss the estimation results from an alternative model under which there are no time-inconsistent preferences and another model assuming consumers are naïve in forming expectation about their consumption choice. Based on the proposed model, we use counterfactuals to show how changing the pricing plan options available to consumers would affect the consumer welfare and the platform’s profit. Finally, we show that a new nonlinear pricing plan with volume surcharge could simultaneously improve the consumer welfare and the platform’s profit. Our findings help shed light on firms’ pricing strategies for product or service categories for which consumers’ strategic self-control behaviors are prevalent.
### 2.5.1 Estimation Results

Table 2.4 Main Estimation Results

beta=0.98  Observation=68,7068

<table>
<thead>
<tr>
<th></th>
<th>Segment 1</th>
<th>s.e.</th>
<th>Segment 2</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_c$</td>
<td>-124.20</td>
<td>1.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{ch}$</td>
<td>0.86</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>7.61</td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>124.36</td>
<td>33.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.57</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\log \sigma_w$</td>
<td>6.57</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\log \sigma_e$</td>
<td>8.24</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Log likelihood: -439259  AIC 878532  BIC 878595

<table>
<thead>
<tr>
<th></th>
<th>Segment 1</th>
<th>s.e.</th>
<th>Segment 2</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_c$</td>
<td>-124.33</td>
<td>2.29</td>
<td>0.06</td>
<td>11.31</td>
</tr>
<tr>
<td>$a_{ch}$</td>
<td>0.68</td>
<td>0.01</td>
<td>0.87</td>
<td>0.01</td>
</tr>
<tr>
<td>$\mu$</td>
<td>9.31</td>
<td>1.72</td>
<td>0.57</td>
<td>0.13</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>124.38</td>
<td>88.71</td>
<td>124.32</td>
<td>153.87</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>2.08</td>
<td>0.21</td>
<td>0.29</td>
<td>0.10</td>
</tr>
<tr>
<td>$\log \sigma_w$</td>
<td>6.52</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\log \sigma_e$</td>
<td>8.24</td>
<td>0.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class size</td>
<td>0.54</td>
<td></td>
<td>0.46</td>
<td></td>
</tr>
</tbody>
</table>

Log likelihood: -430996  AIC 862019  BIC 862138
We use a latent class approach to model the individual heterogeneity in the utility function. We start from a one-class model and keep increasing the number of classes. We find the Bayesian information criterion (BIC) stops increasing when the number of classes reaches three. Furthermore, the size of the third class becomes very small (only about 3%). Therefore, we choose the three-class specification as our main results. For comparison Table 2.4 reports the estimation results from one- to three-class specifications, but we will focus our discussion on the three-class model.

Based on the plan choice and consumption behavior of the three segments implied from the estimation results, we name these three consumer segments as self-controllers, powerless habitual readers, and enthusiastic readers.
There are two large latent segments, 1 and 2, that are similar in size (49% and 47%, respectively), and another very small segment 3 (3%).\(^{33}\) The plan choice and consumption behaviors are significantly different among the three segments. The coefficient \(\alpha_c\) for segment 3 is the largest, indicating consumers of this segment have the highest reading preference. The coefficient \(\alpha_{ch}\) for all three segments is positive, implying the habit stock will enhance consumers’ marginal utility for consumption. The magnitude is also the highest for segment 3.

The price coefficient \(\mu\) is positively significant for all three segments; however, segment 1 is much more price-sensitive than the other two segments. This implies that the reading amount of consumers in this segment will decrease much more than that of the other two segments when facing a plan with a high marginal cost (e.g., pay-per-chapter). Note that \(\mu\) represents the price sensitivity during consumption, which can be lower than that in the pricing plan choice stage. The difference—if it exists—will be captured in a reduced-form way by \(\gamma_0\) and \(\gamma_1\). The larger the difference in the price sensitivity across different habit stock levels, the higher the value of \(\gamma_0\) and \(\gamma_1\). In Table 2.4, the large estimated \(\gamma_0\) across all three segments suggests that the true monetary or non-monetary cost of reading web fictions has been seriously under-evaluated by all consumers during the consumption stage. Worse still, the positive estimated \(\gamma_1\) suggests that as the habit stock increases, the cost will be under-evaluated even more. Therefore, consumers during the pricing plan choice stage have an incentive to self-control their future consumption.

How can they control the consumption level? Obviously choosing the pay-per-chapter plan is a feasible option, yet this option works differently across consumers. For consumers of segment 1, because of their high price sensitivity, the reading amount will significantly drop if they switch

\(^{33}\) The percentages reported are rounded up to the nearest percentage point.
from the monthly subscription to the pay-per-chapter plan. For consumers of the other two segments, however, their much smaller price sensitivities imply that their reading amount will not change much after the switch. As an illustration, we assume that for every consumer, the habit stock is at zero level, and calculate the predicted reading amount of each segment. The average monthly reading amount of a consumer in segment 1 is 173.2 chapters, with a 53% probability of no reading, under the monthly subscription plan. This drops to 133.5 chapters per month, with a 59% probability of no reading, under the pay-per-chapter plan. In contrast, for a consumer of segment 2, the reading amount under monthly subscription is 175.7 chapters per month, with a 52.5% probability of no reading. Switching to the pay-per-chapter plan will only marginally reduce the reading amount.

