2016

Essays on Internet economics: customer reviews, advertising, and technology adoption

Lei, Ying

http://hdl.handle.net/2144/17721

Boston University
ESSAYS ON INTERNET ECONOMICS: CUSTOMER REVIEWS, ADVERTISING, AND TECHNOLOGY ADOPTION

by

YING LEI

B.A., Renmin University of China, 2008
M.A., New York University, 2010

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

2016
Approved by

First Reader
Ching-To Albert Ma, PhD
Professor of Economics

Second Reader
Marc Rysman, PhD
Professor of Economics

Third Reader
Juan Ortner, PhD
Assistant Professor of Economics
Acknowledgments

First and foremost, I would like to thank my main advisor Professor Ching-To Albert Ma. I owe him my greatest gratitude not only for his tremendously helpful comments and suggestions on my research, but more importantly for his invaluable support and encouragement from the very immature stage of my research. He guided me through my hard times in pursuit of the Ph.D. and made this process so much more enjoyable. I learnt from Albert how to do research, how to transform the real-world puzzles to economic models and how to solve them. Besides research, I also benefit a lot from his many suggestions in general.

I also want to thank my committee members Professor Marc Rysman, Professor Juan Ortner and Professor Monic Sun. They spent so much time help me and give advices. I can always receive a variety of comments from them, and this made me able to think from different perspectives and improve a lot. They are all wonderful advisors and mentors, and friends.

I especially want to thank my husband, Ei Yang, for being so supportive and considerate as a peer economist, my best friend and my life partner. We support and accompany each other through the long journey of pursuing knowledge. He made me a better person. And then I want to thank the cutest baby of Ei and I, Ethan Yang. I thank him for being such a wonderful and healthy boy, creating much less hassle than I expected and making it possible for me to be a new mom during the Ph.D. pursuit.

My lots of thanks then go to my parents, Chaoying Zhao and Shenggang Lei. They raised me to be such a person that is always curious and persistent in thinking and hard-working. They are always there for me, with their unconditional love, support and sacrifice. I am very proud to be their daughter, and I am glad that I made them proud too.

Last but not least, I thank the economics department of Boston University. This is a very active research environment with so many great economists and nice mentors. I grew from a student to be a real economist here. The six years that I spent in this building and in this beautiful campus along the Charles River have become an important and precious part of my life.
This dissertation consists of three chapters on the economics of the Internet.

The first chapter begins with presenting the advertising spending patterns of US local restaurants that have different ratings on Yelp.com. Rating information on Yelp includes display ratings and review distributions. The Yelp’s rounding algorithm creates a discontinuity in display ratings. Therefore, I use a regression discontinuity design to identify the effect of a higher display rating on local restaurants’ advertising spendings. I find a significantly negative effect of display rating for highly-rated restaurants on advertising. However, when the display rating is constant between two steps, the relationship between local restaurant advertising spending and average rating is significantly positive.

The second chapter uses a game-theoretic model to analyze competing firms’ advertising and pricing decisions. Here customer reviews are available and firms may build up loyal customer bases. I find that highly-rated firms are more likely to advertise more, i.e., online reviews complement advertising. Comparative static results can explain the results found in the first chapter. Intuitively, when the capacity of a local business becomes limited, a jump in the display rating will reduce the complementary effect of online reviews on advertising. I also analyze an extension of the model, where an entrant and an incumbent interact. I find that customer reviews undo the “fat-cat” effect of a large incumbent with lots of loyal customers.

The third chapter proposes a new explanation for adoption failure or delay in markets with network effects. In the model, consumers and software providers play a dynamic adoption game.
Each group of players choose between two incompatible technologies. Consumers may wait, but firms may not. Although efficiency requires one technology to be adopted by all consumers and firms right away, there is a “market split and adoption delay” equilibrium. In this equilibrium some consumers choose to wait at first and firms split between the two technologies. The model is motivated by the 56K modem market, in which competition between two technologies appears to have led to adoption failure, until an industry standard setting organization coordinated the market on an alternative standard.
Contents

1 Advertising Response to A Better Online Rating: A Regression Discontinuity Design
   with US Local Restaurants and Yelp Reviews ........................................... 1
   1.1 Introduction .............................................................................................. 1
   1.2 Literature Review ...................................................................................... 2
   1.3 Data ........................................................................................................... 3
   1.4 Empirical Analysis and Results .................................................................. 4
      1.4.1 Graphical Analysis .............................................................................. 5
      1.4.2 Regression Analysis and Results ......................................................... 6
   1.5 Robustness ................................................................................................. 9
      1.5.1 Higher-Order Polynomial Regressions ............................................... 9
   1.6 Validation Tests of RD Design ................................................................... 10
      1.6.1 Density of Assignment Variable ......................................................... 10
      1.6.2 Review Characteristics and Covariates .............................................. 13
   1.7 Discussion .................................................................................................. 15
   1.8 Conclusion .................................................................................................. 17
   1.9 Appendix .................................................................................................... 19
      1.9.1 RD Graphs at Lower Thresholds ....................................................... 19

2 How Do Firms Advertise When Customer Reviews Are Available? .................. 20
   2.1 Introduction .............................................................................................. 20
   2.2 Literature Review ...................................................................................... 24
   2.3 Main Model ................................................................................................ 25
      2.3.1 Players and Information ...................................................................... 25
      2.3.2 Belief Updating ................................................................................... 28
   2.4 Equilibrium Analysis .................................................................................. 29
      2.4.1 The Pricing Subgame ......................................................................... 31
List of Tables

1.1 Summary Statistics ......................................................... 4
1.2 The Effect of An Extra Half-Star in Display Rating on Ad Spending ........ 8
1.3 RD Estimates of the Pooled Regression ................................ 9
1.4 RD Estimates With Quadratic Polynomial Regressions .................... 10
1.5 Percentiles of Reviews .................................................... 11
1.6 RD Estimates of Jumps in Percentage of Five-Star Reviews ............... 13
1.7 RD Estimates of Jumps in SD of Ratings ................................ 14
1.8 Pooled RD With Interval Dummies ....................................... 16
2.1 The Normal Form of The Game ........................................... 30
2.2 Payoffs When Condition APC-A Holds .................................. 34
2.3 Firms’ Profits When $s$ is Big, and $L^A > L^B$ .......................... 48
2.4 Summary Statistics for The Graphical Analysis Dataset .................. 52
2.5 Summary Statistics for The Regression Analysis Dataset ................. 54
2.6 RD: Pooled Regression ................................................... 57
List of Figures

1.1 Binned Scatters of Ad Spending by Average Rating around Thresholds 3.25, 3.75, 4.25 and 4.75 ........................................ 6
1.2 Density of Average Rating ................................................. 11
1.3 Density of Average Rating For Two Subsamples ......................... 12
1.4 Binned Scatter Plots of Ad Spending by Average Rating around Lower Thresholds 19

2.1 Three conditions to characterize equilibria ................................. 36
2.2 Multiple Equilibria or Unique Equilibrium? ............................... 38
2.3 Stage 1 and Stage 2 of the entry game ..................................... 41
2.4 Conditions PE and Entry in the entry game ................................. 42
2.5 Condition $APC - A$ in four pricing subgames ............................ 48
2.6 With and without Searchers: When does Firm $A$ always advertise? .... 50
2.7 Equilibrium partition in the Searchers model: Areas of multiple equilibria or unique equilibrium ........................................ 50
2.8 The Advertising Pattern of Local Restaurants in Neighborhoods with Only Two Restaurants (that are both in Kantar Media Database and on Yelp.com) ..... 52
2.9 The Advertising Pattern of Local Restaurants in New York City ........ 53
2.10 The Relationship between Ad Spending and True Rating without Recognition of RD Drops ........................................... 55
2.11 The Relationship between Ad Spending and True Rating with Recognition of RD Drops ........................................... 56
2.12 An Example of Yelp Search Results Page ................................ 59
2.13 Three Conditions for Case $T^A > T^B > 0$ ................................. 66
2.14 The Display Rating and Distribution of Reviews of a Restaurant .......... 67
2.15 Advertising Spending by Local Restaurants’ Average Rating around Thresholds ........ 67
2.16 Best Responses in the Pricing Subgame .................................. 68
Chapter 1

Advertising Response to A Better Online Rating: A Regression Discontinuity Design with US Local Restaurants and Yelp Reviews

This paper presents the advertising spending patterns of US local restaurants that have different ratings on Yelp.com. Rating information on Yelp includes display ratings and review distributions. The Yelp’s rounding algorithm creates a discontinuity in display ratings. Therefore, I use a regression discontinuity design to identify the effect of a higher display rating on local restaurants’ advertising spendings. I find a significantly negative effect of display rating for highly-rated restaurants on advertising. However, when the display rating is constant between two steps, the relationship between local restaurant advertising spending and average rating is significantly positive.

1.1 Introduction

A lot of research have been done on online reviews, but few has noticed the different economic effects of different types of reviews information. This paper studies two rating information on Yelp.com: display rating and average rating, and estimates local restaurants’ advertising responses to these two rating information. Surprisingly, although they are both rating information, advertising changes differently in response to display rating and average rating.

The display rating on Yelp.com is the colored stars displayed below the name of a local business. It rounds the average rating to the nearest half-star.

The average rating of a local restaurant cannot be directly observed. Yet consumers can click the details button to see the entire distribution of the customer reviews of this restaurant. Comparing two restaurants with different distributions of reviews, consumers can get a visual approximation of the average ratings of these two restaurants.

Average rating is a measure of how consumers like the product of a restaurant (or a local business in general). For local businesses providing consumer goods, the likelihood of consumer satisfaction better describes the product quality than the cost of input in producing the good. There-
fore, aggregating user generated reviews, the average rating can represent the quality reputation of a local business among consumers. This paper do not assume that consumers are sophisticated enough to compute the precise number of average rating from the distribution of reviews. Average rating is merely a summary statistic used to represent the quality reputation revealed from the distribution of reviews.

If there is only average rating and no display rating, consumers learn the quality reputation and firms respond to it. With the addition of display rating, in fact no extra information is newly created. Will it change firms’ advertising responses? In this paper we will see the twist brought by the display rating on restaurants’ advertising strategies, and justification will be provided.

Due to the rounding algorithm, the Yelp display rating has several jump thresholds. At each threshold, the display rating jumps by half a star. Restaurants with similar average rating randomly allocate above and below each threshold, and this randomness enable us to identify the effect of display rating on local restaurants’ advertising spending. A sharp RD design is implemented and the validity is justified with a rich set of tests for manipulation on reviews.

The display rating is constant between every two adjacent thresholds. Therefore the pattern in advertising spending on the between-threshold intervals shows the relationship between advertising and average rating, as if there is no display rating.

This paper provides the first detailed comparison between different types of rating information on their relationship with firms’ marketing strategies. And the interesting finding is that, although display rating and average rating are both rating information, local restaurants’ advertising spending responds in entirely opposite directions to them.

1.2 Literature Review

To identify the effect of online reviews on sales or firms’ decisions, we usually need to take differences across review sites or over time or both to eliminate the product quality fixed effects and/or review site fixed effects. Chevalier and Mayzlin [2006] apply a difference-in-difference approach on the book reviews data from Amazon.com and BN.com over three time points. Chen et al. [2007]
use a first-difference approach with book reviews on Amazon.com over a 195 day period.

In recent years, the development in Regression Discontinuity design applications (Imbens and Lemieux [2008]; Lee and Lemieux [2010]) allows us to identify treatment effects without incurring the complexity resulting from comparing with a different website or with a different time point. Anderson and Magruder [2012] use RDD to analyze the effect of a higher Yelp rating on restaurant reservation availability. In this paper, I use RDD to identify local restaurants’ advertising response to a higher Yelp display rating.

1.3 Data

To estimate the effect of Yelp display rating on advertising spending, I combine data on local restaurants from two sources. The first dataset contains the annual advertising spending amount in 2014 of local restaurants in the United States. This dataset contains each local restaurant’s total advertising spending amount as well as their advertising spending in each DMA (Designated Market Area) region. The advertising dataset does not distinguish between different advertising channels, i.e., the advertising spending of each restaurant is the sum of ad spending in all major multimedia channels, including TV, magazines, Internet, newspapers, radio, outdoor, etc.

The second dataset is scripted from Yelp.com and contains the corresponding Yelp reviews and other information of those local restaurants in the advertising dataset. In particular, for each restaurant in this dataset, I recorded the number of reviews of each star-rating (1 to 5), as of January 2014. Other information of restaurants that I collected from Yelp.com include price range, attire, waiter service, delivery, etc.

I merge the two datasets together, therefore the final dataset contains advertising, Yelp reviews and other information of each restaurant. The average rating can be calculated for each restaurant. This calculated average rating is the true average rating of each restaurant, as opposed to their round-off displayed rating on Yelp.

Even though the dataset contains only local restaurants, there are some difference among these

---

1 The local restaurants in this dataset are the restaurants in Kantar Media database with classification code G310, as opposed to national restaurants which have classification code G320.
restaurants. A few restaurants advertise in more than one DMA markets, and a few restaurants are not listed on Yelp.com. In the main body of the analysis in this paper, I focus on the restaurants that advertise in only one DMA markets because the restaurants that advertise in multiple markets are likely to have very different marketing strategies from those entirely local restaurants. I also exclude those restaurants that are not listed on Yelp.com because Yelp rating is crucial in the analysis. Summary statistics for the dataset and some subsets are provided in Table 1.1.

Table 1.1: Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Restaurant On Yelp</th>
<th>Restaurants off Yelp</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Restaurant</td>
<td>Single Market</td>
</tr>
<tr>
<td>Reviews per restaurant</td>
<td>76.996 (180.960)</td>
<td>73.512 (170.971)</td>
</tr>
<tr>
<td>Restaurant average rating</td>
<td>3.579 (0.583)</td>
<td>3.585 (0.584)</td>
</tr>
<tr>
<td>Advertising spending</td>
<td>12.724 (34.290)</td>
<td>9.743 (28.731)</td>
</tr>
<tr>
<td>Observations</td>
<td>13360</td>
<td>11510</td>
</tr>
</tbody>
</table>

Note: “Single Market” denotes the subsample of restaurants that advertise in only one market.

This paper studies the effect of a higher Yelp display rating on local restaurant advertising spending. Therefore, the exclusion of restaurants that are off Yelp.com does not affect the estimates and conclusions here. The effect of the exclusion of restaurants that advertise in multiple markets is unclear at this point and will be checked as a robustness test.

1.4 Empirical Analysis and Results

Yelp display rating rounds off a restaurant’s average rating to the nearest half-star. This algorithm creates discontinuity at each .25 and .75 thresholds. In particular, the correspondence between
display ratings and average ratings is:

\[
\text{Display rating} = \begin{cases} 
1, & \text{if average rating } \in [1, 1.25) \\
1.5, & \text{if average rating } \in [1.25, 1.75) \\
2, & \text{if average rating } \in [1.75, 2.25) \\
2.5, & \text{if average rating } \in [2.25, 2.75) \\
3, & \text{if average rating } \in [2.75, 3.25) \\
3.5, & \text{if average rating } \in [3.25, 3.75) \\
4, & \text{if average rating } \in [3.75, 4.25) \\
4.5, & \text{if average rating } \in [4.25, 4.75) \\
5, & \text{if average rating } \in [4.75, 5] 
\end{cases}
\]

Restaurants around each threshold have similar average ratings, but the ones at and above each threshold have an extra 0.5 star in Yelp display rating. Between every two thresholds, the display rating is constant as the average rating increases. Using average rating as the assignment variable, I implement a sharp RD design to analyze the effect of having an extra half-star in Yelp display rating. Tests of the possibility of manipulation are discussed in Section ???. In these tests we do not see evidence of manipulation, and therefore the RD design is valid.

### 1.4.1 Graphical Analysis

To begin the RD analysis, I present the plots of binned local average of advertising spending around selected thresholds of average rating. Figure 1.1 contains the binned scatter plots around thresholds 3.25, 3.75, 4.25 and 4.75 of average rating. Thresholds are marked by the dashed vertical lines. The upper-left panel is for the average ratings within the range [2.75, 3.75), and in this panel the display rating changes from 3 on the left of the threshold (3.25) to 3.5 on the right of the threshold. Similarly, the upper-right panel shows the average advertising spending as the display rating changes from 3.5 to 4, the lower-left panel is when the display rating changes from 4 to 4.5,
and the lower-right panel is when the display rating changes from 4.5 to 5.

![Figure 1.1: Binned Scatters of Ad Spending by Average Rating around Thresholds 3.25, 3.75, 4.25 and 4.75](image)

From the plots, we can see clear graphical evidence of jumps in advertising spending when the display rating goes up by a half-star, especially at thresholds 3.75, 4.25 and 4.75.

1.4.2 Regression Analysis and Results

Let $X_i$ denote the assignment variable (or forcing variable), i.e., the average rating of each restaurant, and $c$ denote the RD thresholds (1.25, 1.75, ...) as listed above. Around each threshold, I estimate the regression:

$$y_i = \alpha + \tau I(X_i \geq c) + f(X_i - c) + \delta Z_i + \epsilon_i$$ (1.1)

where $y_i$ is the advertising spending level of a restaurant, $f(\cdot)$ is a flexible functional form, and $Z_i$ is a set of additional restaurant covariates. The effect of an extra half-star of display rating on advertising spending is then estimated by the coefficient $\tau$.

Following the suggestions by Gelman and Imbens [2014], I do not use higher order polynomial
regressions. I use local linear regressions in the main body of this paper, and check for robustness with local quadratic polynomial regressions. In particular, around each threshold, I regress

\[ y_i = \alpha + \beta_1(X_i - c) + \tau I(X_i \geq c) + \beta_2(X_i - c)I(X_i \geq c) + \delta Z_i + \epsilon_i \]  

(1.2)

where the coefficient \( \tau \) still estimates the effect of an extra half-star in display rating for restaurants around threshold \( c \). I estimate this local linear regression at each threshold, and then I run a pooled regression to combine all thresholds and estimate an average effect of display rating on advertising spending.

First, I estimate (1.2) separately for each threshold, and the regression results at selected thresholds are given in Table 1.2. The four columns in this table correspond to the four graphs in Figure 1.1. Column (1) estimates the effect of increasing Yelp display rating from 3 to 3.5 on local restaurants’ advertising spending. The regression result tells that if the display rating increases from 3 to 3.5, restaurants on average reduce advertising spending by 3839 dollars. The effect of display rating for restaurants around threshold 3.75, as shown in column (2), is also significantly negative, and the average drop in advertising spending is even larger, i.e., an extra half-star induces an average drop of 5473 dollars in ad spending. The estimate in column (3) is insignificant. Compared to the clear and big jump in the corresponding plot (lower-left panel) in Figure 1.1, reasons for the insignificance of the estimate here might be the smaller number of observations and the relatively big variance across restaurants. I expect to see significant effect of an extra half-star at threshold 4.25 if more observations are available. In column (4), most of the estimates around threshold 4.75 are insignificant due to the small number of observations.

The estimated effect of display rating (i.e., \( \hat{\tau} \)) at thresholds 3.25 and 3.75 are robust to various bandwidth choices (between 0.1 to 0.5). The estimate for threshold 4.25 becomes more significant as the bandwidth increases. At the threshold 4.75, the number of observations is too small even when we change the bandwidth, and the estimated effect of display rating is not significant enough to draw conclusion.

\[ ^2 \text{With bandwidth 0.5, the estimate equals } -2.420 \text{ with standard error 1.481, which is almost significant at the 10\% level.} \]
Table 1.2: The Effect of An Extra Half-Star in Display Rating on Ad Spending

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above 3.25</td>
<td>−3.839***</td>
<td>−3.822</td>
<td>−3.822</td>
<td>−3.822</td>
</tr>
<tr>
<td>Above 3.75</td>
<td>−5.473***</td>
<td>(2.133)</td>
<td>(2.133)</td>
<td>(2.133)</td>
</tr>
<tr>
<td>Above 4.25</td>
<td>−1.097</td>
<td>(2.202)</td>
<td>(2.202)</td>
<td>(2.202)</td>
</tr>
<tr>
<td>Above 4.75</td>
<td>−3.475</td>
<td>(2.579)</td>
<td>(2.579)</td>
<td>(2.579)</td>
</tr>
<tr>
<td>Dressy</td>
<td>2.589</td>
<td>(2.494)</td>
<td>(2.494)</td>
<td>(2.494)</td>
</tr>
<tr>
<td>Waiter</td>
<td>−2.615</td>
<td>(1.666)</td>
<td>(1.878)</td>
<td>(1.878)</td>
</tr>
<tr>
<td>Delivery</td>
<td>−1.216</td>
<td>(1.189)</td>
<td>(1.483)</td>
<td>(1.483)</td>
</tr>
<tr>
<td>Price $$</td>
<td>1.848</td>
<td>(1.379)</td>
<td>(1.651)</td>
<td>(1.651)</td>
</tr>
<tr>
<td>Price $$$</td>
<td>5.417***</td>
<td>(2.305)</td>
<td>(2.481)</td>
<td>(2.481)</td>
</tr>
<tr>
<td>Price $$$$</td>
<td>−0.513</td>
<td>(5.786)</td>
<td>(5.998)</td>
<td>(5.998)</td>
</tr>
<tr>
<td>Constant</td>
<td>13.091***</td>
<td>(2.062)</td>
<td>(2.268)</td>
<td>(2.268)</td>
</tr>
</tbody>
</table>

Notes. The dependant variable is advertising spending (unit: USD Thousands). The variable “Average rating” is normalized by the corresponding threshold in each column, i.e., $X_i - c$. The variable “Above threshold” in the interaction term corresponds to the indicator variable in each column: “Above 3.25”, “Above 3.75”, “Above 4.25” and “Above 4.75” respectively. Bandwidth is 0.25 in all columns. Significance levels: *** 1%, ** 5%, * 10%.

Given the similarities in the estimates for the four thresholds (3.25, 3.75, 4.25 and 4.75), I combine all thresholds in a pooled regression. To pool all thresholds, I normalize each threshold to zero, combine all observations, and estimate

$$ y_i = \alpha + \beta_1 \tilde{X}_i + \tau_{pool}I(\tilde{X}_i \geq 0) + \beta_2 \tilde{X}_i I(\tilde{X}_i \geq 0) + \delta Z_i + \epsilon_i \quad (1.3) $$

where $\tilde{X}_i = X_i - c$ and $c$ is the nearest threshold to $X_i$. The average effect on advertising of an extra half-star in Yelp rating is captured by $\tau_{pool}$. The regression results are given in Table 1.3.

