May 2015

Succeeding in Extremely Competitive Markets: Insights from the Mobile Appmarket

Ruijiao Guo
University of Wisconsin-Milwaukee

SUCCEEDING IN EXTREMELY COMPETITIVE MARKETS: INSIGHTS

FROM THE MOBILE APPMARKET

by

Ruijiao Guo

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Management Science

at

The University of Wisconsin – Milwaukee

May 2015

ABSTRACT
THREE ESSAYS ON EXTREMELY COMPETITIVE MARKETS: INSIGHTS FROM MOBILE APPS MARKET

by

Ruijiao Guo

The University of Wisconsin-Milwaukee, 2015
Under the Supervision of Professor Purushottam Papatla

Some products and services today compete against hundreds or thousands of competitors. Faced with so much competition, developers offer their products and services for free, or at a very low price, to those who are interested in the hope of attracting a large group of users. We label such markets where producers give products away for free or charge a nominal price as extremely competitive markets. Businesses competing in extremely competitive markets need insights regarding how they can increase the interest in, and use of, their products by potential customers. Unfortunately, the literature provides few such insights. This is the gap that I address in this dissertation research using the mobile app category with a specific focus on three questions: (1) what factors affect the number of users who download an app (2) why do some apps generate more interest among their users in terms of the word of mouth that they generate than other apps and (3) why do some apps acquire users faster than other apps?

In the first essay, I propose and empirically verify an implicit assumption to explain why businesses in extremely competitive market charge a zero or a very low price. The assumption is that a product with a large group of users will generate profits in the future through one or more mechanisms. For instance, the large user group could attract advertisers interested in targeting the users with promotions for their products. Alternatively, it may create network effects which could, in turn, increase the willingness of users to buy the product. Finally, a free product with a large user group may increase the developer's ability to target the users with improved versions of that product, or other related new products, at a positive price. Findings from our investigation for essay I on the factors that affect the number of users who download an app suggest that the extent of interest of users in other apps offered by the developer has a significant positive effect on the ability of a currently offered app to attract users. Not surprisingly, charging a price rather than giving it away reduces the number of users and so does an increase in the app's physical size, i.e., the memory that it requires on the phone. In terms of the app's rating, interestingly, we find that apps that either have a low-maturity rating – meaning that they are approved for children as well – or have a high-maturity rating – meaning that their use by kids is not advised – do well in terms of the number of users they attract. Our findings also suggest that apps from some genres such as brain teasers, arcade games and sports gain more users than others. Competition-wise, we find that conversations among users who installed competing apps attract more users for the app while an increase in the price of competing apps that were installed reduces the number of users that an app attracts. Overall, therefore, our results suggest that developers with more experience and awareness among users can attract users more easily for new apps than those with no prior experience thus providing empirical support for one of the three mechanisms that the app industry seems to assume. While they are from one category that involves extreme competition, our results may also apply to other categories such as video channels and blogs in similar markets.

From the first essay, we find that users' discussion of apps developed by a certain developer will help in acquiring users for their future products. In the second essay, therefore, we investigate the issue of the factors that affect consumers' word-of- mouth for apps. Our analysis of the word of mouth for apps also provides some surprising insights into why users discuss some apps more than others. Specifically, we find that users are more likely to post comments, reviews, and discuss apps that they paid for rather than those that they obtained for free. This is clearly a finding with significant implications for the pricing and promotion of apps: apps that are given away are less likely to attract users who are advocates that are willing to promote them to potential users. Developers of apps therefore need to take this into account in their pricing decisions. In addition to this immediate implication for the app category, our finding also raises the possibility that, in general, consumers are more likely to discuss products that they purchased than those that they received as promotional items. Other findings and managerial implications are discussed.

In the third essay, I aim to jointly analyze the customer acquisition reached and the time to get there using a joint ordinal-survival analysis model. The focus in this research is on why, in the face of such extreme competition, some apps acquire customers faster than

others. I investigate this question using data on the number of users acquired, and the acquisition growth, for about 2455 Apps from Google Play. I categorize the number of users acquired into ordered tiers and formulate a joint model of growth and customer acquisition using a survival model for the former and an ordinal logit model for the later. The explanatory variables include price, valence of customer rating, and other product attributes. Additionally, effects of competitive contexts and frames are considered. I also consider the role of information cascades on customer acquisition and growth in extremely competitive markets. The model is calibrated within a Bayesian framework using MCMC methods. Findings for the app category as well as generalizable implications for extremely competitive markets are discussed.

Table of Contents

List of Tables

I. Introduction

Some products and services today compete against hundreds or thousands of competitors. Faced with so much competition, developers offer their products and services for free, or at a very low price, to those who are interested, in the hope of attracting a large group of users. We label such markets where producers give products away for free or charge a nominal price as extremely competitive markets. Businesses competing in extremely competitive markets need insights regarding how they can increase the interest in, and use of, their products by potential customers. Unfortunately, the literature provides few such insights. This is the gap that I address in this dissertation research using the mobile app category with a specific focus on three questions: (1) what factors affect the number of users who download an app (2) why do some apps generate more interest among their users in terms of the word of mouth that they generate than other apps and (3) why do some apps acquire users faster than other apps?

In the first essay, I propose and empirically verify an implicit assumption to explain why businesses in extremely competitive markets charge a zero or a very low price. The assumption is that a product with a large group of users will generate profits in the future through one or more mechanisms. For instance, the large user group could attract advertisers interested in targeting the users with promotions for their products. Alternatively, it may create network effects which could, in turn, increase the willingness of users to buy the product. Finally, a free product with a large user group may increase the developer's ability to target the users with improved versions of that product, or other related new products, at a positive price. Findings from our investigation for essay I on the factors that affect the number of users who download an app, suggest that the extent of interest of users in other apps offered by the developer has a significant positive effect on the ability of a currently offered app to attract users. Not surprisingly, charging a price rather than giving it away reduces the number of users and so does an increase in the app's physical size, i.e., the memory that it requires on the phone. In terms of the app's rating, interestingly, we find that apps that either have a low-maturity rating – meaning that they are approved for children as well – or have a high-maturity rating – meaning that their use by kids is not advised – do well in terms of the number of users they attract. Our findings also suggest that apps from some genres such as brain teasers, arcade games and sports gain more users than others. Competition-wise, we find that conversations among users who installed competing apps attract more users for the app while an increase in the price of competing apps that were installed reduces the number of users that an app attracts. Overall, therefore, our results suggest that developers with more experience and awareness among users can attract users more easily for new apps than those with no prior experience thus providing empirical support for one of the three mechanisms that the app industry seems to assume. While they are from one category that involves extreme competition, our results may also apply to other categories such as video channels and blogs in similar markets.

From the first essay, we find that users' discussion of apps developed by a certain developer will help in acquiring users for their future products. In the second essay, therefore, we investigate the factors that affect consumers' word-of- mouth for apps. This analysis as well provides some surprising insights into why users discuss some apps more than others. Specifically, we find that users are more likely to post comments, reviews, and discuss apps that they paid for rather than those that they obtained for free. This is clearly a finding with significant implications for the pricing and promotion of apps: apps that are given away are less likely to attract users who are advocates that are willing to promote them to potential users. Producers of apps therefore need to take this into account in their pricing decisions. In addition to this immediate implication for the app category, our finding also raises the possibility that, in general, consumers are more likely to discuss products that they purchased than those that they received as promotional items. Other findings and managerial implications are discussed.

In the third essay, I aim to jointly analyze the customer acquisition reached and the time to get there using a joint ordinal-survival analysis model. The focus in this research is on why, in the face of such extreme competition, some apps acquire customers faster than others. I investigate this question using data on the number of users acquired, and the acquisition duration, for about 2455 Apps from Google Play. I categorize the number of users acquired into ordered tiers and formulate a joint model of growth and customer acquisition using a survival model for the former and an ordinal logit model for the later. The explanatory variables include price, valence of customer rating, and other product attributes. Additionally, effects of competitive contexts and frames are considered. I also consider the role of information cascades on customer acquisition and growth in extremely competitive markets. The model is calibrated within a Bayesian framework using MCMC methods. Findings for the app category as well as generalizable implications for extremely competitive markets are discussed.

II. Modeling the Installation Base of Mobile Applications in the Context of Extreme Competition

Abstract

In a marketplace characterized by myriad choices and intense competition, such as the mobile app market, getting consumers to discover and purchase products are probably the biggest challenges facing marketers today. The present study labels such markets as "extremely competitive markets" and aims to uncover the implicit assumptions of the business strategies adopted by marketers in such markets. To achieve the goal, I collect a large dataset from the highly popular mobile app store, googleplay.com, and empirically test the assumptions. Given the nature of the response variable in my data, I rely on the Ordinal Logit Model for my analysis but, to capture the effects of unobserved heterogeneity of developers, I use a hierarchical specification and calibrate the model in the Bayesian Paradigm. My findings suggest that the extent of interest of users in other apps offered by the same developer has a significant positive effect on the ability of a currently offered app to attract users. My empirical results also provide additional insights regarding the marketing of apps in particular and about extremely competitive markets in general. I therefore discuss the managerial implications of my findings and also provide directions for future research.

Key Words: Mobile Apps; Extreme Competition; Hierarchical Ordinal Logit Model;

2.1 Introduction

Some products and services today compete against hundreds or thousands of competitors. For instance, there are more than 1 million apps for products running on the Android platform[1], 900,000 apps are available for Apple's operating system iOS[2], and millions of video channels on YouTube compete for viewers[3]. The competition in other similar markets, such as online Blogs and online radio services, is also intense. Faced with so much competition, marketers usually offer their products and services for free, or at a very low price (for instance, many apps are sold for 99 cents), to those who are interested in the hope of attracting a large group of users.

Markets such as those above would typically be labeled as *perfectly competitive markets*. The literature in economics has a long and rich history of research on perfectly competitive markets (Chamberlain 1933; Robinson 1934; Coase 1937; Stigler 1957; Fama 1972; Wilson 1977; Allen and Hellwig 1986). Robinson (1934) defined perfect competition as "a state of affairs in which the demand for the output of an individual seller is perfectly elastic". Such markets are characterized by low or no barriers to entry of competitors. Many other conditions must be fulfilled for a market to be viewed as perfectly competitive. First, identical products by different competitors must be sold simultaneously at the same price across all segments of the market. Second, the number of firms must be large enough such that when any one firm alters its price there is no

---

[1] On July 24th, 2013, Google Play store officially reached over 1 million apps.

[2] On June, 2013, Apple CEO Tim Cook revealed that there is now 900,0000 apps available in the iOS Apple Store, and that 375,000 of those apps are tailored made to support the iPad.

[3] According to YouTube Statistics, 100 hours of video are uploaded to YouTube every minutes: http://www.youtube.com/yt/press/statistics.html

consequent alteration of the prices charged by the others. Third, the number of buyers should be large enough and have similar preferences. Thus, consumers in perfectly competitive markets would be able to get the product from any of the many competitors at the price set by the market.

Given the emergence of online markets, which are viewed as mirroring perfect competition, marketing scholars have also extensively investigated whether online markets are highly competitive (Lal and Sarvary 1999; Brynjolfsson and Smith 2000; Pan, Shankar and Ratchford 2002; Pan, Ratchford and Shankar 2004; Ratchford 2009). Most of this literature can be organized into two streams: analytical work focusing on characterizing and, analytically predicting, the behavior of competitors and empirical research investigating whether the behavior of prices in perfectly competitive markets is consistent with theoretical predictions, i.e., being uniform across all competitors. Results from the analytical studies suggest that the unique characteristics of the Internet will bring about a nearly perfect market because: (1) Consumers are fully informed of prices and product offerings and (2) The physical location of the Internet marketers is irrelevant. Interestingly, some analytical models suggest that the emergence of perfect markets may, in fact, lead to counter-intuitive behavior. For instance, Lal and Sarvary (1999) find that online price sensitivity could be lower in the online market[4] when some of conditions are met.

[4] Conditions include: (1) there is large enough pool of Internet shoppers; (2) Non-digital attributes are important but not overwhelming; (3) Consumers have a more favorable prior about the brand they currently own; (4) The fixed cost of a shopping trip is higher than the cost of visiting an additional store.

The findings of empirical studies on perfect competition are mixed. For instance, Brynjolfsson and Smith (2000) find that online prices are lower than those products sold in conventional channels, providing some evidence of competition in the Internet market. However, a number of other studies suggest that prices even in perfectly competitive markets are dispersed and do not exhibit uniformity across players in the market (Ratchford et al 2002; Pan et al 2004). It appears that greater information flow and easier entry facilitated by the Internet has not made online markets more competitive and "frictionless" as predicted by theory. Their explanation is that the Internet market is immature and lacks a stable equilibrium in market prices.

The present research makes three contributions to the literature. The first is that we go beyond the conventional notion of *perfection competition* and propose the new concept of *extreme competition*. We believe this is an important distinction since, according to conventional wisdom, marketers in a perfectly competitive markets will at least earn normal profits, i.e., the profits which can cover the marginal cost (Robinson 1934). An extremely competitive market, however, is one that meets the aforementioned typical characteristics of a market with perfect competition, such as a large number of competitors, low entry barriers, and nearly uniform prices, but is characterized by competitor behavior that is atypical of markets regardless of whether they are perfectly competitive or not. Specifically, while competitors in perfectly competitive markets exit when their marginal revenues fall below marginal costs, those in extremely competitive markets continue to operate. Recent examples of such markets include online music services (such as Pandora.com and last.fm), online news services (e.g., oldreader.com) and content providers to video services such as Youtube.com (e.g., content provider allthatglitters21 for Youtube.com) and mobile applications (e.g., the app Spotify[5] on Apple's App Store).

Our second contribution is that we examine a competitive strategy followed by most players in extremely competitive markets and provide empirical insights into whether it is meaningful. Specifically, an implicit assumption of businesses in extremely competitive markets seems to be that continuing to provide the product to customers, even when the marginal costs are higher than marginal revenue, will eventually lead to positive profits through one or more processes. For example, a product that acquires more customers will benefit from network effects leading to greater value of the product and, attract even more users. Thus, for instance, all users of Pandora.com may benefit as more users use the service and rate the music they are listening to thus making it easier for all users to identify music that is liked by more users and, hence, is likely to be of better quality. Another example is that a product that acquires more users is likely to help other products offered by the firm gain those customers as well due to their familiarity with the firm[6]. Interestingly, however, there are few empirical investigations of whether these processes do occur. Our research fills this gap.

---

[5] Spotify's lost $57 million in 2011 based on their report, despite a big increase in revenue, to $236 million. Because they pay a majority of their revenue to music companies. The New York Times, "Pandora and Spotify Rake in the Money and then Send it Off in Royalties", August 24th, 2014: http://mediadecoder.blogs.nytimes.com/2012/08/24/pandora-and-spotify-rake-in-the-money-and-then-send-it-off-in-royalties/

[6] In the case of both routes above, the likelihood of firms eventually realizing profits is likely to increase with the increase in the number of users of their products.

Another noticeable gap both in the economics and marketing literatures is that most research on electronic retailing (Mathwick, Malhotra, and Rigdon 2002; Menon and Kahn 2002; Wallace, Giese, and Johnson 2004; Laroche et al 2005; Yadav and Varadarajan 2005) is limited to cases where the markets mimic the competitive characteristics of traditional brick-and-mortar markets while operating in the online channel. The competition in such markets, therefore, is neither perfect nor extreme. There are therefore no normative insights on how competitors can compete and/or grow in extremely competitive markets. Beginning to address this gap is the third contribution of this research.

Our investigation is based on an analysis of data from 2422 apps for the Android platform using a Hierarchical Bayesian ordinal model. The factors that we consider include: (1) characteristics of the app such as its price, how much space it requires on a user's phone, its genre and rating, (2) extent of interest of users in other apps offered by the developer, (3) experience of the developer with apps (4) promotional support for the app if any, and (5) extent and type of competition for the app in terms of the availability of similar apps and their performance in the market. Findings from our investigation suggest that developers with more experience and awareness among users can attract users more easily for new apps than those with no other experience thus providing empirical support for the second process assumed behind the practice of giving products away for free or at a very low price by firms.

In the next section, I present and discuss a number of streams of previous literatures related to our research. Following this, I describe the data. I then present the empirical approach and results and conclude with a section where I discuss the managerial and research implications of my empirical findings.

2.2 Literature Review

The objective of my research is to empirically investigate the factors that affect success in extremely competitive markets with the mobile app category as my empirical context. I therefore present some findings in the literature on competition and, where relevant, on the mobile app category.

Perfect Competition

From the perspectives of classical and neoclassical economists, markets with perfect competition are basically viewed as transaction arenas in which the sellers are not only numerous but interchangeable entities (Remond 2013). In such markets, firms cannot exert any influence on other players, since: (1) the influence of the product of any one firm upon the price is negligible; (2) The output of any one firm is negligible as compared with the total output; (3) Each firm decides the quantity of production without regard to the effect of its decision on the conduct of its competitors (Stigler 1957).

The mobile app market has many of the above characteristics of perfect competition. The market is characterized by myriad choices offering at very low prices (Racherla, Furner, and Babb, 2012). Additionally, developers may release similar apps in the market at almost the same time and thus compete for users intensely. Some applications may even share the same title and have exactly the same functionalities and benefits while other similar applications may include some additional feature to gain a competitive advantage. For instance, the application, *Temple Run*, which was originally developed by IMANGI STUDIOS, has gained much popularity among users. Consequently, similar applications of *Temple Run* developed by other developers, such as Disney, are burgeoning. Developers in this market, therefore, take their strategic decisions without concern about the effects of their actions on how competitors may react. However, a set of similar options offered at very low prices by competitors may take the shares away from the firm, whom then have difficulties to monetize its products.

Thus, although competition in the mobile app category has many of the characteristics of perfect competition, it has its own characteristics which make it even more competitive. First, a myriad of options without brand reputation leads to low customer loyalty although loyalty is key to success for online businesses (Reichheld and Schefter 2000; Urban et al 2000). Second, given the experiential nature of the category, and the similarity of competitive offerings, consumers may have difficulties to assess the differences before, or even after, they actually experience the products (Dev and Lahiri 2012). This means that firms cannot rely on product differentiation as a competitive advantage as in traditional markets. Third, since information search is facilitated by the digital nature of the products, consumers have very low search costs and can easily compare prices and attributes of different offerings (Pitt et al 2001). Thus, advertising ceases to play the informative role that it typically does (Nelson 1970) and, hence, firms cannot use it to compete. Fourth, customers face very low switching costs since competitors are generally only a “click away”(Fabio Ancarani 2002). It is these four characteristics that force marketers to offer their product at a very low price or even give their product for free.

Another important characteristic of the mobile app market is that due to their digital nature, product attributes can be easily communicated online (Lal and Sarvary 1999). This, combined with the fact that there are a myriad of options for every type of app, results in an information overload for consumers leading them to rely on the opinions of other consumers to understand different offerings and select one that is most appropriate for them. One of our objectives, therefore, is to examine whether the presence of competitors and word of mouth for them affects the acquisition of customers by an app.

Effects of Other Apps

Another objective of the present study is to verify an implicit assumption of businesses in extremely competitive markets, i.e., that other apps offered by the same developer would result in network effects and eventually help the business in acquiring more users for new products that they introduce. The network effects may be realized through one or more routes: (1) the users of other products may possibly became the users of the new product directly (2) the business may gain more experience and skills by developing multiple products and thus increase the user base for their future product (3) current users of the business may spread word-of-mouth which may help other products offered by the firm gain customers. In the case of the three routes above, the network effects may eventually help the firms in acquiring more users for their new products.

Some literature in marketing has indeed investigated the influence of consumer networks on the adoption of products (Bass 1969; Robinson and Lakhani 1975, Kalish 1983 However, the markets they explore are typically oligopolistic (Mahajan, Muller, and Bass 1990). In addition, few studies have explored how firms should manage and strategically influence their customer networks to benefit from them (Godes et al. 2005). Recent studies began to opening this path, focusing primarily on firm's strategies and opportunities to tap into online WOM effects. Dellarocas (2006) inspects how strategic manipulation of online forums can shift the information value of online reviews for customers. Chen and Xie (2009) explore how firms can benefit from establishing an online community where consumers can post reviews. Similarly, Forman et al. (2008) empirically demonstrate that encouraging reviewers to disclose their identity can increase consumer trust in the reviews and, in turn, increase product sales. Some of the other studies in this stream are: (1) Godes and Mayzlin (2009) empirically study how firms should strategically recruit customers for WOM campaigns to increase sales (2) Aral and Walker (2011) who highlight the effectiveness of viral product features in generating social contagion (3) Dou, Niculescu, and Wu (2012) who demonstrate analytically that increasing the strength of network effects can impact the adoption software products. Based on previous literature, we, therefore, expect that, in extremely competitive markets, the number of customers acquired by a product has a significant effect not only on its success but on that of other related products by the same manufacture due to network effects. Specifically, we extend the literature by considering how the firm can strategically engineer the strength of network effects and to empirically verify the implicit assumption of the business strategy of sparking adoption by giving its product away. First, if the product is susceptible to network effects, a larger network may boost the value of the product to each user and implicitly increase the willingness-to-pay of potential adopters. Second, it may induce word-of-mouth effect, leading to faster or more efficient propagation of information about the product thus helping consumers in the valuation learning process. Yet, the extant literature is yet to provide an empirical study to test the above two sub-assumptions.