The difference in how much reading amount can be reduced will have the main impact on the pricing plan choice. We find that the probability of choosing the pay-per-chapter plan is close to 100% among consumers of segment 1. The reading amount is 631 chapters under the monthly subscription and 435 chapters under the pay-per-chapter plan, across consumers under different levels of habit stock in our data. With the 0.1 RMB/chapter price under pay-per-chapter, an average consumer in segment 1 pays 43.5 RMB for his/her monthly reading, while the monthly subscription would cost them 12 RMB. This suggests that an average consumer in segment 1 is willing to pay 31.5 RMB (about 4.6 U.S. dollars) more to reduce their monthly reading amount by 196 chapters. Assuming it takes five minutes to read one chapter, our results imply that consumers of this segment are willing to pay about a little more than four dollars per month to cut their time on web fiction by about 16 hours per month (or about half an hour per day). In contrast, the probability of choosing the pay-per-chapter plan among consumers of segment 2 gradually drops from 55% to 29%, as the habit stock increases from 0 to 2,000 (this is the range among the majority
of consumers estimated from the model), indicating a much weaker willingness to conduct self-control. Given that these consumers are incapable of controlling the reading amount, we find that assuming their habit stock starts from zero, it will increase rapidly each month and eventually stabilize at a level much higher than that of consumers in segment 1. Consequently, their consumption level will also stabilize at a much higher level.

Because of the differences in consumption behaviors and pricing plan choices—and the subsequent changes in the habit stock—between the two segments, we call consumers of segment 1 self-controllers and consumers of segment 2 powerless habitual readers. Due to their large size, these two segments are the most important consumers from the platform’s profit perspective.

To conclude, our estimation results demonstrate that a high cost is associated with reading books, which consumers in the consumption stage do not consider. This leads to the time inconsistency problem that incentivizes consumers to use self-control strategies when they make the plan choice. We find that about half the number of consumers are self-controllers. They choose pay-per-chapter, and most of them overpay. The reason is that this segment is price-sensitive during consumption; therefore, paying for each chapter as a self-control strategy would be effective in curbing future consumption. Consequently, price-sensitive consumers are more likely to overpay. This result is counterintuitive because previous economic and marketing studies have found that, all else being equal, price-sensitive consumers will engage in more cost-saving purchase and consumption strategies (e.g., taking advantage of price promotions, searching more for price information). However, we show that when strategic self-control is an important goal in the purchase decision, the result can reverse.
2.5.2 Model Fit and Alternative Model Specifications
To investigate the model fit, we simulate the plan choice and reading amount for each individual in the data for six time periods, using the estimation results from the proposed model. We repeat the simulation process with 50 different draws from the stationary distribution.

The upper panel of Figure 2.3 illustrates that the simulated average reading amount for all readers is 512.1 chapters per month, whereas it is 506.4 chapters in the actual data. The lower panel shows that the average probability of choosing pay-per-chapter in the simulation is 73.38% across six months, whereas it is 74.51% in the data. These comparisons indicate that our model predictions fit the actual data pattern very well.

Average reading amount
Our model can also replicate the unique overpaying data patterns we discussed in section 2.2. Figure 2.4 compares the overpaying probability under each pricing plan predicted by the model and the actual data. Our model predicts that around 6.4% of consumers overpay under monthly subscription by reading less than 120 chapters, and 45.8% consumers overpay under the pay-per-chapter plan by reading more than 120 chapters. In the data, the overpaying proportions are 6.6% under the monthly subscription and 34.2% under pay-per-chapter. Our model replicates the asymmetrical overpay pattern.
Figure 2.4 Model Fit: overpaying ratio

For comparison purpose, we estimate two alternative models. The first assumes time-consistent agents, that is, the consumers’ reading amount choice during the consumption stage is consistent with the choice during the pricing plan choice stage. This is similar to the empirical implementation of the rational addiction theory proposed in Becker, Grossman, and Murphy (1994). The second one is a model with naive time-inconsistent consumers, that is, during the pricing plan choice stage, they assume their choice during the consumption stage does not deviate (but indeed it does). Estimation results for time-consistent consumers and naive time-inconsistent consumers are reported in Appendix B (Table B1 and Table B2).

We first look at the predictions from the rational model with time-consistent consumers. This model shows worse performance in terms of likelihood and Akaike information criterion/BIC compared to our main estimation results. The parameters and the estimation procedure are the
same except that there are no time-inconsistent parameters ($\gamma_0, \gamma_1$ in our full model). The estimation also gives us three consumer segments. However, this time 94% of the total population has a very low price coefficient of 0.34 in consumer segment 1, while the other two segments with higher price sensitivities only account for 3% of the consumers. Such results are significantly different from our main estimation results where almost half of the consumers have relatively high price sensitivity and thus willing to conduct self-control. Without time-inconsistent preference, there is no incentive for consumers to choose the costlier pricing plan to curb future consumption. Essentially, these estimation results attribute the reason why consumers overpay for their reading to very low price sensitivity, which cannot explain the asymmetry in overpaying pattern presented in the previous section. With 94% consumers caring very little about the monetary cost, this model predicts that the average choice probability of the monthly subscription plan is 62.5%, far higher than 25.5% in the data. The model also predicts the average monthly reading amount to be 743 chapters, well beyond the 506.4 chapters in data. Finally, the predicted proportions of overpaying under the pay-per-chapter plan and monthly subscriptions are 23% and 16%, respectively. It fails to replicate the high overpaying ratio for the pay-per-chapter plan. These results show that, without accounting for the time-inconsistent preferences, a standard rational economic model cannot explain the unique behavioral patterns we observe from data.