The first column pools thresholds 3.25, 3.75, 4.25 and 4.75, which have clear and similar effects of display rating in the above RD estimates. The second column pools all thresholds from 1.25 to 4.75, and uses the entire sample. The pooled RD estimates are not sensitive to the extension of pooling thresholds. We can see from the two columns that, an extra half-star on average reduces advertising spending by more than three thousand dollars.
Table 1.3: RD Estimates of the Pooled Regression

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average rating (normalized)</td>
<td>5.506 (.509)</td>
<td>5.581 (.495)</td>
</tr>
<tr>
<td>Above threshold (extra half-star)</td>
<td>-3.549*** (.252)</td>
<td>-3.188*** (.153)</td>
</tr>
<tr>
<td>Average rating × Above threshold</td>
<td>-3.716 (8.280)</td>
<td>-1.551 (7.500)</td>
</tr>
</tbody>
</table>

Notes. Dependant variable is ad spending for both columns. Column (1) uses the sub-sample of restaurants with average rating above 3. Column (2) uses the full sample of restaurants (with both Ad and Yelp information). Bandwidth used in the pooled regression is 0.25 stars. Significance levels: *** 1%, ** 5%, * 10%.

1.5 Robustness

1.5.1 Higher-Order Polynomial Regressions

Although the plots in Figure 1.1 show mostly linear relationships, to ensure the effects of Yelp rating are estimated correctly, I check the robustness of the estimated effects to second order polynomial regressions. In particular, for thresholds from 3.25 to 4.75, I estimate

\[ y_i = \alpha + \beta_1 (X_i - c) + \gamma_1 (X_i - c)^2 + \tau I(X_i \geq c) + \beta_2 (X_i - c) I(X_i \geq c) + \gamma_2 (X_i - c)^2 I(X_i \geq c) + \delta Z_i + \epsilon_i \quad (1.4) \]

and the regression results are given in Table 1.4.

Since higher-order polynomial regressions work better with wider bandwidths (Lee and Lemieux [2010]), I use bandwidth of 0.3 stars for threshold 3.25 and 0.5 stars for thresholds 3.75, 4.25 and 4.75 in the local quadratic regressions.\(^3\) We find that the estimates for thresholds 3.75 and 4.25 are close to those from the local linear regressions. At the threshold 3.25, the estimated drop in advertising due to an extra half-star is even bigger than that from the local linear regression.

\(^3\)For threshold 3.25, using bandwidth 0.3 gives the best goodness of fit. The estimates for thresholds 3.75 are robust to various bandwidth choices. But the estimates for thresholds 4.25 and 4.75 are still not significant for most bandwidth choices.
Table 1.4: RD Estimates With Quadratic Polynomial Regressions

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above 3.25</td>
<td>−6.604*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.436)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above 3.75</td>
<td>−6.387***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.613)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above 4.25</td>
<td>−0.799</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.178)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average rating</td>
<td>72.625*</td>
<td>15.793</td>
<td>−12.568</td>
<td>5.708</td>
</tr>
<tr>
<td></td>
<td>(43.352)</td>
<td>(13.611)</td>
<td>(12.169)</td>
<td>(20.789)</td>
</tr>
<tr>
<td>Average rating²</td>
<td>260.221*</td>
<td>6.593</td>
<td>−23.975</td>
<td>12.524</td>
</tr>
<tr>
<td></td>
<td>(140.827)</td>
<td>(26.466)</td>
<td>(21.858)</td>
<td>(34.024)</td>
</tr>
<tr>
<td>Average rating × Above threshold</td>
<td>−65.210</td>
<td>−4.293</td>
<td>5.611</td>
<td>−74.733</td>
</tr>
<tr>
<td></td>
<td>(53.296)</td>
<td>(19.045)</td>
<td>(22.673)</td>
<td>(79.644)</td>
</tr>
<tr>
<td>Average rating² × Above threshold</td>
<td>−263.594</td>
<td>−30.368</td>
<td>37.445</td>
<td>258.691</td>
</tr>
<tr>
<td></td>
<td>(172.663)</td>
<td>(38.364)</td>
<td>(56.502)</td>
<td>(277.183)</td>
</tr>
<tr>
<td>Observations</td>
<td>3394</td>
<td>6670</td>
<td>4199</td>
<td>1198</td>
</tr>
</tbody>
</table>

Notes. The dependent variable is advertising spending (unit: USD Thousands). The variable “Average rating” is normalized by the corresponding threshold in each column, i.e., $X_i - c$. The variable “Above threshold” in the interaction term corresponds to the indicator variable in each column: “Above 3.25”, “Above 3.75”, “Above 4.25” and “Above 4.75” respectively. Bandwidth is 0.3 in column 1, and 0.5 in other columns. Significance levels: *** 1%, ** 5%, * 10%.

1.6 Validation Tests of RD Design

To ensure the validity of the RD design, I conduct a rich set of tests, including the routine tests used for RD designs and an additional group of tests for the possibility of reviews manipulation. From these tests, we can see that the RD design in this paper is valid, and therefore we have correct estimates.

1.6.1 Density of Assignment Variable

The first set of tests is the routine check for the discontinuity in the density of the assignment variable, i.e. the average rating. The purpose of checking for the discontinuity in average rating is to ensure that there is no manipulation in the sorting, which would undermine the validity of the RD design. In principle, we do not need the continuity in the assignment variable. But looking at the density and the results from related tests will help us determine the likelihood of manipulation by firms.

The density of average rating is plotted in Figure 1.2. We can see there are a lot of jumps, especially at each half-star and whole-star thresholds. And the jumps at the RD thresholds, i.e.
Figure 1.2: Density of Average Rating

Table 1.5: Percentiles of Reviews

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>Reviews Per Restaurant</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>0</td>
</tr>
<tr>
<td>10%</td>
<td>2</td>
</tr>
<tr>
<td>25%</td>
<td>9</td>
</tr>
<tr>
<td>50%</td>
<td>26</td>
</tr>
<tr>
<td>75%</td>
<td>71</td>
</tr>
<tr>
<td>90%</td>
<td>179</td>
</tr>
<tr>
<td>95%</td>
<td>305</td>
</tr>
<tr>
<td>99%</td>
<td>824</td>
</tr>
<tr>
<td>Observations</td>
<td>13360</td>
</tr>
</tbody>
</table>

the .25 and .75 thresholds, are within the usual range of bin-to-bin jumps in this histogram plot. The nontrivial magnitude and amount of jumps from this graph point to the fact that average rating calculated from a small number of reviews will be very discrete.

Most local restaurants have a small number of customer reviews. As Table 1.5 shows, a large percent of restaurants have a small number of reviews. The average ratings of these restaurants will become several mass points in the density plot. And this discreteness, which is not resulting from manipulation, will interfere with our test for the manipulation. Therefore, I adjust the test on the density to the particular context of this research.

As a first step, I separate the dataset to two equal size parts, one with restaurants that have no
Analyses will be conducted on these two subsamples separately, and compared with the test on the full sample. Histograms for the two subsamples are given in Figure 1.3. It is clear that the big jumps in the density histogram of the full sample (Figure 1.2) are mostly from the subsample with reviews less than 26. The density of average rating in the subsample with reviews more than 26 shows much less discontinuity. In particular, the McCrary test (McCrary [2008]) is passed for the RD thresholds 3.25 and 3.75 where we see significant effects of display rating from the RD estimates in Section 4.

4Note, from Table 1.5, that 26 is the 50% percentile.
Table 1.6: RD Estimates of Jumps in Percentage of Five-Star Reviews

<table>
<thead>
<tr>
<th>Sample</th>
<th>≤ 26 reviews</th>
<th>&gt; 26 reviews</th>
<th>Full sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Pooled thresholds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above threshold</td>
<td>-0.088*** (0.012)</td>
<td>0.014* (0.007)</td>
<td>-0.037*** (0.007)</td>
</tr>
<tr>
<td>Observations</td>
<td>5634</td>
<td>6650</td>
<td>12284</td>
</tr>
<tr>
<td>(b) Around threshold 3.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above threshold</td>
<td>-0.017 (0.012)</td>
<td>0.008 (0.006)</td>
<td>-0.001 (0.006)</td>
</tr>
<tr>
<td>Observations</td>
<td>1472</td>
<td>1866</td>
<td>3338</td>
</tr>
<tr>
<td>(c) Around threshold 3.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above threshold</td>
<td>-0.005 (0.013)</td>
<td>0.001 (0.005)</td>
<td>0.003 (0.005)</td>
</tr>
<tr>
<td>Observations</td>
<td>1449</td>
<td>2716</td>
<td>4165</td>
</tr>
<tr>
<td>(d) Around threshold 4.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above threshold</td>
<td>-0.074*** (0.019)</td>
<td>0.005 (0.006)</td>
<td>-0.044*** (0.009)</td>
</tr>
<tr>
<td>Observations</td>
<td>1194</td>
<td>1334</td>
<td>2528</td>
</tr>
</tbody>
</table>

Notes. This table gives the RD estimates of the jumps in the percentage of five-star reviews from four regressions (a pooled regression and three regressions around separate thresholds) conducted on each of three samples (two subsamples and the full sample). Covariates are used to improve the precision of estimates, but the estimated coefficients to them are not shown here. Significance levels: *** 1%, ** 5%, * 10%.

1.6.2 Review Characteristics and Covariates

For the subsample with less reviews, it is meaningless to conduct the McCrary test of discontinuity because the discontinuity might be entirely caused by the discreteness resulting from averaging with a small number of reviews. But manipulation in reviews can be tested more directly and even easier on this subsample by looking at the percentage of five-star reviews and the standard deviation in review ratings. If a restaurant with a small number of reviews posts fake five-star reviews, the percentage of five-star reviews will be easily driven higher, and the standard deviation in its review ratings will appear much higher than other restaurants that have similar average rating and did not manipulate the reviews. Therefore the next step is to check the percentage of five-star reviews and the standard deviation in review ratings, in both subsamples as well as the full sample.

Table 1.6 shows the RD estimates of the effects of an extra half-star in display rating on the percentage of five-star reviews. The regressions use specification (1.2) or (1.3) with the percent-
Table 1.7: RD Estimates of Jumps in SD of Ratings

<table>
<thead>
<tr>
<th>Sample</th>
<th>≤ 26 reviews</th>
<th>&gt; 26 reviews</th>
<th>Full sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Pooled thresholds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Display rating jumps by half a star)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above threshold</td>
<td>−0.012 (0.021)</td>
<td>−0.007 (0.009)</td>
<td>−0.026∗ (0.012)</td>
</tr>
<tr>
<td>Observations</td>
<td>5634</td>
<td>6650</td>
<td>12284</td>
</tr>
<tr>
<td>(b) Around threshold 3.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Display rating jumps from 3 to 3.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above threshold</td>
<td>−0.127*** (0.033)</td>
<td>0.007 (0.013)</td>
<td>−0.057*** (0.006)</td>
</tr>
<tr>
<td>Observations</td>
<td>1472</td>
<td>1866</td>
<td>3338</td>
</tr>
<tr>
<td>(c) Around threshold 3.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Display rating jumps from 3.5 to 4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above threshold</td>
<td>−0.014 (0.027)</td>
<td>0.002 (0.009)</td>
<td>−0.0002 (0.011)</td>
</tr>
<tr>
<td>Observations</td>
<td>1449</td>
<td>2716</td>
<td>4165</td>
</tr>
<tr>
<td>(d) Around threshold 4.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Display rating jumps from 4 to 4.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above threshold</td>
<td>−0.222*** (0.042)</td>
<td>0.001 (0.013)</td>
<td>−0.143*** (0.021)</td>
</tr>
<tr>
<td>Observations</td>
<td>1194</td>
<td>1334</td>
<td>2528</td>
</tr>
</tbody>
</table>

Notes. This table gives the RD estimates of the jumps in standard deviation of ratings from four regressions (a pooled regression and three regressions around separate thresholds) conducted on each of three samples (two subsamples and the full sample). Covariates are used to improve the precision of estimates, but the estimated coefficients to them are not shown here. Significance levels: *** 1%, ** 5%, * 10%.

regressions are estimated at pooled thresholds or at separate thresholds in each of the two subsamples as well as in the full sample. We again see no significant increase with precise estimates at each threshold (or the pooled threshold).

These tests provide more direct evidence of the absence of reviews manipulation in this dataset than the foregoing density checks. Checking the percentage of five-star reviews and the standard deviation in ratings in the subsample with small number of reviews per restaurant provide a stronger assurance because these small (in terms of the number of reviews) restaurants are the ones that are
easier and more likely to manipulate their ratings by posting fake reviews.

I also checked the continuity in covariates that I used in the RD regressions, including Dressy, Price, Waiter, Delivery and Reservation, and found no discontinuities in these covariates at the RD thresholds. Therefore it is not the case that restaurants of a certain type are grouped above the jump thresholds of display rating.

1.7 Discussion

The rounding algorithm that Yelp uses for display rating creates an exogenous change in display rating as restaurants’ average ratings cross a threshold. Average rating correlates smoothly with a restaurant’s quality reputation, whereas the display rating randomly jumps or drops by half a star around each threshold. Therefore the RD estimates of the drops in advertising spending at each threshold (or the pooled threshold) show us how local restaurants are changing their advertising strategy in response to an increase in Yelp display rating.

Display rating is only part of the rating information on Yelp.com. To have a complete picture of local restaurants’ advertising response to their online ratings, we also need to look at how advertising changes with the rest of rating information on Yelp, i.e., the distribution of reviews. Here I use the mean statistic, i.e., the average rating, to summarize the distribution of reviews. This is not implying that the average rating can fully summarize the distribution, but rather that the average rating reveals the quality reputation which is conveyed through the distribution of reviews.

The contrast between the quality reputation information generated by consumers and the display rating generated by the exogenous rounding algorithm is the key comparison in this research. When display rating jumps, the quality reputation revealed from the distribution reviews is almost constant. Therefore the advertising response to display rating is not responding to consumers’ belief about a restaurant’s quality. On the other hand, by studying how advertising responds to average rating in the following step, we will see how advertising changes with the restaurants’ quality reputation.

\footnotetext{That restaurants have no (complete) control on the display rating is ensured by the validation tests in the preceding section.}
Table 1.8: Pooled RD With Interval Dummies

<table>
<thead>
<tr>
<th>Subsamples</th>
<th>Full Sample</th>
<th>Price $$</th>
<th>Price $$$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average rating (normalized)</td>
<td>2.811***</td>
<td>3.302***</td>
<td>6.660***</td>
</tr>
<tr>
<td>Above threshold</td>
<td>−1.085***</td>
<td>−1.257***</td>
<td>−2.309***</td>
</tr>
<tr>
<td>Average rating × Above threshold</td>
<td>−0.854***</td>
<td>−1.339***</td>
<td>3.407***</td>
</tr>
<tr>
<td>Dressy</td>
<td>0.838***</td>
<td>−0.556</td>
<td>1.144***</td>
</tr>
<tr>
<td>Delivery</td>
<td>−0.931***</td>
<td>−1.023**</td>
<td>−10.686***</td>
</tr>
<tr>
<td>Reservation</td>
<td>0.176</td>
<td>0.345</td>
<td>−0.620</td>
</tr>
<tr>
<td>Price $$</td>
<td>0.384</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price $$$</td>
<td>1.337**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Price $$ $$ $$</td>
<td>0.311</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>6.018***</td>
<td>5.677**</td>
<td>14.235***</td>
</tr>
<tr>
<td>Observations</td>
<td>10197</td>
<td>6996</td>
<td>1076</td>
</tr>
</tbody>
</table>

Two-tail p value

Test: \( \beta_1 + \beta_2 = 0 \)

| p value | 0.27 | 0.37 | 0.09 |

Notes. This table shows the pooled RD estimates using advertising spending as the dependent variable in all three columns. The regressions are on the full sample, the subsample with restaurants of price level $$, and the subsample with restaurants of price level $$. Average rating is normalized to the nearest half-star threshold. Outliers are ruled out. The bottom row gives the p value from the two-sided test of the slope of advertising by average rating above thresholds. Significance levels: *** 1%, ** 5%, * 10%.

Because from Figure 1.1 and Figure 1.4 (in appendix 1.9.1) we can see the between-threshold intervals have similar advertising patterns, I pool all RD thresholds together and estimate the overall trend of advertising spending when average rating is increasing while the display rating is constant. I normalize the average rating by its nearest half-star threshold, and assign a dummy variable for each interval with length 0.5 star around a threshold. Using the specification (1.3) with a group of interval dummies, the estimates are given in Table 1.8.

The first column of Table 1.8 contains the pooled RD estimates using advertising spending as the dependent variable. The sample contains all local restaurants that advertise in only one market. Outliers are excluded. In particular, I exclude the restaurants that advertise more than $50,000 in 2014, and these restaurants constitute about 3% of this sample. Pooling all thresholds together, we can see that the slope of advertising over average ratings to the left of thresholds is significantly positive (2.811). The slope to the right of thresholds has a positive point estimate
(\hat{\beta}_1 + \hat{\beta}_2 = 1.957). From the two-tail test the slope is significant from zero at 73% confidence level, but with the positive point estimate, we can reject against the alternative hypotheses that \( \hat{\beta}_1 + \hat{\beta}_2 > 0 \) at 86.5% confidence level.

Yelp categorizes restaurants to four price ranges: $, $$, $$$, and $$$$$. The price range $$ covers the largest group of restaurants. Price ranges $ and $$$ have similar numbers of restaurants, and the range $$$$ has only a little more than 100 restaurants in this entire sample of over 13,000 restaurants. Column 2 and column 3 provide the pooled RD estimates on the subsample of price range $$ and the subsample of price range $$$.

The big subsample of price range $$ (column 2) has similar regression estimates as those in the full sample (column 1). The slope of advertising over average ratings to the left of thresholds is significantly positive (3.202), but the slope to the right of thresholds is positive at a lower confidence level (81.5%). The subsample of price range $$$ (column 3) has a little different estimates than those in column 1 and 2. The slope of advertising on the left of thresholds is positive (6.660) yet with a not very precise estimate. However, the slope to the right of thresholds is significantly positive with confidence level 95.5%.

The subsamples of price range $ and $$$$ have either too big variance or too small number of observations for us to have precise estimates, and the estimates on these two subsamples are therefore not shown here.

In general we can see from the estimation results that, apart from the jump thresholds, i.e., when the display rating is constant, higher advertising spending is associated with higher average ratings.

1.8 Conclusion

In this paper I study the advertising spending of local restaurants as their Yelp display rating changes or as their average rating changes. Surprisingly the advertising responds in opposite directions to the changes in display rating and in average rating.

---

6Approximate cost per person: $ = under $10, $$ = $11-$30, $$$ = $31-$60, $$$$ = above $61
From the RD design estimates, we find the advertising spending decreases significantly as the display rating jumps at the thresholds above 3. In other words, when the Yelp display rating is above 3, every time the display rating jumps by 0.5 star, local restaurants significantly reduce their advertising spending. The display ratings below 3 do not have significant effect on advertising spending.

From the pooled RD regression with interval dummies, we see that higher advertising spending levels are associated with higher average ratings on the between-threshold intervals, where the display rating is not changing. Average rating summarizes how consumers think about the quality of a restaurant (in their reviews). This finding implies that restaurants with better quality reputation are advertising more.

The key reason that we are seeing this opposite advertising response to the two rating information is the capacity limit of local businesses. A higher display rating increases the click rates of a restaurant and brings more “searchers”, i.e., consumers who search for unknown (nearby) restaurants. Therefore the display rating drives up the “volume”. As analyzed in Chapter 2, a higher average rating, keeping the display rating constant, increases new consumers’ willingness to pay because they can see a better quality reputation from the higher average rating. So the average rating drives up the “margin”.

While both the “volume” increase and the “margin” increase bring higher profits for restaurants, the capacity limit of local restaurants will become binding if the display rating is high enough. To explain using the model in Chapter 2, the benefit of advertising comes from the additional group of new consumers that are attracted by the advertisements. As the “volume” increase becomes big enough (when the display rating is high enough), the benefit of advertising will be crowded out and eventually goes to zero. Intuitively, if the “searchers” are enough to fill the seats in a restaurant, this restaurant would have no benefit from bringing more new consumers by advertising. Note that the “margin” increase will never be bound by the capacity limit, and is therefore “the more the better”.

Therefore as the display rating gets high enough, advertising decreases because it becomes less profitable but is still costly. On the other hand, as the average rating increases, advertising increases because the profit margin increases.
Further investigation will be conducted using a dynamic setting. In a dataset with restaurants’ reviews and advertising spending at different points of time, we can look at how advertising affects the mobility of average rating to higher intervals and complete our study on the interaction between advertising and online ratings.

1.9 Appendix

1.9.1 RD Graphs at Lower Thresholds

Figure 1.4: Binned Scatter Plots of Ad Spending by Average Rating around Lower Thresholds

The four panels in this figure are the RD graphs showing the jumps at thresholds 1.25, 1.75, 2.25 and 2.75. We can see there are no clear discontinuities at these thresholds. And this graphical interpretation also matches the regression results which show that there are insignificant drops in advertising spending at these four thresholds with precise estimates.
Chapter 2

How Do Firms Advertise When Customer Reviews Are Available?

This paper uses a game-theoretic model to analyze competing firms’ advertising and pricing decisions. Here customer reviews are available and firms may build up loyal customer bases. I find that highly-rated firms are more likely to advertise more, i.e., online reviews complement advertising. Comparative static results can explain the results found in the first chapter. Intuitively, when the capacity of a local business becomes limited, a jump in the display rating will reduce the complementary effect of online reviews on advertising. I also analyze an extension of the model, where an entrant and an incumbent interact. I find that customer reviews undo the “fat-cat” effect of a large incumbent with lots of loyal customers.