Effects of Product, Price, and Promotion Cues

Besides two key sets of variables, we aim to provide some normative insights with regard to three of the traditional four Ps for the developers in the markets of extreme competition, i.e., product, price, and promotional tools. Product features include "content rating", "app category", and "file size". The "content rating" resembles MPAA rating in movie industry. Based on the literature on motion pictures, the influence of MPAA rating on box office is mixed (Basuroy, Chatterjee, and Ravid 2003; Boatwright, Basuroy, and Kamakura 2007; and Dellaroacs, Zhang, and Awad 2007). Therefore, the effect of content rating in mobile app industry is unclear. The "File size" resembles the runtime of movies. Moon, Bergy and Iabucci(2010) found a significant positive influence of runtime on box office revenues in the opening week. Guo and Papatla(2012) also find positive effect of runtime on number of user reviews. The possible explanation is that the runtime of the movie is a proxy of quality of motion pictures, since it requires more resources. Although a number of studies explore the effects of product attributes of motion pictures, books, TVs, CDs, etc, no one has yet investigate the influence of app attributes on user adoption.

Pricing is another key aspect of the marketing mix. In extremely competitive markets, a prominent fear among businesses if the likely loss of customers to the many still-free competitors (Pauwels and Weiss 2008). Therefore, products in such markets are generally offered for free or at a very low price. Free products may help the businesses boost product adoptions because, beyond the absence of monetary costs, a free product will reduce consumers' search costs and psychological costs (Ariely and Shampan'er 2004). Therefore, offering products for free may stimulate trial among consumers, which is an important implicit assumption underlie the business strategies currently adopted by companies in extremely competitive markets. However, no one has yet empirically tested the assumption.

Prior work suggests that promotional cues can signal quality of the product, which may ultimately influence consumers' purchase decision. For instance, in Erdem and Keane(1996) and Anand and Shachar(2002), advertising content and user experience provide noisy signals about brand attributes. Ackerbery(2003) also suggests that advertising intensity and use experience signal product quality. In present study, we also aim to test the influence of promotional cues of products on user base.

In summary, the user base of a certain app may be influenced by the aforementioned three major factors: competitive products, firms' previous products, and some of the intrinsic attributes of the product. I next describe the data on which I empirically investigate whether those effects do occur.

2.3 Overview of Data

Our primary interest in this research is what factors expand user base. We, therefore, collected data on mobile apps from a major App Store on Oct. 28, 2012. There are two main features of our data: (1) the dataset has a hierarchical structure. The first level is the application-level. The second level is developer-level. The 2,422 apps are developed by 248 developers. There are 11 Application-level factors and 3 Developer-level factors; (2) the website of data source requires that all the reviewers download the applications before they rate the applications. Therefore, the data avoid the situation of fake reviews. The variable definitions and summary statistics are displayed in Table II-1 to Table II-6.

2.3.1 Dependent Variable

The number of downloads we collected is the range of downloads. For instance, the number of downloads of “Slot Machine” is 5,000,000-10,000,000 as of Oct. 28, 2012. There are 12 categories in total.

Table II-1: Distribution of Downloads

Category	Range of Downloads	Number of Apps	Percentage
1	<100	102	4.21%
2	100 - 500	148	6.11%
3	500 - 1,000	78	3.22%
4	1,000 - 5,000	250	10.32%
5	5,000 - 10,000	160	6.61%
6	10,000 - 50,000	409	16.89%
7	50,000 - 100,000	185	7.64%
8	100,000 - 500,000	440	18.17%
9	500,000 - 1,000,000	174	7.18%
10	1,000,000 - 5,000,000	342	14.12%
11	5,000,000 - 10,000,000	76	3.14%
12	>10,000,000	58	2.39%

Note: This table shows the number of downloads of applications. Originally, there are seventeen categories. We grouped 1-5, 5-10, 10-50, 50-100 as category of “<100”.

Table II-1 demonstrates the distributions of downloads. From the table, we can see that there are 440 apps whose downloads ranged from 100,000-500,000. The second largest number of downloads falls into the range of 10,000-50,000. About 86.56% applications have at least 1,000 downloads. The download of all apps in my dataset is roughly normally distributed.

2.3.2 Independent Variables

Product Attributes, Price, and Promotional cues

The present study includes three major sets of app attributes: (1) Product attributes; (2) Price; (3) Promotional cues.

Product Attributes

The product attributes can be represented by three variables: “Content rating”, “Category”, and “file size”.

Table II-2 Summary Statistics of Product Attributes

Category	Variable name	Number of Apps	Percentage
1	Arcade	386	15.94%
2	Brain	280	11.56%
3	Cards	33	1.36%
4	Casual	373	15.40%
5	Racing	228	9.41%
6	Sport	336	13.87%
7	Wallpaper	360	14.86%
8	Widget	126	5.20%
1	Video-yes	772	31.87%
2	Video-no	1650	68.13%
1	Everyone	1235	50.99%
2	High maturity	223	9.21%
3	Low maturity	523	21.59%
4	Medium maturity	441	18.21%

"***File Size***" measures the installation space that the application needs on the mobile device. The range of file size is 1484 MB. It is right-skewed. The largest file size is more than 1G. File size in this market can be a proxy of the quality of each app: the larger the file size, the more content contained in the app. "***Content Rating***" measures the maturity level of the application content. There are four different levels, everyone, low maturity (***lm***), medium maturity (***mm***), and high maturity (***hm***). 1,235 applications are made for everyone (50.99%). "***Category***" indicates the type of the applications. There are eight types of applications: *arcade*, *brain*, *cards*, *casual*, *racing*, *sport*, *wallpaper*, and *widget*. Arcade and wallpaper are the largest two types of applications. "***video***" means whether there's YouTube Video on the embedded on the webpage of the Apps.

Price

The price of applications is different from other continuous variable. 1487(61.40%) applications are free and therefore the price is "0". The rest of them (935 applications) are paid applications and the range of price for paid ones is $7.99. Since the price dispersion of paid apps is very small, it is more meaningful to treat it as dummy variable. We then recode the free apps as "1s" and paid apps as "0s".   With regard to effect of price, we expect that free apps will be more likely to gain larger user base, based on the opinions of conventional economist that lower price will stimulate demand in the marketplace.

Promotional Cues

The third element – promotional cues – is measured by three variables: "if there's ***promotional video***", "how many screenshots", and "if it has high average rating".

Providing promotional video and screenshots may assist users in choosing apps and, in return, get satisfaction for its product from users. Besides, the promotional video may create awareness of products among YouTube viewers. "***Numscrn***" measures the number of Screenshots of each application displayed on the website of the app. These screenshots can act as advertising by illustrating the app's characteristics to users. "***Average rating***" may reflect users' knowledge of the product and service based on their experience with it and his consequent discovery of its unique features and benefits or drawbacks (Feng and Papatla 2011). Therefore, a higher user rating may help the apps acquire more downloads in the future. The average rating is overwhelmingly positive. Most applications receive an average rating of 4 or above.

Table II-3 Summary Statistics of Other Product Attributes

	Min.	1st Qu	Median	Mean	3rd Qu	Max.
rating	0.000	3.900	4.200	4.001	4.500	5.000
price	0.000	0.000	0.000	0.818	0.990	7.990
numscrn	0.000	3.000	5.000	4.617	5.000	8.000
filesize	0.006	1.900	5.700	16.890	14.000	1843.000

Effects of Competition

Competitive apps in the markets influence the download level of a similar app in terms of price and total number of user reviews. To account for the effects of competition, we include four variables to examine the influence of competitive apps in terms of price and total number user reviews. Due to the large number of competitive applications, we do not include the price and number of user reviews of each competitor separately (Neelamegham and Chintagunta 2004; Gopinath, Chintagunta, and Venkataraman 2010). Similar to Gopinath et al.(2010) and Chintagunta et al.(2010), we create four covariates of competition based on the applications featured on the website of the focal application or online recommendation system of the App Store. The first two variables are average ***price*** and *number of user* ***reviews*** of applications featured as "***viewed*** by the same user". Another two variables are average ***price*** and ***review counts*** of applications featured as "***installed*** by the same user". All apps featured on the webpage can be viewed as competitors of the focal app. And the set of four variables represents the competitive influence on the focal apps. Our expectation is that the higher price of the competitive apps, the more likely that the user downloads the focal app. The number of user reviews, however, represent the past installation base since a user can post a review on the store only after she actually downloads the product. The past installation base may create a network effects among users – the more people use it, the more likely that the app attracts more users (Dube, Hitsch, and Chintagunta 2008). Therefore, we expect that the user reviews of the competitive app will have a negative effect on the download of the focal app.

Table II-4 Description and Summary Statistics of four Competition Variables

Description of Variables	
vvc	Average Counts of the Apps Viewed by similar Users
vvp	Average price of the Apps Viewed by similar Users
ivc	Average Counts of the Apps Installed by similar Users
ivp	Average Price of the Apps Installed by similar Users

Summary Statistics

	Min.	**1st Qu**	**Median**	**Mean**	**3rd Qu**	**Max.**
vvc	0.000	3833	14364	35814	44295	754480
vvp	0.000	0.000	0.740	0.955	1.480	8.310
ivc	0.000	54	121	1256	430	131171
ivp	0.000	0.000	0.000	0.832	1.590	8.070

Table II-4 reports the description of each competition measure. From the summary statistics, we can see the average review counts of apps viewed by the same user are much higher than average review counts installed by the same user. Therefore, the apps featured as "viewed by the same user" are highly popular in the market. However, the average price of the apps viewed by the same user is a little bit higher than the average price installed by the same user, which indicates that users may be interested in popular apps but still want to install those apps with lower price. We, therefore, expect price to play a significant role in generating downloads.

Effects of Other Apps

To examine the above influence of network effects created by other apps developed by the same developer, we create three variables: (1) Total number of apps developed; (2) The average number of user reviews; (3) The average price.

Table II-5 Summary Statistics of Developer-Level Attributes

Variable	Description
dac	the total number of the applications display on Google Play
dvp	the average review counts of applications developed by the same developer
dvp	the average price of applications developed by the same developer

Correlation Matrix				Summary Statistics					
	dac	dvc	dvp	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
dac	1.000	-0.425	-0.046	5.00	12.00	28.00	31.970	49.00	72.00
dvc		1.000	-0.092	1	127	988	11460	8908	294800
dvp			1.000	0.00	0.00	0.37	0.60	0.93	5.37

Table II-5 displays the descriptions and summary statistics of each variable and also presents the correlations between the three variables. "***dac***" measures the total number of the applications developed by the developer before. Therefore, this variable measures, to some extent, the size of each developer. The variable has a range of 67. The distribution of developer size is right-skewed. Most of developers have 28 or less applications. This reflects that the market is dominated by small developers. No one, then, has the power to exert influence on others with regard to price. "***dvc***" measure the average review counts of applications developed by the same developer previous. The range of the average review counts is 294,800. And again, the distribution of it is roughly normally distributed but a little bit right-skewed. Developers with higher average review counts have higher downloads. "***dvp***" measures the average price of applications developed by the same developer. The range of it is 5.37 and it is also roughly normally distributed.

The correlations between the three developer-level variables are relatively low. However, the correlations among the three variables are all negative. Though the correlations are weak, it indicates that more experienced developers are less likely to gain higher average counts and have higher price. In addition, price and review count are negatively correlated.

Correlations Between App-level Attributes.

In Table II-6, correlations among variables are displayed. All correlations lie between -0.5 to 0.5. Therefore, multicollinearity problem would not be a problem in our analysis.

Table II-6 Correlation Matrix of Application-Level Attributes

	dnlds	count	rating	price	nums	video	filesiz	vvc	vvp	ivc	ivp
1	1.000										
2	0.352	1.000									
3	0.362	0.132	1.000								
4	-0.427	-0.127	-0.091	1.000							
5	0.088	0.058	0.155	0.167	1.000						
6	0.141	0.149	0.117	0.130	0.209	1.000					
7	0.043	0.038	0.010	0.193	0.074	0.133	1.000				

8	0.171	0.112	0.085	-0.134	0.031	0.043	0.008	1.000			
9	0.081	0.102	0.022	0.321	0.130	0.192	0.146	-0.147	1.000		
10	0.214	0.280	0.091	-0.089	0.060	0.125	0.069	0.062	0.052	1.000	
11	-0.486	-0.112	-0.150	0.709	0.125	0.107	0.106	-0.144	0.350	-0.094	1.000

Note:

1. “dnlds” represents total number of downloads(dependent variable)
2. “nums” represents number of screen shots
3. We only include application-level covariates in this correlation matrix.

From the correlation table, we found that price and downloads are negatively correlated. Therefore, free apps, have larger user bases compared to paid apps. In addition, the average price of the apps installed by the same users is negatively correlated with downloads. Therefore, it seems users have a budget on expenditures of apps. Price is also negatively correlated with average rating, which indicates that users seem to give lower ratings for free apps.

2.4 Model Specification

Considering the characteristics of our dependent variable, we use the Ordinal Logit Model to investigate the behavior of product adoption in mobile app market. There are several reasons that Ordinal Logit Model is adopted. First, downloads is categorical and ordered; second, the interval is not equal. The adjacent two categories are 10 times difference; third, downloads occur in category j before it can occur in category j+1. However, 12 categories are too many for the model to converge and it’s computationally demanding. We, therefore, reduce the 12 categories to 6 categories by combining the adjacent two categories and code each category as “1”, “2”, “3”, “4”, “5”, and “6”, correspondingly, and implement Ordinal Logit model in Bayesian Paradigm.

Table II-7 Recoded Dependent Variable - Downloads

Ordinal Responses	Range of Downloads	# of Apps	Percentage
“1”	<100; 100 - 500	250	10.32%
“2”	500 - 1,000; 1,000 - 5,000	328	13.54%
“3”	5,000 - 10,000; 10,000 - 50,000	569	23.50%
“4”	50,000 - 100,000; 100,000 - 500,000	625	25.91%
“5”	500,000 - 1,000,000; 1,000,000 - 5,000,000	516	21.30%
“6”	5,000,000 - 10,000,000; >10,000,000	134	5.51%

Basic setting of Ordinal Logit Model:

Assume that the utility of application i is represented by an unobservable latent variable U_i. The downloads of application i will jump to a certain level on the basis of U_i.

$$U_i = \mu_i + \varepsilon$$

The threshold parameters obey the ordering constraint: $\theta_1 < \theta_2 < \theta_3 < \theta_4 < \theta_5$

$$y_i = 1 \xrightarrow{yield} U_i < \theta_1$$
$$y_i = 2 \xrightarrow{yield} \theta_1 < U_i < \theta_2$$
$$y_i = 3 \xrightarrow{yield} \theta_2 < U_i < \theta_3$$
$$y_i = 4 \xrightarrow{yield} \theta_3 < U_i < \theta_4$$
$$y_i = 5 \xrightarrow{yield} \theta_4 < U_i < \theta_5$$
$$y_i = 6 \xrightarrow{yield} U_i > \theta_5$$

Assume that ε follows a logistic distribution, which means the cumulative distribution of ε is $F(\varepsilon) = \exp(\varepsilon)/(1 + \exp(\varepsilon))$. Therefore,

$$Prob("1") = Prob(U_i < \theta_1) = Prob(\mu_i + \varepsilon < \theta_1) = Prob(\varepsilon < \theta_1 - \mu_i)$$
$$Prob("2") = Prob(\theta_1 < U_i < \theta_2) = Prob(\theta_1 < \mu_i + \varepsilon < \theta_2) = Prob(\theta_1 - \mu_i < \varepsilon < \theta_2 - \mu_i)$$
$$Prob("3") = Prob(\theta_2 < U_i < \theta_3) = Prob(\theta_2 < \mu_i + \varepsilon < \theta_3) = Prob(\theta_2 - \mu_i < \varepsilon < \theta_3 - \mu_i)$$
$$Prob("4") = Prob(\theta_3 < U_i < \theta_4) = Prob(\theta_3 < \mu_i + \varepsilon < \theta_4) = Prob(\theta_3 - \mu_i < \varepsilon < \theta_4 - \mu_i)$$
$$Prob("5") = Prob(\theta_4 < U_i < \theta_5) = Prob(\theta_4 < \mu_i + \varepsilon < \theta_5) = Prob(\theta_4 - \mu_i < \varepsilon < \theta_5 - \mu_i)$$
$$Prob("6") = Prob(U_i > \theta_5) = Prob(\mu_i + \varepsilon > \theta_5) = Prob(\varepsilon > \theta_5 - \mu_i)$$

The Utility of application i can be related to two groups of observable variables: the application-level variables and the developer-level variables. For instance, the higher the review counts of an application the more will be the number of downloads. On the other hand, the higher the price of an application the fewer will be the number of users who download it. Therefore, we have our Baseline Model below:

Model 1 (Baseline Model): Ordinal Logit Model

Both application-level variables and developer-level variables were simply included in the basic model. By implementing this model, we'll ignore the hierarchical structure inherent in our data and just fit an ordinary ordinal logit model. The results of this model are displayed in table II-10.

$$\begin{aligned}\mu_{ij} = \beta_1 \log(price_{ij}) + \beta_2 video_{ij} + \beta_3 \log(numscrn_{ij}) + \beta_4 \log(filesize_{ij}) + \beta_5 \mathrm{lm}_{ij} \\ + \beta_6 hm_{ij} + \beta_7 mm_{ij} + \beta_8 arcade_{ij} + \beta_9 brain_{ij} + \beta_{10} cards_{ij} + \beta_{11} casual_{ij} \\ + \beta_{12} racing_{ij} + \beta_{13} sport_{ij} + \beta_{14} wallpaper_{ij} + \beta_{15} \log(rating_{ij}) \\ +\beta_{16} \log(vvc_{ij}) + \beta_{17} \log(vvp_{ij}) + \beta_{18} \log(ivc_{ij}) + \beta_{19} \log(ivp_{ij}) \\ +\alpha_1 \log(dac_j) + \alpha_2 \log(dvc_j) + \alpha_3 \log(dvp_j)\end{aligned}$$

However, some unobserved factors will affect downloads of the application. For instance, the duration since the application was released or the requirement of an android system, could affect downloads? In the previous analysis, we pooled all applications together and made the assumption that one application is independent from another. However, our data structure suggests that applications developed by the same developer may have some common characteristics. To accounts for the multilevel data structure, we introduce a developer-specific parameter η_j to the utility function. Our second model, therefore, is:

Model 2: Ordinal Logit Model with Random Intercepts

In the random intercept model, the intercepts are defined in the developer-level. In the Bayesian approach, each intercept is given an informative prior, or, put another way, the intercepts become modeled parameters. Therefore, random intercept model can capture part of heterogeneity caused by developers. The results of this model are displayed in table II-11.