For the model with naive consumers, we find that the three-class model has a slightly higher log-likelihood value or lower BIC in comparison with the proposed model. Although the model fit is better, we believe that this is unlikely to be the true model that explains the unique behavioral patterns in data. For example, estimation results show a large consumer segment similar to the self-controllers in our proposed model. Using simulations, we find that consumers of this segment when making the pricing plan choice expect to read 220 chapters in each month; however, in
reality, they consistently read more than 800 chapters. Such under-prediction bias also exists for
the other two segments. Overall, the model attributes the overpaying pattern to the fact that more
than 60% consumers systematically underestimate their monthly reading amount by 300 to 800
chapters a month. Given the consistently high consumption level across individuals, and the ease
of tracking the reading history, such a considerable underestimation persisting over many months
is unlikely to happen in our data.

We further test the robustness of the estimation results under other model specifications. We vary the value of $\delta$, the depreciation rate in the habit formation, from 0.1 to 0.9 and re-estimate the proposed model. Though other parameters and model implications remain the same, we find that the lower the value of $\delta$, the larger is the estimated value of $\alpha_{ch}$. However, the multiplication $\delta \cdot \alpha_{ch}$ remains stable. We also test different values for the time discounting factor $\beta$, and extend the range of the state space to [0, 10,000], which covers 99.9% of the data. All the results are similar to the results reported in Table 2.4.

**2.5.3 Counterfactuals**

To illustrate the substantive implications of our study, we investigate four counterfactual scenarios. The first two scenarios restrict the pricing plans that consumers can choose, and the last two scenarios increase consumers’ options by separately introducing a nonlinear pricing plan with volume discount and another with volume surcharge. We use random draws from the stationary distribution calculated from our estimation result as the initial consumption stock state. In each scenario, we simulate the plan and reading choices of consumers from each segment, until the reading distribution becomes stable after the policy change. We then compare the difference in

---

34 We simulate the data for 12 periods under all scenarios and use the last six months for comparison. We find that the reading-amount distribution typically stabilizes after 2 to 4 months.
the monthly average reading amount, consumer welfare, and the platform’s profit (measured by the revenue). It is straightforward to measure the change in reading amount and revenue under new pricing policies. To measure consumer welfare, we follow the “dictator of the present” method (see Cropper and Laibson, 1998; Caplin and Leahy, 2000; Gruber and Köszegi, 2001). We measure a consumer’s welfare as the discounted value of his life-time utility flow under the model equilibrium. Given that a unique stationary distribution exists in our model, we randomly draw the initial state from the stationary distribution for each consumer segment 50 times and calculate the average of the value functions. To compute the change in overall consumer welfare across different consumer segments, we calculate the weighted average of the percentage change for each consumer segments.

<table>
<thead>
<tr>
<th>Table 2.5 Counterfactual Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Monthly Subscription Only</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Reading amount</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Firm revenue</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Consumer welfare</strong></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Pay-per-chapter only (Linear Pricing)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reading amount</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Firm revenue</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Consumer welfare</strong></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
The first counterfactual restricts consumers’ option to monthly subscription only. This scenario is reminiscent of the practice of Netflix, which only offers the monthly subscription plan for watching movies. Consumers choose whether to subscribe or not. The results are in the first panel of Table 2.5. As expected, the reading amount of self-controllers will increase by 44.41% (from 441 chapters to 637 chapters) per month, when they do not have the option of choosing the pay-per-chapter plan to curb consumption. Across all three segments, the reading amount increases by 16.82%. Without the pay-per-chapter plan, the welfare of self-controllers decreases by 76.06%. The platform’s profit from this segment also drops by 79.52%. Overall, the consumer welfare drops by 40.54% and the platform’s profit drop by 76.95%. The reason that higher reading amounts generate much less profit for the platform is that consumers who used to read around 600 to 700
chapters a month and still pay by chapters now have to switch to the monthly subscription. As a result, the average revenue per consumer drops from 60~70 RMB to 12 RMB per month.

In the second counterfactual, the platform only offers the pay-per-chapter plan. This scenario mimics the video game market, in which players have to pay for each game title. The second panel of Table 2.5 reports the results. Because almost all self-controllers choose pay-per-chapter, the policy change has a very limited effect on them. Powerless habitual readers and enthusiastic readers are more affected. Given that the consumers of these two segments are less price-sensitive, they will continue to maintain the high consumption under the pay-per-chapter plan, and therefore, will pay significantly more to the platform. Consequently, the consumer welfare of powerless habitual readers will decrease by 9.74%. Overall, consumer welfare will reduce by 4.51%. The profit of the platform will significantly increase by 46.35%, mostly from the powerless habitual readers.