2.1 Introduction

In recent years, customer reviews have become an important part of consumers’ shopping experience. The percentage of consumers who read reviews (occasionally or regularly) before purchase to determine the quality of a business has been steadily increasing, and in 2015 this number reached 92%.\(^1\) When consumers can get information about firms’ qualities from online customer reviews, how do firms compete when they have different reviews? In particular, my paper studies competing firms’ advertising strategies when customer reviews are available, and combining with data, offers new insights on the interaction between online customer reviews and informative advertising: Are they substitutes or complements?

Previous research finds that reviews, specifically good reviews, are a substitute for advertising, and in particular, that firms with better reviews advertise less. However, I introduce two realistic features that complicate this story. The first is loyal customers. In particular, consumers that are satisfied with a local business not only leave good reviews, and they will also return and become

---

\(^1\)Data source: The Local Consumer Review Survey by BrightLocal. This survey was conducted over 2 weeks in July each year, starting from 2010. The 2015 survey has 2,354 entries with 90% of respondents coming from the US and 10% from Canada.
loyal customers. Second, in many cases, advertising and reviews offer different types of information. Whereas advertising informs consumers about the existence of a product or raises consumer awareness, reviews offer a more credible source of information about quality. I show these features interact to generate surprising results about the relationship between reviews and advertising.

I consider a non-durable experience good. For such a good, consumers’ individual values are initially unknown and can be learned only after purchase, and each consumer may purchase more than once. Advertising for an experience good, as Nelson [1970] points out, cannot convey direct information of product quality because consumers will not trust such information that is not verifiable before purchase. Indirect information that advertising may carry includes a product’s existence and price.

By advertising, a firm informs new consumers about the existence of its product, and these informed consumers will have access to the firm’s price and reviews before purchase. A familiar example is consumers see various types of advertisements of a local restaurant, and then go to Yelp.com to check this restaurant’s reviews and menu (including prices). There are also some other less obvious forms of advertising. Consider the example of Amazon.com: firms need to pay a fee ($39.99) per month to be listed and sell on Amazon. Shoppers visiting Amazon can see all listed firms’ prices and reviews. Therefore, advertising in my paper can be interpreted more generally as a marketing tool that raises awareness of consumers about a product, and consequently makes the price and reviews of this product observable to those consumers.

Customer reviews in my model are reports of individual satisfaction from previous buyers. Consumers have heterogeneous preferences, and each consumer may be satisfied or dissatisfied with a product. Before purchase, individual satisfactions are unknown and all consumers are ex ante homogeneous. Consumers who see reviews of a product, see how previous buyers were satisfied with it, and will use these reviews to update their belief about the likelihood of themselves being satisfied with the product. I assume that consumers report truthfully, but this assumption can be relaxed and does not affect the conclusions.

I define a product’s quality to be its ex ante satisfaction likelihood for a new consumer. That is, a higher quality product has a higher probability, ex ante, to make a random new consumer
satisfied. A new firm’s product quality is unknown, and consumers and firms hold a common prior belief about it. As a firm receives reviews from its buyers, the belief about its product quality will be updated according to Bayes’ rule.

In this model, two firms with unknown qualities compete with each other. Firms first choose whether or not to advertise, and after observing each other’s advertising decision, firms choose prices. There are a finite number \( n \) of consumers randomly drawn from the population to become “shoppers”. (Justifications can be the group of consumers who consider having lunch at a local restaurant today in a certain neighborhood, or the group of consumers who visit Amazon.com to search for a specific product in a month.) Only firms that choose to advertise are known to the \( n \) shoppers. If only one firm advertises, it can charge a monopoly price. But if both firms advertise, they need to compete for the shoppers in price.

I consider the competition between two firms that have operated for some periods, and hence have already received some reviews. The good reviews that a firm receives come from its previous buyers who were satisfied with their purchase. I assume satisfied previous buyers will repeat purchase and become “loyal customers” of this firm.\(^2\) Firms know about their own and each other’s loyal customer base (because they are automatically aware of each other and can see each other’s reviews). If a firm advertises, the new shoppers will be able to see this firm’s reviews and use them to update the belief about this firm’s product quality. The shoppers make purchase decisions based on both prices and reviews of the advertised firm(s).

For a firm with some reviews and a nonzero loyal customer base, in choosing a price, this firm trades off between charging a high price to sell only to loyal customers, and lowering price to attract new shoppers. In the price competition between two advertised firms, there is no pure strategy equilibrium if at least one firm owns nonzero loyal customers. However, there is a unique mixed strategy equilibrium, in which both firms randomize pricing, and the firm with better reviews and more loyal customers randomizes over a higher range of prices.

In the subgame perfect Nash equilibrium of the two-stage advertising-pricing game, firms’ advertising strategies will depend on how differently the two firms are rated in their reviews. If

\(^2\)Repeat purchase from loyal customers is a very important part of local restaurants’ business.
firms are similarly rated, there will be multiple equilibria and one firm advertises only if the rival does not. In other words, advertising is a strategic substitute for similarly rated firms. However, if one firm is rated much better than the rival, the better-rated firm is the only one that advertises. Therefore, having a relatively higher rating helps a firm to be dominant in advertising.

In an extension, I show that these conclusions still hold if I allow a group of consumers to search for firms. The “Searchers” do not need to receive advertisements to be informed about firms, and they can see all firms’ prices and reviews. This robustness check makes the implications of my model applicable to a more general case, where consumers may search but firms can advertise to reach more consumers.

In the main body of this paper, I consider the case where two firms have the same number of previous buyers. But the conclusions also hold if two firms have different numbers of previous buyers, i.e. asymmetric firm histories. In particular, I extend the main model to an entry deterrence problem which can be interpreted as an extreme case of asymmetric firm histories. An incumbent firm interacts with a potential entrant, and entry happens if the entrant chooses to advertise. The incumbent already has some reviews from previous buyers, and owns a loyal customer base. In this extension, I find that entry can be deterred if the incumbent was successful enough in the past, i.e. has a high ratio of good reviews. There is an interesting interpretation of my model vis--vis the “fat-cat” effect as in Fudenberg and Tirole [1984]. When the incumbent has a big group of loyal customers (a “fat cat”), in Fugenberg and Tirole’s model, it should be weak in the competition with potential entrants. However, with the availability of customer reviews, if the incumbent has a large enough percentage of good reviews, it will successfully deter entry and is therefore a “fat-but-strong cat”.

Finally, I use data on local restaurants’ advertising spending and Yelp reviews to test the model prediction and find supporting evidence on the positive effect of a higher average rating on restaurants’ advertising spending. In addition, by using a Regression Discontinuity Design, I successfully separate the effect of display rating from the effect of average rating.3  In the discussion section,

3See Figure [fig:Yelp page example] in Appendix 2 for an example Yelp web page showing the display rating and the average rating of a restaurant.
I explain why we observe these two effects having opposite directions in their relationship with advertising.

2.2 Literature Review

My paper first contributes to the literature of advertising (for a review, see Bagwell [2007]), and in particular relates to the papers on informative advertising for experience goods. Nelson’s seminal paper (Nelson [1970]) differentiates between search goods and experience goods, and starts a discussion about indirect informative advertising for experience goods Nelson [1974]. An important benefit of such advertising is creating repeat purchases. Following Nelson’s conjecture, Schmalensee [1978], Kihlstrom and Riordan [1984], Milgrom and Roberts [1986] and Hertzendorf [1993] have formally studied this repeat-purchase effect of advertising. Because of the various restrictions these papers put on firms’ or consumers’ dynamic decisions, the repeat-purchase effect has not been fully investigated. Although creating repeat purchases in the future is a long-term benefit of advertising (the repeat-purchase effect), the current repeated purchases from previous buyers reduces a firm’s incentive to advertise in the current period (the loyal-customer effect). In my paper, I take the loyal-customer effect into account to study firms’ advertising strategies.

There have been several papers investigating the Bertrand price competition between firms with loyal customers, and showing that such Bertrand games have no pure strategy equilibrium. Varian [1980] and Narasimhan [1988] study the competition between firms when each firm has an exogenous group of loyal customers (“uninformed consumers” in Varian [1980], and analyze the mixed strategy equilibrium of the Bertrand game. In McGahan and Ghemawat [1994] and Chioveanu [2008], firms are allowed to invest first in building their loyal customer bases. They both assume that the size of a firm’s loyal customer base is determined with certainty by firms’ decisions (service in McGahan and Ghemawat [1994] or persuasive advertising in Chioveanu [2008]. In this paper, I study the two-stage advertising-pricing game, and the mixed strategy equilibrium only happens in a subgame and actually serves as a threat to the firm without advantage in the price competition.
Another literature that my paper contribute to is the one on customer reviews. There have been many great papers showing the importance of online customer reviews. Using data on Amazon.com and BN.com, Chevalier and Mayzlin [2006] show that online reviews are very influential to consumers’ purchase decisions about books. Luca [2011] uses data from Yelp.com to study how reviews affect firms’ sales and how consumers learn from reviews. Sun [2012] demonstrates how the variance and the average of ratings interact in affecting firms’ sales.

Within this literature, my paper is closely related to a specific group of papers that studies the interaction between firms’ marketing strategies and product reviews. Chen and Xie [2005] study two competing firms’ advertising and pricing strategies in response to third-party product reviews when a lot of consumers have strong preference on horizontal attributes. Mayzlin [2006] talks about when firms post fake reviews in response to different ratings from customer reviews. Another paper by Chen and Xie [2008] studies how a monopoly firm 1) chooses how much product attribute information to reveal, 2) decides whether to make previous customer reviews available to future consumers, and 3) proactively control the informativeness of customer reviews, for different types of products. In Chen and Xie [2008], all customer reviews are the same and give one signal, match or mismatch. My paper takes into account the heterogeneity in customer reviews and uses Bayesian learning to model new consumers’ belief updating using these reviews. Therefore, when firms are rated differently, their advertising strategies in competition will be different.

The rest of this paper proceeds as follows: Section 3 gives the setting of the main model, and Section 4 contains the equilibrium analysis, Section 5 contains two extensions: entry game and Searchers model. In Section 6 I test the model prediction with data, Section 7 discusses the empirical findings and Section 8 concludes.

2.3 Main Model

2.3.1 Players and Information

There is a continuum of consumers in the population, and a finite number \( n \in \mathbb{N} \) of consumers are randomly drawn to be shoppers. Each would like to buy at most one unit of a good. The good
is an *experience* good. After purchase, a consumer derives either a value of 0 or 1 from consuming the good, but this individual match value is unknown before purchase, and consumers are *ex ante* homogeneous. I also assume that consumers are initially unaware about the availability of the good, but they can be informed by firms through advertising.

Firm A and Firm B sell the good, and compete for only one period. Marginal cost of production is assumed to be the same for both firms, and normalized to 0. An advertisement by a firm informs shoppers about the availability of the firm’s product. If both Firms A and B advertise, the n shoppers are informed of both firms’ goods. Advertising is a discrete-choice variable. A firm can choose either to advertise (\( M = 1 \)) or not to (\( M = 0 \)). The cost of advertising is fixed and denoted by \( c \in \mathbb{R}_+ \).

Let \( \theta^A \) and \( \theta^B \) denote, respectively, the probabilities that products of Firms A and B will yield a value 1 to a shopper. Neither firms nor consumers know these probabilities *ex ante*. However, it is common knowledge that \( \theta^A \) and \( \theta^B \) are drawn independently and identically from a uniform distribution on \([0, 1]\). We may call \( \theta^A \) and \( \theta^B \) firms’ product qualities, or their types.

There is a special group of consumers, called “loyal customers”. If a previous buyer derives a value of 1 from a firm’s product, this consumer will repeat purchase and become a “loyal customer” of this firm thereafter. Assume that, at the beginning of this model, each firm has already built up a loyal customer base, denoted by \( L^k, k = A, B \). Let \( T^A \) and \( T^B \) denote the total number of previous buyers of Firm A and Firm B respectively.\(^4\) Assume \( T^A \) and \( T^B \) are both finite numbers, i.e. Firms A and B have not been operated for infinite periods, so that the ratio \( L^k / T^k \) cannot predict precisely the value of \( \theta^k \).

Customer reviews are defined to be consumers’ truthful reports of their satisfaction with a product after consuming it. Reviews can be good (taking value 1), or bad (taking value 0). Assume that all previous buyers wrote reviews.\(^5\) Firm A’s good reviews come from those previous buyers of Firm A who derived value 1, and these consumers now constitute Firm A’s loyal customer base (\( L^A \)).

\(^{4}\text{For this customer review model to be meaningful, assume at least one of } T^A \text{ and } T^B \text{ is nonzero. In other words, assume at least one firm has some reviews.}\)

\(^{5}\text{This assumption is not crucial and can be relaxed, because new consumers learn about a product’s quality only through the available reviews. The shopping experience, good or bad, of those previous buyers who have not left reviews failed to be conveyed to new consumers. What matters is the available reviews.}\)
Firm A’s bad reviews come from the dissatisfied previous buyers of Firm A, \( T^A - L^A \). Similarly, for Firm B, the good reviews come from its satisfied previous buyers who now constitute Firm B’s loyal customer base \((L^B)\), and bad reviews come from its dissatisfied previous buyers, \( T^B - L^B \).

Once a consumer is aware of a firm’s existence, she will have free access to all of the firm’s previous customer reviews, \( L^k \) good and \( T^k - L^k \) bad, and the firm’s price \( p^k \). Firms can always see each other’s previous customer reviews.

The firms and consumers interact over two stages in the single period.

**Stage 1 (Advertising)** Each firm chooses whether to advertise. The new consumers \((n)\) are informed by firms’ advertisements, and become aware of the advertised firm(s).

**Stage 2 (Pricing)** Observing each other’s advertising decision, firms now choose prices simultaneously. The new consumers have access to the advertised firms’ customer reviews and prices, and then make purchase decisions. Each firm’s loyal customers repeat purchase.

If both firms advertise, the \( n \) shoppers choose between two firms based on their reviews and prices. Loyal customers do not consider switching because they are already enjoying the highest possible value – 1. I assume that a firm cannot discriminate between its loyal customers and new shoppers. The same price is charged to all buyers.

If Firm A’s quality \( \theta^A \) is known, after purchasing from Firm A, a consumer gets value 1 with probability \( \theta^A \); analogously with probability \( \theta^B \) if the purchase is from Firm B. However, two firms’ qualities are unknown, to consumers and to firms themselves.\(^6\) The probability of being satisfied (i.e., getting value 1) with a product is unknown, and everyone learns from firms’ previous customer reviews.

\(^6\)This assumption can be interpreted as the uncertainty about whether consumers will like the food of a restaurant. If restaurant owners know exactly what consumers like, they will all provide the most favorable food, and we won’t see bad reviews at all. However, obviously this is not true.
2.3.2 Belief Updating

I assume consumers are rational. First, before knowing the individual value of a product, a rational consumer updates her belief of the product quality according to Bayes’ rule. Second, a consumer values a product by the expected quality in the initial purchase, and by her individual value of this product in repeated purchases, and a consumer is rational in that she will not pay for prices above her value of a product.

Recall that a firm’s product quality is defined to be the probability \( \theta \in [0, 1] \) that its product will yield a value 1 to a randomly chosen shopper. A belief about a firm’s quality is therefore a probability distribution on \([0, 1]\). I use the Beta distribution to model beliefs.\(^7\) Good and bad reviews of a firm can be viewed as successes and failures of Bernoulli trials, and all serve as signals to update belief. For a prior belief that is described by a Beta distribution, after updating with the Bernoulli trials, the posterior belief again follows Beta distribution, only with updated parameters.\(^8\) Specifically, the belief-updating process is as follows.

For a new firm with no reviews, the common prior belief of its quality is the Beta distribution with parameters 1 and 1, \( \text{Beta}(1, 1) \), which is equivalent to the uniform distribution on \([0, 1]\). Let \( \tilde{\theta}_0 \) denote the expected quality of a new firm, then \( \tilde{\theta}_0 = 1/2 \). Therefore, all new firms share the same prior belief and expected quality, 1/2, even though they may have different true qualities.

As firms start receiving reviews, the beliefs about their qualities will be updated. If a firm has received a total number \( T \) of reviews, among which there are \( L \) good reviews, using Bayes’ rule, the updated expected quality will be \( \tilde{\theta} = \frac{1+L}{2+T} \).

Specifically, in this model, Firm \( A \) has, in total, \( T^A \) reviews, and \( L^A \) out of \( T^A \) are good reviews. Update the common prior belief, \( \text{Beta}(1, 1) \), with these good and bad reviews, and the posterior belief will be distributed as \( \text{Beta}(1 + L^A, 1 + T^A - L^A) \). Therefore, when consumers see Firm \( A \’s \)

\(^7\)The Beta distribution, \( \text{Beta}(a, b) \) (\( a \) and \( b \) are parameters), is a continuous distribution on \([0, 1]\), and the expectation is \( \frac{a}{a+b} \).

\(^8\)For more details about the Beta distribution and its property of being a conjugate prior distribution, please refer to DeGroot and Schervish [2011] (specifically p. 327-333 and Theorem 7.3.1 on p. 394).
reviews, their belief is updated such that the expected quality of Firm A becomes

\[ \tilde{\theta}_A = \frac{1 + L_A}{2 + T_A} \]  

(2.1)

Similarly, Firm B has \( L_B \) good reviews, and \( T_B - L_B \) bad reviews. Therefore, updated with these reviews, Firm B’s expected quality is

\[ \tilde{\theta}_B = \frac{1 + L_B}{2 + T_B} \]  

(2.2)

Only new consumers need the reviews to update their beliefs. Loyal customers of each firm have already learned their personal match value, which is 1, with the product they are buying.

### 2.4 Equilibrium Analysis

In this model, firms’ actions in Stage 1 are publicly observable in Stage 2, and once a firm has advertised, new consumers and firms will have symmetric information about the advertised firm’s (expected) product quality. Therefore, I solve by backward induction for the subgame perfect Nash equilibrium of this two-stage game.

At the beginning of Stage 1, the state of the game is described by two firms’ total reviews, \( T_A \) and \( T_B \) respectively, and good reviews, \( L_A \) and \( L_B \) respectively. For the following analysis, I use a special case where \( T_A = T_B = T > 0 \). The analyses for the other two cases, \( T_A > T_B \geq 0 \) and \( T_B > T_A \geq 0 \), will be essentially the same, and are briefly explained in the Appendix.\(^9\)

Each firm owns a loyal customer base, the size of which equals to the number of the firm’s good reviews, \( L_k, k = A, B \). These loyal customers are willing to pay price \( 1 \) for the firm’s product. If the firm advertises, new consumers read its reviews and are willing to pay \( \tilde{\theta}_k = \frac{1 + L_k}{2 + T_k} < 1 \) for the firm’s product.

By the equilibrium analysis of this two-stage game, we want to see firms’ advertising strategies \( (M^k) \) in competition when they have different ratios of good reviews \( (L_k/T_k) \).\(^{10}\) Advertising is

---

\(^9\) An extreme case of the asymmetric previous-buyers setting is an entry game with an established incumbent and a new entrant, \( T_A > T_B = 0 \), which is analyzed in Section 5 as an extension.

\(^{10}\) The ratio of good reviews can be roughly corresponded to a firm’s rating.
Table 2.1: The Normal Form of The Game

<table>
<thead>
<tr>
<th>Firm B</th>
<th>Firm A</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M^B = 1$</td>
<td>$M^A = 1$</td>
</tr>
<tr>
<td>$M^B = 0$</td>
<td>$(L^A + n)\hat{\theta}^A - c, L^B$</td>
</tr>
</tbody>
</table>

costly (fixed cost $c$), and the benefit it brings is “expansion”: to expand a firm’s customer base and sell to more consumers.

For a firm with nonzero loyal customers, if it does not expand (by advertising), it always has a “secured profit” because this firm can charge price 1 to its loyal customers. So Firm A’s secured profit is $L^A$, and Firm B’s is $L^B$. The existence of secured profit reduces a firm’s incentive to advertise. When Firm A owns a group of loyal customers $L^A > 0$, if the highest profit that Firm A can obtain from expansion is lower than its secured profit ($L^A$), Firm A will never choose to expand (by advertising). The highest profit Firm A can get from expansion is when the opponent does not advertise, and Firm A sells to all new consumers ($n$) and its loyal customers ($L^A$) at the monopoly price $\tilde{\theta}^A$: $(L^A + n)\tilde{\theta}^A - c$. In other words, it is profitable for Firm A to expand through advertising only if

$$(L^A + n)\tilde{\theta}^A - c \geq L^A \quad (2.3)$$

We say Firm A satisfies the “Profitable Expansion” (PE) condition if 2.3 is satisfied. Analogously, Firm B satisfies the PE condition if

$$(L^B + n)\tilde{\theta}^B - c \geq L^B \quad (2.4)$$

If no firms satisfy the PE condition, the equilibrium will be trivial: no firm advertises and each firm sells to its loyal customer base at price 1. If there is only one firm, say Firm A, that satisfies the PE condition, the equilibrium will be that only Firm A advertises. And in this equilibrium, Firm A charges a price $p^A = \tilde{\theta}^A$ to both new and loyal customers, while Firm B charges 1 and earns its secured profit $L^B$ from loyal customers.

Suppose the state variables are such that both firms satisfy PE, i.e., both (2.3) and (2.4) hold,
then the normal form of the game is as shown in Table 2.1. For subgames where at most one firm advertises, the payoffs are straightforward. If both firms advertise, the \( n \) shoppers are aware of both firms and can see all customer reviews. A pricing subgame follows, and I show next how two established firms compete for new shoppers by price.

### 2.4.1 The Pricing Subgame

Firms have different expected qualities unless they have exactly the same number of good reviews.\(^{11}\) The pricing subgame is therefore a Bertrand competition between firms producing goods of different expected qualities (at the same cost 0). New shoppers choose the product from Firm A if and only if

\[
\tilde{\theta}^A - p^A > \tilde{\theta}^B - p^B
\]  

Let \( d = \tilde{\theta}^A - \tilde{\theta}^B \). Therefore, Firm A tends to undercut Firm B’s price \( p^B \) by charging just below \( p^B + d \), and similarly, Firm B tends to undercut Firm A by charging just below \( p^A - d \).