Level 1:

$$\begin{aligned}\mu_{ij} = \beta_1 \log(price_{ij}) + \beta_2 video_{ij} + \beta_3 \log(numscrn_{ij}) + \beta_4 \log(filesize_{ij}) + \beta_5 \mathrm{lm}_{ij} \\ + \beta_6 hm_{ij} + \beta_7 mm_{ij} + \beta_8 arcade_{ij} + \beta_9 brain_{ij} + \beta_{10} cards_{ij} + \beta_{11} casual_{ij} \\ + \beta_{12} racing_{ij} + \beta_{13} sport_{ij} + \beta_{14} wallpaper_{ij} + \beta_{15} \log(rating_{ij}) \\ + \beta_{16} \log(vvc_{ij}) + \beta_{17} \log(vvp_{ij}) + \beta_{18} \log(ivc_{ij}) + \beta_{19} \log(ivp_{ij}) + \eta_j\end{aligned}$$

Level 2:

$$\eta_j \sim normal(\mu.\eta_j, \sigma.\eta)$$
$$\mu.\eta_j = \alpha_1 \log(dac_j) + \alpha_2 \log(dvc_j) + \alpha_3 \log(dvp_j)$$

To further account for developer heterogeneity, we allow random effects in both intercepts and application-level slopes. Therefore, we have our third model:

Model 3: Ordinal Logit Model with Random Coefficients

In the Random Coefficient Model, we assume there are j separate slopes for each variable and each estimate follows a normal distribution, where j represent a developer. By allowing each estimate to vary across different developers, we can control the influence of unobserved heterogeneity of developers on each variable. Besides, we add a random component δ_{ij}, which can capture unobserved heterogeneity among applications. The results of this model are displayed in the table II-12.

$$\mu_{ij} = \beta_{j,1}\log\left(price_{ij}\right) + \beta_{j,2}video_{ij} + \beta_{j,3}\log(numscrn_{ij}) + \beta_{j,4}\log\left(filesize_{ij}\right) + \beta_{j,5}\mathrm{lm}_{ij}$$
$$+ \beta_{j,6}hm_{ij} + \beta_{j,7}mm_{ij} + \beta_{j,8}arcade_{ij} + \beta_{j,9}brain_{ij} + \beta_{j,10}cards_{ij}$$
$$+ \beta_{j,11}casual_{ij} + \beta_{j,12}racing_{ij} + \beta_{j,13}sport_{ij} + \beta_{j,14}wallpaper_{ij}$$
$$+ \beta_{j,15}\log\left(rating_{ij}\right) + \beta_{j,16}\log\left(vvc_{ij}\right) + \beta_{j,17}\log\left(vvp_{ij}\right) + \beta_{j,18}\log\left(ivc_{ij}\right)$$
$$+ \beta_{j,19}\log\left(ivp_{ij}\right) + \eta_j$$
$$\beta_{jk} \sim normal(\mu.\beta_k, \tau.\beta_k), k = 1,2,\dots 19$$
$$\mu.\beta_k \sim normal(0, 0.001)$$
$$\tau.\beta_k \sim gamma(0.001,0.001)$$

$$\eta_j \sim normal(\mu.\eta_j, \sigma.\eta)$$
$$\mu.\eta_j = \alpha_1\log\left(dac_j\right) + \alpha_2\log(dvc_j) + \alpha_3\log(dvp_j)$$
$$\alpha_m \sim normal(0,0.001)$$
$$\sigma.\eta \sim gamma(0.001,0.001)$$

To fit the proposed models, we take a Bayesian Approach to estimate parameters and the four cutoff points. For all models, we assume diffuse priors and run a Markov Chain Monte Carlo sampler for 5,000 iteration which serves as a burn-in period. We then obtain inferences from posterior samples from the next 20,000 iterations.

2.5 Empirical Results

2.5.1 Model Comparison and Unobserved heterogeneity

Table II-8 illustrates the DIC scores, the random error, and the cutoffs of the three models.

Table II-8 Results of random effects, model fit, and Cutoffs

	Ordinal Logit Model			**Random Intercepts**			**Random Coefficients**		
	mean	2.5%	97.5%	mean	2.5%	97.5%	mean	2.5%	97.5%
σ_η				1.570	0.153	1.300	1.790	1.397	2.260

σ							0.452	0.098	0.278
	Dbar	Dhat	DIC	Dbar	Dhat	DIC	Dbar	Dhat	DIC
	5308	5282	**5333**	4618	4393	**4842**	3786	3081	**4490**
	mean	2.5%	97.5%	mean	2.5%	97.5%	mean	2.5%	97.5%
cutoff1	0.009	0.000	0.032	0.016	0.001	0.058	0.019	0.001	0.073
cutoff2	1.565	1.419	1.720	2.061	1.856	2.296	2.631	2.364	2.958
cutoff3	3.859	3.654	4.073	4.861	4.561	5.184	6.120	5.707	6.608
cutoff4	6.546	6.251	6.840	8.075	7.638	8.517	10.020	9.446	10.720
cutoff5	9.796	9.386	10.230	11.820	11.260	12.420	14.770	13.980	15.770

We can see that the DIC score of random coefficient model is much lower than the other two models, thus provides clear evidence that it is critical to account for heterogeneity among developers in an analysis of the effects of app attributes on downloads. Moreover, the error itself is significantly positive, which indicates that there's substantial unobserved heterogeneity among developers ($\sigma_\eta = 1.79$). The random error represents unobserved heterogeneity among apps ($\sigma = 0.452$) which is also significantly positive.

Table II-9 illustrates the impact of unobserved developer heterogeneity on each variable. From the table, we can see the unobserved variance of average rating, price, and racing, are the largest three, which indicates that the app difference on rating, price, and Racing are mostly attributed to unknown heterogeneity among developers. The unobserved heterogeneity of developer plays an important role on the effects of each factor.

Table II-9 Radom effects of each factor

Unobserved Variance	mean	val2.5pc	val97.5pc
rating	1.107	0.490	2.109
video	0.944	0.555	1.414
price	1.246	0.753	1.746
numscrn	0.851	0.539	1.240
filesize	0.479	0.313	0.687
arcade	0.989	0.516	1.526

brain	0.742	0.385	1.266
cards	0.875	0.379	1.645
casual	0.706	0.357	1.181
racing	1.623	0.921	2.393
sport	0.788	0.389	1.347
wallpaper	0.974	0.409	1.766
lm	0.690	0.386	1.093
hm	0.811	0.390	1.526
mm	0.738	0.365	1.278
vvc	0.459	0.302	0.656
vvp	0.573	0.391	0.767
ivp	0.575	0.367	0.819
ivc	0.484	0.317	0.674

2.5.2 Effects of App Attributes

Table 11-10 demonstrates the posterior means of intrinsic app attributes by implementing the three different models.

Table II-10 Results of Model with App-Level Attributes

	Ordinal Logit Model			**Random Intercepts**			**Random Coefficients**		
	mean	2.5%	97.5%	mean	2.5%	97.5%	mean	2.5%	97.5%
β_1	-2.166**	-2.509	-1.825	-3.586**	-4.040	-3.126	-4.163**	-4.737	-3.545
β_2	0.180**	0.000	0.371	0.801**	0.532	1.088	1.010**	0.611	1.394
β_3	-0.005	-0.100	0.095	0.052	-0.077	0.189	0.113	-0.120	0.357
β_4	-0.166**	-0.249	-0.083	-0.186**	-0.300	-0.077	-0.220**	-0.423	-0.027
β_5	0.366**	0.161	0.561	0.492**	0.213	0.778	0.591**	0.209	0.984
β_6	0.852**	0.543	1.171	0.776**	0.264	1.308	0.822**	0.113	1.505
β_7	0.236	-0.023	0.491	0.487**	0.140	0.852	0.599**	0.162	1.107
β_8	4.703**	4.350	5.098	4.985**	4.443	5.494	5.834**	4.727	6.823
β_9	4.732**	4.375	5.124	5.271**	4.701	5.798	6.100**	5.069	7.109
β_{10}	3.950**	3.573	4.360	4.080**	3.443	4.677	4.495**	3.374	5.443
β_{11}	4.305**	3.969	4.667	4.722**	4.193	5.235	5.450**	4.434	6.449
β_{12}	4.290**	3.938	4.677	4.461**	3.905	4.989	4.695**	3.711	5.614
β_{13}	4.211**	3.883	4.576	4.513**	3.940	4.998	4.972**	4.018	5.834
β_{14}	2.948**	2.586	3.318	1.952**	1.453	2.427	1.624**	0.921	2.345
β_{15}	0.932**	0.559	1.447	0.916**	0.505	1.481	1.987**	1.067	3.346

"Price", as we expected, has significantly negative influence on the number of downloads ($\beta_1 = -2.166$). The magnitude of the mean estimates on price is increasing

as the model accounting for more unobserved heterogeneity among developers. The absolute magnitude of price estimates using random intercepts model ($\beta_1 = -3.586$) and random coefficient model ($\beta_1 = -4.163$) becomes greater. "*Youtube Video*" has significantly positive influence on the number of downloads across the three models. It indicates that providing video may help to generate more downloads for the app ($\beta_2 = 1.010$). "*File size*", as a proxy of App quality, surprisingly, has significantly negative influence on the number of downloads ($\beta_4 = -0.220$). Thus, smaller apps are more likely to be downloaded by users. "*Number of Screenshots*", however, do not have any significant an effect on number of downloads.

The estimates of β_{5-7} represent the effects of content rating on the number of downloads. The base level is "Everyone", which means the application is appropriate for everyone to use. Compared to base level "Everyone", all the other three levels, low maturity ($\beta_5 = 0.5907$), high maturity ($\beta_6 = 0.8218$), and medium maturity ($\beta_7 = 0.5990$), have significantly positive influence on number of downloads. Apps of high maturity level are among the most popular ones.

The estimates of β_{8-14} represent effects of the category of applications. Compared to base level "*widget*", all the other seven levels, arcade, brain, cards, casual, racing, sports, and wallpaper, have significantly positive influence on number of downloads. The category of "brain" exert largest influence on the number of downloads.

2.5.3 Effects of Other Apps

The estimates of α_{1-3} represent the influence of other apps developed by each developer. The set of variables are the key findings of present research.

Table II-11 Results of Effects of Other Products of the Developer

	Ordinal Logit Model			**Random Intercepts**			**Random Coefficients**		
	mean	2.5%	97.5%	mean	2.5%	97.5%	mean	2.5%	97.5%
α_1	0.006	-0.132	0.144	0.989**	0.562	1.496	1.505**	0.958	2.253
α_2	1.274**	1.147	1.392	1.781**	1.501	2.110	2.404**	2.009	2.857
α_3	0.058	-0.062	0.173	-0.064	-0.306	0.181	-0.234	-0.559	0.090

From table II-11, we can see the total number of applications developed by the developer positively influence the number of downloads. As the unobserved heterogeneity among developers is controlled, the influence becomes significantly positive ($\alpha_1 = 1.505$), which provides support to the industry's assumption that acquiring users for one application will create network effects among current users and help in acquiring users for future applications.

The average number of user reviews of the applications developed by the developer, as we expected, has a significant positive effect on the number of downloads as well ($\alpha_2 = 2.404$), thus supporting the industry's second assumption that enhancing word of mouth among current users will lead to even larger user base. Further, as developers' unobserved heterogeneity is controlled, the positive influence becomes more salient.

The average price of the applications developed by the developer, interestingly, does not have any significant effect on the number of downloads. However, we do find that average price has negative effects on the number of downloads once a new application is released into the market. The reason might be the price of the new app is also relatively high compared to the similar apps in the marketplace.

2.5.4 Effects of Competition

The estimates of β_{16-19} indicate the influence of competition in the market. The set of variables again reflects the key findings of present research.

Table II-12 Results of Competition Effects

	Ordinal Logit Model			**Random Intercepts**			**Random Coefficients**		
	mean	2.5%	97.5%	mean	2.5%	97.5%	mean	2.5%	97.5%
β_{16}	-0.027	-0.110	0.054	0.037	-0.059	0.139	0.068	-0.083	0.216
β_{17}	0.241**	0.153	0.325	0.297**	0.197	0.395	0.348**	0.196	0.499
β_{18}	0.632**	0.525	0.741	0.776**	0.651	0.906	1.013**	0.818	1.211
β_{19}	-0.905**	-1.061	-0.748	-0.751**	-0.926	-0.555	-0.991**	-1.274	-0.728

First, we found that the average price of the apps viewed by the same user is positive and significant ($\beta_{17} = 0.348$). Therefore, the higher the price of those competitive apps, the more likely is that the users download this app. The estimate provides some evidence of the existence of extreme competition in such markets. If the developers expect to acquire a large number of users, they must offer their products at a lower price compared to competitive apps.

Second, the average review counts of the applications installed by the same user again have significantly positive effects on the number of downloads ($\beta_{18} = 1.013$) of similar apps. Since, as what we observed from our data, the apps featured as "installed by the same users" includes both competitive and complementary apps. The implication is that users tend to install apps similar/complement to the apps which are currently popular in the market.

Third, the average price of the applications installed by the same user has significantly negative effects on the number of downloads ($\beta_{19} = -0.991$). Therefore, if the users installed many higher priced apps on their device, it is less likely for them to download a similar or complementary app. The explanation could be that the users have a budget constraint on the apps installed on their device.

Finally, we do not find any significant effects of the average number of user reviews of the apps featured as “also viewed by the same user”.

2.6 Conclusion

2.6.1 A Summary of Results

In the present research, we propose and empirically test many assumptions to explain what drives businesses to succeed in extremely competitive markets. Currently, many new markets develop rapidly even though they operate their businesses even when the marginal revenue is lower than the marginal cost. In the present research, we label such markets as extremely competitive. The issues examined in this research aim to answer the question of why firms continue in market with extreme competition while struggling to be profitable. There could be two explanations. The first explanation is that by offering their products for free or at a very low price, marketers expect to acquire a large group of users in a relatively short time span, in the hope of making profits from those advertisers who are interested in the use groups. The second explanation is that the marketers expect the users will eventually purchase their premium products after they try their product for free. However, the key assumption underlying both outcomes is that the marketers need to obtain a large group of user base. A further question that this brings up is: what factors contribute to the acquisition of a large user group by businesses? We propose and test two explanations. Specifically, our first explanation is that acquiring users for a product will help in acquiring users for other products since the user base already obtained by the firm will create network effects. To be specific, an app that acquired users may become the source of users for the products that follow. Once the developer releases another product into the market, therefore, the user base of its other products may start using those products as well. The second assumption is that multiple products with acquired users should help even more in acquiring users for new product, or conversely, we can say that firms can enhance WOM for their current products, which may help them attract more users once they release a new product into the market. However, few academic studies have formally tested the two assumptions so far. The objective, therefore, of the present research is to explicitly verify the two assumptions and provide some normative insights for the businesses in the markets of extreme competition.

To empirically examine the effects of those factors on installation base, we collect 2,422 observations across 248 developers from a large App Store. Three set of variables were included in our research: (1) The characteristics of apps; (2) The competition of the market; (3) The proxies of the influence of other apps. The success of each app is measured by the number of downloads. We develop a hierarchical ordinal logit model and calibrate it using Bayesian methods. Empirical results support the two assumptions proposed in our research. Specifically, we find that apps developed by the same developer help in acquiring users for new products thus providing empirical evidence for our two assumptions. In other words, developers with a product released into the marketplace are more likely to acquire users once they release another product. There are many explanations of the mechanism. The first explanation could be that the users of other products will directly become users of their new product. Furthermore, the users of the developer will spread word of mouth, which may indirectly help the developers acquiring users for their new products. The third explanation is that developers with more products will have more exposures among the users which may lead to liking and ultimate behavior of install.

We found evidence of the existence of extreme competition in such markets. First, we found that users, facing a myriad of similar choices, are more likely to choose the product with the lowest price, which explain the source of extreme competition. Developers are keen to offer products with lower price in the hope of expanding their user base at the sacrifice of profits. Second, our research revealed that apps installed on users' device are widely discussed among consumers. In addition, users are more likely to install a similar of a complementary app on their device. Third, users seem to have a budget constraint on the apps installed on their device. If they already installed a number of apps with a relatively high average price, they are less likely to install a similar one.

In addition to the two set of variables mentioned above, our research provides many normative insights into the industry. First, we find that free apps indeed have larger user bases thus confirming extreme competition in the market. The efforts that marketers put on advertising their products by providing YouTube video do help in acquiring users. Higher user rating is an important indicator of larger user base. However, the pictures displayed on their website make no difference in expanding user group. In addition, apps requiring larger storage space will reduce the number of downloads. Originally, we believe that the larger the file size, the more content included in the apps. However, we did not consider other two issues in our conceptual framework. First, the capacity of mobile device may limit downloads of applications with large file size. Second, some other psychological considerations of consumers may also influence consumers' decision to download applications of large size. Apps with higher content rating will have more people to download compared to those apps made for "Everyone". Some of app categories are more popular in the market than others, for instance, "Brain", "Arcade", and "Casual". Specifically, free applications are more likely to gain downloads than paid applications. However, since users can download free applications without any cost, downloads for the applications can not represent the quality of it. And thus, price cannot serve as indicator of quality for this special product. Another search attribute, file size, also negatively influence downloads, which is on the contrary of our expectation.

2.6.2 Contributions

The research makes three contributions to the academic research. First, our present study focuses on the markets where firms face such intense competition that the marginal revenue they earn cannot even compensate the marginal cost. We then label such markets as extremely competitive and found some evidence of "extreme competition". Second, our present study investigates the business strategies of the marketers used in such market and try to verify the implicit assumptions behind the different business strategies. Third, our present study aims to provide some normative insights for such market, which has never been examined empirically by previous studies. Fourth, given the hierarchical nature of our data, unobserved developer heterogeneity is considered in our framework. Results demonstrate that the unobserved developer heterogeneities do play a very important role in generating downloads.

2.6.3 Managerial Implications

Marketers and application developers may find some important managerial implications from our findings. First, free applications with smaller file size may generate more downloads in the marketplace. However, displaying more screenshots does not help app developers acquiring more users. Applications of higher content rating have more download rates, comparing to applications made for everyone. Brain, arcade, and casual applications have significantly more downloads than widgets. Second, competition from similar applications is significant. Our research suggests that a good strategy for developers is to create similar apps which are very popular in the marketplace, since those apps are more likely to acquire user base. Third, developers play a very important role in generating large number of downloads. After controlling for the unobserved heterogeneity of developers, we found that the more apps that the developers released into the market, the more likely that they acquire a large user base for their future products. In addition, users' discussion of their application will increase user base for their future products.

2.6.4 Limitations and Future directions

While our research provides some instructions for the apps market, there are some limitations. First, the apps data we collected are more game-like applications. Another type of applications, more informational and utilitarian applications, exists in the marketplace. The influence of attributes of this type of applications may be quite different from game-like applications.

Second, many other important factors such as the duration that the application has been in the marketplace for, may have significant influence on the number of downloads. Therefore, we expect that the longer the applications in the marketplace, the more customer reviews they can get.

Third, an important mechanism which is probably a bigger driver of app discovery is offline WOM and this face-to-face mechanism is not easily understood nor can easily be influenced. Many consumers instantly download applications that their friends/acquaintances are using when this discovery happens face-to-face as a friend can convey why he or she likes the applications and uses it. Knowing what friends or family members are using has a greater influence on the eventual choice. For example, if most friends of a consumer are using a particular location-sharing, photo-sharing, social networking, or other apps, know that can and probably will influence that consumer's choice of an app. An important area for future research is to not only understand the varying effects of reviews on apps sales but also the review generation process that can determine how many reviews an app gets.

2.7 References of Essay I

Ackerberg (2003), "Advertising, Learning, and Consumer Choice in Experience Good Markets: an empirical Examination", *International Economic Review*, Vol 44, Issue 2, Pages 1007-1040, August

Anand and Shachar (2002), "Brands as Beacons: A New Source of Loyalty to Multiproduct Firms", *Journal of Marketing Research*, Vol. 41, No. 2(may, 2004), pp 135-150

Allen and Hellwig (1986), "Price-Setting Firms and the Oligopolistic Foundations of Perfect Competition", *The American Economic Review*, Vol. 76, No. 2

Aral and Walker (2011), "Creating social Contagion Through Viral Product Design: A Randomized Trial of Peer Influence in Networks", *Management Science*, September 2011 vol. 57. No. 9 1623-1639

Ariely and Shampaner (2006), "How Small is Zero Price? The true Value of Free Products", SSRN, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=951742

Basuroy, Chatterjee, and Ravid (2003), "How critical are critical reviews? The box office effects of film critics, star power, and budgets", *Journal of Marketing*, vol. 67, No.4(Oct. 2003), pp. 103-117

Boatwright, Basuroy, and Kamakura (2007), "Reviewing the reviewers: the impact of individual film critics on box office performance", *Quant Market Econ*, 2007, 5: 401-425

Brynjolfsson, Erik, and Michael Smith (2000). Frictionless Commerce? A Comparison of Internet and Conventional Retailers. *Management Science,* 46(4) 563-585.

Bass (1969), "A New Product Growth Model for Consumer Durables", *Management Science*, 15(January 1969), 215-227

Chamberlain (1933), "The Theory of Monopolistic Competition", Cambridge, MA: Harvard University Press, 1933.

Chen, Y., and Xie, J. (2008). "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix," *Management Science* (54:3), pp. 477-490

Coase (1937), "The Nature of the Firm", *Economica*, Volume 4, Issue16, pages386 - 405, November 1937

Chintagunta, Gopinath, Venkataraman (2010), "The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets", *Marketing Science*, September/October 2010, Vol. 29, No. 5, Page 944-957

Dellarocas, Zhang, and Awad (2007), “Exploring the Value of Online Product Reviews in Forecasting Sales: The Case of Motion Pictures”, *Journal of Interactive Marketing*, Vol.21, Issue 4, Pages 23-45

Godes and Mayzlin (2004), “Using Online Conversations to Study WOM Communication”, *Marketing Science*, 23(4), 545-60

Dou, Niculescu, and Wu (2012), “Engineering Optimal Network Effects via Social Media Features and Seeding in Markets for Digital Goods and Services”, *Information Systems Research*, March 2013 vol. 24 no. 1 164-185

Erdem and Keane (1996), “Decision-Making Under Uncertainty: Capturing Dynamics Brand Choice Processes in Turbulent Consumer Goods Markets”, *Marketing Science*, Winter 1996, vol. 15 no 1 1-50

Fama (1972), “Agency Problems and the Theory of the Firm”, *Journal of Political Economy*, Vol.88, No. 2(Apr., 1980), pp. 288-307

Fabio Ancarani (2002), “Price Levels and Price Dispersion Within and Across Multiple Retailer Types: Further Evidence and Extension”, *Journal of the Academy of Marketing Science,* April 2004 vol. 32 no. 2 176-187

Forman, Ghose, and Wiesenfeld (2008), “Examining the relationship between Reviews and sales: The Role of Reviewer Identity Disclosure in Electronic Markets”, *Information Systems Research* (19:3), pp. 291-313

Gopinath, Chintagunta, and Venkataraman (2010), “Blogs, Advertising and Local-Market Movie Box-Office Performance”, *SSRN*, September 22, 2010.