The above two counterfactuals suggest that taking away options from consumers will result in a decrease in consumer welfare. We further investigate if there exists a pricing policy that could improve both the platform’s profit and consumers’ well-being. In the next two counterfactuals, we introduce a step-wise pay-per-chapter pricing structure, where the price is a nonlinear function of the reading amount. In the third counterfactual, we introduce a novel volume discount plan with a nonlinear structure. Under this plan, consumers pay 0.12 RMB for the first 400 chapters every month, after which the price decreases by 0.04 RMB. The 400-chapter threshold is chosen based on the median reading amount under the pay-per-chapter scenario. As the pay-per-chapter plan charges 0.1 RMB, the new plan is more expensive if a consumer reads below 400 chapters. The

---

35 For enthusiastic readers, they are very price insensitive and have a high reading preference, hence, their welfare is not much affected.
third panel of Table 2.5 reports the results. Offering the additional pricing plan, the platform’s profit will decrease by 45.48%. This is because consumers can switch from the pay-per-chapter plan to the volume discount plan and pay less for reading. Though consumers pay less, consumer welfare will reduce by 17.54%. This is due to the higher non-monetary cost as they consume more.

Finally, we introduce a novel volume surcharge plan. Under this plan, consumers are charged 0.08 RMB for the first 400 chapters they read and 0.12 RMB afterwards. Such a pricing scheme is similar to the anti-addiction law for minors playing video games in China, wherein the reward from the game (such as virtual coins) decrease or are forfeited by the system when a player spends a certain amount of time playing video games. If the player keeps playing, he/she will be forced to quit the game for a certain period (usually 24 hours) before being allowed to log back into the game. Under such a pricing scheme, our results are reported in the bottom panel of Table 2.5. With this additional pricing plan, self-containers’ welfare will increase by 14.47%, as many of these consumers will switch from the pay-per-chapter plan to the new plan, and thus, reduce their reading amount by 7.48%. The welfare of powerless habitual readers, on the other hand, will decrease because they pay more for reading (though their reading amount will marginally reduce).

Overall, the aggregate consumer welfare will increase by 2.48%, while the platform’s profit also increases by 46.02%. This profit increase is only 0.33% less than the profit increase generated when the platform only offers a pay-per-chapter plan in the second counterfactual.

The findings from the four counterfactuals are astonishing. Taking away the monthly subscription option in the second counterfactual implies the platform forces consumers to restrict their consumption. Our results show that consumers who would have chosen monthly subscription will still read a lot, and thus, overpay under the pay-per-chapter plan. Therefore, choosing a pricing
strategy that helps prevent consumption in this empirical setting can benefit the platform. A pricing plan with volume surcharge, which will restrict consumers’ overall consumption, can simultaneously improve the platform’s profit and consumer welfare. By charging a higher price when reading above a certain level, the plan helps self-control to further control their consumption and also generate more revenue for the platform from powerless habitual readers, who already have a strong habit of reading web fictions. In contrast, pricing strategies that encourage more consumption, such as offering the monthly subscription only, will hurt not only the consumer welfare but also the platform’s profit. The managerial implications will be completely the opposite if one does not take time-inconsistent preferences and consumers’ strategic self-control into account. This highlights the substantive contribution of our study.

2.6 Conclusion
In this study, we examine how consumers overpay for reading web fiction as a means of strategic self-control when time-inconsistent preferences exist during consumption. Using data from one of China’s largest digital book platforms, we find a large percentage of consumers consistently choose pay-per-chapter even when the monthly subscription plan would be less costly. To explain this behavior, we construct a dynamic structural model featuring habit formation and time-inconsistent preferences and demonstrate that choosing a costlier pricing plan to curb consumption can be optimal for consumers. We apply our model to the data. Estimation results suggest that the market has three segments of consumers. Self-control overpay for the pay-per-chapter plan because the high cost of reading can effectively work as a “commitment device” to restrict future consumption. In the counterfactuals, we show that eliminating the pay-per-chapter plan would hurt consumer welfare and the platform’s profit. Eliminating the monthly subscription plan, however,
would increase the platform’s profit but reduce consumer welfare. We introduce a novel nonlinear pricing plan with volume surcharge and show how it can simultaneously improve the platform’s profit and consumer welfare.

Our study contributes both theoretically and empirically to the literature on consumer time-inconsistent preferences and strategic self-control behaviors. Findings from our structural model help shed light on firms’ pricing strategies in the digital content market. Despite its contributions, this research has limitations that call for future study. First, the lack of price variation in our data limits our ability to investigate the optimal prices the platform should charge for various pricing plans. Second, due to the lack of data, our study abstracts away from platform competition, which may bias our counterfactual results. For example, if removing the monthly subscription would push heavy readers to switch to another competing platform, the increase in the focal platform’s profit would be more limited. Finally, we acknowledge that strategic self-control may not be the only mechanism that can explain the overpaying behavior. One can use surveys to test what is the underlying mechanism that drives the observed consumer behaviors.

References


Taxes," Marketing Science (34:3), pp. 452-470


Appendices

Appendix A. Additional Proofs for Chapter 2

A1. Proof of Proposition 1 in Chapter 2

In this section, we provide the detailed proof of proposition 1, showing the analytical solution to our baseline model. Before we start the formal proof, we first restate the problems to be solved:

\[
V(h) = \max_s u_p(c, s, h) - \mu p_s 1\{s = 1\} + \beta V(h')
\]  

\[
u_p(c, s, h) = (\alpha_c + \alpha c h) c - \alpha_{cc} c^2 - \mu p_c 1\{s = 0\} c - \gamma c
\]

\[
h' = (1 - \delta) h + c
\]

\[
c^*(s, h) \equiv \frac{\alpha_c + \alpha c h - \mu p_c 1\{s = 0\}}{2\alpha_{cc}} \text{if } c^* \geq 0, \text{and } 0 \text{ otherwise,}
\]

where equation 4 is easily derived from the maximization of the utility function in equation 2 in terms of consumption c. \(^{36}\) For notation simplicity, let \(u_s(h) \equiv u_p(c^*(s, h), s, h) - \mu p_s 1\{s = 1\},\) we can rewrite the problems above as:

\[
V(h) = \max_s \{V_0(h), V_1(h)\}
\]

\[
V_0(h) = u_0(h) + \beta \ast \max \{V_0(h^0), V_1(h^0)\}
\]

\[
V_1(h) = u_1(h) + \beta \ast \max\{V_0(h^1), V_1(h^1)\}.
\]