Unlike in common Bertrand games, firms here are unwilling to undercut each other all the way down to the marginal cost of production (here it is 0). Because of loyal customers (and hence the secured profit), there is a lowest price that a firm is willing to charge in the pricing subgame, which I call the firm’s “reservation price” in the price competition. Firm A does not want to charge any price that yields a lower profit from the price competition than Firm A’s secured profit \( L_A \), and the lowest price that Firm A is willing to charge in the price competition satisfies \((L^A + n)p^A = L^A\). So I call Firm A’s reservation price

\[
\gamma^A = \frac{L^A}{L^A + n}
\]  

Similarly, Firm B’s reservation price in the price competition is

\[
\gamma^B = \frac{L^B}{L^B + n}
\]

The existence of loyal customers creates jump discontinuities in firms’ best response functions.

\(^{11}\)Recall that I assume two firms have the same total number of reviews for now.
in the pricing subgame. For each firm, the jump happens at the point of the firm’s reservation price. If undercutting the rival requires Firm A to charge a price below its reservation price $\gamma^A$, Firm A would rather charge price 1 to its loyal customers and do not sell to new shoppers. The discontinuity at firms’ reservation prices, caused by the existence of loyal customers, leads to the following lemma (see all proofs in the Appendix).

**Lemma 1.** When both firms advertise and compete in price for new shoppers, if at least one firm has a nonzero loyal customer base, there will be no pure strategy equilibrium in this pricing subgame.

The existence of loyal customers creates asymmetric information among consumers. Like papers in the literature of price dispersion (Varian 1980; Chioveanu 2008), this asymmetry in information among consumers leads to a mixed strategy equilibrium. How firms act in the mixed strategy equilibrium depends on the relationship between $\gamma^A$ and $\gamma^B + d$ (recall that $d = \tilde{\theta}^A - \tilde{\theta}^B$). Intuitively, it depends on which firm can undercut the opponent further in the price competition. In particular, if $\gamma^A$ and $\gamma^B$ satisfy

$$\gamma^A < \gamma^B + d, \quad (2.8)$$

when Firm B charges its reservation price $p^B = \gamma^B$, Firm A can undercut it by charging just below $\gamma^B + d$, which is still above Firm A’s reservation price $\gamma^A$. I say Firm A satisfies the “Advantage in Price Competition” (APC) condition if (2.8) is satisfied. Similarly, Firm B satisfies the APC condition if

$$\gamma^A > \gamma^B + d. \quad (2.9)$$

How the mixed strategy equilibrium depends on this inequality is shown in the next proposition.

**Proposition 1.** Consider the pricing subgame when Firm A and Firm B both advertise. Suppose that at least one of them has a nonzero loyal customer base. In particular, suppose Firm A has a loyal customer base $L^A$ and expected quality $\tilde{\theta}^A$, and analogously Firm B has $L^B$ and $\tilde{\theta}^B$. This pricing subgame has a unique mixed strategy equilibrium.

Let $d = \tilde{\theta}^A - \tilde{\theta}^B$, $\gamma^A = \frac{L^A}{1 + \nu}$, $\gamma^B = \frac{L^B}{1 + \nu}$. 

(1) If Firm A has advantage in competition, i.e., (2.8) holds, in the mixed strategy equilibrium $(F_A(p), F_B(p))$, Firm A gets an expected profit of $(L^A + n)(\gamma^B + d)$, Firm B gets an expected profit of $L^B$, and $F_A(p)$ first order stochastically dominates $F_B(p)$. The distribution functions in the equilibrium are

$$ F_A(p) = \begin{cases} 
0 & p \leq \gamma^B + d \\
1 - \frac{L^B}{n(p-d)} + \frac{L^B}{n} & \gamma^B + d \leq p \leq \tilde{\theta}^A \\
1 & p \geq \tilde{\theta}^A 
\end{cases} $$

and

$$ F_B(p) = \begin{cases} 
0 & p \leq \gamma^B \\
1 - \frac{(L^A + n)(\gamma^A - d)}{m(p+d)} + \frac{L^A}{n} & \gamma^A \leq p \leq \tilde{\theta}^B \\
1 & p \geq \tilde{\theta}^B 
\end{cases} $$

(2) If Firm B has advantage in price competition, i.e., (2.9) holds, in the mixed strategy equilibrium $(F_A(p), F_B(p))$, Firm A gets an expected profit of $L^A$, Firm B gets expected profit $(L^B + n)(\gamma^A - d)$, and $F_B(p)$ first order stochastically dominates $F_A(p)$. The distribution functions in the equilibrium are

$$ F_A(p) = \begin{cases} 
0 & p \leq \gamma^A \\
1 - \frac{(L^A + n)(\gamma^A - d)}{m(p+d)} + \frac{L^A}{n} & \gamma^A \leq p \leq \tilde{\theta}^A \\
1 & p \geq \tilde{\theta}^A 
\end{cases} $$

and

$$ F_B(p) = \begin{cases} 
0 & p \leq \gamma^A - d \\
1 - \frac{L^A}{m(p+d)} + \frac{L^A}{n} & \gamma^A - d \leq p \leq \tilde{\theta}^B \\
1 & p \geq \tilde{\theta}^B 
\end{cases} $$

If a firm has advantage in price competition, it gets an expected profit from the pricing subgame higher than its secured profit. The firm without advantage in price competition only earns the same
Table 2.2: Payoffs When Condition APC-A Holds

<table>
<thead>
<tr>
<th>Firm B</th>
<th>Firm A</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M^B = 1$</td>
<td>$M^A = 1$</td>
</tr>
<tr>
<td>$M^B = 0$</td>
<td>$M^A = 0$</td>
</tr>
</tbody>
</table>

(in expectation) as its secured profit. In general, a firm satisfies the APC condition if it has a relatively higher ratio of good reviews. See the green area in Figure 2.1.

The APC condition only predicts the winner (in terms of expected profit) in the pricing subgame if both firms have advertised. However, it does not give any information about whether the expected winner might want to enter this pricing subgame.

2.4.2 Advertising Game

Without loss of generality, in the following analysis, I suppose Firm A satisfied the APC condition, i.e., condition (2.8) holds.

From Proposition 1 and the fact that (2.8) is true (APC-A), in the pricing subgame after both firms advertise, Firm A gets an expected payoff of $(L^A + n)(\gamma^B + d) - c$, and Firm B gets an expected payoff of $L^B - c$. We now have payoffs in the pricing subgame, and the advertising game with payoffs when APC-A holds is shown in Table 2.2.

Equilibria in cases where at least one firm do not satisfy the PE condition, i.e., (2.3) or (2.4) does not hold, have been analyzed earlier, and here I study the equilibrium for the case where it is profitable for both firms to expand through advertising. That is, (2.3) and (2.4) hold simultaneously.

From Table 2.2, we can see that Firm B (the firm that does not satisfy the APC condition) advertises only if Firm A does not advertise. Firm A also advertises if Firm B does not, but it may still choose to advertise even if Firm B advertises. Therefore, the relationship between Firm A’s expected payoff from the pricing subgame, $(L^A + n)(\gamma^B + d) - c$, and Firm A’s secured profit, $L^A$, determines how firms advertise in the equilibrium.

**Proposition 2.** Consider the competition between two firms, Firm A with a loyal customer base $L^A$ and expected quality $\tilde{\theta}^A$, Firm B with loyal customer base $L^B$ and expected quality $\tilde{\theta}^B$. Suppose
Firm A satisfies the APC condition, i.e., (2.8) holds, and suppose both (2.3) and (2.4) are satisfied, then,

1) If

\[(L^A + n)(\gamma^B + d) - c \geq L^A, \quad (2.10)\]

there is a unique equilibrium in which only Firm A advertises \((M^A = 1, p^A = \tilde{\theta}^A; M^B = 0, p^B = 1)\).

2) If

\[(L^A + n)(\gamma^B + d) - c < L^A, \quad (2.11)\]

there are multiple equilibria: i) only Firm A advertises \((M^A = 1, p^A = \tilde{\theta}^A; M^B = 0, p^B = 1)\); ii) only Firm B advertises \((M^A = 0, p^A = 1; M^B = 1, p^B = \tilde{\theta}^B)\); and iii) both firms randomize advertising with probabilities respectively \(\lambda^A = \frac{L^B - (L^B + n)\tilde{\theta}^B + c}{L^B - (L^B + n)\tilde{\theta}^B}, \quad \lambda^B = \frac{L^A - (L^A + n)\tilde{\theta}^A + c}{L^A - (L^A + n)\tilde{\theta}^A}.\)

To summarize the equilibrium analysis, the advertising equilibrium is characterized by three conditions on state variables \((L^A, L^B, T^A = T, T^B = T)\), or equivalently \((L^A, L^B, \tilde{\theta}^A, \tilde{\theta}^B)\). The three conditions are Profitable Expansion (PE), Advantage in Price Competition (APC), and Uniqueness (U), and are summarized below.

**Profitable Expansion (PE)**

- (PE – Firm A) \((L^A + n)\tilde{\theta}^A - c > L^A\)
- (PE – Firm B) \((L^B + n)\tilde{\theta}^B - c > L^B\)

**Advantage in Price Competition (APC)**

- (APC – Firm A) \(\gamma^A < \gamma^B + d\)
- (APC – Firm B) \(\gamma^A > \gamma^B + d\)

**Uniqueness (U)**

- (U – Firm A) \((L^A + n)(\gamma^B + d) - c \geq L^A\)
As Proposition 2 shows, the condition of equilibrium uniqueness for Firm A \((U - A)\) is relevant only when both firms satisfy the PE condition \((PE - A, \text{ and } PE - B)\), and Firm A satisfies the APC condition \((APC - A)\). Similarly, \(U - B\) is relevant only when \(PE - A, PE - B\) and \(APC - B\) are satisfied at the same time.

The interaction of the three conditions is shown in Figure 2.1. When two firms share the same number of previous buyers, \(T^A = T^B = T\), the state variables can be summarized by \(L^A\) and \(L^B\), which are the axes in Figure 2.1. Note that expected qualities \((\bar{\theta}^A, \bar{\theta}^B)\) are increasing linear functions of the corresponding loyal customer bases \((L^A, L^B)\): The bigger a firm’s loyal customer base is, fixing \(T\), the higher is the firm’s expected quality (or rating).

Figure 2.1 shows that, the condition Uniqueness is satisfied only when the firm with advantage in price competition (APC) has a much bigger loyal customer base than the rival. In other words, if Firm A and Firm B have similar numbers of loyal customers (i.e. \(|L^A - L^B|\) is small), advertising is a strategic substitute between two firms. In this case, there are multiple equilibria, and one firm chooses to advertise only when the rival does not. However, if firms differ a lot in their loyal
customer bases (i.e., $|L^A - L^B|$ is big), advertising is a dominant strategy for the firm with more loyal customers, and the firm with less will never advertise.

We can see from Figure 2.1 that the area of each Uniqueness condition is a strict subset of the area of the corresponding APC condition. In particular, $U - B$ is a strict subset of $APC - B$, and $U - A$ is a strict subset of $APC - A$. This observation leads to the following proposition.

**Proposition 3.** If advertising cost is positive, $c > 0$, there exists a set of state variables $(L^A, L^B, T^A, T^B)$ such that multiple equilibria exist in the competition between two established firms.

The area of multiple equilibria is marked by black dashed lines in Figure 2.1. Proposition 3 is saying that this area is nonempty as long as advertising cost is positive ($c > 0$). Therefore, if firms are rated similarly, multiple equilibria always exist, and one case we might see is that only the worse-rated firm advertises.

If one firm does not satisfy the PE condition, i.e. expansion through advertising is not profitable for this firm, whereas the other firm satisfies the PE condition, there is also a unique equilibrium in which only the firm that satisfied PE advertises. If neither firm satisfies the PE condition, the unique equilibrium is that no firm advertises. We can see from Figure 2.1 that a firm finds it profitable to expand through advertising if its loyal customer base exceeds a certain level (fixing $T$).

Combining all cases in the entire set of state variables ($0 \leq L^A \leq T, 0 \leq L^B \leq T$), I use Figure 2.2 to illustrate the areas of different competition results. In the lower-left white block, expansion is not profitable for either firm, so neither firm advertises. In the lower-right blue area, Firm A has a much higher ratio of loyal customers (or good reviews) than Firm B, there is a unique equilibrium and only Firm A advertises. Correspondingly, in the upper-left red area, there is a unique equilibrium, and only Firm B advertises. In the upper-right green area, firms both have a high ratio of good reviews, and their ratings differ a little, there are multiple equilibria, and either firm might be the one that is advertising. We can see that a firm will be dominant in advertising if it has a much higher ratio of good reviews.
2.5 Extensions

In this section, I give two extensions of the main model. First, I consider an extreme case of the main model in an entry setting. In this entry game, the incumbent is an established firm facing potential entry, the incumbent and the entrant sequentially make advertising decisions, and if entry takes place, they compete in price. Customer reviews interact with advertising in the entry deterrence problem.

In the second part, I consider an extension where a group of consumers are allowed to search for firms and can see the reviews of each firm. Such consumers are called “Searchers”, and they do not need to see advertisements to be informed about the existence of products. I check robustness of the conclusions from the main model to this extension.

2.5.1 An Entry Game

The settings on the product and on consumers remain the same as the main model.

On the firm side there are two firms, an incumbent, Firm A, and a potential entrant, Firm B. Firms’ true qualities, respectively $\theta^A$ and $\theta^B$, are unknown.
Firm A has operated for several periods, had $T \in \mathbb{N}$ previous buyers and $L^A \leq T$ loyal customers. Assume that $T \geq n$. Firm A has a secured profit equal to $L^A$. All previous buyers have written reviews, so Firm A has $T$ reviews in total, and $L^A$ out of them are good reviews (1’s).

Firm B, the entrant, has no previous buyers and thus no loyal customers and no reviews. As a result, Firm B has no secured profit.

The incumbent and the entrant interact in three stages. In the first (pre-entry) stage, the incumbent, Firm A, chooses whether to advertise ($M^A = 1, 0$). In the second (entry) stage, after observing the incumbent’s advertising decision, the entrant, Firm B, chooses whether to enter and advertise ($M^B = 1, 0$). In the third (post-entry) stage, $n$ shoppers are randomly drawn to receive advertisements and see customer reviews (if exist) of the advertised firms, and firms simultaneously choose prices.

Customer reviews of Firm A are observable to both firms and all informed shoppers. It is also common knowledge that Firm B is a new firm and has no customer reviews nor loyal customers. Therefore, firms and informed shoppers share the same belief that Firm A’s expected quality is $\tilde{\theta}^A = \frac{1+L^A}{2+T}$ and Firm B’s expected quality is $\tilde{\theta}^B = \frac{1}{2}$. Let $d$ denote the difference between firms’ expected qualities, $d = \tilde{\theta}^A - \tilde{\theta}^B = \tilde{\theta}^A - \frac{1}{2}$.

Using backward induction, I solve for the subgame perfect Nash equilibrium of the three-stage game. As in the main model, the equilibrium here is characterized by three conditions, Profitable Expansion (PE), Advantage in Price Competition (APC), and Uniqueness (U).

Before analyzing the equilibrium of the entry game, we need first to look at the PE condition. If it is not profitable for the incumbent to expand, entry deterrence will never happen. If the entrant does not satisfy the PE condition, entry will never happen.

The PE condition is defined in the same way as in the main model. Firm A satisfies the PE condition if its loyal customer base $L^A$ and expected quality $\tilde{\theta}^A$ satisfy (2.3). Firm B has $L^B = 0$ and $\tilde{\theta}^B = \frac{1}{2}$, and satisfies PE if $\frac{1}{2}n \geq c$. I assume that advertising is not too costly, specifically $c < \frac{1}{2}n$, so that an entrant is willing to enter the market and advertise if there is no competition. In other words, the entrant, Firm B, always satisfies the PE condition.

If expansion is not profitable for at least one firm, there will be no entry game. In the following
study of entry deterrence, I consider only the case where both firms find it profitable to expand. That is, suppose (2.3) is satisfied.

**Third (post-entry) stage:**

In the third stage, firms carry out pricing decisions. If entry did not happen in the second stage, i.e., Firm B did not advertise, the \( n \) shoppers will be aware only of Firm A. In this case, Firm A selects its monopoly price, \( \tilde{\theta}^A \), and sells to both loyal customers (\( L^A \)) and new shoppers (\( n \)). Firm B does not move in this stage because it did not enter the market.

If entry occurred (i.e., the entrant – Firm B – advertised), and Firm A also advertised, the \( n \) shoppers will be aware of both firms, and Firm A and Firm B need to compete in price for these new shoppers. Lemma 1 implies that as long as Firm A has a nonzero loyal customer base, there is no pure strategy equilibrium of the price competition, and firms’ expected payoffs in the mixed strategy equilibrium is determined by the condition “Advantage in Price Competition” (APC). In the competition between an established firm and a new firm, we have the following lemma about the APC condition.

**Lemma 2.** If the incumbent (Firm A) and the entrant (Firm B) have both advertised, then in the pricing subgame of the third stage, Firm B has advantage in price competition (APC-B), i.e., \( \gamma^B < \gamma^A - d \), for all values of \( L^A \) and the corresponding \( \tilde{\theta}^A \).

Lemma 2 and Proposition 1 together provide us with the firms’ expected payoffs in the pricing subgame after entry occurred. Firm A gets \( L^A \) in expectation, and Firm B gets \( n(\gamma^A - d) \) in expectation.

In an extreme case that the incumbent never satisfied consumers: if Firm A has no loyal customers (\( L^A = 0 \) and \( \tilde{\theta}^A = \frac{1}{2\gamma^A} \)), there is a pure strategy equilibrium of the pricing competition, and firms get the same payoffs equal to the expected payoff in the mixed strategy equilibrium, with certainty. That is, Firm A does not sell and gets zero profit, and Firm B wins all new shoppers and gets \( n(\tilde{\theta}^B - \tilde{\theta}^A) = n(\gamma^A - d) \).
First and second (pre-entry and entry) stage:

Given the (expected) payoffs in the third stage, I analyze how the incumbent and the entrant make advertising decisions sequentially in Stage 1 and Stage 2. The game tree is shown in Figure 2.3.

If Firm A did not advertise in Stage 1, entry is accommodated and Firm B will advertise in Stage 2. If Firm A has advertised in Stage 1, Firm B may still enter and advertise if \( n(\gamma^A - d) - c \geq 0 \). If Firm B enters, Firm A eventually gets a payoff \( (L^A - c) \) less than its secured profit \( L^A \).

Therefore, entering is a dominant strategy for Firm B if \( n(\gamma^A - d) - c \geq 0 \) is satisfied. This is exactly the “Uniqueness” condition in the main model. I call it the “Entry” condition here. If the “Entry” condition is satisfied, Firm A will not advertise in Stage 1, and Firm B will enter and sell to all new shoppers in the equilibrium. If the “Entry” condition is not satisfied, Firm A, having the first-move advantage, will advertise, and Firm B will not enter. In this case, entry is successfully deterred.

To sum up, there are two conditions relevant in this entry game:

**\((PE – Firm A)\):** \( (L^A + n)\tilde{\theta}^A - c \geq L^A \)

**\((Entry)\):** \( n(\gamma^A - d) - c \geq 0 \)

The entry game is defined by two variables, Firm A’s loyal customer base \( L^A \) and expected quality \( \tilde{\theta}^A = \frac{1 + L^A}{2 + T} \). These two variables can be summarized by \( L^A \) alone. There are three parame-
Horizontal axis: advertising cost \( c \)
Vertical axis: Firm A’s loyal customer base \( L^A \)

Figure 2.4: Conditions PE and Entry in the entry game

ters, previous buyers of Firm A (\( T \)), new shoppers (\( n \)), and advertising cost (\( c \)). Firm A has operated alone for more than one periods, therefore, \( T \) is assumed to be greater than \( n \). Figure 2.4 shows how the two conditions interact for different values of \( L^A \) and the parameter \( c \), fixing the values of \( T \) and \( n \).

We can see from Figure 2.4 that, 1) if advertising cost \( c \) is too small (\( c \leq c_1 \)), entry always occurs for all values of \( L^A \); 2) if advertising is too costly (\( c \geq c_2 \)), whenever the incumbent (Firm A) is willing to expand through advertising (i.e., satisfies the PE condition), entry will not occur; and 3) for moderate advertising cost (\( c_1 < c < c_2 \)), entry may still occur even when Firm A satisfies the PE condition (\( L^A \geq L^* \)), and only if Firm A has a big enough loyal customer base (\( L^A \)) relative to its total previous buyers (\( b \)), then entry will be successfully deterred.

From this entry extension, I show that for an entry game with a moderate advertising cost (case 3), entry can be deterred only if the incumbent has a high enough ratio of loyal customers, or in other words, only if the incumbent has a big percentage of good reviews.

Comparing this to what Fudenberg and Tirole [1984] show: when there are no customer reviews, a big incumbent is weak in the competition with entrant and cannot deter entry. If customer
reviews do not exist, even if the incumbent advertises, it will lose the price competition with the entrant. Therefore, in a world where customer reviews do not exist, the incumbent will not advertise and cannot deter entry.

We see that the availability of customer reviews undoes the “fat-cat” effect of big incumbents, and strengthens incumbents with a high ratio of good reviews.

2.5.2 Extension: Searchers

Now I consider an extension of the main model where “Searchers” are allowed. Searchers are a group of consumers who are used to searching for firms instead of watching advertisements, like the tech-savvy consumers. Searchers are not affected by firms’ advertisements, and can always search for all firms and see each firm’s reviews and price.

Assume that there are two types of new consumers, the traditional new consumers and the Searchers. Traditional consumers are the same as the consumers in the main model: they need to see a firm’s advertisements to be informed about the firm’s existence. Searchers are as defined above. Besides searching, there is no other difference between traditional new consumers and Searchers. In other words, if a product has true quality $\theta$, a traditional new consumer and a searcher have the same probability ($\theta$) of being satisfied in consuming this product.

In this model, there are $n \in \mathbb{N}$ traditional shoppers and $s \in \mathbb{N}$ Searchers. Each consumer chooses between Firm $A$ and Firm $B$, and purchase, at most, one unit of the good.

Firm $A$ has $L^A$ loyal customers and an expected quality $\tilde{\theta}^A = \frac{1 + L^A}{2 + T^A}$, and Firm $B$ has $L^B$ loyal customers, and an expected quality $\tilde{\theta}^B = \frac{1 + L^B}{2 + T^B}$.