Godes and Mayzlin (2009), “Firm-created Word-of-Mouth Communication: Evidence from a field test”, *Marketing Science*, 28(4), 721-39

Kalish (1983), “Monoplist Pricing with Dynamic Demand and Production Cost”, *Marketing Science*, Vol. 2, No. 2(Spring, 1983), pp. 135-159

Lal and Sarvary (1999), “When and How is the Internet likely to Decrease Price Competition?” *Marketing Science*, 18.4(1999): 485-503

Mathwick, Malhotra, and Rigdon (2002), “The Effect of Dynamic Retail Experiences on Experiential Perceptions of Value: an Internet and Catalog Comparison”, Journal of Retailing, 78(2002) 51-60

Menon and Kahn (2002); Cross-category effects of induced arousal and pleasure on the internet shopping experience, *Journal of Retailing,* Volume 78, Issue 1, Spring 2002, Pages 31–40

Moon, Bergy and Iabucci (2010), “Dynamic Effects among Movie Ratings, Movie Revenues, and Viewer Satisfaction”, *Journal of Marketing*, 2010, Vol. 74(January 2010), 108-121

Neelamegham and Chintagunta (2004), “Modeling and Forecasting the Sales of Technology Products”, *Quantitative Marketing and Economics*, 2, 195-232, 2004

Pauwels and Weiss (2008), “Moving from Free to Fee: How online Firms Market to Change Their Business Model Successfully”, *Journal of Marketing*, May 2008, Vol. 72, No. 3, pp. 14-31

Pan, Shankar and Ratchford(2002), “Price Competition between Pure Play versus Bricks-and-Clicks E-Tailers: Analytical Model and Empirical Analysis”, *Advances in Applied Microeconomics: Economics of Internet & E-Commerce*, 11 (2002), pp. 29–62

Pan, Ratchford and Shankar (2004), “Price Dispersion on the Internet: A Review and Directions for Future Research”, *Journal of Interactive Marketing*, 18 (4) (2004), pp. 116–135

Ratchford (2009), “Online Pricing: Review and Directions for Research”, *Journal of Interactive Marketing*, Vol 23, Issue 1, Feb 2009, Pages 82-90

Racherla, Furner, and Babb (2013), “Conceptualizing the Implications of Mobile App Usage and Stickness: A research Agenda”, electronic version available at: http://ssrn.com/abstract=2187056

Redmond (2013), “Three Modes of Competition in the Marketplace”, *The American Journal of Economics and Sociology*, Vol. 72, No. 2 (April 2013)

Robinson and Lakhani (1975), “Dynamic Price Models for New-Product Planning”, *Management Science,* June 1975vol. 21 no. 10 1113-1122

Reichheld and Schefter (2000), “E-Loyalty”, *Harvard Business Review*, 2000

Stigler (1957), “Perfect Competition, Historically Contemplated”, *Journal of Political Economy*, Vol. 65, No. 1(Feb., 1957), pp. 1-17

Sinha (2000), “Consumers' Preferences of Promotional Framing of Price”, *Psychology & Marketing*, Mar 2000, Vol. 17 Issue 3, p257-275, 19p, 3 Graphs

Urban, G.L., Sultan, F. and Qualls, W.J. (2000), “Placing trust at the center of your Internet Strategy”, *Sloan Management Review*, Vol. 42 No. 1, pp. 39-48

Wilson (1977), “A bidding model of Perfect Competition”, *The Review of Economics Studies*, Vol. 44, No. 3(Oct., 1977), pp. 511-518

Wallace, Giese, and Johnson (2004), “Customer retailer loyalty in the context of multiple channel strategies”, *Journal of Retailing*, Volume 80, Issue 4, 2004, Pages 249–263

Yadav and Varadarajan (2005), “Interactivity in the Electronic Marketplace: An Exposition of the Concept and Implications for Research”, *Journal of the Academy of Marketing Science*, Fall 2005 vol. 33 no. 4 585-603

III. Price acts as a Proxy of Advertising? Explore the Role of Freemium Business Strategy in the Context of Extreme Competition

Abstract

Many prior studies show that online Word-Of-Mouth (WOM) affects sales, but the issue that why consumers would like to discuss the product online receives relatively less attention from marketers and researchers. The objective of the present study is to examine the drivers of online WOM of a category of unique products, mobile apps, under the context of extreme competition. We are especially interested in the role of Freemium business strategy adopted in such markets where products are mostly offered for free. The implicit assumption behind the strategy is that free products will create more WOM and thus increase user base. However, the role of price in generating online WOM is mixed according to previous studies. The present study aims to examine the question and provide some insights into the industry. Results show that, contrary to the intuitive assumption of the markets that free apps will increase online WOM, premium products are more likely to generate discussion among users. Effects of other factors on online WOM are found. Conclusions and managerial implications are discussed.

Key Words: Online WOM; Freemium Business Strategy; Extreme Competition, Mobile Applications

3.1 Introduction

Give your service away for free, possibly ad supported but maybe not, acquire a lot of customers very efficiently through WOM, referral networks, organic search marketing, etc., then offer premium priced value added services or an enhanced version of your service to your customer base.

------ Fred Wilson[7]

In recently years, “Freemium” become a highly popular business model by which a product or service, such as software, media, games, web services, and so on, is provided free of charge, but a premium is charged for advanced features, functionality, or virtual goods[8]. The business model has notably been in use for markets, such as mobile app market, online radio services, online blogs, etc. Businesses in those markets are mostly operated based on their intuition, as Fred Wilson suggested that most marketers today adopt the Freemium business strategy under which they give their products or services for free, in the hope of acquiring a lot of customers very efficiently through WOM, referral networks, organic search marketing, etc. The strategy indicates that businesses expect that online WOM will help in generating larger user base wherein Freemium is the driving force of online WOM. However, the freemium business strategy is questionable in two aspects (a) whether online WOM will help in generating larger user bases (b) Whether Freemiums will generate more online WOM.

---

[7] http://www.avc.com/a_vc/2006/03/my_favorite_bus.html
[8] http://link.springer.com/content/pdf/10.1007/978-0-387-85895-1_6

Prior studies investigated the question from several perspectives. First, the products sold in those markets, such as mobile apps, blogs, videos, radios, news, etc, are almost purely digital products. Lal and Sarvary (1999) suggest that products and services with primarily digital attributes can be easily communicated online and a number of empirical studies suggested that online WOM has positive influence on sales revenues in markets such as motion picture, books, CDs, and so on(Godes and Mayzlin, 2004; Chevalier and Mayzlin, 2006; Feng and Papatla, 2011; Moe and Trusov, 2011; Godes and Silva, 2012). Second, digital products are a special type of experiential products. Studies show that a consumer tends to rely more on others' recommendations and product experience, when he considers and experiences product than a search product (Bearden and Etzel 1982; Childers and Rao 1992; King and Balasubramanian 1994; Klein 1998; Park and Lee 2009; Senecal and Nantel 2004). Compared to other physical experience goods, digital attributes are much more difficult to evaluate before or even after they actually experience the product. Online WOM, therefore, is critical for the products to succeed in the market, rather than product differentiation. Third, markets of pure digital products are characterized by myriad choices and intense competition (Racherla, Furner, and Babb, 2012). In these markets, hundreds of thousands of individual and small developers compete with others at a very low price and none is sufficiently large so as to exercise any influence whatsoever on prices. Therefore, price competition is not enough to help the businesses to stand out. This again explains why businesses strive to enhance online WOM for their products and services. Fourth, from the previous study, we do find that a larger user base not only contributes users to new products, but also spreads WOM which may ultimately lead to more users. The above characteristics of markets of extreme competition indicate the importance of creating online WOM for products and services if businesses want to succeed in such markets.

The second question, i.e., what factors help in generating WOM in such markets, however, is yet to be addressed. Although prior studies suggest that free applications will indeed acquire more downloads than paid ones, getting consumers to discover and download an app does not mean that consumers will stick to the app and ultimately post reviews online or share the product with their friends. By using Freemium business model, consumers generally can acquire the products or services for free or at a very low price and easily get rid of them. Freemium strategy, therefore, may help businesses to expand user base, but not necessarily help them generate interests in products and services. If consumers demonstrate no interest in the product, there are no opportunities for them to purchase premium products, refer the products to others, or click the ads embedded in the products. Eventually, businesses adopting Freemium strategy will still have difficulties in monetizing their products. Therefore, a major concern to these marketers is whether the Freemium business model will stimulate users to express their opinions online in this hyper-competitive landscape with a wide range of choices. The present study aims to address this question.

By empirically examining the data collected from a major Mobile Apps Store, the present study found that paid apps generate significantly more WOM than free applications when original user base considered. Specifically, we first explore the role of free apps on the volume of WOM without considering the influence of user base. Results show that free applications will significantly boost the volume of online WOM, which is consistent with the results of Berger and Schwartz (2011). We then, add the size of user base as an explanatory variable in our model. Since product diffusion theory suggests that internal influence is mainly determined by the number of consumers who have experienced the products (Mahajan, Muller, and Wind 2000), we expect that the user base acquired by the product would be interested in discussing it. Results still indicate that free apps will boost the volume of online WOM but with a smaller magnitude. Subsequently, rather than relying on the size of the user base as an explanatory variable, we examine the effect of free apps on the likelihood that users spread WOM online after they download the applications. This investigation reveals that the effects of Freemium indicated by the previous two models are spurious. Paid applications, rather than free apps, will significantly boost volume of online WOM. Effects of other factors are also found.

We organize the remainder of this article as follows: in the next section, we provide an overview of previous literature on the relationship between price and online WOM. We then describe the data and estimation method and present the empirical results. Finally, we discuss the contribution of this research and the managerial implications of the results.

3.2 Literature Review

The predominant research focus of online WOM has been on the effects on sales of paid goods (Chevalier and Mayzlin 2006; Dellarocas et al. 2004; Duan et al. 2008; Godes and

Mayzlin 2004). Few studies, however, examine how to stimulate consumers to spread WOM online. In particular, as a key element of market strategy, the effect of price on WOM has received very little attention from scholars. The conventional wisdom of practitioners, however, suggests that consumers like getting products and services for free. Interestingly, willingness to try free products does not mean that consumers like to discuss the products and ultimately buy premium products.

3.2.1 Freemium and Online WOM

Does the Freemium Strategy Increase Online WOM?

Most WOM campaigns involve sending consumer promotional giveaways to encourage them to talk about the product. Several previous studies suggest that giveaways will indeed generate positive effects on online WOM. Holmes and Lett (1977) suggest that product sampling may stimulate WOM among users. A recent study by Berger and Schwartz (2011) find that users who receive giveaways talk more than those who doesn't. Specifically, they find that giving away the product itself or nonproduct extras (e.g., logo hats, recipes) are positively linked to more overall WOM. However, neither samples nor coupons and rebates were linked to an increase of WOM, though they may be useful for increasing other outcomes, such as sales, and quality of conversations. They then suggest that businesses, that aim to generate more WOM, may send consumers the full product or related extras to try (Berger and Schwartz 2011). Their explanation is that giveaways provide product experience which may boost information and reduce uncertainty (Hoch and Ha, 1986) thus making it easier for people to learn about the product and have an opinion to share. Other studies examine the effects of free sample promotions on measures such as belief strength and attitude (Marks and Kamins 1988), perceptions of the brand (Hamm et al, 1969), purchase event feedback, i.e., brand loyalty (Gedenk and Neslin 1999), and reciprocity (Cialdini 2001). The results of these studies may give further support to the positive effects of free giveaways on online WOM found by Berger and Schwartz(2011). However, these studies neither examine the effects on online WOM directly nor measure any long-term effects, they only suggest that free samples could generate positive brand attitudes towards brand. Although Berger and Schwartz (2011) find that free giveaways do help in generating more discussions among the users, the findings may not be applied in pure digital markets due to the extreme competition and other uniqueness of the products.

Does charging a price for products increase online WOM?

The above literature suggests that free giveaways will stimulate online WOM, but some other studies suggest that premium products are more likely to generate online WOM, though the effects hasn't verified directly so far. A stream of research suggests that price can be interpreted as a cue for product quality (Gerstner 1985; Tao and Monroe 1989). The research indicates that, when faced with quality uncertainty, consumers are likely to use price as a signal of quality before they make purchase (Dodds et al, 1991; Grewal 1995; Kirmani and Rao 2000; Mitra 1995; Rao and Monroe 1988, 1989). A high customer satisfaction will be generated if the product performance is consistent with the premium price. Several studies are found in this area. One notable study by Voss et al(1998) suggests that performance expectations will have a positive effect on satisfaction when there is price-performance consistency and will have no effect when price-performance are inconsistent. Grewal, Krishnan, Baker, and Borin (1998) indicates that the influence of price discount on a brand's perceived quality was minimal but exerted significant positive influence on perceived value since prices paid less than an individual's reference price enhance buyers' value perceptions (Grewal et al 1998). Based on the aforementioned literature, a premium price may signal a high product performance. If users perceive that the price-performance is consistent, their expectations are met and satisfaction increased. Consumers, therefore, are likely to spread WOM, since consumer satisfaction or dissatisfaction is an important determinant of WOM (Yi 1990). Some of the quality expectation may be disconfirmed by actual experience, and ultimately lead to dissatisfaction (Cadotte et al. 1987; Churchill and Surprenant 1982; Spreng et al. 1996; Rust et al. 1999). Consumers may therefore engage in complaint behavior and spread negative WOM online. Hence, higher product price will generate more WOM regardless consumers' level of satisfaction.

However, previous literature on the relationship of price premium and WOM is scarce, though these are a few exceptions. Richins (1983b) examined negative WOM by dissatisfied consumers (telling others about their unsatisfactory experience) and indicated that negative WOM occurred when the problem was severe[9]. In addition, consumers seem to give more weight to negative information than to positive information (Lutz 1975). Curren and Folkes(1987) expanded on Richins'(1983b) work by examining whether communications for product performance influenced consumers' positive as well as negative communications about products. Li and Hitt(2010) suggests

---

[9] High price is one of indicators of problem severity (Richins 1983b).

that price has positive effects on consumer ratings since it not only reflects the perceived product quality but also demonstrates the perceived value, i.e., the difference between price and perceived quality. However, they did not examine the influence of price on WOM volume.

In the present study, we aim to examine the influence of price on WOM volume in the markets of pure digital products, where others' opinions are especially important. In the markets of pure digital products, although most of products are offered for free, a group of consumers who are seeking for high quality products believe that high price signal high quality, facing a myriad of options and a difficulty to differentiate product attributes. These consumers may put more cognitive effort on high-priced items (Wathieu and Bertini 2007), and seek for product information before they purchase the product and examine the benefits of the product during the process of experiencing the product. Thus they have a need to share their knowledge of the product based on their experience with it and their consequent discovery of its unique features and benefits or drawbacks (Feng and Papatla 2011). Therefore, premium price can be considered as another vehicle of advertising demonstrating the product quality which helps the businesses attract more less-price sensitive users and, hence, share their opinions online. Along these lines, the present paper proposes that products sold at a premium price are more likely to spread their WOM online.

Premium and Positive/Negative WOM

Based on aforementioned literature, products sold at a premium price will generate online WOM because of (1) satisfaction due to price-performance consistency (2) dissatisfaction due to price-performance inconsistency. Specifically, if the consumers perceived the price is confirmed by the high performance, they will have a need to share the good experience with others. However, if the product performance does not meet their expectations, they will vent their dissatisfaction by spreading negative WOM. In this sense, premium price will help in generating either positive WOM when consumers are satisfied with the product or negative WOM when consumers are dissatisfied with the product.

3.2.2 Effects of Other product attributes on online WOM

Average rating

The average rating is another indicator of product quality, perhaps a more accurate quality indicator than price, since average rating reflects the actual experience and evaluation of a group of users. A number of studies indicate that higher average rating will lead to increased sales revenues (Forman et al. 2008; Chen et al. 2007; Dellarocas et al 2008; Clemons et al. 2006). However, few studies examine the influence of WOM valence on WOM volume. A notable exception is that Dellarocas, Awad, and Zhang (2008) considers the interplay between WOM valence and WOM volume. They find a positive influence of WOM valence on WOM volume, which in turn influences retail sales. The explanation is that higher average rating indicates more of the community members agree with the consumer’s assessment of the product, which encourages him or her to distribute WOM to enhance self-esteem (Sundaram et al. 1998; Wangenheim et al.2003). However, Dellarocas et al. (2008) did not consider the effect of valence on positive WOM volume and negative WOM volume separately. Previous studies suggest that consumers’ online WOM are overwhelmingly positive (Chevalier and Mayzlin 2004). Therefore, we expect that a higher average rating will lead to more positive and overall WOM volume and less negative WOM volume. However, a consumer adopt a product with low average rating will be less likely to post review online due to self-esteem.

Promotion cues: video and screenshots

Consumers who search and shop for products are generally exposed to a number of promotional cues. In the case of mobile apps store, most of developers will provide some screenshots or a YouTube video or both to show the content of their applications. These promotional cues can activate associated concepts in consumers' memory and making them more accessible (Higgins, Rholes, and Jones 1977; Lynch and Scrull 1982; Brakus, Bernd and Zarantonello, 2009) during the process of making purchase decision or spreading WOM. In other words, these cues are designed to increase the likelihood of retrieval of contents of the ad memory trace during brand decisions. If positive, these cues should result in more favorable brand evaluations and an increased likelihood of purchase (Keller, 2009) and eventually lead to more online WOM volume. In addition, a product with more promotional cues may indicate that the marketers are more devoted to their products and have more resources to create a better product. Therefore, the promotional cues, to some extent, can also act as cues of product quality. More promotional cues may thus result in high customer satisfaction, which, hence, encourages the discussion among users.

3.2.3 Effects of Other Apps

In choosing among competing products, consumers are faced with uncertainty of product performance. Although price and average rating can serve as important proxies of product quality, brand reputation could also signal the quality of products (Dawar and Parker 1994; Rao and Monroe 1989; Grewal, Krishnan, Baker, and Borin 1998). A mobile developer with more products is more likely to develop a better product in the future, since they may gain lots of experience and skills from the process of developing other products. In addition, the more products the developer release into the market, the larger user base they have. Current user base may spread WOM, which may increase brand awareness, and eventually encourage more people to discuss it. Moreover, consumers may perceive a product to have higher performance if the developers' products are generally sold at a premium price.

3.2.4 User Base and online WOM

A salient characteristic of the data used by our study is that the reviews and ratings are provided by actual users of applications. In other words, if users want to express their opinions in the app store, they must first download the applications. Duan, Gu, and Whinston (2008) suggests that the larger the pool of consumers who have experienced a movie, the more WOM will be generated, which is, to some extent, consistent with studies of product diffusion, which indicate that internal influence is mainly determined by the number of consumers who have experienced the products (Mahajan, Muller, and Wind, 2000). In the study of Duan et al. (2008), movie sales are used as a proxy for the number of consumers. The results suggest that the number of consumers has positive influence on the volume of online WOM. We, therefore, expect that use base will have significant positive influence on the volume of online WOM.

To empirically test the effect of price on online WOM in pure digital markets under the context of extreme competition, we collect a dataset from a famous App Store and a log-normal regression model was implemented in Bayesian paradigm. A number of other factors, such as the categories of apps, app content rating, brand reputation, and competition from other similar applications were controlled. To control unobserved heterogeneity of developers, random intercept and random coefficient model were implemented in Bayesian paradigm. Results indicate that the negative effect of price on online WOM is spurious without considering the influence of user base. Contrary to intuition apparent among practitioners, free apps are less likely to generate interest among users. Users are more interested in discuss the products that they actually paid for. Effects of other factors are found.

3.3 Data description

2,422 applications were collected from Google Play Store on Oct. 28, 2012. There are two main features of our data: (1) The data has a hierarchical structure. The first level is the application-level. The second level is developer-level. The total 2,422 apps are nested under 248 developers. There are 11 Application-level factors and 3 Developer-level factors. (2) The website of data source requires that all the reviewers download the applications before they rate the applications. The variable definitions and summary statistics were displayed in Table III-1 to Table III-3.