\(^{36}\) It is straightforward to see that if \(s = 1, c^* > 0\); if \(s=0\), then if \(c^*=0\) at time \(t\), all consumers’ future decisions will be reduced to the trivial equilibrium where \(s=0\) and \(c^*=0\) for all future periods. So we focus on the case where \(c^* > 0\).
\[
\begin{align*}
    h^0 &= \left[(1 - \delta) + \frac{a_{ch}}{2a_{cc}}\right]h + \frac{a_c - \mu_p}{2a_{cc}} \\
    h^1 &= \left[(1 - \delta) + \frac{a_{ch}}{2a_{cc}}\right]h + \frac{a_c}{2a_{cc}}.
\end{align*}
\]  

\[6\]

Where \(V_0(h)/V_1(h)\) and \(h^0/h^1\) represent the value function and the next-period consumption stock when the consumer chooses \(s=0/s=1\) for the current period.

Notice that for a given plan choice \(s\), the value function \(V\) is continuous in terms of the state variable \(h\). Given the state space is a compact set (a closed interval on \(\mathbb{R}^1\)), the contraction mapping theorem holds and there exists a unique fixed point for value function \(V\) for any \(h\).

From equation 6, we can solve for two candidate steady states:

\[
\begin{align*}
    h_1^{ss} &= \frac{a_c - \mu_p}{2\delta a_{cc} - a_{ch}}, \\
    h_2^{ss} &= \frac{a_c}{2\delta a_{cc} - a_{ch}}.
\end{align*}
\]

Because consumption is nonnegative, \(h\) is also nonnegative. So we further impose an additional assumption.

Assumption I

\[7\]

\(2\delta a_{cc} - a_{ch} > 0\)

Assumption I guarantees at least one of the steady states is positive. This assumption also give us an ideal property of the model: The “slope” in the linear equation 6 is strictly less than 1 under assumption I, which ensures both steady states as convergent steady state for the model. Before we prove the existence of the equilibria, we first consider a simple case in which the consumers always choose \(s=1\) or \(s=0\). Denote the value under these scenarios as \(W_s(h)\), we have

\[
W_0(h) = u_0(h) + \beta W_0(h^0)
\]
\[ W_1(h) = u_1(h) + \beta W_1(h^1). \]  

[8]

Because \( u_s(h) \) is a quadratic form in \( h \) and equation 6 is linear in \( h \), we know \( W \) must also be quadratic functions of \( h \). Plugging equation 2 and 6 into equation 8, we can solve for the \( W_s(h) \).

Assume

\[ W_0(h) = a_0 h^2 + b_0 h + c_0 \]  

[9]

\[ W_1(h) = a_1 h^2 + b_1 h + c_1. \]

To keep the notation from being overcomplicated, we introduce a new set of utility parameter notations such that

\[ u_0(h) = A h^2 + B_0 h + C_0 \]  

[10]

\[ u_1(h) = A h^2 + B_1 h + C_1 \]

\[ h^0 = e h + f_0 \]

\[ h^1 = e h + f_1. \]

Now, we can use equation 8 to solve for the parameters in equation 9 in terms of these simplified notations for utility parameters:

\[ a_0 h^2 + b_0 h + c_0 = A h^2 + B_0 h + C_0 + \beta [a_0 (e h + f_0)^2 + b_0 (e h + f_0) + c_0], \forall h. \]

We get
Similarly, we can solve for $W_1(h)$ in the same way and get

\[
\begin{align*}
    a_1 &= \frac{A}{1-\beta e^2} \\
    b_1 &= \frac{B_1+2\beta e f_0 a_1}{1-\beta e} \\
    c_1 &= \frac{c_1+\beta a_1 f_0^2+\beta b_1 f_1}{1-\beta}
\end{align*}
\]

where

\[
A = \frac{\alpha_{ch}^2}{4\alpha_{cc}}; \quad e = (1-\delta) + \frac{\alpha_{ch}}{2\alpha_{cc}};
\]

\[
B_0 = \frac{(\alpha_c-\mu_p_c-\gamma)\alpha_{ch}}{2\alpha_{cc}}; \quad C_0 = \frac{\left(\alpha_c-\mu_p_c(\alpha_c-\mu_p_c-2\gamma)\right)}{4\alpha_{cc}}; \quad f_0 = \frac{\alpha_c-\mu_p_c}{2\alpha_{cc}};
\]

\[
B_1 = \frac{(\alpha_c-\gamma)\alpha_{ch}}{2\alpha_{cc}}; \quad C_1 = \frac{\alpha_c(\alpha_c-2\gamma) - \mu p}{4\alpha_{cc}}; \quad f_1 = \frac{\alpha_c}{2\alpha_{cc}}.
\]

Notice $a_1 = a_0$, so $W_1(h) - W_0(h) = (b_1 - b_0)h - (c_0 - c_1)$, which is linear in $h$ and thus guarantees the “single-crossing” property of $W_1(h)$ and $W_0(h)$. Equipped with these results, we start the formal proof of proposition 1. Our method is by guess and verify of the value function and then solve for the policy function.