The timing of the model is:

The $s$ Searchers are aware of both firms, and have access to their customer reviews. Firms $A$ and $B$ first make advertising decisions simultaneously. The $n$ traditional shoppers are informed by firms’ advertisements, and have access to the advertised firm’s customer reviews. After observing each other’s advertising decision, firms choose prices. Traditional shoppers and Searchers then make purchase decisions based on firms’ reviews and prices, and loyal customers of each firm repeat purchases.
The key difference of having Searchers is the option to compete for the group of Searchers even if a firm does not advertise. Therefore, in this extension model, there is a pricing subgame for every combination of firms’ advertising decisions, whereas in the main model (without Searchers), firms compete in price only when both firms advertise. The four combinations are \((M^A = 1, M^B = 1)\), \((M^A = 1, M^B = 0)\), \((M^A = 0, M^B = 1)\) and \((M^A = 0, M^B = 0)\).

Recall that in the pricing subgame, the condition “Advantage in Price Competition” (APC) determines which firm wins the pricing subgame (in terms of expected profit). Here in the extension with Searchers, we will have four APC conditions, one for each pricing subgame.

**Pricing Subgame 1: \(M^A = 1, M^B = 1\)**

This case is very close to the price competition in the main model. When both firms choose to advertise, the traditional consumers are informed about the existence of both firms. Therefore, now the traditional consumers and Searchers have exactly the same information, and can be viewed as one group in this pricing subgame. Firm A and Firm B compete in price for these \(n + s\) new consumers.

Given Firm A’s loyal customer base \(L^A\), the reservation price, \(\gamma^A_{11}\), of Firm A in this subgame \((M^A = 1, M^B = 1)\) is the lowest price that it is willing to charge in the price competition:

\[
\gamma^A_{11} = \frac{L^A}{L^A + n + s} \tag{2.12}
\]

Similarly, Firm B has \(L^B\) loyal customers, and the reservation price \(\gamma^B_{11}\) of Firm B in this pricing subgame is

\[
\gamma^B_{11} = \frac{L^B}{L^B + n + s} \tag{2.13}
\]

The APC condition in this subgame is denoted as \(APC_{11}\). The condition of Firm A having advantage in price competition is \(APC_{11} - A\):

\[
\gamma^A_{11} < \gamma^B_{11} + d \tag{2.14}
\]
where \( d \) is the difference in ratings of two firms, \( \bar{\theta}_A - \bar{\theta}_B \). Firm B has APC if (2.14) does not hold.

If Firm A has advantage in this pricing subgame, i.e. \( APC_{11} - A \) holds, Firm A wins this pricing subgame (in terms of expected profit), and two firms expected profit will be:

\[
\pi^A_{11} = (L^A + n + s)(\gamma^B_{11} + d) - c \\
\pi^B_{11} = L^B - c
\]

**Pricing Subgame 2:** \( M^A = 1, M^B = 0 \)

If only Firm A advertises, those \( n \) traditional consumers will not be aware of Firm B, and they only consider buying from Firm A. However, two firms may compete for Searchers \( (s) \). Reservation prices will be different because of the existence of Searchers.

For Firm A, it now has two reservation options in pricing. It may charge price \( p^A = 1 \), and sell only to its loyal customers \( L^A \). Or, it may charge price \( p^A = \tilde{\theta}_A \), and at least loyal customers \( L^A \) and traditional shoppers \( n \) will buy from Firm A for certain. Therefore, for any price \( p \) that Firm A charges in pricing competition, it has to satisfy

\[
p(L^A + n + s) \geq \max\{L^A, (L^A + n)\tilde{\theta}_A\}
\]

Then Firm A’s reservation price in this pricing subgame is

\[
\gamma^A_{10} = \frac{\max\{L^A, (L^A + n)\tilde{\theta}_A\}}{L^A + n + s}
\]

(2.16)

Firm B’s reservation option in pricing is still only one: charging \( p^B = 1 \) and sell only to its loyal customers \( L^B \). Therefore, Firm B’s reservation price in this pricing subgame is

\[
\gamma^B_{10} = \frac{L^B}{L^B + s}
\]

(2.17)

Note that if \( s \) is small, \( \gamma^B_{10} \) will be close to 1, and higher than \( \tilde{\theta}_B \), which is the highest price that
Searchers would accept for product $B$. Therefore, if only Firm $A$ advertises, there will be competition for Searchers only when

$$\gamma^B_{10} \leq \tilde{\theta}^B$$  \hspace{1cm} (2.18)

The APC condition in this pricing subgame is denoted as $APC_{10}$. And $APC_{10} - A$ is satisfied if

$$\gamma^A_{10} < \gamma^B_{10} + d$$  \hspace{1cm} (2.19)

And again, $APC_{10} - B$ is satisfied if (2.19) does not hold.

If Firm $B$ is willing to compete for Searchers (i.e., (2.18) holds), and Firm $A$ satisfies the APC condition (i.e., (2.19) holds), two firms’ expected profits will be

$$\pi^A_{10} = (L^A + n + s)(\gamma^B_{10} + d) - c$$

$$\pi^B_{10} = L^B$$

**Pricing Subgame 3:** $M^A = 0, M^B = 1$

This subgame is similar to subgame 2, only with firms switching roles, so I will just list the conditions for this subgame.

Firm $A$’s reservation price in this subgame is $\gamma^A_{01} = \frac{L^A}{L^A + n + s}$, and Firm $B$’s reservation price in this subgame is $\gamma^B_{01} = \frac{\max(L^B, (L^B + n + s)\tilde{\theta}^B)}{L^B + n + s}$.

Firm $A$ is willing to compete for Searchers only if

$$\gamma^A_{01} \leq \tilde{\theta}^A$$  \hspace{1cm} (2.20)

The APC condition for this subgame is denoted as $APC_{01}$, and $APC_{01} - A$ is satisfied if

$$\gamma^A_{01} < \gamma^B_{01} + d$$  \hspace{1cm} (2.21)
If both (2.20) and (2.21) are satisfied, two firms’ expected profits will be

\[
\begin{align*}
\pi^A_{01} &= (L^A + s)(\gamma^B_{01} + d) \\
\pi^B_{01} &= \max\{L^B, (L^B + n)\tilde{\theta}^B\} - c
\end{align*}
\]

**Pricing Subgame 4: \( M^A = 0, M^B = 0 \)**

If neither firm advertises, this pricing subgame will again be similar to the pricing subgame in the main model, except that firms are now competing for Searchers, not traditional consumers.

The two firms’ reservation prices are \( \gamma^A_{00} = \frac{L^A}{L^A + s}, \gamma^B_{00} = \frac{L^B}{L^B + s} \). And they will compete for Searchers only if

\[
\begin{align*}
\gamma^A_{00} &\leq \tilde{\theta}^A \quad (2.22) \\
\gamma^B_{00} &\leq \tilde{\theta}^B \quad (2.23)
\end{align*}
\]

are both satisfied.

In this subgame, Firm A satisfies the APC condition, i.e. \( APC_{00} - A \) holds, if

\[
\gamma^A_{00} < \gamma^B_{00} + d \quad (2.24)
\]

If (2.22), (2.23) and (2.24) all hold, the two firms’ expected profits will be

\[
\begin{align*}
\pi^A_{00} &= (L^A + s)(\gamma^B_{00} + d) \\
\pi^B_{00} &= L^B
\end{align*}
\]

Next, we will see how the four \( APC - A \) conditions interact with each other. For simplicity of analysis, here I focus on the case where two firms have the same number of previous buyers: \( T^A = T^B = T > 0 \).

As shown in Figure 2.5, all four \( APC - A \) conditions are satisfied when Firm A has a higher ratio of good reviews than Firm B. That is, if Firm A has a much higher ratio of good reviews
Figure 2.5: Condition APC – A in four pricing subgames

Table 2.3: Firms’ Profits When $s$ is Big, and $L^A > L^B$

<table>
<thead>
<tr>
<th>$M^A$</th>
<th>$M^B = 1$</th>
<th>$M^B = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M^A = 1$</td>
<td>$(L^A + n + s)(\gamma_{11}^B + d) - c$</td>
<td>$(L^A + n + s)(\gamma_{10}^B + d) - c$</td>
</tr>
<tr>
<td>$M^A = 0$</td>
<td>$(L^A + s)(\gamma_{01}^B + d)$</td>
<td>$(L^A + s)(\gamma_{00}^B + d)$</td>
</tr>
<tr>
<td></td>
<td>$\max{L^B, (L^B + n)\tilde{\theta}^B}$</td>
<td>$L^B$</td>
</tr>
</tbody>
</table>

than Firm B, Firm A will have advantage in price competition (APC) in all four pricing subgames. And as the value of $s$ increases, the four conditions converge to be the same area. Therefore, if the group of Searchers ($s$) is very big, a firm either has advantage in price competition (APC) in all four pricing subgames, or has APC in no pricing subgame.

Suppose the group of Searchers ($s$) is very big, and Firm A is the one with a higher ratio of good reviews, and hence satisfies the APC conditions of all four pricing subgames. The profit functions for each combination of advertising strategies by Firms A and B are provided in Table 2.3. We can see from the table that when Firm A has advantage in price competition, if Firm A advertises, Firm B never wants to advertise; but if Firm A does not advertise, Firm B would want to advertise. This is the same as in the main model.

Next, let us see how Firm A’s advertising strategy here differs with the main model where there are no Searchers. Recall that in the main model, the Profitable Expansion (PE) condition is defined
such that a firm is willing to advertise when the opponent does not advertise. Here, the PE condition for Firm A (PE-A) in the extension model with Searchers is

\[(L^A + n + s)(\gamma_{10}^B + d) - c > (L^A + s)(\gamma_{00}^B + d)\]  

(2.25)

which can be simplified to: \[n\left(\frac{L^B}{L^A + s} + d\right) > c\]. If Firm A wants to advertise even when Firm B chooses to advertise, we say the condition Uniqueness (U-A) is satisfied, that is, there will be a unique equilibrium where only Firm A advertises. The condition (U-A) is

\[(L^A + n + s)(\gamma_{11}^B + d) - c > (L^A + s)(\gamma_{01}^B + d)\]  

(2.26)

When (2.25) and (2.26) are simultaneously satisfied, Firm A always advertises, regardless of Firm B’s advertising decision, and the unique equilibrium is that only Firm A advertises. The intersection of these two conditions (PE-A and U-A) is illustrated by the red-contoured area in Figure 2.6.

In this area, Firm A advertises no matter what.

Compare the extension model with Searchers to the main model. The area of Firm A always advertising in the main model is contoured by blue curves. We can see that Firm A (the one with higher ratio of good reviews) has a bigger chance to be dominant in advertising when there are Searchers. However, if the opponent (Firm B) has a very high ratio of good reviews already, it is harder for Firm A to be dominant in advertising in the Searchers model than in the main model.

By this comparison, we see that the spirit of the main model still holds: having a high percentage of good reviews is important for a firm to win the competition, and the areas of unique equilibrium and multiple equilibria all exist when there are Searchers. In particular, the equilibrium partition in the Searchers extension model is given in Figure 2.7.

\[\text{Note that conditions } APC_{11} - A, APC_{10} - A, APC_{01} - A \text{ and } APC_{00} - A \text{ are already satisfied, as we are considering the case when } s \text{ is big and } L^A > L^B.\]
Figure 2.6: With and without Searchers: When does Firm A always advertise?

Figure 2.7: Equilibrium partition in the Searchers model: Areas of multiple equilibria or unique equilibrium
2.6 Data Evidence

In this section I use advertising data and reviews data of local restaurants to test the main prediction of my theoretical model: a firm with a relatively higher average rating, in general, advertises more.

I combine two datasets. The first dataset is obtained from the advertising spending database of Kantar Media; it contains local restaurants’ advertising spending in the year 2014. The advertising spending amount in this dataset is the total amount of ad spending in all channels: TV, magazines, Internet, newspapers, radio, outdoor, etc.

The second dataset is scripted from Yelp.com and contains the corresponding Yelp reviews and location information of those local restaurants in the first (advertising) dataset. I only took those reviews posted before January 1st, 2014. Merging two datasets together, then we have the total advertising spending amount and Yelp reviews of these local restaurants.

In this section, I first use a small dataset to graphically show the distribution of local restaurants’ advertising spending levels in the two dimensional space of their total number of reviews and the ratio of good reviews. And in the second subsection, I use a large dataset with Regression Discontinuity Design to find the relationship between advertising spending and average rating on Yelp.

2.6.1 Graph Illustration

The graphical analysis in this section can be directly linked to the prediction of the model, and in particular corresponds to the red-contoured area (with Searchers) in Figure 2.6.

The small dataset contains only the local restaurants that advertise in the New York City market, one of the DMA (designated marketing area) regions, in Q1 of 2014. To repeat here, only the Yelp reviews posted by Jan 1 of 2014 are collected for these local restaurants. Consumers in this region share the same access for advertisements in all channels.

Filtering out restaurants that are not listed on Yelp, I have 553 local restaurants left in this dataset. Summary statistics are provided in Table 2.4.

To match with the simplifying setting of my model, I take the four-star and five-star reviews
Table 2.4: Summary Statistics for The Graphical Analysis Dataset

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good reviews (4/5 stars)</td>
<td>35.84991</td>
<td>85.78242</td>
</tr>
<tr>
<td>All reviews</td>
<td>57.7613</td>
<td>122.1084</td>
</tr>
<tr>
<td>Advertising spending (000)</td>
<td>5.622604</td>
<td>11.34311</td>
</tr>
</tbody>
</table>

Notes. Observations: 553

Figure 2.8: The Advertising Pattern of Local Restaurants in Neighborhoods with Only Two Restaurants (that are both in Kantar Media Database and on Yelp.com)

to be the good reviews ($L$). Consumers who give four stars or five stars to a local restaurant are highly likely to come back and purchase again.\textsuperscript{13} Therefore, I approximate a local restaurant’s loyal customer base by the group of consumers who rated this restaurant four or five stars.

I use contour plots to visualize the advertising pattern of restaurants with different levels of good reviews ($L$) and all reviews ($T$). I divide the dataset into several subsets and plot the advertising pattern for each subset. Each subset contains restaurants that have the same number of opponents in its neighborhood.\textsuperscript{14} That is, restaurants that are the only restaurant in its neighborhood belong to one subset, restaurants that locate in the neighborhoods with only two local restaurants are in another subset, and so on.\textsuperscript{15}

Using loyal customer base (i.e., $L$ in the model) as the vertical axis, and total number of reviews

\textsuperscript{13}See an evidence for this from the word clouds of Yelp reviews analyzed by Max Woolf at http://minimaxir.com/2014/09/one-star-five-stars/.

\textsuperscript{14}This neighborhood concept is defined by the “city” information of restaurants on Yelp pages.

\textsuperscript{15}Note that, here I am not saying that these neighborhoods really have only one or two local restaurants, but that they only have one or two local restaurants that are both advertising in the New York market and are listed on Yelp.
Figure 2.8 shows the advertising spending levels of the 90 restaurants that are in neighborhoods with only two restaurants, and different colors indicate different levels. From this graph, restaurants with a very high ratio of good reviews (or loyal customers) generally have very high level of advertising spending. In the middle range (around the 45 degree line), advertising spending levels are mixed: some restaurants advertise a lot while some others advertise a little. Then for restaurants with a very low ratio of good reviews, their advertising spending level is really low.

A more extreme subset is the one that contains the local restaurants located in the neighborhood of New York City, and there are in total 107 restaurants in this subset. The contour plot is shown in Figure 2.9. We can again see a similar pattern.

From the contour plots of advertising spending levels for local restaurants, we can see that firms with a high ratio of good reviews are indeed more likely to be dominant in advertising. Next, I will provide supporting evidence from regression analysis as well.

2.6.2 Regression Analysis

To do regression analysis, I use a larger dataset that contains all US local restaurants that advertised in 2014 and are listed on Yelp. To be specific, this dataset is merged from an advertising dataset
Table 2.5: Summary Statistics for The Regression Analysis Dataset

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of one-star reviews</td>
<td>9.9%</td>
<td>11.3%</td>
</tr>
<tr>
<td>Percentage of two-star reviews</td>
<td>11.6%</td>
<td>10.3%</td>
</tr>
<tr>
<td>Percentage of three-star reviews</td>
<td>17.2%</td>
<td>12%</td>
</tr>
<tr>
<td>Percentage of four-star reviews</td>
<td>33.1%</td>
<td>15%</td>
</tr>
<tr>
<td>Percentage of five-star reviews</td>
<td>28.1%</td>
<td>19.1%</td>
</tr>
<tr>
<td>Advertising spending ($ 000)</td>
<td>12.724</td>
<td>44.290</td>
</tr>
</tbody>
</table>

*Notes: Observations: 13,360

with 13,360 local restaurants’ total advertising spending in 2014, and a Yelp information dataset with reviews (by Jan 1, 2014) and other information of those local restaurants. Summary statistics of this dataset is given in Table 2.5.

In the regression analysis, I regress with restaurants’ average rating instead of the ratio of 4- and 5-star reviews (i.e., the ratio of good reviews). This is because restaurants do not make decisions based on the simplification concept “the ratio of good reviews” but rather on their average rating, and average rating and the ratio of good reviews are not one-to-one corresponding to each other, so running regression directly with the ratio of good reviews includes too much unnecessary noise. More importantly, the ratio of good reviews exactly corresponds to the true average rating in the theoretical model setting with only 1 and 0 reviews. Therefore, using data from a five-star review system, we should regress with the true average rating.

To analyze the relationship between local restaurants’ advertising spending and their average rating, I use the following specification:

$$ Ad_i = \alpha + \beta \cdot R_i + \gamma^{DR} \cdot dressy_i + \gamma^{DE} \cdot deliver_i + \gamma^W \cdot waiter_i + \Gamma^P \cdot P_i + \epsilon_i $$ (2.27)

where $R_i$ is the average rating of a restaurant, $dressy_i$ is an indicator variable that equals to 1 if the dressing code is “Dressy”, $deliver_i$ is an indicator variable that equals 1 if delivery is available, $waiter_i$ is an indicator variable of the availability of waiter service, and $P_i$ consists of three price dummy variables indicating the price level ($$$, $$, or $$). If we simply run such a regression with all restaurants, we will get a regression result indicating a negative relationship between restaurants’ average rating and advertising spending amount. See
Figure 2.10: The Relationship between Ad Spending and True Rating without Recognition of RD Drops

column (1) in Table 2.6. The estimated coefficient of variable $R_i$ is significantly negative (-0.807). However, this regression is incorrect since there is a complication caused by the discrepancy between the display rating and the true rating of a restaurant on Yelp. It is out of the scope of the current paper to go into all the details, so I will use some graphs to briefly show the effect on advertising caused by such discrepancy. More detailed discussion can be found in another paper of mine that investigates the effect of a higher display rating on restaurants’ advertising spending using a Regression Discontinuity design.

According to the empirical findings from the empirical RDD paper, there exist significant drops in advertising spending when the true average rating crosses the thresholds of 3.25, 3.75 and 4.25. That is, for relatively higher-rated restaurants, a higher display rating induces drops in advertising spending of local restaurants. This effect of display rating needs to be separated from the effect of true rating in order for us to learn the real relationship between a local restaurant’s (true) average rating and its advertising spending.

Figure 2.10 is a binned scatterplot of advertising spending for restaurants with different true average ratings. Each dot represents the average level of advertising spending of the restaurants within that bin. Each bin contains about 130 restaurants. We can see that, when the average rating goes above 3, advertising spending...
Figure 2.11: The Relationship between Ad Spending and True Rating with Recognition of RD Drops

decreases with the average rating, and this seems consistent with the traditional conclusions in the literature of product reviews. However, once we look in detail at the plot with the recognition of the effects of display rating (i.e., the drops at RD thresholds), we can see from Figure 2.11 that the downward trend is entirely caused by the drops at the RD thresholds (where the display rating jumps by 0.5 star), and in each interval between two thresholds advertising spending in fact goes up with the average rating. This can be seen more clearly from the pooled RD regression results as shown in Table 2.6.

Column (2) in Table 2.6 provides the estimates of the relationship between advertising spending and average rating to the left and to the right of the RD thresholds, where the display ratings are constant. I pool all RD thresholds together and assign a dummy for each interval (that contains a threshold and has length 0.5). The average rating of each observation is normalized by its nearest threshold.