3.3.1 Dependent Variable

Review Counts

"Total Review Counts" measures the total review counts of each application. From Table III-2, we can see that the range of review counts is very large – 1,429,000, and it is heavily right-skewed. "Positive Review Counts" measures the volume of positive total WOM, including all the ratings of "star 5". The positive review counts are also strongly right-skewed. The range is 1,131,000. "Negative Review Counts" measures the volume of negative total WOM, including all the ratings of "star 1" and "star 2". The negative counts are strongly right skewed as well. However, the range is smaller than that of "positive review counts" and "total review counts" – 89,640. We are more conservative in the designation of a review as positive because empirical evidence (Chevalier and Mayzlin, 2006) suggests that consumer reviews tend to be overwhelmingly positive. We, therefore, designate a review as positive only if the reviewer gives the highest possible rating to an application. On the other hand, ratings that are extremely negative (for example, a rating of "star 1"), or close to being extremely negative (e.g., a rating of "star 2"), are designate as negative since there is little empirical evidence of consumers being overly negative. Hence, extremely negative, or close to being so, are both designate as negative ratings.

Downloads

The number of downloads is obtained as the range of downloads rather than the absolute number. For instance, the number of downloads in last 30 days of "Slot Machine" is 5,000,000-10,000,000 up to Oct. 28, 2012. There are 12 categories in total. From lower right corner of Table III-2, we can see the distributions of downloads. There are 440 apps whose downloads ranged from 100,000-500,000. And the second largest number of installs falls into the range of 10,000-50,000 – 409 apps. The distribution of installs is close to a normal distribution.

3.3.2 Independent Variables

Price

The price of application is different from other continuous variable. 1487(61.40%) applications are free and therefore the price is “0”. The rest of it (935 applications) are paid applications and the range of price for paid ones is $7.99. The distribution of price for paid applications is roughly normally distributed. Given the intrinsic attributes of price, we recode the measures of price as dummy variable. The free apps are coded as “zeroes” and paid applications as “ones”.

Average rating:

“*Average Rating*” measures the valence of online review for each application. The average rating is overwhelmingly positive. Most applications receive an average rating of 4 or above.

Promotional cues: Number of Screenshots and Video

“*numscrn*” measures the number of Screenshots of each application. If you go the webpage of a certain application, it usually will display several screenshots to demonstrate the content of the product. The range of the number of screenshots is 8 and it is roughly normal distributed. “*video*” means whether there’s YouTube Video embedded on the webpage of the App, acting as a promotional cue for it. There are 772 applications YouTube video on their website.

Other attributes of product

"*Content Rating*" measures the maturity level of the application content. There are four different levels, everyone, low maturity (*lm*), medium maturity (*mm*), and high maturity (*hm*). 1,235 applications are made for everyone (50.99%). "*Category*" indicates the type of the applications. There are eight types of applications: *arcade*, *brain*, *cards*, *casual*, *racing*, *sport*, *wallpaper*, and *widget*. Arcade and wallpaper are the largest two types of applications. "*File Size*" measures the installation space needed by the applications. Applications of larger file size usually contain more features and functionalities.

Effects of Competition

On the webpage of each application, Google Play will feature some other applications on the left-hand side. It usually lists four applications viewed by the same users and four other applications installed by the same users. Those applications are competing users with the focal app. Therefore, we create four variables based on the information of the above two types of products. Specifically, we calculate the average counts and price for both groups of applications and lead to the following four variables.

Table III-1 Description of four Competition Measures

Variable Name	Description
vvc	Average Counts of the Apps Viewed by similar Users
vvp	Average price of the Apps Viewed by similar Users
ivc	Average Counts of the Apps Installed by similar Users
ivp	Average Price of the Apps Installed by similar Users

Table III-2 Summary Statistics of Application-Level Attributes

	Min.	**1st Qu**	**Median**	**Mean**	**3rd Qu**	**Max.**
Total review counts	1	54	450	14217	4243	1429241
Positive review	0	32	263	10110	2506	1131000
Negative review	0	7	60	868	469	89640

rating	0.000	3.900	4.200	4.001	4.500	5.000
price	0.000	0.000	0.000	0.818	0.990	7.990
numscrn	0	3	5	4.617	5	8
filesize	0.006	1.900	5.700	16.890	14.000	1843.000
vvc	0	3833	14364	35814	44295	754480
vvp	0.000	0.000	0.740	0.955	1.480	8.310
ivc	0	54	121	1256	430	131171
ivp	0.000	0.000	0.000	0.832	1.590	8.070

Categorical Variables			**Downloads (Dependent Variable)**		
Variable name	**# of Apps**	**Percentage**	**Range of Downloads**	**# of Apps**	**Percentage**
Arcade	386	15.94%	<100	102	4.21%
Brain	280	11.56%	100 - 500	148	6.11%
Cards	33	1.36%	500 - 1,000	78	3.22%
Casual	373	15.40%	1,000 - 5,000	250	10.32%
Racing	228	9.41%	5,000 - 10,000	160	6.61%
Sport	336	13.87%	10,000 - 50,000	409	16.89%
Wallpaper	360	14.86%	50,000 - 100,000	185	7.64%
Widget	126	5.20%	100,000 - 500,000	440	18.17%
Video-yes	772	31.87%	500,000 - 1,000,000	174	7.18%
Video-no	1650	68.13%	1,000,000 - 5,000,000	342	14.12%
Everyone	1235	50.99%	5,000,000 - 10,000,000	76	3.14%
High maturity	223	9.21%	>10,000,000	58	2.39%
Low maturity	523	21.59%			
Medium maturity	441	18.21%			

Note: “nums” represents number of screen shots

Effects of Other Apps

To account for the heterogeneity of developers, we create three developer-level attributes:

“*dac*” is the total number of the applications display on Google Play and installed by the same this developer. Therefore, this variable measures, to some extent, the size of each developer. We can see the range is 67. The distribution of developer size is right-skewed. Most of developers has 28 or less applications.

"*dvc*" is the average number of review counts of those applications developed by the same developer. The range of the average review counts is 294,800. And again, the distribution of it is roughly normally distributed but a little bit right-skewed.

"*dvp*" is the average price of applications developed by the same developer. The range of it is 5.37 and it is also roughly normally distributed.

The correlations between the three developer-level variables are relatively low so we don't have to worry about multicollinearity problem.

Table III-3 Summary Statistics of Developer-Level Attributes

Correlation Matrix				Summary Statistics					
	dac	dvc	dvp	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
dac	1.000	-0.425	-0.046	5.00	12.00	28.00	31.970	49.00	72.00
dvc		1.000	-0.092	1	127	988	11460	8908	294800
dvp			1.000	0.00	0.00	0.37	0.60	0.93	5.37

3.4 Model Specification

Our dependent variable is the number of online reviews. Since the variability of counts is huge, the ordinary count models, such as Poisson regression, Binomial regression, and Negative Binomial regression are not appropriate. In addition, the dependent variable is strongly right-skewed. Therefore, we adopt a log-normal regression to empirically test the issue:

Model 11: Log-Normal regression

The dependent variable in this model is the total number of online reviews. Since the measures are positive and strongly right-skewed, we use the log-transformed version of the variable. Both application-level variables and developer-level variables were simply included in the basic model. The results of this model were displayed in table III-4.

$$\begin{aligned}\text{Log}(counts_{ij}) &= \beta_1 \log(price_{ij}) + \beta_2 video_{ij} + \beta_3 \log(numscrn_{ij}) + \beta_4 \log(filesize_{ij}) + \beta_5 \mathrm{lm}_{ij} + \\ &\beta_6 hm_{ij} + \beta_7 mm_{ij} + \beta_8 arcade_{ij} + \beta_9 brain_{ij} + \beta_{10} cards_{ij} + \beta_{11} casual_{ij} + \\ &\beta_{12} racing_{ij} + \beta_{13} sport_{ij} + \beta_{14} wallpaper_{ij} + \beta_{15} \log(rating_{ij}) \\ &+\beta_{16} \log(vvc_{ij}) + \beta_{17} \log(vvp_{ij}) + \beta_{18} \log(ivc_{ij}) + \beta_{19} \log(ivp_{ij}) \\ &+\alpha_1 \log(dac_j) + \alpha_2 \log(dvc_j) + \alpha_3 \log(dvp_j)\end{aligned}$$

Where,

i represents application 1 to application 2422;

j represents developer 1 to develop 248;

In the above equation, the log of the total review counts is related to an applications intrinsic attributes, average rating, competition, and developers' experience and skills. Note that our data has a hierarchical structure (i.e., applications are nested under developers). However, this model did not consider the heterogeneity of developers. We assumed independence of applications and pooled all the applications together.

Model 12: Log-Normal regression with User base

In the above model, we ignore the influence of user base on the number of online reviews. According to previous studies, a larger use base would generate more sales revenues and online WOM volume for the marketers in the future. To account for the effect of user base, we add the number of downloads to the model as a continuous explanatory variable. The results of this model were displayed in table III-5.

$$\text{Log}(counts_{ij})$$
$$= \beta_1 \log(price_{ij}) + \beta_2 video_{ij} + \beta_3 \log(numscrn_{ij}) + \beta_4 \log(filesize_{ij}) + \beta_5 \mathrm{lm}_{ij} +$$
$$\beta_6 hm_{ij} + \beta_7 mm_{ij} + \beta_8 arcade_{ij} + \beta_9 brain_{ij} + \beta_{10} cards_{ij} + \beta_{11} casual_{ij} +$$
$$\beta_{12} racing_{ij} + \beta_{13} sport_{ij} + \beta_{14} wallpaper_{ij} + \beta_{15} \log(rating_{ij})$$
$$+\beta_{16} \log(vvc_{ij}) + \beta_{17} \log(vvp_{ij}) + \beta_{18} \log(ivc_{ij}) + \beta_{19} \log(ivp_{ij}) +$$
$$\beta_{20} \log(downloads_{ij})$$
$$+\alpha_1 \log(dac_j) + \alpha_2 \log(dvc_j) + \alpha_3 \log(dvp_j)$$

Where,

i represents application 1 to application 2422;

j represents developer 1 to develop 248;

Model 13: Log-Normal Regression on Rates

There is an important characteristic of online reviews for applications posted by users. Users have to first download the application if they want to express their opinions online. In this regard, it is different from movies or other entertainment goods, where some people may post reviews based on their impressions about the movie trailers or commercials before they actually experience the product. In mobile app market, online WOM is generated by the actual user of the application. Therefore, a relevant question is: what is the probability that a user posts a review for the product after using it? Therefore, instead of modeling total number of reviews directly, we regress on the ratio of the total number of reviews posted for the product to the number of downloads. The model can describe how the ratio depends on the explanatory variables, such as price, competition, and developers' experience and skills. We then specify the following model, where the dependent variable is the ratio of total review counts and total downloads of applications. The results of this model are displayed in table III-6.

$$\text{Log}(\frac{volume_{ij}}{downloads_{ij}})$$
$$= \beta_1 \log(price_{ij}) + \beta_2 video_{ij} + \beta_3 \log(numscrn_{ij}) + \beta_4 \log(filesize_{ij}) + \beta_5 \text{lm}_{ij} +$$
$$\beta_6 hm_{ij} + \beta_7 mm_{ij} + \beta_8 arcade_{ij} + \beta_9 brain_{ij} + \beta_{10} cards_{ij} + \beta_{11} casual_{ij} +$$
$$\beta_{12} racing_{ij} + \beta_{13} sport_{ij} + \beta_{14} wallpaper_{ij} + \beta_{15} \log(rating_{ij})$$
$$+\beta_{16} \log(vvc_{ij}) + \beta_{17} \log(vvp_{ij}) + \beta_{18} \log(ivc_{ij}) + \beta_{19} \log(ivp_{ij})$$
$$+\alpha_1 \log(dac_j) + \alpha_2 \log(dvc_j) + \alpha_3 \log(dvp_j)$$

Where,

i represents application 1 to application 2422;

j represents developer 1 to develop 248;

Model 22: Log-Normal Regression Model with Random Intercepts

In the previous analysis, we pooled all applications together, and made the assumption that one application is independent from another. However, our data structure suggests that applications developed by the same developer may have some common characteristics. To account for the hierarchical data structure, we introduce a developer-specific parameter η_j to the utility function. Therefore, we have our second model:

Level 1:

$$\text{Log}(\frac{volume_{ij}}{downloads_{ij}})$$
$$= \beta_1 \log(price_{ij}) + \beta_2 video_{ij} + \beta_3 \log(numscrn_{ij}) + \beta_4 \log(filesize_{ij}) + \beta_5 \text{lm}_{ij} +$$
$$\beta_6 hm_{ij} + \beta_7 mm_{ij} + \beta_8 arcade_{ij} + \beta_9 brain_{ij} + \beta_{10} cards_{ij} + \beta_{11} casual_{ij} +$$
$$\beta_{12} racing_{ij} + \beta_{13} sport_{ij} + \beta_{14} wallpaper_{ij} + \beta_{15} \log(rating_{ij})$$
$$+\beta_{16} \log(vvc_{ij}) + \beta_{17} \log(vvp_{ij}) + \beta_{18} \log(ivc_{ij}) + \beta_{19} \log(ivp_{ij})$$
$$+\alpha_1 \log(dac_j) + \alpha_2 \log(dvc_j) + \alpha_3 \log(dvp_j)$$
$$+\eta_j$$

Level 2:

$$\eta_j \sim normal(\mu.\eta_j, \sigma.\eta)$$

$$\mu.\eta_j = \alpha_1 \log(dac_j) + \alpha_2 \log(dvc_j) + \alpha_3 \log(dvp_j)$$

To further account for developer heterogeneity, we allow random effects in both intercepts and application-level slopes. Therefore, we have our third model:

Model 32: Log-Normal Regression Model with Random Coefficients (without correlation)

In the Random Coefficient Model, we assume there are J separate slopes for each variable which follow a normal distribution. Therefore, random coefficient model can capture heterogeneity among developers as well. We also add a random component δ_{ij}, which can capture unobserved heterogeneity among applications. The results of this model are displayed in table III-7.

Level 1:

$$Log(\frac{volume_{ij}}{downloads_{ij}})$$

$$= \beta_{j,1} \log(price_{ij}) + \beta_{j,2} video_{ij} + \beta_{j,3} \log(numscrn_{ij}) + \beta_{j,4} \log(filesize_{ij})$$

$$+ \beta_{j,5} \mathrm{lm}_{ij} + \beta_{j,6} hm_{ij} + \beta_{j,7} mm_{ij} + \beta_{j,8} arcade_{ij} + \beta_{j,9} brain_{ij} + \beta_{j,10} cards_{ij}$$

$$+ \beta_{j,11} casual_{ij} + \beta_{j,12} racing_{ij} + \beta_{j,13} sport_{ij} + \beta_{j,14} wallpaper_{ij}$$

$$+ \beta_{j,15} \log(rating_{ij})$$

$$+\beta_{j,16} \log(vvc_{ij}) + \beta_{j,17} \log(vvp_{ij}) + \beta_{j,18} \log(ivc_{ij}) + \beta_{j,19} \log(ivp_{ij})$$

$$+\alpha_1 \log(dac_j) + \alpha_2 \log(dvc_j) + \alpha_3 \log(dvp_j)$$

$$+\eta_j$$

Level 2:

$$\beta_{jk} \sim dnorm(\mu.\beta_k, \tau.\beta_k), k = 1,2, \dots 20$$

$$\eta_j \sim normal(\mu.\eta_j, \sigma.\eta)$$

$$\mu.\eta_j = \alpha_1 \log(dac_j) + \alpha_2 \log(dvc_j) + \alpha_3 \log(dvp_j)$$

Model 42: Log-Normal Regression Model with Random Coefficients (with correlation)

In the above specification, although we assume random coefficients, we allow the coefficients to be independent, thus ignoring the correlation between variables. In the next specification, we assume that the coefficients are correlated. Therefore, the model can not only capture part of heterogeneity among developers, but also the correlation between variables.

Level 1:

$$\text{Log}\left(\frac{volume_{ij}}{downloads_{ij}}\right)$$
$$= \beta_{j,1}\log(price_{ij}) + \beta_{j,2}video_{ij} + \beta_{j,3}\log(numscrn_{ij}) + \beta_{j,4}\log(filesize_{ij}) + \beta_{j,5}\text{lm}_{ij}$$
$$+ \beta_{j,6}hm_{ij} + \beta_{j,7}mm_{ij} + \beta_{j,8}arcade_{ij} + \beta_{j,9}brain_{ij} + \beta_{j,10}cards_{ij}$$
$$+ \beta_{j,11}casual_{ij} + \beta_{j,12}racing_{ij} + \beta_{j,13}sport_{ij} + \beta_{j,14}wallpaper_{ij}$$
$$+ \beta_{j,15}\log(rating_{ij})$$
$$+\beta_{j,16}\log(vvc_{ij}) + \beta_{j,17}\log(vvp_{ij}) + \beta_{j,18}\log(ivc_{ij}) + \beta_{j,19}\log(ivp_{ij})$$
$$+\eta_j + \delta_{ij}$$

$$\overleftarrow{\overrightarrow{\beta_{jk}}} \sim dmnorm\left(\overrightarrow{\mu.\beta_k}, T\right), k = 1,2, \dots 19$$
$$T \sim wishart(R, v)$$
$$\overrightarrow{\mu.\beta_k} \sim dmnorm(\overrightarrow{mn.\beta_k}, B)$$

$$\eta_j \sim normal(\mu.\eta_j, \tau_\eta)$$
$$\mu.\eta_j = \alpha_1\log(dac_j) + \alpha_2\log(dvc_j) + \alpha_3\log(dvp_j)$$
$$\alpha_m \sim normal(0{,}0.001)$$
$$\tau_\eta \sim gamma(0.001{,}0.001)$$
$$\delta_{ij} \sim normal(0, \sigma)$$
$$\sigma \sim gamma(0.001{,}0.001)$$

To fit the proposed models, we take a Bayesian Approach to estimate parameters. For all models, we assume diffuse priors and run a Markov Chain Monte Carlo sampler for 5,000 iteration which serves as a burn-in period. We then obtain inferences from posterior samples from the next 20,000 iterations.

3.5 Empirical Results

Table III-4 reports posterior mean estimates of implementing model on total number of reviews. There are many findings. First, without considering the influence of user base, price ($\beta_1 = -0.3243$) and file size ($\beta_4 = -0.1126$) have significantly negative influence on total number of reviews. In other words, consumers are more likely to post reviews for free applications and applications of smaller size. Second, having YouTube video ($\beta_2 = 0.3994$) and number of screenshots ($\beta_3 = 0.1592$) will significantly boost volume of online WOM. Third, number of applications developed by the developer ($\alpha_1 = 0.1528$), average review counts ($\alpha_2 = 1.5970$) of the applications developed by the developer, average rating counts ($\beta_{15} = 0.3534$), and having higher maturity content ($\beta_5 = 0.8174$; $\beta_6 = 1.3460$; $\beta_7 = 0.7749$), will significantly increase total number of reviews. However, applications with high maturity content turn to be not significant after controlling for developer heterogeneity. Fourth, after controlling for developer heterogeneity, average rating ($\beta_{16} = 0.0821$) and price ($\beta_{17} = 0.1885$) of the applications featured as "viewed by other users", average counts ($\beta_{18} = 0.5375$) featured as "installed by other users", have significantly positive influence on total review counts. However, average price ($\beta_{19} = -0.8873$) of applications featured as "installed by other users" has significantly negative influence on total review counts.

Table III-5 reports posterior mean estimates of implementing the model on total number of reviews. However, in this model, the number of downloads was included as an explanatory variable. According to DIC score, we know the second model fits better than the first one. By simply adding the variable of user base, we can see many of our findings of model 1 are spurious.