1. When $h_2^{ss} > h_1^* > h_1^{ss}$

\[
V = \begin{cases} 
    ah^2 + b_0 h + c_0, h < h_1^* \\
    ah^2 + b_1 h + c_1, h \geq h_1^* 
\end{cases}
\]
\[ h_1^* = \frac{c_0 - c_1}{b_1 - b_0}. \]

Without loss of generality, assume the consumer starts with some initial consumption stock state \( h > h_1^* \). If the above value function is correct, after one iteration, \( V \) will remain identical. In other words, because \( h > h_1^* \), we need to show that for any \( h \) after the consumer makes the optimal plan-choice decision to maximize the value function, \( V \) will still be equal to \( ah^2 + b_1 h + c_1 \) after it is plugged into equation 5. Given assumption I, we know that if the consumer chooses monthly subscription \( s=1 \), \( h_1^* \) is lower than her next-period consumption stock state, \( h_1^1 \). However, it is undetermined whether \( h^0 > h_1^* \). So, we discuss two cases here.

**Case I  \( h^0 < h_1^* \)**

\[
V(h) = \max (V_0(h), V_1(h))
\]
\[
= \max \left\{ Ah^2 + B_0 h + C_0 + \beta \left( a h^{0^2} + b_0 h^0 + c_0 \right), Ah^2 + B_1 h + C_1 \right\}
\]
\[
= \max\left\{Ah^2 + b_0 h + c_0, ah^2 + b_1 h + c_1 \right\}
\]
\[
= ah^2 + b_1 h + c_1.
\]

where the second equation utilizes the “fixed-point” property of \( W_1 \) and \( W_2 \), and the third equation is simply algebra with the fact the \( b_1 > b_0 \) and \( ah^2 + b_1 h + c_1 > ah^2 + b_0 h + c_0 \) with any \( h > h_1^* = (c_0 - c_1)/(b_1 - b_0) \).

**Case II  \( h^0 \geq h_1^* \)**
\[
V(h) = \max \left\{ Ah^2 + B_0 h + C_0 + \beta \left( ah_0^2 + b_1 h^0 + c_1 \right), \right. \\
\left. Ah^2 + B_1 h + C_1 + \beta \left( ah_1^2 + b_1 h^1 + c_1 \right) \right\}
\]

\[
= \max \left\{ ah^2 + b_0 h + c_0 + \beta [(b_1 - b_0)h^0 + (c_1 - c_0)], \right. \\
\left. ah^2 + b_1 h + c_1 \right\}
\]

\[
= \max \{ ah^2 + b_0 h + c_0 + \beta [(b_1 - b_0)h^0 + (c_1 - c_0)], \\
(ah^2 + b_1 h + c_1) \}.
\]

Because \( h > h_1^* > h_2^* \), by equation 6, we know that \( h > h^0 \) so we have

\[
[(b_1 - b_0)h^0 + (c_1 - c_0)] > \beta [(b_1 - b_0)h^0 + (c_1 - c_0)] > \beta [(b_1 - b_0)h^0 + (c_1 - c_0)],
\]

and

\[
V(h) = ah^2 + b_1 h + c_1.
\]

Similarly, we can prove that for any \( h < h_1^* \), \( V(h) = ah^2 + b_0 h + c_0 \).

Given the functional form of the value function, it is straightforward to solve for the optimal policy

function \( s(h) = 1 \cdot 1(h > h_1^*) + 0 \cdot 1(h < h_1^*) \).

2. When \( h_1^* > h_2^* > h_1^* \),

\[
V(h) = \begin{cases} 
\sum_{t=0}^{T_h} \beta^t u_1(h_t) + \beta^{T_h+1}(a_0 h_{T_h+1}^2 + b_0 h_{T_h+1}^0 + c_0) & \text{if } h \geq h_2^*, \\
 a_0 h^2 + b_0 h + c_0 & \text{Otherwise}
\end{cases}
\]

where

\[
h_0 = h
\]

\[
h_{t+1} = eh_t + f_t; t = 0,1,2,\ldots, T_h
\]

\[
T_h = \left\lfloor \log \left( \frac{(1-e)h_2^* - f_t}{(1-e)h - f_t} \right) \right\rfloor
\]

\[
= 109
\]
\[ h_2^* = \frac{(c_0 - c_1) + \beta[(b_1 - b_0)f_1 + c_1 - c_0]}{(1 - \beta e)(b_1 - b_0)}. \]

Proof:

First, we define the “difference function” \( Z(h) \) as

\[ Z(h) \equiv u_1(h) + \beta W_0(h^1) - W_0(h) = u_1(h) + \beta W_0(\epsilon h + f_1) - W_0(h). \]