We can see directly from the estimates in column (2) that the relationship between advertising spending and average rating is significantly positive (2.811) to the left of the thresholds. I also test for the significance of the slope to the right of the thresholds, and find that it is significantly positive at 87% confidence level.\footnote{The two-tail p value is 26%. Because the point estimate is positive (2.811 - 0.954 = 1.957), therefore we can reject against the hypothesis that it is positive at 87% confidence level.}
### Table 2.6: RD: Pooled Regression

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average rating</td>
<td>$-0.807^{***}$</td>
<td>$2.811^*$</td>
</tr>
<tr>
<td></td>
<td>(0.150)</td>
<td>(1.551)</td>
</tr>
<tr>
<td>Above threshold</td>
<td>$-1.085^{***}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.351)</td>
<td></td>
</tr>
<tr>
<td>Average rating $\times$ Above threshold</td>
<td>$-0.854^{**}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.349)</td>
<td></td>
</tr>
<tr>
<td>Rating $[1.5, 2)$</td>
<td>$2.965$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.240)</td>
<td></td>
</tr>
<tr>
<td>Rating $[2, 2.5)$</td>
<td>$1.299$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.930)</td>
<td></td>
</tr>
<tr>
<td>Rating $[2.5, 3)$</td>
<td>$1.922$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.890)</td>
<td></td>
</tr>
<tr>
<td>Rating $[3, 3.5)$</td>
<td>$1.446$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.876)</td>
<td></td>
</tr>
<tr>
<td>Rating $[3.5, 4)$</td>
<td>$1.201$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.875)</td>
<td></td>
</tr>
<tr>
<td>Rating $[4, 4.5)$</td>
<td>$0.295$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.877)</td>
<td></td>
</tr>
<tr>
<td>Rating $[4.5, 5]$</td>
<td>$-0.187$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.907)</td>
<td></td>
</tr>
<tr>
<td>Dressy</td>
<td>$0.819^*$</td>
<td>$0.838^*$</td>
</tr>
<tr>
<td></td>
<td>(0.454)</td>
<td>(0.454)</td>
</tr>
<tr>
<td>Waiter service</td>
<td>$-0.920^{***}$</td>
<td>$-0.931^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.301)</td>
<td>(0.301)</td>
</tr>
<tr>
<td>Delivery</td>
<td>$-0.505^{**}$</td>
<td>$-0.511^{**}$</td>
</tr>
<tr>
<td></td>
<td>(0.236)</td>
<td>(0.236)</td>
</tr>
<tr>
<td>Reservation</td>
<td>$0.195$</td>
<td>$0.176$</td>
</tr>
<tr>
<td></td>
<td>(0.213)</td>
<td>(0.213)</td>
</tr>
<tr>
<td>Price level $$</td>
<td>$0.466^*$</td>
<td>$0.384$</td>
</tr>
<tr>
<td></td>
<td>(0.262)</td>
<td>(0.262)</td>
</tr>
<tr>
<td>Price level $$$</td>
<td>$1.437^{***}$</td>
<td>$1.337^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.424)</td>
<td>(0.424)</td>
</tr>
<tr>
<td>Price level $$$$</td>
<td>$0.353$</td>
<td>$0.311$</td>
</tr>
<tr>
<td></td>
<td>(0.898)</td>
<td>(0.898)</td>
</tr>
<tr>
<td>Constant</td>
<td>$9.287^{***}$</td>
<td>$6.018^{***}$</td>
</tr>
<tr>
<td></td>
<td>(6.331)</td>
<td>(1.905)</td>
</tr>
</tbody>
</table>

Observations | 10197 | 10197 |

Notes. The dependent variable in both specifications is the total amount of advertising spending (unit: USD Thousands). Average rating is normalized by the nearest threshold. Regress only with observations that advertise in only one market, and outliers are ruled out. All RD jump thresholds of display rating are pooled together, and interval dummies are used. Significance levels: *** 1%, ** 5%, * 10%.

In summary, controlling for the disturbing effect of Yelp display ratings on advertising, there is in fact a significantly positive relationship between advertising spending and the local restaurants’ average rating. Therefore, we have seen supporting empirical evidence for the model prediction: Higher-rated restaurants advertise more.\(^{18}\)

\(^{18}\)It is out of the scope of the current paper to analyze the effect of Yelp display rating at each threshold, but graphs showing better details about the drops at thresholds above 3 and below 3 are provided in Figure ?? and Figure ?? in Appendix 3.
2.7 Discussion: The Capacity Limit

We have seen from the empirical findings that advertising of local restaurants goes up with average ratings away from RD thresholds (where the display rating is constant) and that advertising goes down with display ratings above 3. The positive relationship between advertising and average ratings is consistent with the theory prediction in this paper that better online reviews increase the benefit of advertising and therefore complement advertising. However, the negative response in advertising to display ratings above 3 seems surprising and counterintuitive because display rating is also an information on reviews, only coarser than the information of average rating.

The essential difference between display rating and the average rating is the way of their interaction with firm profits. Display ratings of restaurants are shown on the search results page if a consumer search for nearby restaurants. See Figure 2.12. An increase (by 0.5 stars) in the display rating of a restaurant will increase the click rates of this restaurant. In other words, the group size of Searchers for a restaurant will increase every time the display rating jumps. On the other hand, when average rating increases between two adjacent jump thresholds of display rating, the display rating is constant, therefore even though the reviews become better (which is the reason of the increase in average rating), the number of Searchers will not change because any change in reviews other than the display rating is unobservable from the search results page. However, for the consumers who have opened the Yelp page of a restaurant and see the reviews, the increase in average rating is observable and will raise consumers’ belief about the quality. Apply the prediction of the theory model in this paper, a higher average rating complement advertising by increasing the benefit of advertising.

In short, a higher display rating increases a firm’s profit by raising the number of consumers, i.e., this is a “volume increase” that leads to a higher profit; a higher average rating increases a firm’s profit by raising consumers’ willingness to pay, i.e., this is a “margin increase” that leads to a higher profit. Both ways work to increase a firm’s profit, but they have different interactions with advertising. The benefit of advertising is the part of the profit that comes from the group “traditional new consumers” who are attracted by advertisements. The margin increase will increase the firm’s
profit margin over every single consumer, including the ones that are attracted by advertisements, i.e., the traditional new consumers. Therefore, the margin increase resulting from a higher average rating will complement the benefit from advertising and make advertising more profitable, but the volume increase resulting from a higher display rating has almost no interaction with advertising and the benefit of advertising basically remains the same.

A key thing to notice is that a volume increase might be bound by the capacity limit of a local business, but a margin increase will never be bound and is always “the more the better”.

Most local businesses have the concern of capacity limits. A local restaurant, a hair salon, or a hotel, cannot accommodate an unlimited number of customers. In the rest of this section, I use comparative statics of the Searchers model (presented in Subsection 5.2) to show the volume increase and the margin increase in profit, and the effect of capacity limits on them.

A capacity limit is an upper bound on the total number of consumers that a restaurant can accommodate. Let $\bar{n}_{k=A,B}$ denote the capacity limits of Firm $A$ and Firm $B$. The sum of loyal customers, traditional new consumers and Searchers of a firm cannot exceed its capacity limit: $L_k + n_k + s_k \leq \bar{n}_k$.
2.7.1 Without Capacity Limit

First, suppose there is no capacity limit, and a firm may accommodate as many customers as there will be. Because Firm A and Firm B are symmetric, here I discuss only Firm A.

If Firm A’s average rating increases between two jump thresholds of display rating, namely when the display rating remains constant while average rating increases, there is no change in $s$: $\Delta s = 0$, but only $\tilde{\theta}_A$ increases. Firm A’s profit always increases with $d = \tilde{\theta}_A - \tilde{\theta}_B$, which in turn increases with $\tilde{\theta}_A$. This increase in Firm A’s profit caused by a higher average rating is the “margin increase”.

If Firm A’s average rating increases across a jump threshold of display rating and causes its display rating on Yelp to jump by half a star, then besides the increase in average rating $\tilde{\theta}_A$, there is also an increase in $s$: $\Delta s > 0$, more Searchers find Firm A. We can see from Table 2.3 that Firm A’s profit always increase with $s$. This increase in profit caused by the jump in display rating is the “volume increase”. At the same time, due to the increase in average rating $\tilde{\theta}_A$, there is also a “margin increase”.

In terms of a firm’s total profit, we can see that an increase in average rating increases a firm’s profit by a larger amount, i.e., margin increase + volume increase, if this increase goes across a jump threshold of display rating and makes the display rating to increase by 0.5 stars. But if an increase in the average rating happens between two jump thresholds of display rating, where the display rating remains constant, the firm’s profit will increase by a smaller amount, i.e., margin increase only.

A firm’s total profit increases more if its display rating increases together with its average rating. To see the interaction between rating and advertising, however, we need to find the change in the benefit and cost of advertising resulting from a change in rating. An increase in total profit might come with a decrease in the benefit of advertising.

The cost of advertising is always $c > 0$.

When Firm A advertises and has advantage in price competition (i.e., Firm A has a higher ratio of loyal customers), see Table 2.3, the benefit of advertising (denoted by $R$) for Firm A is
\( \mathcal{R}_A(M_B = 1) = (L^A + n + s)(\gamma^B_{11} + d) - (L^A + s)(\gamma^B_{01} + d) \) if Firm B advertises, and is \( \mathcal{R}_A(M_B = 0) = (L^A + n + s)(\gamma^B_{10} + d) - (L^A + s)(\gamma^B_{00} + d) \) if Firm B does not advertise.

If Firm A’s average rating, i.e., \( \tilde{\theta}_A \), increases by \( \epsilon > 0 \) but its display rating remains the same, and hold all else constant, the change in \( \mathcal{R}_A \) will come only from \( d = \tilde{\theta}_A - \tilde{\theta}_B \). In particular, \( d \) will increase by \( \epsilon \). The change in the benefit of advertising resulting from the change in Firm A’s average rating is

\[
\Delta \mathcal{R}_A(M_B = 1; d + \epsilon) = \Delta \mathcal{R}_A(M_B = 0; d + \epsilon) = n\epsilon \tag{2.28}
\]

Therefore a higher average rating unambiguously increases a firm’s benefit from advertising by raising consumers’ willingness to pay: the profit margin on each traditional new consumer attracted by advertisement increases by \( \epsilon \). The change in advertising benefit is also a margin increase and matches the change in Firm A’s total profit when only the average rating goes up.

If Firm A’s average rating \( \tilde{\theta}_A \) increases by \( \epsilon > 0 \) and crosses a jump threshold of display rating, i.e., its display rating jumps by 0.5 star, two things will be changing, \( d = \tilde{\theta}_A - \tilde{\theta}_B \) and \( s \). The change in \( d \) is again \( \epsilon \), and the change in \( s \) (\( \Delta s > 0 \)) is the number of extra Searchers that are attracted by the jump in Firm A’s display rating. Let \( s' = s + \Delta s \), and recall that \( \gamma^B_{11} = \frac{L^B}{L^B + n + s^B} \), \( \gamma^B_{01} = \frac{\max(L^B, (L^B + n)\tilde{\theta}_B)}{L^B + n + s^B} \), and \( \gamma^B_{10} = \gamma^B_{00} = \frac{L^B}{L^B + n} \).19 The changes in Firm A’s benefit of advertising resulting from the change in both average rating and display rating are

\[
\Delta \mathcal{R}_A(M_B = 1; d + \epsilon, s' = s + \Delta s) = \mathcal{R}_A(M_B = 1; d + \epsilon, s') - \mathcal{R}_A(M_B = 1; d, s)
= n\epsilon + \Delta s \cdot \frac{L^B - \max(L^B, (L^B + n)\tilde{\theta}_B)}{L^B + n + s^B} \tag{2.29}
\]

and

\[
\Delta \mathcal{R}_A(M_B = 0; d + \epsilon, s' = s + \Delta s) = \mathcal{R}_A(M_B = 0; d + \epsilon, s') - \mathcal{R}_A(M_B = 0; d, s)
= n\epsilon \tag{2.30}
\]

19Note that here I use \( s^B \) instead of \( s \) to denote the group of Searchers for Firm B and \( s^B \) is held constant, because the number of Searchers that can find Firm B will not be affected by a change in Firm A’s rating.
First, note that \( \Delta R_A(M_B = 0; d + \epsilon, s' = s + \Delta s) = \Delta R_A(M_B = 0; d + \epsilon) \), that is the increase in \( s \) does not change Firm A’s benefit of advertising if Firm B does not advertise. If Firm B advertises, the increase in \( s \) brings a new part in \( \Delta R_A: \Delta s \cdot \frac{L^B_{\max}(L^B_{\max}(L^B_{\max} + n)\bar{d}^B)}{L^B_{\max}n + s^B} \), which equals 0 if \( L^B \geq (L^B + n)\bar{d}^B \) and is negative otherwise. If \( L^B < (L^B + n)\bar{d}^B \) is satisfied, intuitively it means that the competitor of Firm A has a very good rating (\( \bar{d}^B \)) or the advertising is very effective in reaching and bringing new consumers (\( n \)). In this case, \( \Delta s \cdot \frac{L^B_{\max}(L^B_{\max}(L^B_{\max} + n)\bar{d}^B)}{L^B_{\max}n + s^B} = \Delta s \cdot \frac{L^B_{\max}(L^B_{\max} + n)\bar{d}^B}{L^B_{\max}n + s^B} < 0 \), i.e., the increase in \( s \) reduces the benefit of advertising.

The effect of \( \Delta s \) enters \( \Delta R_A \), i.e., \( \Delta s \cdot \frac{L^B_{\max}(L^B_{\max}(L^B_{\max} + n)\bar{d}^B)}{L^B_{\max}n + s^B} < 0 \), if either \( \bar{d}^B \) or \( n \) is large. In this model, the effectiveness of advertising (\( n \)) must be large enough for the costly advertising to be ever profitable, and the group of Searchers (\( s \)) is assumed to have a nontrivial size to make a difference. And in the real world, take Yelp as an example, a high enough \( \bar{d}^B \) comes with a large \( s^B \), a restaurant with a high rating will be found by a large number of Searchers. Therefore the magnitude of \( \frac{L^B_{\max}(L^B_{\max}(L^B_{\max} + n)\bar{d}^B)}{L^B_{\max}n + s^B} \) is very small. Rewrite (2.29) as \( n(\epsilon + \frac{\Delta s}{\epsilon}) \cdot \frac{L^B_{\max}(L^B_{\max}(L^B_{\max} + n)\bar{d}^B)}{L^B_{\max}n + s^B} \). Therefore if \( \Delta s \) is less than or only slightly larger than \( n \), \( \Delta R_A \) would still be positive and small. In case of \( \Delta s \) being very large and exceeding \( n \) a lot, \( \Delta R_A \) might become negative, but because both \( n \) and \( s^B \) are nontrivial, the change in the marginal benefit of advertising will be small.

We can see that, without capacity limit, if a jump in display rating comes with the increase in average rating, the resulting increase in \( s \) (Searchers) increases Firm A’s total profits unambiguously, but in most cases it has no effect on Firm A’s benefit of advertising (\( R_A \)). In the case that \( \Delta s \) does change \( R_A \), it reduces the margin increase \( \epsilon \) that comes with the higher average rating \( \bar{d}^A \). But the reduction force from the jump in display rating is only of a small magnitude.

### 2.7.2 With Capacity Limit

Now I consider the case that only Firm A’s capacity limit becomes binding, i.e., \( L^A + n + s = \bar{n}^A \), when Firm A advertises and wins the Searchers.

Firm A’s benefit of advertising when capacity limit is binding is \( R_{A\text{CL}}(M_B = 1; \bar{n}^A) = (L^A + n + s)(\gamma_{11}^B + d) - (L^A + s)(\gamma_{01}^B + d) = \bar{n}^A(\gamma_{11}^B + d) - (L^A + s)(\gamma_{01}^B + d) \) if Firm B advertises, and is \( R_{A\text{CL}}(M_B = 0; \bar{n}^A) = (L^A + n + s)(\gamma_{10}^B + d) - (L^A + s)(\gamma_{00}^B + d) = \bar{n}^A(\gamma_{10}^B + d) - (L^A + s)(\gamma_{00}^B + d) \) if
Firm $B$ does not advertise.

If Firm $A$’s average rating $\tilde{\theta}^A$ increases by $\epsilon$ between two jump thresholds of display rating, i.e., the display rating remains constant, then only $d = \tilde{\theta}^A - \tilde{\theta}^B$ increases (by $\epsilon$) in $R^{CL}_A$. In particular, the changes in Firm $A$’s advertising benefit when its average rating increases under a binding capacity limit are

$$
\Delta R^{CL}_A(M_B = 1; d + \epsilon, \bar{n}^A) = \Delta R^{CL}_A(M_B = 0; d + \epsilon, \bar{n}^A) = n\epsilon
$$

(2.31)

Therefore with capacity limit the change in advertising benefit is still a margin increase ($\epsilon$) that raises Firm $A$’s profit margin from each traditional new consumer ($n$) attracted by Firm $A$’s advertisements. A higher average rating always increases a firm’s benefit and also its willingness to advertising.

If Firm $A$’s average rating $\tilde{\theta}^A$ increases by $\epsilon > 0$ and also crosses a jump threshold of display rating, i.e., its display rating jumps by 0.5 star, again both $d = \tilde{\theta}^A - \tilde{\theta}^B$ and $s$ will be increasing. But now with capacity limit, if $s$ increases by $\Delta s > 0$, the number of traditional new consumers that are attracted by advertisements and can be accommodated will change as well, and $\Delta n = -\Delta s$ to make $L^A + n' + s' = \bar{n}^A$ still hold. So if Firm $A$’s display rating changes with its average rating, the changes in Firm $A$’s benefit of advertising when capacity limit is binding are

$$
\Delta R^{CL}_A(M_B = 1; d + \epsilon, s' = s + \Delta s, \bar{n}^A) = R^{CL}_A(M_B = 1; d + \epsilon, s', \bar{n}^A) - R^{CL}_A(M_B = 1; d, s, \bar{n}^A) = n\epsilon - \Delta s(\gamma_{01}^B + d + \epsilon)
$$

(2.32)

and

$$
\Delta R^{CL}_A(M_B = 0; d + \epsilon, s' = s + \Delta s, \bar{n}^A) = R^{CL}_A(M_B = 0; d + \epsilon, s', \bar{n}^A) - R^{CL}_A(M_B = 0; d, s, \bar{n}^A) = n\epsilon - \Delta s(\gamma_{00}^B + d + \epsilon)
$$

(2.33)

We can see that, when capacity limit becomes binding, the jump in display rating (that leads to $\Delta s > 0$) that comes with the increase in average rating ($\Delta \tilde{\theta}^A = \epsilon$) will always reduce the advertising benefit. In particular, if $\Delta s$ is big enough, the jump in display rating will overturn the margin
increase caused by the higher average rating, and results in a margin decrease in advertising benefit after the average rating has increased. For example, if $\Delta s$ is close to the size of $n$, then $\Delta R_A(M_B = 1) = -\Delta s(\gamma_{01} + d)$ and $\Delta R_A(M_B = 0) = -\Delta s(\gamma_{00} + d)$. If $\Delta s$ is larger than $n$, the margin decrease in advertising benefit will be larger than $\frac{\Delta s}{n}(\gamma + d)$.\(^{20}\)

2.8 Conclusion

In this paper, I use game theory and Bayesian learning to study how competing firms with different customer reviews choose their advertising strategies, with and without Searchers. As an extension, I also study how the availability of customer reviews changes the advertising strategies of the incumbent and entrant firms in an entry game.

The key question that my paper answers is “Do online customer reviews complement or substitute firms’ advertising?” In the literature of online customer reviews, good reviews and advertising are often thought to be substitutes, since a high rating can improve the effectiveness of advertising and can even directly substitute advertising when people can search for ratings. However, findings from the RDD analysis (see Figure 2.11) show that local restaurants’ advertising spending goes up with their average rating on Yelp, but drops with display ratings above 3. This opposite pattern in advertising implies that online reviews are in fact complements to advertising, and display ratings above 3 work as substitutes for advertising.

The RDD analysis enables us to distinguish between the effect of average rating and the effect of display rating on advertising spending. It also provides an explanation of why we have been seeing a negative correlation between rating and advertising spending in regressions all the time: the display rating (above 3) has a strong negative effect and the downward trend comes entirely from this negative effect of display rating. By controlling for the effect of display rating using RDD, I find a significantly positive relationship between advertising and average rating, i.e., if consumers consider only the average rating in their purchase decision, a higher rated firm will be advertising more.

\(^{20}\)Take (2.33) for example. Rewrite it as $n[\epsilon - \frac{\Delta s}{n}\epsilon - \frac{\Delta s}{n}(\gamma_{00} + d)]$. If $\Delta s > n$, then $n[\epsilon - \frac{\Delta s}{n}\epsilon - \frac{\Delta s}{n}(\gamma_{00} + d)] < 0$ and the marginal benefit of advertising decreases by $|\epsilon - \frac{\Delta s}{n}\epsilon - \frac{\Delta s}{n}(\gamma_{00} + d)| = \frac{\Delta s}{n}\epsilon - \epsilon + \frac{\Delta s}{n}(\gamma_{00} + d) > \frac{\Delta s}{n}(\gamma_{00} + d)$.
The reason for display rating (above 3) to have a negative effect on advertising is the capacity limits of local businesses. A higher display rating increases the click rates and, therefore, increases the number of Searchers visiting the restaurant. However, when the display rating is high enough, the capacity limit becomes binding and the increased Searchers will crowd out the benefit from those new consumers that are attracted by advertising, thus the benefit of advertising starts to decrease. Comparative statics of the theory model, presented in Section 7, show that the jump in display rating cannot cause a significant change in advertising benefit if there is no capacity limit.

Applying the findings of this research to other industries with online reviews and capacity limits, if the rounding algorithm is less discrete (for example, Expedia and Priceline use a rounding algorithm to the nearest tenth), we would expect to see smaller drops and more increasing trend as the average rating increases. The effect of jumps in display rating should be smaller if the rounding algorithm rounds to the nearest tenth.

Another interesting prediction comes from the extension to an entry game. An incumbent firm with a long history and a big loyal customer base is not necessarily intimidating. If the incumbent does not have a high enough ratio of good reviews, i.e., the ratio of its loyal customers to all of its previous buyers is not high enough, this big incumbent is weak in the competition with the entrant, and therefore entry is profitable and cannot be deterred. Such incumbent is a “fat cat” as in Fudenberg and Tirole [1984]. But the difference caused by the availability of customer reviews is that a big incumbent with a high ratio of good reviews is very strong and is able to deter entry. Such application in entry deterrence problem might be tested in future research.

In summary, the existence of customer reviews provides an effective information channel for consumers to learn about firms’ qualities, so firms should evaluate the effect of their marketing strategies in this new environment, and adjust their marketing strategies accordingly. We have seen that the availability of customer reviews brings big changes to the traditional predictions on firms’ advertising strategies. A lot of other traditional topics about firms can be revised in the “Age of the Internet”, and we might get many interesting new results.
2.9 Appendix

2.9.1 Equilibrium Analysis for The Asymmetric Previous Buyers Cases

Since Firm A and Firm B are symmetric, here I only study the case $T^A > T^B > 0$. The three conditions characterizing the equilibrium for this case are shown in Figure 2.13. We can see they are essentially the same as the symmetric case ($T^A = T^B = T$) in the main model, a firm needs to have a high enough ratio of good reviews to be the dominant firm in advertising. What is different here is that, when Firm A has more previous buyers than Firm B, it is harder for Firm A to reach the “high enough ratio” of good reviews.