Table III-4 Posterior means of the Lognormal Regression

	Lognormal regression			**random intercept**			**random coefficient - uncorrelated**			**random coefficient - correlated**		
	mean	2.50%	97.50%	mean	2.50%	97.50%	mean	2.50%	97.50%	mean	2.50%	97.50%
α_1	0.1528**	0.0308	0.2784	4.0320**	3.3950	4.672	3.1450**	2.5070	3.8400	3.8960**	3.2220	4.5470
α_2	1.5970**	1.4970	1.6950	3.0930**	2.6340	3.556	2.7300**	2.3130	3.1870	3.1060**	2.6340	3.5800
α_3	0.0248	-0.0806	0.1303	0.4520	-0.0135	0.9299	0.1477	-0.2486	0.5517	0.3324	-0.1098	0.7645
β_1	-0.3243**	-0.6109	-0.0354	-2.1400**	-2.3740	-1.894	-1.9800**	-2.2880	-1.6700	-2.0300**	-2.3160	-1.7260
β_2	0.3994**	0.2280	0.5716	0.6624**	0.4918	0.8425	0.6986**	0.4541	0.9593	0.6689**	0.4388	0.8969
β_3	0.0840	-0.0035	0.1710	0.1592**	0.0732	0.2528	0.1770**	0.0348	0.3249	0.1789**	0.0458	0.3116
β_4	-0.1126**	-0.1866	-0.0378	-0.0840**	-0.1544	-0.01054	-0.0426	-0.1531	0.0775	-0.0434	-0.1596	0.0660
β_5	0.8174**	0.6350	1.0010	0.2866**	0.1075	0.4723	0.3260**	0.1053	0.5471	0.3238**	0.1309	0.5117
β_6	1.3460**	1.0680	1.6230	0.3642	-0.0175	0.7385	0.3266	-0.1443	0.8229	0.3183	-0.1267	0.7550
β_7	0.7749**	0.5384	1.0100	0.3827**	0.1461	0.6102	0.4281**	0.1329	0.7140	0.4145**	0.1456	0.6857
β_8	5.4820**	5.1890	5.7720	1.8810**	1.4520	2.307	3.0000**	2.3590	3.6320	1.9650**	1.5390	2.4380
β_9	5.5310**	5.2510	5.8090	1.8750**	1.4060	2.336	2.9490**	2.2700	3.6040	1.9410**	1.4710	2.4410
β_{10}	4.6430**	4.3560	4.9400	1.0660**	0.5863	1.542	2.1440**	1.4130	2.8200	1.2180**	0.7452	1.8970
β_{11}	5.2440**	4.9850	5.5080	1.6130**	1.1740	2.055	2.7530**	2.0590	3.3990	1.6860**	1.2500	2.2370
β_{12}	4.8890**	4.6030	5.1780	1.2190**	0.7631	1.664	2.0630**	1.4360	2.7120	1.1660**	0.7546	1.6180
β_{13}	4.7450**	4.4860	5.0030	1.2220**	0.7606	1.66	2.2390**	1.6160	2.8630	1.2890**	0.8521	1.7750
β_{14}	3.9170**	3.6460	4.1920	0.3399	-0.0134	0.7106	0.7623**	0.2296	1.3440	0.3282	-0.0513	0.7482
β_{15}	0.3534**	0.2861	0.4217	0.2035**	0.1438	0.2593	0.3555**	0.1851	0.5983	0.3346**	0.1699	0.5272
β_{16}	0.0294	-0.0446	0.1030	0.0821**	0.0203	0.1423	0.1028**	0.0224	0.1854	0.1107**	0.0319	0.1922
β_{17}	0.1459**	0.0673	0.2236	0.1885**	0.1267	0.2479	0.1817**	0.0954	0.2671	0.1917**	0.1082	0.2744
β_{18}	0.5982**	0.5157	0.6817	0.5375**	0.4677	0.6078	0.5817**	0.4776	0.6890	0.5805**	0.4800	0.6807
β_{19}	-0.8837**	-1.0150	-0.7523	-0.3172**	-0.4294	-0.2115	-0.3203**	-0.4458	-0.1832	-0.2986**	-0.4270	-0.1720
σ	1.7840	1.7330	1.8340	1.2450	1.2070	1.284	1.0180	0.9762	1.0630	1.0240	0.9839	1.0690
σ_η				3.4840	3.1000	3.916	2.7250	2.2540	3.2330	3.3400	2.9160	3.7910
DIC			39220			37730			37320			37280

First, the user base ($\beta_{20} = 0.6312$) obviously has significantly positive influence on the volume of online WOM, which is consistent with the results of Duan, Gu, and Whinston(2008), i.e., the larger the pool of user base, the more WOM will be generated. Second, the negative influence of price on volume of online WOM became smaller. Additionally, after controlling for both observed and unobserved heterogeneity of developers, price ($\beta_1 = 0.0979$) does not have significant negative impact on online WOM anymore. In this sense, free applications may attract many people to install but not necessarily generate more online WOM. Third, the influence of file size ($\beta_4 = 0.1791$) became positive in this model, which is consistent with our expectation. The larger the file size, the more interesting it is. Based on previous research, consumers are more likely express their opinions for interesting products. Fourth, the influence of number of applications ($\alpha_1 = -0.1538$) developed by the developer turns out to significantly reduce volume of online WOM. In this sense, developers' reputation and experience suppress users' intention to spread their WOM online. However, users are more likely to talk about those applications developed by developers with less reputation or public visibility.

The finding that the more users the application has, the more posts it will receive, seems to be supported at this point. However, it cannot fully account for the unique attribute of our dataset, i.e., the volume of online WOM comes from the actual user of the applications. Therefore, we are more interesting the probability that the user post their opinions online after they experience the product. We therefore, take the probability of posting opinions online as outcome and implement the lognormal regression in Bayesian paradigm.

Table III-5 Posterior means of the Lognormal Model with downloads included as an IV

	Lognormal Regression			With Random Intercept			With Random Coefficient			
	mean	val2.5pc	val97.5pc	mean	val2.5pc	val97.5pc	mean	val2.5pc	val97.5pc	σ_β
α_1	-0.1538**	-0.2075	-0.1009	-0.2041**	-0.3066	-0.1051	-0.7231**	-0.9423	-0.4960	
α_2	0.7168**	0.6709	0.7630	0.6657**	0.5901	0.7366	0.3116**	0.1526	0.4658	
α_3	0.2242**	0.1790	0.2702	0.2613**	0.1953	0.3273	0.2880**	0.1245	0.4526	
β_1	-0.2136**	-0.3335	-0.0897	-0.1790**	-0.2999	-0.0559	-0.0657	-0.2316	0.0979	0.3957
β_2	0.1203**	0.0905	0.1496	0.0740**	0.0447	0.1033	0.2787**	0.1050	0.6112	0.3761
β_3	0.0439**	0.0075	0.0805	0.0789**	0.0363	0.1216	0.0955**	0.0195	0.1673	0.2888
β_4	0.1791**	0.1053	0.2527	0.1666**	0.0853	0.2508	0.1020	-0.0203	0.2330	0.3529
β_5	0.0166	-0.0148	0.0479	0.0120	-0.0226	0.0472	0.0596	-0.0093	0.1336	0.2421
β_6	0.0601**	0.0288	0.0910	0.0498**	0.0188	0.0813	0.0469	-0.0003	0.0913	0.1966
β_7	0.0699**	0.0372	0.1033	0.0618**	0.0303	0.0936	0.0514**	0.0071	0.0979	0.1936
β_8	0.2786**	0.2433	0.3146	0.2215**	0.1851	0.2578	0.1704**	0.1112	0.2263	0.2092
β_9	0.0865**	0.0274	0.1448	0.0650**	0.0067	0.1226	0.0598	-0.0115	0.1379	0.2258
β_{10}	0.0614	-0.0189	0.1405	0.0040	-0.0857	0.0912	0.0011	-0.1260	0.1194	0.2956
β_{11}	0.1206**	0.0189	0.2201	0.0534	-0.0563	0.1648	0.0355	-0.1219	0.1945	0.3650
β_{12}	0.1907**	0.0699	0.3099	0.0065	-0.1639	0.1772	-0.0476	-0.3189	0.2180	0.4422
β_{13}	-0.1857**	-0.3440	-0.0246	-0.4139**	-0.5805	-0.2380	-0.5066**	-0.7519	-0.2561	0.3493
β_{14}	-0.3025**	-0.4667	-0.1426	-0.5213**	-0.6933	-0.3418	-0.6775**	-0.9358	-0.4046	0.3986
β_{15}	-0.3187**	-0.4706	-0.1723	-0.5413**	-0.7051	-0.3703	-0.6671**	-0.9251	-0.4198	0.3641
β_{16}	-0.5344**	-0.6919	-0.3755	-0.7327**	-0.9188	-0.5458	-0.8719**	-1.1550	-0.5773	0.4865
β_{17}	-0.6706**	-0.8281	-0.5115	-0.8340**	-1.0060	-0.6598	-0.8559**	-1.0870	-0.6115	0.3758
β_{18}	-0.7406**	-0.8883	-0.5937	-0.8925**	-1.0550	-0.7260	-0.9267**	-1.1520	-0.6913	0.3919
β_{19}	-0.6153**	-0.7571	-0.4753	-0.5446**	-0.6937	-0.3947	-0.5144**	-0.7589	-0.2955	0.4318
β_{20}	0.6312**	0.6194	0.6426	0.6612**	0.6483	0.6743	0.6950**	0.6720	0.7168	0.0922
σ	0.7557	0.7346	0.7770	0.6741	0.6540	0.6946	0.5419	0.5190	0.5640	
σ_η				0.3815	0.3309	0.4371	0.3262	0.2396	0.4431	
DIC		35060			34290			33430		

Table III-6 reports the posterior mean estimates of this model. By considering the effect of total downloads in this way, the results turn out to be of much difference from that of simply implementing on the total number of review counts(table III-4) and adding user base as an explanatory variable (table III-5). The DIC score has been largely reduced for this model. Many important findings were demonstrated in table III-6.

First, after controlling for developer heterogeneity, price ($\beta_{1,2} = 1.0140$) has significant positive influence on total review counts, holding downloads constant. Compared to the effect of price of simply implementing on total review counts, we can see the effect of price has been totally reversed. Price in this model, on the contrary of model 1, will significantly increase the probability of users posting their reviews online. In other words, users are more likely to talk about those applications that they actually bought from the store instead of free applications. Although applications having larger installation base will have more volume of online WOM based on results of model 2, in this sense, we argue that the freemium business model will not boost installation base through online WOM, but on the contrary, it boosts volume of online WOM through installation base. Results of model 3 further confirm this proposition. From the results of table III-6, we can see that users who download the applications are more likely to talk about paid applications instead of free ones. The result supports the idea that price determines the evaluation effort invested by consumers. A price premium can stimulate consumers to revisit their perception of benefit relevance, and thus, spread their opinions online (Wathieu and Bertini, 2007). Effects of other factors also are found different from that of model 1.

Table III-6 Posterior means of the Lognormal density regression

	Lognormal regression			random intercept			random coefficient - uncorrelated			random coefficient - correlated		
	mean	2.50%	97.50%	mean	2.50%	97.50%	mean	2.50%	97.50%	mean	2.50%	97.50%
α_1	-0.3287**	-0.4115	-0.2434	-3.4640**	-3.9520	-2.9820	-3.1450**	-3.6300	-2.6800	-3.3550**	-3.8370	-2.8890
α_2	0.2053**	0.1377	0.2719	-1.0790**	-1.4500	-0.7067	-0.9476**	-1.3050	-0.6028	-1.0460**	-1.4040	-0.6954
α_3	0.3400**	0.2685	0.4116	-0.0109	-0.3784	0.3558	0.0946	-0.2425	0.4357	0.0342	-0.3259	0.3902
β_1	-0.1508	-0.3454	0.0452	1.0140**	0.8830	1.1420	0.9487**	0.7669	1.1240	0.9318**	0.7602	1.1130
β_2	0.0468	-0.0696	0.1636	0.0643	-0.0324	0.1595	-0.0234	-0.1515	0.0980	-0.0009	-0.1214	0.1208
β_3	0.0217	-0.0378	0.0807	0.0744**	0.0240	0.1250	0.0828**	0.0013	0.1627	0.0826**	0.0060	0.1592
β_4	0.0925**	0.0422	0.1432	0.0672**	0.0285	0.1072	0.1105**	0.0282	0.1973	0.1112**	0.0317	0.1937
β_5	-0.3738**	-0.4976	-0.2490	0.0239	-0.0760	0.1277	-0.0490	-0.1874	0.0865	-0.0214	-0.1481	0.1025
β_6	-0.4832**	-0.6720	-0.2955	0.0007	-0.2173	0.2132	-0.1058	-0.4186	0.2082	-0.0475	-0.3165	0.2336
β_7	-0.2589**	-0.4194	-0.0997	0.0654	-0.0666	0.1958	0.0092	-0.1691	0.1795	0.0477	-0.1144	0.2162
β_8	-3.5040**	-3.7030	-3.3070	-0.4824**	-0.7126	-0.2506	-0.8751**	-1.1810	-0.5663	-0.6047**	-0.8657	-0.3429
β_9	-3.7140**	-3.9040	-3.5250	-0.5623**	-0.8030	-0.3274	-1.0270**	-1.3640	-0.6897	-0.7419**	-1.0340	-0.4710
β_{10}	-3.5630**	-3.7580	-3.3620	-0.6023**	-0.8723	-0.3204	-1.0160**	-1.4450	-0.6191	-0.7024**	-1.0270	-0.3717
β_{11}	-3.5740**	-3.7500	-3.3950	-0.4715**	-0.7063	-0.2405	-0.9240**	-1.2460	-0.5993	-0.6378**	-0.9234	-0.3671
β_{12}	-3.9240**	-4.1170	-3.7270	-0.7480**	-0.9830	-0.5122	-1.0500**	-1.3550	-0.7340	-0.8121**	-1.0760	-0.5493
β_{13}	-3.9490**	-4.1240	-3.7740	-0.7134**	-0.9452	-0.4847	-1.0780**	-1.3820	-0.7618	-0.8118**	-1.0740	-0.5426
β_{14}	-3.2640**	-3.4480	-3.0770	-0.2083**	-0.3966	-0.0257	-0.3847**	-0.6489	-0.1235	-0.2654**	-0.4927	-0.0547
β_{15}	-0.0151	-0.0608	0.0313	0.0036	-0.0270	0.0352	0.1937**	0.0462	0.4325	0.1563**	0.0268	0.3072
β_{16}	0.0775**	0.0273	0.1274	0.0391**	0.0056	0.0724	0.0366	-0.0112	0.0850	0.0373	-0.0102	0.0854
β_{17}	0.0261	-0.0272	0.0789	-0.0172	-0.0509	0.0170	-0.0098	-0.0605	0.0395	-0.0107	-0.0577	0.0366
β_{18}	0.0919**	0.0359	0.1485	0.0417**	0.0036	0.0789	-0.0125	-0.0722	0.0461	-0.0045	-0.0617	0.0521
β_{19}	0.6564**	0.5675	0.7455	0.1569**	0.0972	0.2175	0.1582**	0.0811	0.2439	0.1727**	0.0927	0.2562
τ	0.6830	0.6452	0.7228	2.1110	1.9830	2.2420	2.9490	2.7320	3.1850	2.9110	2.7000	3.1320
τ_η				0.1230	0.0996	0.1490	0.1572	0.1232	0.1955	0.1369	0.1097	0.1688
DIC			-11860			-14350			-14560			-14590

Second, the number of screenshot ($\beta_{32} = 0.0744$), average rating ($\beta_{15,3} = 0.1937$), and file size($\beta_{42} = 0.0672$) have significantly POSITIVE influence on the probability of users posting their reviews, which are consistent with our expectation.

Third, after controlling for developer heterogeneity, number of applications developed by the developer ($\alpha_{11} = -0.3287$) will significantly decrease users' probability of posting their reviews online. The average review counts($\alpha_{21} = 0.2053$) of the applications developed by the developer will significantly increase users' probability of posting their reviews online.

Fourth, applications with high maturity content will not help to significantly boost users' probability of expressing their opinions online after controlling for developer heterogeneity.

Fifth, after controlling for developer heterogeneity, average price of applications featured as "installed by other users" ($\beta_{19,3} = 0.1582$) has significantly POSITIVE influence on total review counts. However, the effect of three competition variables, i.e., average counts and price of applications featured as "viewed by other users" and average counts of applications featured as "installed by other users" become not significant.

Sixth, the effect of average rating remains positive and significant after controlling for developers' heterogeneity. In this sense, applications with higher average rating are more likely to be talked by users. Therefore, it is of much importance for us to explore why users will give positive ratings. Since the more positive ratings the application receives, the higher the average rating. In addition, based on previous research, negative

WOM has more impact than that of positive WOM. Therefore, examining the influence of negative volume of online WOM is even more important. Table III-7 and table III-8 report the results of the effects of factors on both negative and positive volume of online WOM.

Table III-7 Posterior means of the model on log(negative counts/downloads)

	Lognormal regression			random intercept			random coefficient - uncorrelated			random coefficient - correlated		
	mean	2.50%	97.50%	mean	2.50%	97.50%	mean	2.50%	97.50%	mean	2.50%	97.50%
α_1	-0.3298**	-0.4350	-0.2214	-5.1370**	-5.8330	-4.4470	-4.5940**	-5.2370	-3.9570	-4.7530**	-5.4210	-4.1080
α_2	-0.1403**	-0.2262	-0.0554	-2.1040**	-2.6510	-1.5800	-1.7470**	-2.2510	-1.2580	-1.8360**	-2.3510	-1.3380
α_3	0.4248**	0.3336	0.5159	-0.1648	-0.7072	0.3777	-0.0642	-0.5491	0.4222	-0.0737	-0.5678	0.4136
β_1	-0.3569**	-0.6039	-0.1078	1.3900**	1.2500	1.5300	1.2610**	1.0670	1.4520	1.2280**	1.0430	1.4130
β_2	-0.1999**	-0.3478	-0.0513	-0.1440**	-0.2464	-0.0384	-0.1262**	-0.2520	-0.0025	-0.1160	-0.2337	0.0049
β_3	-0.1060**	-0.1816	-0.0309	0.0143	-0.0381	0.0689	0.0173	-0.0604	0.0976	0.0150	-0.0605	0.0865
β_4	0.1242**	0.0603	0.1887	0.0515**	0.0078	0.0943	0.1096**	0.0271	0.1979	0.1087**	0.0272	0.1893
β_5	-0.5250**	-0.6822	-0.3664	0.0494	-0.0565	0.1574	-0.0353	-0.1695	0.1005	-0.0165	-0.1476	0.1104
β_6	-1.0470**	-1.2870	-0.8079	0.0508	-0.1787	0.2776	-0.0441	-0.3499	0.2598	-0.0127	-0.2859	0.2897
β_7	-0.4323**	-0.6363	-0.2297	0.1315	-0.0027	0.2688	0.0486	-0.1300	0.2297	0.0722	-0.1035	0.2450
β_8	-5.1380**	-5.3900	-4.8870	-0.5530**	-0.7877	-0.3131	-0.7396**	-1.0220	-0.4475	-0.5919**	-0.9166	-0.2998
β_9	-5.7060**	-5.9480	-5.4660	-0.7985**	-1.0600	-0.5486	-1.0120**	-1.3190	-0.7077	-0.8595**	-1.2060	-0.5659
β_{10}	-5.1340**	-5.3820	-4.8780	-0.6252**	-0.8939	-0.3625	-0.7869**	-1.1840	-0.4103	-0.6362**	-1.0060	-0.2914
β_{11}	-5.1660**	-5.3900	-4.9380	-0.4812**	-0.7225	-0.2427	-0.7336**	-1.0310	-0.4251	-0.5790**	-0.9115	-0.2753
β_{12}	-5.2280**	-5.4740	-4.9790	-0.5630**	-0.8108	-0.3244	-0.7456**	-1.0250	-0.4523	-0.6362**	-0.9422	-0.3355
β_{13}	-5.2850**	-5.5080	-5.0630	-0.5830**	-0.8361	-0.3474	-0.7937**	-1.0630	-0.4982	-0.6508**	-0.9568	-0.3403
β_{14}	-4.7230**	-4.9570	-4.4860	-0.1783	-0.3764	0.0261	-0.1820	-0.4335	0.0760	-0.1691	-0.4844	0.0813
β_{15}	-0.2511**	-0.3092	-0.1921	-0.1601**	-0.1937	-0.1267	-3.0760**	-3.7540	-2.4280	-2.5280**	-3.1370	-1.9420
β_{16}	0.0741**	0.0103	0.1376	0.0126	-0.0243	0.0494	0.0089	-0.0393	0.0580	0.0085	-0.0392	0.0579
β_{17}	0.0208	-0.0470	0.0878	-0.0252	-0.0614	0.0104	-0.0072	-0.0580	0.0415	-0.0080	-0.0549	0.0370
β_{18}	0.0760**	0.0048	0.1480	-0.0018	-0.0426	0.0392	-0.0390	-0.0996	0.0223	-0.0307	-0.0880	0.0289
β_{19}	0.9174**	0.8044	1.0310	0.1891**	0.1241	0.2516	0.1506**	0.0730	0.2365	0.1632**	0.0833	0.2492
σ	1.5390	1.4950	1.5830	0.7324	0.7107	0.7549	0.5717	0.5491	0.5938	0.5782	0.5565	0.6021
σ_η				4.2370	3.8660	4.6630	3.6920	3.3460	4.0780	3.8690	3.4840	4.2780
DIC			-20120			-23470			-24020			-24030

Table III-7 reports the posterior mean estimates of implementing the models on ratio of total NEGATIVE review counts over lower bound of downloads.