\( Z(h) \) measures the lifetime utility difference between when consumers always choose pay-per-chapter and when consumers purchase monthly subscription in the current period and choose to pay-per-chapter starting next period. When \( Z(h)=0 \), the consumer receives same utility value no matter what price plan he chooses in the next time period, given that he keep pay-per-chapter since the third time period. With some algebra, we can show

\[ Z(h) = (B_1 + \beta(2ae f_1 + b_0 e) - b_0)h + (1 - \beta)(c_0 - c_1) + \beta(b_1 - b_0)f_1 \]

\[ = (B_1 - B_0 + \beta(2ae(f_1 - f_0)))h + (1 - \beta)(c_0 - c_1) + \beta(b_1 - b_0)f_1, \]

where the second equation uses the fixed-point property of \( W_0(h) \). Because \( B_1 > B_0 \) and \( f_1 > f_0 \) (the utility for monetary cost is always negative), \( Z(h) \) is a linear monotone increasing function in \( h \), and \( h_2^* \) is the unique solution to \( Z(h)=0 \). So,

\[ u_1(h) + \beta w_0(h^1) \geq w_0(h) \forall h \geq h_2^* \tag{11} \]

\[ u_1(h) + \beta w_0(h^1) < w_0(h) \forall h < h_2^*, \tag{12} \]

and vice versa.

Next, we show that if \( h_1^* > h_2^{ss} \), then \( h_2^* > h_2^{ss} \);
\[ h^*_2 = \frac{(1 - \beta)(c_0 - c_1)}{b_1 - b_0} + \beta f_1 \]

\[ = \frac{(1 - \beta)h^*_1 + \beta f_1}{1 - \beta e} \]

\[ > \frac{(1 - \beta)h^{ss}_2 + \beta f_1}{1 - \beta e} \]

\[ = \frac{(1 - \beta)f_1}{1 - e} + \beta f_1 \]

\[ = \frac{f_1}{1 - e} = h^{ss}_2, \]

where the third equation is derived from the expression of \( h^{ss}_2 \) in the form of \( f_1 \) and \( e \).

Finally, we show the value function above is the fixed point of equation 5. Without loss of generosity, assume some consumer with initial consumption stock state \( h > h^*_2 \) at the initial period \( t=0 \). We first show by induction that for any \( T \) (for any \( h \)), when the consumer starts with any consumption stock state \( h \geq h^*_2 \), the optimal plan choice is always \( s=1 \).

by equation 11, we know that for \( T_h = 0 \),

\[ u_1(h) + \beta(a_0 h^2 + b_0 h + c_0) = u_1(h) + \beta w_0(h^1) \geq w_0(h) = a_0 h^2 + b_0 h + c_0. \]

If for \( T_h = n \), the following inequality holds:

\[ \sum_{t=0}^{n} \beta^t u_1(h_t) + \beta^{n+1} w_0(h_{n+1}) \geq w_0(h_t). \]

Then, for \( T_h = n + 1 \), the following also holds:
\[
\begin{align*}
\sum_{t=0}^{n+1} & \beta^t u_1(h_t) + \beta^{n+2}w_0(h_{n+2}) \\
= & \sum_{t=0}^{n} \beta^t u_1(h_t) + \beta^{n+1}u_1(h_{n+1}) + \beta^{n+2}w_0(h_{n+2}) \\
= & \sum_{t=0}^{n} \beta^t u_1(h_t) + \beta^{n+1}(u_1(h_{n+1}) + \beta w_0(h_{n+2})) \\
\geq & \sum_{t=0}^{n} \beta^t u_1(h_t) + \beta^{n+1}w_0(h_{n+1}) \\
\geq w_0(h).
\end{align*}
\]

The proof above establishes that when the consumer starts with any consumption stock state \( h \geq h_2^* \), the optimal decision is always to choose \( s=1 \) at the current period. We now show that after one iteration, the value function remains the same functional form. Because the next-period consumption stock states will be \( h^1 \) under \( s=1 \), again we discuss two cases in which \( h^1 \geq h_2^* \) and \( h^1 < h_2^* \):

**Case 1.** \( h^1 \geq h_2^* \)

\[
V(h) = u_1(h) + \beta V(h^1)
\]

\[
= u_1(h) + \beta \left( \sum_{t=0}^{T_{h^1}} \beta^t u_1(h_{t+1}) + \beta^{T_{h^1}+1}(a_0 h_{T_{h^1}+2}^2 + b_0 h_{T_{h^1}+2} + c_0) \right)
\]

Because \( h^1 = eh + f_1, T_{h^1} = T_h - 1 \),

\[
= \sum_{t=0}^{T_h} \beta^t u_1(h_t) + \beta^{T_{h}+1}(a_0 h_{T_{h}+1}^2 + b_0 h_{T_{h}+1} + c_0).
\]

112
Case II. $h^1 < h_2^*$

When $h'' < h_2^*$, it is straightforward to see for any time period we have

$$V(h) = u_1(h) + \beta V(h^1)$$

$$= u_1(h) + \beta(a_0 h^1 + b_0 h + c_0)$$

$$= \sum_{t=0}^{T_h} \beta^t u_1(h_t) + \beta^{T_h+1}(a_0 h_{T_h+1} + b_0 h_{T_h+1} + c_0).$$

The above case confirms that under either case, the value function always remains the same functional form. Use equation 12 instead of equation 11, we can repeat the same proof above and show for any $h < h_2^*, V(h) = a_0 h^2 + b_0 h + c_0$.