Figure 2.13: Three Conditions for Case $T^A > T^B > 0$

2.9.2 The Yelp Web Page of A Restaurant: Average Rating and Display Rating

See Figure 2.14. When consumers open the web page of a restaurant on Yelp.com, they directly see the display rating: the colored stars (4 stars here) displayed right below the restaurant name. If they click on the button “Details” beside the display star rating, they can see the full distribution of reviews. The display rating rounds the average rating calculated from the distribution to the nearest half star. The distribution of reviews allow consumers to visually approximate and compare between restaurants with the same display star rating.
2.9.3 RD Graphs Showing The Effects of Yelp Display Rating on The Advertising Spending

In Figure 2.15, the top panel shows the drops at the thresholds above 3, including thresholds 3.25, 3.75, 4.25 and 4.75; the bottom panel shows the insignificant drops at the thresholds below 3, including thresholds 1.25, 1.75, 2.25 and 2.75. Empirical estimates show that there are insignificant drops at the thresholds below 3 with precise estimates.
2.9.4 Proofs

Proof of Lemma 1. (No pure strategy equilibrium)

Proof. The best responses of two firms are shown in Figure 2.16. There are no intersection of two firms’ best response functions, hence no pure strategy equilibrium in the Bertrand game.

![Figure 2.16: Best Responses in the Pricing Subgame](image)

Red: Firm A’s best response function; Blue: Firm B’s best response function. This case shown in the graph is when Firm A has more loyal customers: \( L^A > L^B \), i.e. \( d = \bar{\theta}^A - \bar{\theta}^B > 0 \).

Figure 2.16: Best Responses in the Pricing Subgame

Proof of Proposition 1. (The mixed strategy equilibrium of the pricing subgame)

Proof. It has been shown that the pricing subgame has no pure strategy equilibrium. Here I show how to construct the mixed strategy equilibrium for the case where firm A has advantage, and the analysis when firm B has advantage is the same.

In the mixed strategy equilibrium, both firms mix over a range of prices. We know that a firm would never charge a price below its reservation price \( \gamma^k \), so the prices that firm \( k \) mixes must be above \( \gamma^k \). Also, it is not optimal for a firm to charge a price higher than 1, since no consumer would purchase at that price.
Here I give the proof only for the case where firm A has advantage, i.e. $\gamma_A < \gamma_B + d$. The proof for the case of firm B having advantage is the same.

If firm A has advantage in price competition, by charging a price of $\gamma_B + d$, firm A can force firm B out of the competition. Therefore, the highest profit that firm A could secure is $(L_A + n)(\gamma_B + d)$, which is higher than $L_A$ since $\gamma_A < \gamma_B + d$; the highest profit that firm B could secure is $L_B$. In the mixed strategy equilibrium, firms get an expected profit equal to their highest secured profit. Let $F_A(p)$ and $F_B(p)$ denote two firms’ equilibrium mixed strategies, then we have

$$L_A p_A + [1 - F_B(p_A - d)]n p_A = (L_A + n)(\gamma_B + d) \quad \gamma_B + d \leq p_A \leq 1 \quad (2.34)$$

$$L_B p_B + [1 - F_A(p_B + d)]n p_B = L_B \quad \gamma_B \leq p_B \leq 1 \quad (2.35)$$

Note that firm A does not mix below $\gamma_B + d$ (> $\gamma_A$) as firm B will never choose prices below $\gamma_B$, and $p_A = \gamma_B + d$ is already enough to beat the lowest price of firm B.

We can therefore calculate each firm’s mixed strategy distribution function from (2.34) and (2.35):

Before calculate for the distributions of all mixed prices, let us look at the special interval ($\tilde{\theta}_B, 1$). For $p_A > \tilde{\theta}_A$, no new consumers purchase product A (even if $p_A < p_B + d$), so we have $1 - F_B(p_A - d) = 0$ for $p_A > \tilde{\theta}_A$ (or $p_A - d > \tilde{\theta}_B$). Therefore, $F_B(p) = 1$ for $p > \tilde{\theta}_B$, which means $F_B$ has no mass point at 1, nor at any price between $\tilde{\theta}_B$ and 1. Similarly, $1 - F_A(p_B + d) = 0$ for $p_B > \tilde{\theta}_B$ (or $p_B + d > \tilde{\theta}_B + d = \tilde{\theta}_A$). So $F_A(p) = 1$ for $p > \tilde{\theta}_A$, $F_A$ has no mass point at 1 (or any other price above $\tilde{\theta}_A$) either.

Then we get the distribution functions ($F_A$ and $F_B$) in the mixed strategy equilibrium as follows.

$$F_A(p) = \begin{cases} 
0 & p \leq \gamma_B + d \\
1 - \frac{L_B}{n(p-d)} + \frac{L_B}{n} & \gamma_B + d \leq p \leq \tilde{\theta}_A \\
1 & p \geq \tilde{\theta}_A 
\end{cases}$$
and

\[
F_B(p) = \begin{cases} 
0 & p \leq \gamma^B \\
1 - \frac{(L^A + n)(\gamma^B + d)}{\alpha p + d} + \frac{L^A}{n} & \gamma^B \leq p \leq \tilde{\theta}^B \\
1 & p \geq \tilde{\theta}^B 
\end{cases}
\]


\[F_A\] first order stochastically dominates \(F_B\). See Figure 2.17.

---

Figure 2.17: The distributions of firms’ pricing strategies in the pricing subgame

Each firm has a mass point. Denote the probability at mass points by \(m(A)\) and \(m(B)\). For firm A, there is a mass point at \(\tilde{\theta}^A\), and

\[m(A) = 1 - F_A(\tilde{\theta}^A) = \frac{L^B}{\pi} \left( \frac{1}{\tilde{\theta}^B} - 1 \right) > 0.\]

For firm B, the mass point is \(\tilde{\theta}^B\), and

\[m(B) = 1 - F_B(\tilde{\theta}^B) = \frac{(L^A + n)(\gamma^B + d)}{m\bar{\theta}^B} - \frac{L^A}{\pi}.\]

Recall that this equilibrium is under the case where \(\gamma^B + d > \gamma^A\), so we have

\[m(B) > \frac{(L^A + n)\gamma^A}{m\bar{\theta}^A} - \frac{L^A}{\pi} = \frac{L^A}{m\bar{\theta}^A} - \frac{L^A}{\pi} > 0.\]

Next I give an intuitive proof of the uniqueness of this mixed strategy equilibrium \((F_A, F_B)\).

Suppose there exists another mixed strategy equilibrium \((F'_A, F'_B)\), Firm B must get a higher expected profit than in the above equilibrium. This is because that Firm B does not accept any expected profit lower than \(L^B_2\), which equals its secured profit. If Firm B’s expected profit remains the same, the equilibrium \((F'_A, F'_B)\) will be the same as the above one.

If Firm B gets a higher expected profit than \(L^B\), the distribution \(F'_B\) must shift probabilities to higher prices than \(F_B\). Then Firm A can undercut by shifting probabilities to prices just below Firm B’s, and extract all the increased profits of Firm B. Therefore, Firm B cannot get any higher expected profit than \(L^B\) in the mixed strategy equilibrium, and \((F_A, F_B)\) is the unique mixed strategy

\[\footnote{According to the assumption of Beta-distributed beliefs, the mean \(\bar{\theta}^B = \frac{L^B}{m\bar{\theta}^B}\) is always strictly less than 1.}\]
Proof of Proposition 2. (Condition of Uniqueness)

Proof. When both Firm A and Firm B satisfy PE, and Firm A has advantage in price competition against Firm B, the stage game is given in Table 2.2.

When \((L^A + n)(\gamma^B + d) - c \geq L^A\) is satisfied, advertising is a dominant strategy for Firm A. Firm B chooses not to advertise whenever Firm A advertises. Therefore, the unique equilibrium is that only Firm A advertises.

If \((L^A + n)(\gamma^B + d) - c < L^A\), both firms only advertise when the rival does not. Therefore, there are three equilibria, which are, only Firm A advertises, only Firm B advertises, and each firm randomizes advertising with a probability \((\lambda^A, \lambda^B)\).

\(\lambda^A\) and \(\lambda^B\) are such that 
\[\gamma^A - \tilde{\theta}^A < \gamma^B - \tilde{\theta}^B,\]
which is the condition \(APC - A\).
As long as \(c > 0\), there always exist \(L^A\) and \(L^B\) such that \((L^A + n)(\gamma^B + d) > L^A + (L^A + n)(\gamma^B + d) - c\), which means that Firm A satisfies \(APC - A\), i.e. \(\gamma^A < \gamma^B + d\), but Uniqueness condition does not hold.

Therefore, the set \(\{U - A\}\) is a strict subset of \(\{APC - A\}\). □

Proof of Proposition 3: (Non-empty set of multiple equilibria)

Proof. Here I only prove \(U - A \Rightarrow APC - A\), and it is similar for \(U - B \Rightarrow APC - B\).

The condition \(U - A\) is \((L^A + n)(\gamma^B + d) - c > L^A\).

\[\Rightarrow (L^A + n)(\frac{L^B}{L^B + n} + \tilde{\theta}^A - \tilde{\theta}^B) > L^A\]

\[\Rightarrow \gamma^A - \tilde{\theta}^A < \gamma^B - \tilde{\theta}^B,\]
which is the condition \(APC - A\)

As long as \(c > 0\), there always exist \(L^A\) and \(L^B\) such that \((L^A + n)(\gamma^B + d) > L^A > (L^A + n)(\gamma^B + d) - c\), which means that Firm A satisfies \(APC - A\), i.e. \(\gamma^A < \gamma^B + d\), but Uniqueness condition does not hold.

Therefore, the set \(\{U - A\}\) is a strict subset of \(\{APC - A\}\). □

Proof of Lemma 2. (A new firm always has advantage in price competition against an established firm.)

Proof. WLOG, here I prove for the case that Firm A is the established firm, or the incumbent. Consider the pricing subgame after both firms advertise.
Firm $B$ has advantage in price competition if and only if $\gamma^B + d < \gamma^A$. We have $\gamma^B = 0$, $d = \frac{1 + L^A}{2 + T} - \frac{1}{2}$, and $\gamma^A = \frac{L^A}{L^A + n}$.

Compare $LHS = \frac{1 + L^A}{2 + T} - \frac{1}{2}$ and $RHS = \frac{L^A}{L^A + n}$. Note that $T \geq n$.

At end points, $L^A = 0$ and $L^A = T$, we have $LHS < RHS$.

Both LHS and RHS are monotonically increasing: $\frac{\partial LHS}{\partial L^A} > 0$, $\frac{\partial RHS}{\partial L^A} > 0$.

Therefore, for all values of $L^A$, we have $LHS < RHS$. □
Chapter 3

Delay in Platform Adoption: A Dynamic Adoption Model \textit{(with Marc Rysman)}

This paper proposes a new explanation for adoption failure or delay in markets with network effects. In the model, consumers and software providers play a dynamic adoption game. Each group of players choose between two incompatible technologies. Consumers may wait, but firms may not. Although efficiency requires one technology to be adopted by all consumers and firms right away, there is a “market split and adoption delay” equilibrium. In this equilibrium some consumers choose to wait at first and firms split between the two technologies. The model is motivated by the 56K modem market, in which competition between two technologies appears to have led to adoption failure, until an industry standard setting organization coordinated the market on an alternative standard.

3.1 Introduction

In a classic two-sided platform or technology adoption problem in markets with network effects, consumers and firms need to adopt one of two (incompatible) platforms or technologies. In the literature, most papers study the adoption game in a static setting and each player has to make her adoption decision right away given the available information on two platforms. However, no study has been done to allow players to “delay” their adoption.

This paper is based on the model used in Church and Gandal [1992]. We extend Church and Gandal’s static adoption game to a dynamic version which has two periods. Consumers may wait and delay their adoption decision, while firms may not and have to choose a platform (technology) to operate with. The dynamic setting and the freedom to “delay” brings significant difference in the results.

Consider a market with two incompatible technology platforms. To use the product, consumers need to purchase a hardware with one of the two technologies and then purchase softwares from firms. Each firm has to choose one and only one platform to provide its software product.
Consumers enjoy higher utility from a larger variety and a lower price of software products. So consumers in general would like to adopt a platform with a larger number of firms, and this is the network effect of consumers. Firms, however, do not always benefit from adopting a bigger platform. As the number of firms increases on a platform, there will be more consumers adopting this platform, firms will get higher profits, this is the network effect of firms. But at the same time, as the number of firms increases, the competition becomes more intense and the price will be lower, profit then decreases, and this is the competition effect of firms. Therefore, firms do not always adopt the bigger platform.

We show that, in the parameter space that only has standardization equilibrium (i.e. all firms and consumers adopt one platform) in Church and Gandal [1992], if we have two periods instead, there exist a “market split and adoption delay” equilibrium where, in period 1, some consumers choose to adopt A and the same number of consumers adopt B, the rest of the consumers choose to wait, and the period-1 firms split equally between two platforms, then in period 2, there are two possible equilibria, either all waiting consumers and most period-2 firms adopt A (i.e., platform A becomes the dominant platform), or all waiting consumers and most period-2 firms adopt B (i.e., platform B becomes dominant).

When standardization equilibrium happens, it means that the network effect of firms dominate the competition effect. After we add a second period, the “market split and adoption delay” equilibrium exists because, when certain conditions hold, the dynamic competition effect dominates the dynamic network effect.

Consumers are differentiated in their intrinsic preferences for two platforms. The consumers who have strong preferences toward a platform are more likely to adopt right away (in the first period), and the consumers who are more indifferent between two platforms are more likely to wait and adopt in the second period after the dominant platform reveals.
3.2 Literature Review

There is a huge literature on technology adoption. Consumers or firms choose between two available technologies, i.e., one-sided adoption, or both consumers and firms choose between two technologies, i.e., two-sided adoption. An important feature of most papers on technology adoption is network effect. (See the handbook by Farrell and Klemperer [2007] for a summary on competition with network effects.) There are two types of network effects, direct network effect and indirect network effect.

Direct network effect (also called the bandwagon effect) describes the situation when the size of peer users (network size) enters directly into the utility function of each user, i.e., there exists network externality of adoption. Papers with direct network effects include but are not limited to: Katz and Shapiro [1985], Farrell and Saloner [1986], Arthur [1989], Bassanini and Dosi [1999], Ellison and Fudenberg [2003], among which Arthur [1989] and Bassanini and Dosi [1999] study competing technology dynamics with one-sided adoption model.

Indirect network effect does not have the network size in users’ utility function, but the network size affects the factors in the utility function and indirectly improves utility. As Chou and Shy [1990] show, assuming increasing returns to scale and that consumers prefer the variety of products are sufficient for the existence of (indirect) network effect. Chou and Shy [1990] and Church and Gandal [1992] study a one-period two-sided adoption model with indirect network effects, and Jeffrey et al. [2008] study a one-period one-sided adoption model with indirect network effects.

This paper falls into the class of indirect network effect models, and in particular is based on the model used in Church and Gandal [1992]. What we add to the literature are: 1) we study two-sided adoption with indirect network effects in a dynamic setting; and 2) consumers’ choice set extends from \{A, B\} to \{A, B, wait\} in the first period. With these new features, we are able to find a new type of equilibrium, which we call “market split and adoption delay” equilibrium where standardization does not happen because of the dynamics, some consumers choose to wait until period 2 to adopt, and in the end both platforms are viable with nonzero consumers and firms.
3.3 Setting

The consumer preferences are the same as in Church and Gandal [1992], i.e. modeled by the CES utility function.

\[ U(x_1, x_2, ..., x_N) = \left( \sum_i x_i^{1/\beta} \right)^\beta + \phi, \ 1 < \beta \leq 2 \]  \hspace{1cm} (3.1)

where \( x_i \) is the consumed amount of software good \( i \), \( N \) is the number of software products.

There are two platforms, \( A \) and \( B \). To use the softwares, a consumer has to purchase a unit of hardware for one platform. The hardware product itself does not provide benefit in consumer utility, it only grants access to the software products on the same platform. Assume that the two platforms are incompatible: softwares for platform \( A \) cannot be used on a hardware of platform \( B \), and vice versa.

There are a continuum of consumers. Same as Church and Gandal, we use a linear Hoteling model to represent consumers’ preferences for two platforms. Assume that these consumers are distributed uniformly over a unit interval, \( m \in [0, 1] \), according to their tastes for two platforms. Platform \( A \) locates at 0, and \( B \) locates at 1. The utility of consumer of type \( m \) from products on a platform \( (h=A, B) \) is:

\[ U_h(x_1, x_2, ..., x_N, m) = \left( \sum_i x_i^{1/\beta} \right)^\beta + \phi - km, \ 1 < \beta \leq 2 \]  \hspace{1cm} (3.2)

where \( k \) is a measure of the degree of differentiation between platforms.

The representative consumer has a total income \( y \) to spend on the hardware and software products. Let \( p_h \) denote the unit price of hardware product for platform \( h \). After adopting a platform, type \( m \) consumer maximizes her utility \( U_h(x_1, x_2, ..., x_N, m) \) under the budget constraint

\[ \sum_i \rho_i x_i = y - p_h \]  \hspace{1cm} (3.3)

where \( \rho_i \) is the unit price of software product \( i \) (on platform \( h \)).
The indirect utility of consumer locating a distance \( m \) away from platform \( h \) is therefore

\[
V(q_h, p_h, N, m) = (y - p_h)/q_h + \phi - km
\]  

(3.4)

where \( q_h \) is the aggregate price of softwares for platform \( h \). When all software providers charge the same price \( \rho \), we have

\[
q_h = \rho N(1-\beta)
\]

(3.5)

Therefore, the indirect utility from platform \( h \) that locates a distance of \( m \) away is now:

\[
V(\rho, p_h, N, m) = N^{\beta-1}(y - p_h)/\rho + \phi - km
\]

(3.6)

As in Church and Gandal [1992], the hardware technologies are non-proprietary, and therefore are sold at the hardware marginal cost \( c \), i.e. \( p_h = c \). The software products share the same constant marginal cost \( s \) and the same fixed cost \( F \). In the equilibrium of the monopolistic competition between software firms, the unit price of software is \( \rho = \beta s \). So consumers’ indirect utility from platform \( h \) is further simplified to

\[
V(\rho = \beta s, p_h = c, N, m) = \frac{N^{\beta-1}(y - c)}{\beta s} + \phi - km
\]

(3.7)

### 3.4 The Adoption Game

#### 3.4.1 Timing

The adoption game has two periods, a set of new firms enter in each period but there are no new consumers. There are two stages in each period, firm adoption stage and consumer adoption stage. The timing of the two-period game is:

**Period 1**

- *Firm Adoption*: There are \( n \) software firms enter the industry, and each firm needs to choose a platform (\( A \) or \( B \)) to provide product for. (Let \( N_1^A \) denote the number of firms that adopt...
platform $A$, and $N_1^B$ denote firms that adopt platform $B$, then $N_1^A + N_1^B = n$.)

- **Consumer Adoption:** After software firms adopt platform and start providing software products, consumers choose between two platforms, or choose to wait. (Let $m_1^A$ and $m_1^B$ denote the measure of consumers that adopt $A$ and $B$ respectively, and $1 - m_1^A - m_1^B$ is the number of consumers that choose to wait.)

### Period 2

- **Firm Adoption:** Network sizes of two platforms (i.e. $N_1^A, m_1^A$ and $N_1^B, m_1^B$) are observable to all. There are another $n$ new firms enter the industry, and these new firms also choose between platforms $A$ and $B$. (Let $N_2^A$ and $N_2^B$ denote firms that adopt $A$ and $B$ respectively, and $N_2^A + N_2^B = n$.)

- **Consumer Adoption:** Consumers who have adopted a platform in period 1 simply continue purchasing from all software firms on the same platform. All waiting consumers choose between two platforms and purchase both hardware and software products. (Let $m^*$ denote all the consumers that have adopted $A$ by the end of period 2, and $1 - m^*$ consumers have adopted $B$ by the end of period 2.)

We solve for the SPNE (Subgame Perfect Nash Equilibrium) of this two-period game by backward induction. An equilibrium is described by the adoption decisions by firms entering in period 1 ($N_1^A, N_1^B$), by firms entering in period 2 ($N_2^A, N_2^B$), and by consumers in two periods ($m_1^A, m_1^B, m^*$).

In Church and Gandal [1992], when the total number of entrant firms ($N$) exceeds a certain level, i.e.

$$N^{\beta-1} > 2^{\beta-1} k \beta s / [2(\beta - 1)(y - c)]$$ (3.8)

, only standardization equilibria exist.¹ That is, if (3.8) is true, the only equilibria are that, either all firms and all consumers adopt $A$, or all firms and consumers adopt $B$ in the equilibrium. In our

¹See Proposition 2 (case 2) in Church and Gandal (1992).
paper, we focus on the subset of parameter space where

\[(2n)^{\beta-1} > 2^{\beta-1}k\beta s/[2(\beta - 1)(y - c)]\]  \hspace{1cm} (3.9)

, where \(n\) is the number of entrant firms in each period.

We will show the significant effect of dynamic tradeoffs on firms’ and consumers’ adoption strategies by proving the existence of a new type of equilibrium where standardization does not happen.

### 3.4.2 Period 2

Adoption decisions from period 1 are observable in period 2. There have been \(m_A^1\) consumers adopted \(A\), \(m_B^1\) consumers adopted \(B\), and \(1 - m_A^1 - m_B^1\) consumers waiting. Firms are not allowed to wait, and it is now observable that \(N_A^1\) firms adopted \(A\) and \(N_B^1 = n - N_A^1\) firms adopted \(B\) in period 1.

Consumers won’t choose to wait again in period 2 since the game ends in period 2. The marginal consumer \(m^*\) is indifferent between adopting \(A\) and adopting \(B\):

\[V(\rho = \beta s, p_A = c, N_A^1 + N_A^2, m^*) = V(\rho = \beta s, p_B = c, N_B^1 + N_B^2, 1 - m^*)\] \hspace{1cm} (3.10)

Then the marginal consumer \(m^*\) as a function of \(N_A^1 + N_A^2\) can be expressed as

\[m^*(N_A^1, N_A^2) = \frac{(N_A^1 + N_A^2)^{\beta-1}(y - c) - (2n - N_A^1 - N_A^2)^{\beta-1}(y - c) + k\beta s}{2k\beta s}\] \hspace{1cm} (3.11)

In period 2, \(N_A^1\) is already given and fixed, therefore \(m^*\) changes with \(N_A^2\), i.e. how firms in period 2 allocate between two platforms. We can see that \(m^*\) changes monotonically with \(N_A^2\):

\[\frac{dm^*}{dN_A^2} = \frac{(\beta - 1)(N_A^1 + N_A^2)^{\beta-2}(y - c) + (\beta - 1)(2n - N_A^1 - N_A^2)^{\beta-2}(y - c)}{2k\beta s} > 0\] \hspace{1cm} (3.12)

The more firms adopt \(A\) in period 2, the more waiting consumers will adopt \(A\).
However, different from a static adoption model, \( m^* \) will not change freely from 0 to 1 as \( N_2^A \) varies; there are upper and lower bounds to the range of \( m^* \): \( m_1^A \leq m^* \leq 1 - m_1^B \). Therefore, by (3.12), we know that there will be two thresholds of \( N_2^A \), say \( N_2^* \) and \( N_2^{**} \), such that when \( N_2^A \) exceeds \( N_2^{**} \), \( m^* \) will always equal to \( 1 - m_1^B \), and when \( N_2^A \leq N_2^* \), \( m^* \) will equal to \( m_1^A \). The bounds of \( m^* \) are exactly the important changes that lead to our different equilibrium result, and this importance will be illustrated later by graphs.