First, average rating ($\beta_{15} = -0.2511$) will significantly decrease the total number of negative review counts, even after controlling developer heterogeneity. Number of applications developed by the developer ($\alpha_1 = -0.3289$), average counts of applications developed by the developer ($\alpha_2 = -0.1403$), and having YouTube video ($\beta_2 = -0.1999$), will also significantly decrease negative review counts.

Second, after controlling for developer heterogeneity, price ($\beta_{1,2} = 1.3900$), file size ($\beta_{4,2} = 0.0515$), and average price of applications featured as "installed by other users" ($\beta_{19,2} = 0.1891$), will significantly increase the number of negative review counts.

Table III-8 reports the posterior mean estimates of implementing the models on ratio of total POSITIVE review counts over lower bound of downloads.

First, price ($\beta_1 = 1.0420$), average rating ($\beta_{15,3} = 2.7900$), number of screenshots ($\beta_3 = 0.0834$), and file size ($\beta_4 = 0.0898$), and average price of applications featured as "installed by other users"($\beta_{19,3} = 0.1431$), will significantly increase the total number of positive review counts, after controlling developer heterogeneity.

Second, the number of applications developed by the developer ($\alpha_1 = -3.8470$) and average counts of applications developed by the developer ($\alpha_2 = -1.2740$) will significantly decrease positive review counts, even after controlling for developer heterogeneity.

Table III-8 Posterior means of the model on log(positive counts/downloads)

	Lognormal regression			random intercept			random coefficient - uncorrelated			random coefficient - correlated		
	mean	2.50%	97.50%	mean	2.50%	97.50%	mean	2.50%	97.50%	mean	2.50%	97.50%
α_1	-0.3616**	-0.4536	-0.2669	-3.8470**	-4.4010	-3.3370	-4.1380**	-4.7260	-3.5590	-4.1740**	-4.7670	-3.6110
α_2	0.1500**	0.0749	0.2241	-1.2740**	-1.6750	-0.8842	-1.4850**	-1.9360	-1.0520	-1.4900**	-1.9500	-1.0470
α_3	0.3280**	0.2482	0.4076	-0.0702	-0.4938	0.3508	-0.0018	-0.4340	0.4279	-0.0307	-0.4680	0.3908
β_1	-0.2540**	-0.4699	-0.0363	1.0420**	0.9008	1.1890	1.1130**	0.9241	1.2980	1.1240**	0.9270	1.3170
β_2	0.1081	-0.0211	0.2379	0.0963	-0.0087	0.2060	-0.0350	-0.1579	0.0914	-0.0251	-0.1538	0.1177
β_3	0.0649	-0.0012	0.1305	0.0834**	0.0299	0.1409	0.0595	-0.0197	0.1407	0.0665	-0.0152	0.1410
β_4	0.1107**	0.0549	0.1671	0.0898**	0.0465	0.1347	0.1347**	0.0483	0.2289	0.1382**	0.0485	0.2270
β_5	-0.4185**	-0.5559	-0.2799	0.0088	-0.1013	0.1227	-0.0248	-0.1623	0.1133	-0.0092	-0.1392	0.1184
β_6	-0.6169**	-0.8268	-0.4084	-0.0753	-0.3097	0.1546	-0.0722	-0.3775	0.2436	-0.0166	-0.3237	0.2855
β_7	-0.3365**	-0.5148	-0.1595	0.0203	-0.1235	0.1621	0.0481	-0.1242	0.2198	0.0680	-0.0980	0.2331
β_8	-3.8380**	-4.0580	-3.6190	-0.4835**	-0.7367	-0.2307	-0.7306**	-1.0260	-0.4386	-0.5674**	-0.9110	-0.2465
β_9	-4.1390**	-4.3500	-3.9290	-0.5965**	-0.8699	-0.3302	-0.8991**	-1.2210	-0.5882	-0.7563**	-1.1310	-0.4130
β_{10}	-3.8300**	-4.0460	-3.6060	-0.5880**	-0.8715	-0.3139	-0.8224**	-1.2240	-0.4421	-0.6502**	-1.1030	-0.2705
β_{11}	-3.9400**	-4.1360	-3.7410	-0.4858**	-0.7409	-0.2354	-0.7501**	-1.0540	-0.4454	-0.6354**	-0.9726	-0.3119
β_{12}	-4.3570**	-4.5730	-4.1400	-0.8015**	-1.0690	-0.5507	-0.8796**	-1.1630	-0.5880	-0.7780**	-1.1170	-0.4581
β_{13}	-4.3320**	-4.5270	-4.1380	-0.7409**	-1.0090	-0.4942	-0.8931**	-1.1780	-0.5962	-0.7741**	-1.1280	-0.4478
β_{14}	-3.5190**	-3.7230	-3.3110	-0.2174	-0.4267	0.0011	-0.2647**	-0.5121	-0.0175	-0.2441	-0.5052	0.0010
β_{15}	-0.0647**	-0.1155	-0.0132	-0.0124	-0.0489	0.0220	2.7900**	2.1770	3.4340	2.3460**	1.8090	2.9100
β_{16}	0.0830**	0.0273	0.1385	0.0437**	0.0058	0.0804	0.0263	-0.0227	0.0759	0.0286	-0.0191	0.0798
β_{17}	0.0095	-0.0498	0.0680	-0.0272	-0.0645	0.0093	-0.0145	-0.0647	0.0333	-0.0160	-0.0636	0.0304
β_{18}	0.1108**	0.0486	0.1737	0.0621**	0.0197	0.1052	-0.0148	-0.0754	0.0468	-0.0024	-0.0619	0.0564
β_{19}	0.7242**	0.6255	0.8233	0.1844**	0.1167	0.2491	0.1431**	0.0636	0.2291	0.1535**	0.0710	0.2453
σ	1.3450	1.3070	1.3830	0.7598	0.7371	0.7831	0.5741	0.5519	0.5961	0.5874	0.5651	0.6103
σ_η				3.1710	2.8730	3.4960	3.2570	2.9360	3.6110	3.3210	2.9490	3.7070
DIC			-13510			-16030			-16740			-16700

3.6 Conclusion

3.6.1 A Summary of Results

In this study, we developed models to test the effects of price on the volume of online consumer reviews for mobile applications. We first implement the model on the total volume of online WOM (model 1). Results show that free applications will significantly increase the volume of online WOM, which, to some extent, confirms the practitioner assumption that the Freemium business model will help marketers gain online WOM. However, our further exploration proved that this effect is spurious after considering the influence of the size of the user base.

More than half of the mobile applications in the marketplace are free giveaways. In some cases, developers may release both free and paid versions with different features. Marketers of mobile applications offer free giveaways in the hope of increasing sales through boosting online WOM. The assumption here is that the more user bases that a product acquires, the more value that users perceive. This should increase the volume of online WOM and therefore increase acquisition of new customers. We test this idea in our second model by adding the variable of user base as an explanatory variable.

However, simply treating user base as an explanatory variable cannot exclude the possibility that current users may attract more future users. Given the unique attributes of our dataset that the ratings and reviews of each application were posted by actual users, we empirically implement a third model. In this model, the outcome is not the absolute volume of online WOM. Instead, we explore what factors influence the probability of users posting their reviews after they actually experience the product. Our results suggest that Freemium business model does not help to boost online WOM at all. Instead, users who adopt paid applications are most likely to express their opinions online. The explanation is that users downloading paid applications may put more effort on evaluating the potential usage value of the benefit (Wathieu and Bertini, 2007) and thus are more engaged in behaviors such as spreading WOM online.

In addition to this important finding, this study reveals the influence of many other factors on users' probability of posting their WOM online. First, mobile applications of larger file size will be more likely to be talked about. Second, number of screenshots, as an advertising cue, helps to boost users' probability of posting their ratings and reviews online after accounting for unobserved heterogeneity of developers. Third, mobile applications with more mature content will less likely to stimulate users to spread their WOM online. Fourth, users will be more likely to talk about mobile applications with higher average rating after accounting for unobserved heterogeneity of developers. Fifth, competition from similar applications actually seems not so important after accounting for unobserved heterogeneity of developers. Finally, however, developers' experience is important. An interesting point found in our research is that the more applications developed by the developer, the less likely that the focal applications to be talked about. In other words, public visibility in this sense does not help to simulate online WOM. But on the contrary, users like to talk about the applications developed by unknown developers, which is consistent with the idea of Hughes (2005) that unusual, outrageous, or remarkable things generate conversation. People love to talk about things that are different and surprising (Rosen, 2009; Knox 2010; Nulman, 2009).

Moreover, given the positive influence of average rating on volume, it is important to examine the key factors to generate both positive and negative volume of online WOM. There are many important findings. First, low average rating stimulates negative online WOM. On the positive side, however, higher average rating will simulate more positive volume of online WOM. Second, having YouTube Video will decrease the number of negative WOM.

3.6.2 Managerial Implications

The findings of the present study have important managerial implications for mobile applications and other digital products, such as online blogs, online videos, etc. In the marketplace of digital products, since online WOM has dominant influence on boosting user bases, it is important for marketers to stimulate users spreading their WOM in order to acquire more users. However, an intuitive assumption in the marketplace is that free giveaways may attract more users to install the applications and therefore increase online WOM. Our empirical results support this assumption. However, a further question faced by current developers is that of how to monetize their products. The findings suggest that free mobile applications are double-edged swords. On the one hand, they may help the developers gain users and increase online WOM. On the other hand, free mobile applications may cannibalize the market share of paid application and thus hurt the profitability of this product. Our research, however, shows that among those users who download the applications, users who actually spend money to buy the application will be more likely to spread WOM. To this point, although free giveaways may help to boost installation base, it may not necessarily create interests among users. Users are more likely to discuss apps that they actually paid for. To generate more online WOM, developers have to make their applications more interesting. In addition, applications with less mature content will help to boost online WOM. High average ratings will definitely increase the volume of online WOM. But the influence of average rating is extreme. Higher average rating will increase later on positive volume of online WOM, but lower average rating will increase later on negative volume of online WOM.

3.6.3 Limitations and Future Directions

As a new industry, in just 4 years, mobile apps have overtaken the web and are beginning to challenge television, the top media channel. Yet, academic research hasn't caught up with this domain. In this hyper-competitive landscape with a wide range of choices, mobile apps consumption of mobile apps gives rise to a whole host of research areas that straddle both marketing and technologies. In our study, we explore the issue that what factors generate online WOM. However, an important mechanism which is probably a bigger driver of app discovery and usage is offline WOM. To understand what factors influence these face-to-face mechanisms are important because many consumers instantly download applications that their friends/acquaintances are using when this discovery happens face-to-face as a friend can convey why he or she likes the application and uses it. In addition, for apps with strong network effects, knowing what friends and family members are using has a greater influence on the eventual choices. Therefore, an important area for future research is to not only understand the varying effects of reviews on apps sales but also the review generation process that can determine how may reviews an app gets.

In our current study, we only account for both observed and unobserved heterogeneity of mobile application developers. Another type of heterogeneity may be even more important in terms of online WOM – the user characteristics. Users differ in their ability and motivations to use applications in the ways they are intended, and as a consequence, have varying perceptions of app features. Mobile apps developers, therefore, must understand the characteristics of users to improve the design and other features of their products.

3.7 References of Essay II

Bearden and Etzel (1982), "Reference Group Influence on Product and Brand Purchase Decisions", *Journal of Consumer Research*, Vol. 9, No. 2(Sep, 1982), pp. 183-194

Berger and Schwartz (2011), "What Drives Immediate and Ongoing WOM?", *Journal of Marketing Research*, vol. XLVIII (October 2011), 869-880

Brakus, Bernd and Zarantonello (2009), "Brand Experience: What is it? How is it measured? Does it Affect Loyalty?" *Journal of Marketing*, Vol.73 (May 2009), 52-68

Chevalier and Mayzlin (2006), "The Effect of WOM on Sales: Online Book Reviews", *Journal of Marketing Research*, 43(August), 345-54.

Cialdini (2001), "Harnessing the Science of Persuasion", *Harvard Business Review*, 79(October), 72-79

Cadotte, Ernest R. Robert B. Woodruff and Roger L. Jenkins. (1987). "Expectations and Norms in Models of Consumer Satisfaction." *Journal of Marketing Research*, 24, August, 305-314.

Churchill and Surprenant (1982), "An Investigation Into the Determinants of Customer Satisfaction", *Journal of Marketing Research*, Vol. XIX (November 1982), 491-504

Curren, Mary T. and Valerie S. Folkes (1987), "Attributional Influences on Consumers' Desires to Communicate About Products", *Psychology and Marketing*. 4 (Spring), 31-45.

Chen, Y., and Xie, J. (2008). "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix," *Management Science* (54:3), pp. 477-490.

Duan, Gu, and Whinston (2008), "The Dynamics of Online WOM and Product Sales – An Empirical Investigation of the Movie Industry", *Journal of Retailing*, 84(2, 2008) 233-242

Dellarocas, C, Awad, N. F., and Zhang, X. (2004). "Exploring the Value of Online Product Ratings in Revenue Forecasting: The Case of Motion Pictures," *Proceedings of the 25th International Conference on Information Systems*, Washington, DC, December 12-15, pp. 379-386.

Dawar and Parker (1994), "Marketing Universals: Consumers' User of Brand Name, Price, Physical Appearance, and Retailer Reputation as Signals of Product Quality", *Journal of Marketing*, April

Feng and Papatla (2011), Advertising: Stimulant or Suppressant of Online Word of Mouth? *Journal of Interactive Marketing*, 25 (2) (2011), pp. 75–84

Gedenk and Neslin (1999), "The Role of Retail Promotion in Determining Future Brand Loyalty: Its effect on Purchase Event Feedback", *Journal of Retailing*, 75(4), 433-459

Gerstner (1985), “Do Higher Prices Signal Higher Qualities?” *Journal of Marketing Research*, 22(May): 209-215

Grewal ,Krishnan, Baker, and Borin (1998), “The effect of store name, brand name and price discounts on consumers' evaluations and purchase intentions”, *Journal of Retailing*, Volume 74, Issue 3, Autumn 1998, Pages 331–352

Godes and Mayzlin (2004), “Using Online Conversations to Study WOM Communication”, *Marketing Science*, 23(4), 545-60

Higgins, Rholes, and Jones (1977), “Category Accessibility and Impression Formation”, *Journal of Social Psychology*, 13(March), 141-54

Hoch and Ha (1986), “Consumer Learning: Advertising and the Ambiguity of Product Experience”, *Journal of Consumer Research*, 13(2), 221-33.

Holmes and Lett (1977), “Product sampling and word of mouth”, *Journal of Advertising Research*, Vol 17(5), Oct 1977, 35-40.

Hamm and Cundiff (1969), “Self-Actualization and Product Perception”, *Journal of Marketing Research*. 6 (November), 470-472

Keller (1987), “Memory Factors in Advertising: The Effect of Advertising Retrieval Cues on Brand Evaluations”, *Journal of Consumer Research*, Vol.14, December (1987)

Kirmani, A., and Rao, A. R (2000) “No Pain, No Gain: A Critical Review of the Literature on Signaling Unobservable Product Quality”, *Journal of Marketing* (64:2), pp. 66-79.

Klein (1998), “Evaluating the Potential of Interactive Media through a New Lens: Search versus. Experience Goods,” *Journal of Business Research*, 41: 195-203

King and Balasubramanian (1994), “The Effects of Expertise, End Goal, and Product Type on Adoption of Preference Formation Strategy”, *Journal of the Academy of Marketing Science*, Vol. 22, No. 2, Pages 146-159

Lynch and Scrull (1982), “Memory and Attentional Factors in Consumer Choice: Concepts and Research Methods”, *Journal of Consumer Research*, 9(1), 18-37

Li and Hitt (2008), “Self-Selection and Information Role of Online Product Review”, *Information Systems Research*, Forthcoming

Lutz (1975), “Changing Brand Attitudes through Modification of Cognitive Structure”, *Journal of Consumer res*earch, vol. 1, 1975, page 45

Li and Hitt (2010), “Price Effects in Online Product Reviews: An Analytical Model and Empirical Analysis”, *MIS Quarterly*, 2010, Vol. 34 No. 4 pp. 809-831 / December 2010.

Lal, R. and M. Sarvary (1999), "When and How Is the Internet Likely to Decrease Price Competition?" *Marketing Science* 18 (4): 485-503.

Mahajan, Muller, and Wind (2000), "New Product Diffusion Models", Boston, MA: Kluwer Academic Publisher.

Marks and Kamins (1988), The Use of Product Sampling and Advertising: Effects of Sequence of Exposure and Degree of Advertising Claim Exaggeration on Consumers' Belief Strength, Belief Confidence, and Attitudes, *Journal of Marketing Research*, Vol. 25, No. 3 (Aug., 1988), pp. 266-281

Mitra, A (1995), "Price Cue Utilization in Product Evaluations: The Moderating Role of Motivation and Attribute Information," *Journal of Business Research* (33), pp. 187-195.

Park and Lee (2009), "Information Direction, Website Reputation and eWOM effect: a moderating role of product type", *Journal of Business Research*, 2009; 61-7.

Racherla, Furner, and Babb (2012), "Conceptualizing the Implications of Mobile App Usage and Stickiness: A research Agenda", electronic version available at: http://ssrn.com/abstract=2187056

Redmond (2013), "Three Modes of Competition in the Marketplace", The American *Journal of Economics and Sociology*, Vol. 72, No. 2 (April 2013)

Rao, A. R., and Monroe, K. B (1989), "The Effect of Price, Brand Name, and Store Name on Buyers' Perceptions of Product Quality: An Integrative Review," *Journal of Marketing Research* (25:3), pp. 351-357.

Richins, M. L (1983), "Negative Word-of-Mouth by Dissatisfied Consumers: A Pilot Study," *Journal of Marketing* (47:1), pp. 68-78.

Richins, Marsha L. (1983b), "Negative-Word-of-Mouth By Dissatisfied Consumers," *Journal of Marketing*, 47 (Winter), 68-78.

Rao, A. R., and Monroe, K. B (1988), "The Moderating Effect of Prior Knowledge on Cue Utilization in Product Evaluations," *The Journal of Consumer Research* (15:2), pp. 253-264.Voss et al(1998)

Yi (1990), "A critical Review of Consumer Satisfaction", Review of Marketing, 1990

Spreng R. A, MacKenzie, S. B., and Olshavsky, R. W (1996), "A Reexamination of the Determinants of Consumer Satisfaction*", Journal of Marketing*, (60:3), pp. 15-32.

Sundaram, D.S., Mitra, K., & Webster, C (1998), "Word-of-Mouth Communications: A Motivational Analysis", *Advances in Consumer Research*, 25, 527–531.

Senecal and Nantel (2004), "The Influence of Online Product Recommendations on Consumers' online choices", *Journal of Retailing*, 2004; 80:159-169

Wathieu and Bertini (2007), “Price as a Stimulus to Think: The Case for Willful Overpricing”, *Marketing Science*, 26(1, January-February), 118-129

Wangenheim, F.V. and Bayon, T. (2002), “The Economic Value of Customer Word of Mouth Referrals: Conceptualization and Empirical Results”, IU School of Business Administration, Bruchsal.

IV. Customer Acquisition and Growth in Extremely Competitive Markets: Insights from the Mobile App Category

Abstract

In essay III, I aim to jointly analyze the customer acquisition reached and the time to get there using a joint ordinal-survival analysis model. The focus in this research is on why, in the face of such extreme competition, some apps acquire customers faster than others. I investigate this question using data on the number of users acquired, and the acquisition growth, for about 2455 Apps from Google Play. I categorize the number of users acquired into ordered tiers and formulate a joint model of growth and customer acquisition using a survival model for the former and an ordinal logit model for the later. The explanatory variables include price, valence of customer rating, and other product attributes. Additionally, effects of competitive contexts and frames are considered. I also consider the role of information cascades on customer acquisition and growth in extremely competitive markets. The model is calibrated within a Bayesian framework using MCMC methods. Findings for the app category as well as generalizable implications for extremely competitive markets are discussed.

Key Words: Customer Acquisition; Growth; Extremely Competitive Markets; Mobile App; Joint Ordinal-Survival Analysis Model

4.1 Introduction

Many categories today are extremely competitive due to a plethora of similar products being continuously offered in the markets. For instance, mobile app users have more than 1.3 million available apps to choose from Google Play and 1.2 million apps on Apple's app store as of July 2014[10]. Similarly, more than a billion subscriptions were spread across 250, 000 unique podcasts in more than 100 languages on Apple's iTunes Store as of July 2013[11]. Individuals can also write and self-publish eBook or articles through e-book publishing platforms such as Amazon Kindle Direct Publishing[12]. Additional examples of extremely competitive categories include, online microblog subscriptions (e.g., Twitter), online video subscriptions (e.g., YouTube), online question-and-answer services (e.g., Quora), and so on. The format that platforms allows anybody to launch products exaggerates the competition in these categories.