Thus, we have proved that for any $h$ when $h_1^* > h_2^* > h_1^{ss}$, the policy function is

$$s(h) = 1 \cdot (h > h_2^*) + 0 \cdot (h < h_2^*).$$

Symmetrically, when $h_2^{ss} > h_1^{ss} > h_1^{ss}$, we can show $s(h) = 1 \cdot (h > h_2^{**}) + 0 \cdot (h < h_2^{**})$, where $h_2^{**}$ is the unique solution to

$$u_0(h) + \beta w_1(h^0) - w_1(h) = 0.$$

Solving the equation above gives us $h_2^{**} = \frac{(1-\beta)(c_1-c_0)+\beta(b_1-b_0)f_0}{(1-\beta)(b_1-b_0)}$. Q.E.D.

A2. Proof of Proposition 2 in Chapter 2

Because it is obvious that all states for $h$ are aperiodic given $p_{m,m,}(\theta) > 0$ for any $m$. To prove the existence of the limiting distribution, we only need to show that all states for $h$ are irreducible, which means for any state $i,j = 1,2,...,H, \exists n_0$ s.t. $p_{ij}^{n_0} > 0$. Together with aperiodic and
irreducible states, the Markov chain for $h$ is ergodic. Thus, the ergodicity theorem guarantees the existence of the limiting distribution for $h$ specified in proposition 2.

For any $i$, we first show that if $j \geq i$, then $\exists n_0 = 1$ s.t. $p^1_{ij} > 0$: without loss of generosity, we start from any $h$ in state $i$, which means $h \in \left[\frac{i-1}{N}H, \frac{i}{N}H\right)$. The c.d.f. for the next period of consumption stock state $h'$ follows the c.d.f $f(h'|h, \theta)$, which is specified in the main text right before the proposition 2. Because the consumption is nonnegative, it is straightforward to see that $f(h'|h, \theta) > 0$ on the support of $[(1 - \delta)h, H]$. Also, we have $\frac{j}{N}H > \frac{i-1}{N}H > h > (1 - \delta)h$, so

$$p^1_{ij} = \int_{\frac{i}{N}H}^{\frac{j-1}{N}H} f(h'|h \in i, \theta) dh' > 0.$$ 

Similarly, we can show that for any $j$ that $\frac{j-1}{N}H \geq (1 - \delta)h$, $p^1_{ij} > 0$. Because $\lim_{n \to \infty} (1 - \delta)^n h = 0$, for any $j$ that $\frac{j-1}{N}H \geq 0$, there always exists a finite $n_0$ so that $p^1_{ij} > 0$.

Now that we have established that any $i,j=1, 2, \ldots, N, \exists n_0$ s.t. $p^{n_0}_{ij} > 0$, together with the facts that the Markov chain is aperiodic, we have proved the $\{h_t\}$ is an ergodic Markov chain that has a unique stationary distribution. Q.E.D.
Appendix B. Additional Estimation Results for Chapter 2

B1. Estimation results (time-consistent consumers)

beta=0.98  Observation=68,706

<table>
<thead>
<tr>
<th></th>
<th>Segment 1</th>
<th>s.e.</th>
<th>Segment 2</th>
<th>s.e.</th>
<th>Segment 3</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_c$</td>
<td>-113.41</td>
<td>0.41</td>
<td>-25.16</td>
<td>0.84</td>
<td>760.39</td>
<td>31.25</td>
</tr>
<tr>
<td>$a_{ch}$</td>
<td>0.34</td>
<td>37.56</td>
<td>557.86</td>
<td>0.69</td>
<td>0.63</td>
<td>0.02</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.88</td>
<td>418.03</td>
<td>0.65</td>
<td>623.96</td>
<td>1.52</td>
<td>2.23</td>
</tr>
<tr>
<td>$\log \sigma_w$</td>
<td>6.38</td>
<td>3.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\log \sigma_e$</td>
<td>8.49</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class size</td>
<td>0.94</td>
<td>0.03</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-439711</td>
<td>848697</td>
<td>BIC</td>
<td>848870</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

115
## B2. Estimation results (Naive time-inconsistent consumers)

**beta=0.98  Observation=68,706**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Segment 1</th>
<th>s.e.</th>
<th>Segment 2</th>
<th>s.e.</th>
<th>Segment 3</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_c)</td>
<td>-37.17</td>
<td>6.61</td>
<td>-32.78</td>
<td>6.25</td>
<td>627.14</td>
<td>31.25</td>
</tr>
<tr>
<td>(a_{ch})</td>
<td>0.82</td>
<td>0.01</td>
<td>0.79</td>
<td>2.43E-04</td>
<td>1.47</td>
<td>0.02</td>
</tr>
<tr>
<td>(\mu)</td>
<td>160.37</td>
<td>5.49</td>
<td>1.40</td>
<td>0.08</td>
<td>0.07</td>
<td>2.23</td>
</tr>
<tr>
<td>(\gamma_0)</td>
<td>649.33</td>
<td>72.65</td>
<td>20.52</td>
<td>0.30</td>
<td>100.30</td>
<td>31.82</td>
</tr>
<tr>
<td>3-Class</td>
<td>(\gamma_1)</td>
<td>63.27</td>
<td>57.33</td>
<td>0.09</td>
<td>7.81E-04</td>
<td>14.30</td>
</tr>
<tr>
<td>(\log \sigma_w)</td>
<td>8.54</td>
<td>3.19E-02</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\log \sigma_e)</td>
<td>6.39</td>
<td>1.38E-04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class size</td>
<td>0.44</td>
<td>0.41</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-424329</td>
<td>AIC 848697</td>
<td>BIC 848870.44</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>