Knowing the waiting consumers’ adoption strategies, described by \( m^*(N_1^A, N_2^A) \), firms’ profits of adopting each platform are:

\[
\pi_A^2[m^*(N_1^A, N_2^A), p_A = c, N_1^A + N_2^A] = \frac{(\beta - 1)(1 - m^*(N_1^A, N_2^A))(y - c)}{\beta(N_1^A + N_2^A)} - F \tag{3.13}
\]

\[
\pi_B^2[m^*(N_1^A, N_2^A), p_B = c, N_1^B + N_2^B] = \frac{(\beta - 1)(1 - m^*(N_1^A, N_2^A))(y - c)}{\beta(N_1^B + N_2^B)} - F = \frac{(\beta - 1)(1 - m^*(N_1^A, N_2^A))(y - c)}{\beta(2n - N_1^A - N_2^A)} - F \tag{3.14}
\]

Substituting (3.11) in here, \( \pi_A^2 \) and \( \pi_B^2 \) are then both functions of \( N_2^A \).

If there are no boundaries of \( m^* \) as in a static setting, or equivalently \( m_1^A = m_1^B = 0 \), the profit functions are similar to those in Church and Gandal (1992), and will only generate standardization equilibria in the parameter space (3.9) that we chose. This can be shown in a few simple steps. Within this parameter space (3.9), \( \pi_A^2 \) and \( \pi_B^2 \) are monotone functions of \( N_2^A \), and will intersect for once and only once. This intersection point is \( N_2^A = n - N_1^A \). See Figure 3.1, since \( \pi_A^2 \) is increasing in \( N_2^A \) and \( \pi_B^2 \) is decreasing in \( N_2^B \), the intersection point \( N_2^A = n - N_1^A \) is not an equilibrium, and there will be no equilibrium at any point except the two ends of the range of \( N_2^A \), i.e. \( N_2^A = 0 \) and \( N_2^A = n \).

However, if at least one of \( m_1^A \) and \( m_1^B \) is nonzero, firms’ profit function would look a lot different. For values of \( N_2^A \) above \( N_2^{**} \), the profit of adopting platform A is

\[
\pi_A^2[m^* = 1 - m_1^B, p_A = c, N_1^A + N_2^A] = \frac{(\beta - 1)(1 - m_1^B)(y - c)}{\beta(N_1^A + N_2^A)} - F
\]
which is decreasing in $N_2^A$, and for values of $N_2^A$ below $N_2^*$, the profit of adopting $A$ is

$$\pi_A[m^* = m_1^A, p_A = c, N_1^A + N_2^A] = \frac{(\beta - 1)m_1^A(y - c)}{\beta(N_1^A + N_2^A)} - F$$

which is also decreasing in $N_2^A$. Depending on the values of $m_1^A$ and $m_1^B$, there can be three cases.

**Case 1.** $0 < m_1^A < 1/2, 0 < m_1^B < 1/2$. See Figure 3.2. There are two possible equilibria: left intersection – all waiting consumers adopt $B$, $m^* = m_1^A, N_2^A = 2nm_1^A - N_1^A$, and right intersection– all waiting consumers adopt $A$ $m^* = 1 - m_1^B, N_2^A = 2n(1 - m_1^B) - N_1^A$. The right intersection is within the range of $N_2^A$, $[0, n]$, if $N_1^A > n(1 - 2m_1^B)$. And if $N_1^A \leq n(1 - 2m_1^B)$ is true, there is not much difference in the result, and the right-hand equilibrium will simply be $N_2^A = n, m^* = 1 - m_1^B$. 

---

**Figure 3.1:** Firms’ profit of adopting each platform when $m_1^A = m_1^B = 0$

**Figure 3.2:** Case 1. $0 < m_1^A < 1/2, 0 < m_1^B < 1/2$
If less than \( N^* \) period-2 firms adopt platform \( A \) (i.e., \( N_2^A \leq N^* \)), platform \( B \) becomes the dominant platform and all waiting consumers will adopt \( B \). The consumer network size of platform \( A \) will then become constant at \( m_A^1 \) as \( N_2^A \) continues to decrease below \( N^* \). Therefore, as \( N_2^A \) decreases below \( N^* \), the network effect is fixed and the competition effect becomes smaller (less firms are competing on platform \( A \)), and we can see the profit of adopting platform \( A \) increases as \( N_2^A \) decreases. If \( N_2^A \geq N^{**} \), all waiting consumers will adopt platform \( A \) which is the dominant platform. As \( N_2^A \) increases above \( N^{**} \), the consumer network size is fixed at \( m_B^1 \) so the network effect is fixed, and the competition effect increases on platform \( A \), therefore, the profit of adopting \( A \) decreases as \( N_2^A \) increases.

Similarly for the profit from adopting platform \( B \): When \( N_2^A \) decreases below \( N^* \), the network effect is fixed and the competition effect on platform \( B \) increases (more firms adopt \( B \)), so the profit from adopting \( B \) decreases as \( N_2^A \) decreases. When \( N_2^A \) increases above \( N^{**} \), the network effect is again fixed and the competition effect on platform \( B \) decreases (less firms adopt \( B \)), therefore, the profit from adopting \( B \) increases as \( N_2^A \) increases.

**Case 2.** \( m_1^A > 1/2, 0 < m_1^B < 1/2 \). See Figure 3.3. There is only one equilibrium: all waiting consumers adopt \( A \), \( m^* = 1 - m_1^B, N_2^A = 2n(1 - m_1^B) - N_1^A \). Again, the intersection will be smaller than \( n \) if \( N_1^A > n(1 - 2m_1^B) \), and if not, the equilibrium will be \( m^* = 1 - m_1^B, N_2^A = n \).

The shapes of period-2 firms’ profits and the intuitions are in general similar to those in case
1. The difference here is that, when there have already been a large number of consumers adopting platform A in period 1, the period-2 firms that adopt A will enjoy a larger network effect compared to those adopting B. But because the competition effect quickly increases as the network effect becomes fixed, as long as $N_1^A > n(1 - 2m_1^B)$, there will still be some period-2 firms adopting platform B to enjoy the lesser competition on the smaller platform.

If more than half consumers adopt platform A in period 1, then everyone can predict that the dominant platform will be A.

**Case 3.** $0 < m_1^A < 1/2, m_1^B > 1/2$. See Figure 3.4. There is only one equilibrium: all waiting consumers adopt B, $m^* = m_1^A, N_2^A = 2nm_1^A - N_1^A$. If $N_1^A > 2nm_1^A$, there will be no period-2 firms adopting A ($N_2^A = 0$).

This case is symmetric to case 2: If more than half consumers adopt platform B in period 1, then it is certain that the dominant platform will be B.

### 3.4.3 Period 1

Consider case 1 ($0 < m_1^A < 1/2, 0 < m_1^B < 1/2$), which has two equilibria in period 2. Because firms get the same profit at both equilibria and everything is symmetric, each of the two possible equilibria in case 1 happens with probability 1/2. Back at period 1, in any equilibrium with consistent beliefs, consumers and firms will have the same belief that the two equilibria in case 1 each has
probability 1/2. For ease of notation, denote the left-hand equilibrium with «, and the right-hand equilibrium with ».

Consumers make adoption decisions in period 1, choosing between A, B and waiting. Let \( m \) denote the location of a consumer on the unit interval, then \( V_A(m) \) denote her total expected value of adopting A, \( V_B(m) \) her total expected value of adopting B, and \( V_0(m) \) denote her expected value of waiting.

\[
V_A(m) = V(\rho = \beta s, p_A = c, N_1^A, m) + \frac{1}{2} V(\rho = \beta s, p_A = c, N_1^A + N_2^A, m|s)
+ \frac{1}{2} V(\rho = \beta s, p_A = c, N_1^A + N_2^A, m|\bar{s})
\]
\[
= \frac{(y - c)(N_1^A)^{\beta - 1}}{\beta s} + \phi - km + \frac{1}{2} \left[ \frac{(y - c)(2nm_1^A)^{\beta - 1}}{\beta s} + \phi - km \right]
+ \frac{1}{2} \left[ \frac{(y - c)(2n - 2nm_1^B)^{\beta - 1}}{\beta s} + \phi - k(1 - m) \right] \tag{3.15}
\]

\[
V_B(m) = V(\rho = \beta s, p_B = c, n - N_1^A, 1 - m) + \frac{1}{2} V(\rho = \beta s, p_B = c, 2n - N_1^A - N_2^A, 1 - m|s)
+ \frac{1}{2} V(\rho = \beta s, p_B = c, 2n - N_1^A - N_2^A, 1 - m|\bar{s})
\]
\[
= \frac{(y - c)(N_1^B)^{\beta - 1}}{\beta s} + \phi - k(1 - m) + \frac{1}{2} \left[ \frac{(y - c)(2n - 2nm_1^B)^{\beta - 1}}{\beta s} + \phi - k(1 - m) \right]
+ \frac{1}{2} \left[ \frac{(y - c)(2n - 2n\bar{m}_1^B)^{\beta - 1}}{\beta s} + \phi - k(1 - m) \right] \tag{3.16}
\]

\[
V_0(m) = \frac{1}{2} V(\rho = \beta s, p_B = c, 2n - N_1^A - N_2^A, 1 - m|s) + \frac{1}{2} V(\rho = \beta s, p_A = c, N_1^A + N_2^A, m|\bar{s})
\]
\[
= \frac{1}{2} \left[ \frac{(y - c)(2n - 2nm_1^A)^{\beta - 1}}{\beta s} + \phi - k(1 - m) \right] + \frac{1}{2} \left[ \frac{(y - c)(2n - 2nm_1^B)^{\beta - 1}}{\beta s} + \phi - km \right] \tag{3.17}
\]

The marginal consumer between adopting A and waiting is \( m = m_1^A \), and is obtained by solving \( V_A(m_1^A) = V_0(m_1^A) \). Similarly, the marginal consumer between adopting B and waiting is \( m = 1 - m_1^B \), and is obtained by solving \( V_B(1 - m_1^B) = V_0(1 - m_1^B) \). If \( m_1^A \) and \( m_1^B \) are both between 0 and 1/2, then we are consistent with Case 1.
To ease the computational burden, I pick the parameter value $\beta = 2$ to show the existence of a “market split and adoption delay” equilibrium.

The marginal consumers when $\beta = 2$ are:

$$m^A_1(N^A_1) = \frac{-N^A_1(y - c) + n(y - c) - ks - 2s\phi}{2[(y - c)n - 2ks]} \quad (3.18)$$

$$m^B_1(N^A_1) = \frac{N^A_1(y - c) - ks - 2s\phi}{2[(y - c)n - 2ks]} \quad (3.19)$$

When $\beta = 2$, the sum of $m^A_1(N^A_1)$ and $m^B_1(N^A_1)$ is independent of the value of $N^A_1$.

Knowing consumers’ adoption strategies ($m^A_1(N^A_1)$ and $m^B_1(N^B_1)$), the total profit of a period-1 firm from adopting platform $A$ is

$$\pi^A(N^A_1) = \pi^A_1(N^A_1) + \frac{1}{2}\pi^A_2(N^A_1; \prec) + \frac{1}{2}\pi^A_2(N^A_1; \succ)$$

$$= \frac{m^A_1(y - c)}{2N^A_1} - F + \frac{1}{2}\left(\frac{m^A_1(y - c)}{4N^A_1} - F\right) + \frac{1}{2}\left(\frac{(1 - m^A_1)(y - c)}{4(1 - m^A_1)} - F\right)$$

$$= \frac{m^A_1(y - c)}{2N^A_1} + \frac{y - c}{4n} - 2F$$

$$= \frac{ks + 2s\phi + N^A_1(y - c) - n(y - c)}{2[2ks - n(y - c)]} \cdot \frac{y - c}{2N^A_1} + \frac{y - c}{4n} - 2F \quad (3.20)$$

Similarly, the total profit of a period-1 firm from adopting platform $B$ is

$$\pi^B(N^A_1) = \pi^B_1(N^A_1) + \frac{1}{2}\pi^B_2(N^A_1; \prec) + \frac{1}{2}\pi^B_2(N^A_1; \succ)$$

$$= \frac{m^B_1(y - c)}{2(n - N^A_1)} - F + \frac{1}{2}\left(\frac{(1 - m^B_1)(y - c)}{4n(1 - m^A_1)} - F\right) + \frac{1}{2}\left(\frac{m^B_1(y - c)}{4nm^B_1} - F\right)$$

$$= \frac{m^B_1(y - c)}{2(n - N^A_1)} + \frac{y - c}{4n} - 2F$$

$$= \frac{ks + 2s\phi - N^A_1(y - c)}{2[2ks - n(y - c)]} \cdot \frac{y - c}{2(n - N^A_1)} + \frac{y - c}{4n} - 2F \quad (3.21)$$
Figure 3.5: Period 1 - Total Profits ($\beta = 2$)

The first-order derivatives are

$$\frac{d\pi^A(N^A_1)}{dN^A_1} = \frac{(y - c)(ks + 2s\phi - n(y - c))}{4(N^A_1)^2(n(y - c) - 2ks)}$$  (3.22)

and

$$\frac{d\pi^B(N^A_1)}{dN^A_1} = \frac{(y - c)(n(y - c) - ks - 2s\phi)}{4(n - N^A_1)^2(n(y - c) - 2ks)}.  \quad (3.23)$$

So $\pi^A(N^A_1)$ decreases in $N^A_1$ if 1) $0 < k < 2\phi$, and $0 < y - c < \frac{2ks}{n}$ or $y - c > \frac{ks + 2s\phi}{n}$; or 2) $k > 2\phi$, and $0 < y - c < \frac{ks + 2s\phi}{n}$ or $y - c > \frac{2ks}{n}$. And when $\pi^A(N^A_1)$ decreases in $N^A_1$, $\pi^B(N^A_1)$ will be increasing in $N^A_1$. If $\pi^A(N^A_1)$ increases and $\pi^B(N^A_1)$ decreases in $N^A_1$, there will be no market split equilibrium and all firms will adopt platform $A$ or all firms adopt platform $B$. If this case happens, we will no longer have $0 < m^A_1, m^B_1 < 1/2$.\(^2\)

The profit functions are shown in Figure 3.5. The intersection of two profit functions happens exactly at $N^A_1 = \frac{n}{2}$, i.e., half of the period-1 firms adopt platform $A$ and the other half adopt platform $B$.

**Proposition 4.** If the parameters satisfy any one of the three conditions,

\(^2\)In particular, there does not exist any parameter values such that both $\frac{d\pi^A(N^A_1)}{dN^A_1} > 0$ and $0 < m^A_1(N^A_1 = 0), m^B_1(N^A_1 = 0) < 1/2$ hold, or both $\frac{d\pi^A(N^A_1)}{dN^A_1} > 0$ and $0 < m^A_1(N^A_1 = n), m^B_1(N^A_1 = n) < 1/2$ hold.
1) $0 < k \leq 2\phi$, and $y - c > \frac{2ks + 4s\phi}{n}$; or

2) $2\phi < k \leq 6\phi$, and $0 < y - c < \frac{2ks - 4s\phi}{n}$ or $y - c > \frac{2ks + 4s\phi}{n}$; or

3) $k > 6\phi$, and $0 < y - c < \frac{ks - 2s\phi}{n}$ or $y - c > \frac{2ks + 4s\phi}{n}$.

Then there exists a “Market-Split-and-Adoption-Delay” equilibrium where

\[ \frac{1}{4} \left( \frac{s\phi}{n(y - c) - 2ks} \right) \text{ consumers adopt platform A and the same number of consumers adopt platform B (i.e., } m_A^1 = m_B^1 = \frac{1}{4} - \frac{s\phi}{n(y - c) - 2ks} \text{) and the period-1 firms split evenly between two platforms (i.e., } N_A^1 = N_B^1 = \frac{n}{2} \text{). There are } 1 - m_A^1 - m_B^1 = \frac{1}{4} + \frac{2s\phi}{n(y - c) - 2ks} > 0 \text{ consumers choosing to wait in period 1 and adopt in period 2 after the dominant platform reveals.} \]

If one of the three conditions in proposition 4 holds, the profit functions from adopting either platform as shown in Figure 3.5 imply that the competition effect dominates the network effect. For example, as the number of firms that adopt platform A increases, the profit from adopting A decreases.

When there is no dynamics in the adoption problem, as in Church and Gandal [1992], if parameters fall into the range (3.9), firms’ profit from adopting either platform shows that the network effect dominates the competition effect when (3.9) holds.\footnote{When $\beta = 2$, the parameter range (3.9) can be reduced to $y - c > \frac{L_k}{n}$, which overlaps a lot with the three conditions in Proposition 4.} See Figure 3.6 in the appendix. In that case (i.e., the “Case 2” in Church and Gandal [1992]), there are only standardization equilibria, i.e., all firms and consumers adopt one platform.

When we have dynamics in the adoption process and allow consumers to wait, this will change firms’ profits and make the competition effect dominate the network effect, as Figure 3.5 shows. Then we will have this “market split and adoption delay” equilibrium as in Proposition 4.

### 3.5 Conclusion

We extend the static adoption model in Church and Gandal [1992] to a two-period dynamic adoption model. The key elements are that 1) consumers may choose to wait in period 1 instead of adopting right away; and 2) there is a group of new firms enter the market in each period, therefore the uncertainty about future will enter earlier entrants’ adoption decision.
We find that the dynamics may cause significant difference in the results. In a parameter space that only has standardization equilibria in the static setting, we now have, in the dynamic adoption model, a “market split and adoption delay” equilibrium where some consumers choose to wait and firms split evenly between two platforms in period 1. In this equilibrium, at the end of the two periods, we will have both platforms viable with nonzero consumers and firms.

The key reason in the emergence of this new equilibrium is that the network effect is no longer certain in the first period. When the first-period firms make their adoption decisions, they cannot predict which platform will become dominant in period 2, therefore, the uncertainty will dampen the dynamic network effect and the dynamic competition effect will become relatively stronger.

3.6 Appendix

3.6.1 Profit functions in a static setting

When (3.9) holds, the profit functions of adopting either platform will be as shown in Figure 3.6. This is the Figure 3 on page 96 of Church and Gandal [1992]. We can see that, in a static adoption setting as in Church and Gandal [1992], the profit from adopting platform A increases with $N_A$, the network effect dominates the competition effect.

![Figure 3.6: Profit functions from Church and Gandal [1992]](image)
Bibliography


Bassanini, A. and Dosi, G. (1999). Competing technologies, technological monopolies and the role of convergence to a stable market structure. LEM Papers Series 1999/03, Laboratory of Economics and Management (LEM), Sant’Anna School of Advanced Studies, Pisa, Italy.


Curriculum Vitae

EDUCATION

Ph.D., Economics, Boston University, Boston MA, May 2016 (expected)
   Dissertation Title: Essays on Internet Economics: Customer Reviews, Advertising, and Technology Adoption
   Dissertation Committee: Ching-To Albert Ma, Marc Rysman, Juan Ortner and Monic Sun

M.A., Economics, New York University, New York NY, 2010
B.A., Economics, Renmin University of China, Beijing, China, 2008

FIELDS OF INTEREST

Industrial Organization, Quantitative Marketing, Applied Game Theory, Applied Econometrics

FELLOWSHIPS AND AWARDS

Special Research Fellowship, Boston University, Fall 2013, Fall 2014
Summer Research Grant, Boston University, Summer 2014
Teaching Fellowship, Boston University, 2011-2015
Excellent Student Award, Renmin University of China, 2006-2008

WORKING PAPERS

“How Do Firms Advertise When Customer Reviews are Available?” (Job Market Paper), October 2015.

WORK IN PROGRESS

“Delay in Platform Adoption” (joint with Marc Rysman)
“Customer Reviews and Quality Disclosure” (joint with Jacopo Bizzotto)
   (joint with Mengxi Zhang)

TEACHING EXPERIENCE

Instructor, Intermediate Microeconomics, Department of Economics, Boston University, Fall 2015
Teaching Assistant, Empirical Economics, Department of Economics, Boston University, Spring 2015
Teaching Assistant, Statistics for Economists (graduate), Department of Economics, Boston University, Spring 2015
Teaching Assistant, Economics of Information (graduate), Department of Economics, Boston University, Spring 2014, Fall 2011
Teaching Assistant, Behavioral Economics, Department of Economics, Boston University, Spring 2014, Fall 2011
Teaching Assistant, Game Theory (undergraduate and graduate), Department of Economics, Boston University, Spring 2012
Teaching Assistant, Microeconomics, Stern School of Business, New York University, Spring 2010

Work Experience

Research Assistant, Questrom School of Business, Boston University, Professor Monic Sun, Spring 2015, Spring 2016
Research Assistant, School of Law, Boston University, Professor Keith Hylton, Spring 2016
Research Assistant (Data Analysis), Department of Economics, Boston University, Professor Daniele Paserman, Fall 2012-Spring 2013

Referee Experience


Conferences

INFORMS Marketing Science Conference, Shanghai, 2016 (scheduled)
10th Economic Graduate Student Conference, WashU in St. Louis, 2015
8th Workshop on Economics of Advertising and Marketing (Invited), Oxford, UK, 2015

Computer Skills: R, MATLAB, STATA, Mathematica, \LaTeX, Microsoft Office

Citizenship: China