Firms typically rely on three revenue models to monetize their products under the extreme competition. First, a majority of firms in the markets adopt a freemium business strategy, offering products in the hope of expanding customer acquisition. Thus, advertisers may seek to promote their products through built-in advertisements in products in extremely competitive markets. Second, some firms may also provide a version of the product with more features offered at a price while a free version can be

[10] http://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/
[11] Trevor Mogg (July, 2013), "Apple Hits One Billion Podcast Subscriptions Via iTunes Store", http://www.digitaltrends.com/mobile/apple-hits-one-billion-podcast-subscriptions/
[12] David Carnoy(2012), "How to self-publish an ebook", http://www.cnet.com/how-to/how-to-self-publish-an-ebook/

acquired, which may bring some revenue to the company. Third, other companies may adopt a base and add-on pricing strategy (Bertini, Ofek and Ariely 2009; Erat and Bhaskaran 2012), offering a basic version of the product for free or at a very low price, but charging fees for additional product features, which are referred as in-app products.

Despite the widespread use of the three revenue models, a vast majority of businesses make little profits, which is especially true in mobile application market (Forbes 2013). An average iOS developer may earn a dime for every one of the 40,000 potential app downloads[13]. An Android developer, however, makes substantially lower revenue, with the average app download bringing in around 2 cents to its developer (Forbes 2013). Therefore, there is much work to be done to increase monetization of products in extremely competitive markets. One way is to increase the unit revenue per purchase. In other words, businesses should either increase unit prices of the base product and add-ons or provide a paid-version of the product for customers, i.e., by utilizing the last two revenue models mentioned above.

Another way to increase revenue, giving the universally low margin in these categories, is to increase the customer acquisition and growth rate. The developers may, thus, make decent profits based on a large user base through the first revenue model, though the unit margin is relatively low. The businesses, however, vary substantially in the customer base acquired and the time taken to reach the level of customer acquisition. For

[13] In the article of "How Much Do Average Apps Make?" (Forbes 2013), Tristan Louis computed the revenue per download based on the total downloads and thus the revenue earned in the two platforms.

instance, in the mobile app market, based on a sample I collected, when Rovio Mobile Ltd. launched a new app, “Angry Birds Star War”, on Nov 8, 2012, it quickly became a hit - acquired more than 10 million users in less than six months (161 days). However, another product in the sample, “Neon Smoke”, which was released on Mar 6, 2012, only acquired more than 10 thousand users over 14 months (408 days).

The substantial variation in customer acquisition and growth rate permits us to explore the question of what factors help to attract more users rapidly. The issue has not yet been investigated in previous literature. This is the focus of the research. Specifically, the study aims to provide insights into why some apps grow faster than others regardless of what level they grow to. For instance, there would be two apps which both grow to the 5000-10,000 level but one gets there within 20 days while the other take 200 days. Likewise, in another case where two apps both reached the million download mark, one app would have reached there in 60 days while the other took 100 days. By jointly analyzing both the level reached, and the time taken to get there, across these very different pairs, one can get overall insights regarding when apps grow rapidly. To investigate the issue, I developed a conceptual framework in which both customer acquisition and growth rate are affected by context effects, framing effects, and information cascade effects. The data of 2454 mobile apps were collected for the study and a simultaneous equation model with joint random effects is implemented in Bayesian Paradigm by using MCMC methods.

The research setting is the mobile app category. This is a relevant category to investigate for multiple reasons. First, this is a typical market with extreme competition. As of

October 18, 2014, the total number of mobile apps in Android market is 1,376,761. When users search for a certain type of product, the distributor will present a choice set with many competitive options. Therefore, businesses in the market compete for customers fiercely. Second, the market grew rapidly in recent years. According to Gartner (2013)[14], the total number of downloads in 2012 was 63 billion, and it is projected to reach 269 billion in 2017 – about 4.3 times. Third, developers vary substantially in customer acquisition and growth rate. In the sample of over two thousand apps that I investigate, for instance, the customer base varied from 0 to more than 10 million and the time taken to reach a certain customer acquisition level varies substantially across different products. Finally, the product attributes, competitive context, and developers' abilities and experiences vary substantially as well. For instance, the file size, representing product complexity, varied from 0.001 MB to 1024 MB. The average volume of customer ratings for competitive products varied from 0 to 697185. The product assortment size varied from 5 to 67. Such substantial variations in the product attributes, competitive context, and developers' abilities and experiences and corresponding customer acquisition and growth rate permit us to investigate the effects of those factors on the two.

Next I review the literature and develop the conceptual framework. I then describe the data and the empirical model for the investigation. Following this, I present the empirical results for the empirical investigation. Finally, I discuss the implications of the findings

[14] Mobile App Store Downloads, Worldwide, 2010-2016 (Millions of Downloads). Source: Gartner (September 2013)

for the customer acquisition and growth rate in extremely competitive categories and provide directions for additional research.

4.2 Conceptual Framework

In extremely competitive markets, customers start making purchase decisions by searching for product information through big platforms. For instance, customers acquire mobile apps primarily through the five largest platforms, Google Play Store, Apple's App Store, Amazon's App Store, and Windows App Store. The platforms provide information of a large number of products in a certain similar format. Taking Google Play Store as an example, a search of products on the platform results in icons of hundreds of similar options. An icon displays the app name, the developer's name, and the price. By clicking the icon, customers can obtain more information of the product, including screenshots, valence and volume of customer ratings, product description, file size in megabytes, content rating, customer acquisition in the form of ranges, etc.

Additionally, distributors will display a set of competitive products on the same page in which four of them are apps *viewed* by similar users and another set of four apps are apps *installed* by the users who also installed the app being evaluated. The app name, developer name, price, and volume of customer ratings used to be displayed on each icon. Distributors thus define the competitive context (Tversky and Simonson 1993) of each item in the choice set.

Moreover, an assortment of products created by the same developer would also be displayed on the website. By clicking the link of "more from developer", customers can find icons of all other products created by the developer on the platform. A click on an icon would reveal its price, valence and volume of customer ratings, etc. The set of products created by the same developer, therefore, provides a frame (Kahneman and Tversky 1986) for each product presented in the choice set.

In extremely competitive markets such as mobile apps, customers may rely on price to evaluate the alternatives and select the best option. However, faced a large number of similar options, customers may be susceptible to choice overload (Iyengar and Lepper 2003) and desire to simplify the choice decisions (Dhar, Nowlis and Sherman 2000). Thus, we suggest that the customer acquisition and growth rate are also subject to context effects (Prelec, Wernerfelt and Zettelmeyer 1997; Bertini, Wathieu and Iyengar 2012) and framing effects (Tversky and Kahneman 1981; Kahneman and Tversky 1986; Levin, Schneider and Gaeth 1998) which can help simplify the choice.

Additionally, an increase in the number of customers may also lead to informational cascades (Anderson, 2001; Bikhchandani, et al., 1992; Bikhchandani, et al., 1998; Walden and Browne, 2002) and thus increase the growth rate. The theoretical framework, thus, is based on the above four components.

Weighted Additive Utility Model

Facing a wide selection of similar products in extremely competitive markets, consumers may need to take considerable cognitive effort on evaluating the product attributes to find an optimal choice. To achieve the goal, we suggest that customers may rely on the weighted additive utility model (Bettman, Luce and Payne 1998) to assess the utility of a specific product.

There is evidence in the literature that consumers use the weighted additive-utility model where they evaluate the attributes of an alternative in detail in order to select the best option. For instance, Bettman, Luce and Payne (1998) suggest task complexity can influence information processing. Specifically, as the complexity of the product increases, consumers are likely to resort to simpler heuristics. Swaminathan (2003) suggests customers are more likely to rely on recommendation agents to select the “best” option while product complexity is low. In the case of products in mobile application market, a product with larger file size may contain more features and provides more in-app products for customers, and thus increase the product complexity, which may require more cognitive efforts. Therefore, I suggest that a product with a high level of complexity may discourage customers to select it due to information overload (Jacoby 1977), and thus, attract less customers and decreases growth rate.

Previous literature also suggest that product category may affect customers’ choice decision. For instance, Park and Lessig (1981) suggest that consumers with greater knowledge are better able to distinguish between attribute levels than those with lower levels of knowledge. In addition, consumers with higher knowledge levels are able to make trade-offs between various attribute levels more easily than consumers with less knowledge. Swaminathan (2003) indicates that when consumers have less category knowledge, to increase decision quality, consumers are more likely to rely on recommendation agents. When it comes to products in extremely competitive markets, knowledge of product categories becomes even more important, since the themes of each category vary substantively. Therefore, I also expect that knowledge of the product category of a product can help to attract more customers and thus increase growth rate. However, whether customers are familiar with a certain product category is an empirical results.

Additionally, on the one hand, previous studies suggest the product price can affect the perceived quality of products (Gabor and Granger 1966; Dodds, Kent and Grewal 1991; Wathieu and Bertini 2007; Bertini, Wathieu and Iyengar 2012; Lalwani and Shavitt 2013). The reliance on price to assess a product's quality or performance is particularly likely in extremely competitive categories (Bertini, Wathieu and Iyengar 2012). In these categories, most firms give their products away for free or sell at very low prices. Thus, a higher price may signal better product quality. Wathieu and Bertini (2007) suggests higher prices in such categories could be "thought provoking and enhance the perception of relevance" thus increasing the likelihood of choice. On the other hand, price also plays an allocative role due to budget constraints (Becker 1965, Friedman 1957) faced by consumers who need to allocate available monetary resources across multiple products. Thus, as the price of a product increases, less would be available to allocate to other products (Erickson and Johansson 1986) if the product is purchased. In extremely competitive markets, we suggest that the allocative role of app price is stronger than its informative role since the price is too low to signal product quality. Thus, a high product price is expected to decrease customer acquisition and growth rate.

Advertising can also affect the sales and shares of the products they sell (Tellis 1988; Erdem, Keane and Sun 2008). For instance, Tellis (1998) suggests advertising exposures will reinforce preference for brands. Erdem, Keane and Sun (2008) also imply that both advertising frequency and content can signal product quality. In extremely competitive categories, some developers will provide promotional videos for customers to increase the visibility of the product, and/or to signal better product quality (Erdem, Keane and Sun 2008) and reinforce preference (Tellis 1988), which would eventually increase customer acquisition and growth rate.

Finally, previous studies suggest that the valence of customer ratings can affect product sales (Liu 2006; Dellarocas, Awad and Zhang 2005; Duan, Gu and Whinston; Hao, Li, Tan and Xu 2011). For instance, Liu (2006) suggests positive WOM enhances perceived value and thus has either a direct or an indirect recommendation for product purchase. In extremely competitive markets, products are a type of experience goods. It is difficult for customers to evaluate the true quality before actually using the product. Additionally, a considerable heterogeneity exists among a mixture of individual and organizational businesses which make it even more difficult for users to distinguish the quality of products *ex ante* (Hao, Li, Tan, and Xu 2011). Thus, the valence of customer ratings, i.e., average ratings, displayed by distributors, provides good source for customers to form initial evaluations for products and thus helps them to make final decisions. Specifically, we expect that the positive valence of customer ratings would encourage future users to purchase it and thus increase the growth rate.

Therefore, we suggest customers may evaluate the above attributes based on the weighted additive-utility model. However, although a reliance on weighted additive utility model may lead to relatively better decisions, the proliferation of choice may demotivate the shoppers (Iyengar and Lepper 2000), who thus may desire to simply their choices (Dhar, Nowlis and Sherman 2000) by examining contexts and frames of the products. Specifically, in extremely competitive markets, customers are expected to utilize the information in the context of the alternatives provided by other firms (Prelec, Wernerfelt and Zettelmeyer 1997; Bertini, Wathieu and Iyengar 2012). Additionally, they may evaluate the frames (Tversky and Kahneman 1981; Kahneman and Tversky 1986; Levin, Schneider and Gaeth 1998) to further assess the evaluated alternative based on the product assortment of a firm.

Context Effects

There is extensive evidence in the literature (Tversky and Simonson 1993[15]; Prelec, Wernerfelt and Zettelmeyer 1997; Kivetz, Netzer and Srinivasan 2004; Sela, Berger and Liu 2009; Simonson 2008[16]; Bertini, Wathieu and Iyengar 2012) that customers rely partly on context to evaluate alternatives. Context effects are likely to occur when consumers face uncertain decision environments either due to product proliferation and clutter (Bertini, Wathieu and Iyengar 2012) or uncertainty of product attributes (Prelec, Wernerfelt and Zettlrmeyer 1997; Simonson 2008). Both conditions are likely to occur in

[15] Tversky and Simonson (1993) define the context as “the set of options under consideration” (p.1181).
[16] Simonson (2008) refer to the preference that are context-dependent as “constructed preferences” in contrast to “inherent preferences”.

extremely competitive markets. First, in extremely competitive markets, products being offered for customers have increased rapidly over the past several years. For a single product category, customers may find hundreds of competitive options in the markets when making a purchase decision. Second, since products in these markets are experiential, customers lack of the knowledge to distinguish the quality of product attributes until they actually use them (Hao, Li, Tan, and Xu 2011). When customers face options that can provide very limited or no experience (Kivetz, Netzer and Srinivasan 2004), the constructed preference based on context effects should dominate the overall evaluation (Simonson 2008). I, therefore, expect context effects would play a big role in customer acquisition and growth in extremely competitive markets since products in these markets are displayed with a set of competitive options selected by the platforms which forms a competitive context for the product being evaluated.

To construct the context for products in extremely competitive categories, we rely on average price of the competitive products to capture the influence of price in the context and the average volume of customer ratings to represent the popularity of products in the context. Two types of context effects have been widely discussed in previous literature. One type of context effect is the attraction effect – adding an inferior product to the choice set may increase the attraction of the current options. Another consequence of the context is the compromise effect which leads customers to select a middle of the road product to avoid extremes. Both effects suggest that price and quality of other products can serve as important components of context for the one being evaluated but which effect should dominate is an empirical result. If attraction effect dominates, we expect that a product with a context of lower-priced items is more likely to attract customer acquisitions and increase the growth rate. In addition, a product displayed within a context of widely discussed items are less attractive and thus decrease the customer acquisitions and growth rate. However, if compromise effect dominates, adding a high-priced and more popular item to the context will increase the acquisition of the product and its growth rate.

Framing Effects

Although context effects arise from the relative differences between an option being considered and other competing options available in the choice set, the past and present context of experience may serve as frames during a decision making process. Individuals tend to choose alternatives framed as gains to avoid any risks based on previous economic literature (Tversky and Kahneman 1981; Kahneman and Tversky 1986). The behavioral literature identified three types of framing effects (Levin, Schneider and Gaeth 1998): (1) risky choice framing based on levels of risk described; (2) Attribute framing based on some characteristics of an object or event; (3) Goal framing based on how is the goal of an action or behavior is framed. Additionally, they find that consumers respond more favorably to positive frames than negative frames, since positive frames evoke favorable associations in memory and encourage the retrieval of positive information (Levin, Schneider and Gaeth 1998; Janiszewski, Silk and Cooke 2003). Similar evidence can be found in literature. For instance, Berger, Draganska and Simonson (2007) suggests the assortment demonstrates variety and indicates experience in developing similar products may increase the valuation of its alternative. Additionally, adding more options to existing product lines will increase the valuation of the alternative being evaluated (Bertini, Ofek and Ariely 2007).

Therefore, we expect a product's frames based on other products created by the firm may affect the acquisition decision and growth rate in extremely competitive categories. To operationalize the framing effects, we construct three variables (1) the average price of the products created by the developer; (2) the average volume created by the developer; (3) the total number of products created by the developer. First, higher average price is expected to provide a favorable frame due to an association of better product quality and thus increase customer acquisition and growth rate. Second, higher volume of positive customer ratings[17] can also result in associations with positive information and thus provide favorable frames for customers. Additionally, firms with more products may indicates that more resources can be provided to improve product design and implement effective marketing strategies, hence, serving as favorable frames.

Information Cascades

An informational cascade occurs when it is optimal for an individual, having observed the actions of those ahead of him, to follow the behavior of those individuals without regard to the private information (Bikhchandani, Hirshleifer and Welch 1992). The potential for observational learning and herding behavior arising from informational cascades has been discussed widely in the economic literature (Bikhchandani, Hirshleifer and Welch

[17] On Google Play Store, the competitive apps displayed have high volume of positive ratings.

1992; Anderson and Holt 1997; Hung and Holt 2001; Celen and Kariv 2004; Alevy, Haigh and List 2006; Bowden 2013).

The effect of information cascades has also studied in two streams of marketing literature (Golder and Tellis 2004; Godes and Mayzlin 2004; Watts and Dodds 2007; Zhang 2009; Lee, Tan, and Hosanagar 2009; Zhu and Zhang 2010; Chen, Wang and Xie 2011; Godes and Silva 2012). One stream of literature explored the effect of information cascades is consumers' product adoption behavior (Golder and Tellis 2004; Zhang 2009). Another stream of marketing literature exploring the Word-of-Mouth behavior has both empirically and experimentally tested the influence of information cascades (Godes and Mayzlin 2004; Watts and Dodds 2007; Zhu and Zhang 2010; Chen, Wang and Xie 2011; Godes and Silva 2012).

According to Bikhchandani, Hirshleifer and Welch (1998), an informational cascade are likely to arise because individuals obtain similar information, face similar alternatives, and face similar payoffs, and as a result, individuals tend to make similar choices. Thus, the behavior resulted from information cascade is especially likely to occur in extremely competitive markets due to following reasons: (1) Distributors in mobile application market, such as Google Play Store, provide similar information of products for customers, including product attributes, information of other similar products, and information of other products created by the same developer; (2) The competitive products existed in the market can be viewed as similar due to uncertainties of the product attributes and difficulties of evaluate all possible options; (3) Customers do not have explicit preferences towards the options, the payoffs obtained from the products they bought, thus, can be viewed as homogeneous. Therefore, in extremely competitive markets, I expect the information cascade would increase growth rate.

Additionally, in the early stage of customer acquisition, the effect of product attributes may outweigh information cascades effect. For instance, Golder and Tellis (2004) suggest that the effect of early sales of a new product could be suppressed (Golder and Tellis 1997; Tellis et al. 2003), only a small number of customers adopts a new product based on its product quality. However, as the number of new adopters increases, it provides increasingly strong signal to the non-adopters, who then adopt in increasing numbers. Therefore, I expect that the same pattern of customer acquisition would follow in extremely competitive markets. In its initial stage, a small group of customers acquire the products based on private information (Bikhchandani, Hirshleifer and Welch 1998), i.e., the information of product attributes. As the number of customer acquisition increases, the information derived from the decisions of others begins to outweigh an individual's private evaluation. The informational cascades effects tend be more salient due to the convergence in actions as acquiring more customers.

However, the cascade of customers to acquire a new product is likely to end in a certain stage which usually happens at the onset of maturity due to a decline of marginal benefits from current product and the announcement of new competitors. Thus, the growth rate may suffer. Therefore, in the present study, I include the categories of customer acquisition as indicators in the growth rate equation. I expect that a higher category of customer acquisition would increase the growth rate in the early and growth stage. However, as the customer acquisition reaches a certain level, the growth rate would decline.

4.3 Research Setting and Data Description

I use data from the mobile app category for the empirical investigation. Most apps in the market are developed to be used with Google's Android operation system, Apple's iOS, or both, and are distributed through large distributors like Google Play and Apple's App Store. Further, the number of apps carried by each store is very high. The Google Play store, for instance, had more than 30,000 apps offered for users across different product categories. The mobile app category is therefore extremely competitive and serves as an appropriate setting for this research.

I collected the data on February 12, 2014, on all the apps available in the six categories of games, such as Arcade, Brain, and Cards for Google Play Store. Of the 2937 mobile applications that were available on that date, 483 apps were deleted either due to replicates or missing values. Therefore, the data of 2454 mobile apps were used for analysis. The mobile apps in the dataset were created by 272 developers. Hence, the data has a hierarchical structure. Further, a variable of primary interest in this research is the growth of the apps. I, therefore, identified the release date of each app in the dataset based on the information provided by App Annie[18].

[18] App Annie is a business intelligence company which tracks over 6,400,000 mobile applications across the iOS, Mac, Google Play, Amazon App Store, Windows Phone Store, and Windows 8 Store: http://www.appannie.com/

Dependent Variables:

The two dependent variables are the customer acquisition and growth rate. For each app, Google Play Store provides a measure of customer acquisition in terms of the tier of total downloads reached. A total of twelve tiers was collected from the store. However, the customer acquisition is grouped into six different categories, because it would be computationally demanding if we implement the issue using Ordinal Logit Model with twelve categories, especially when there is a hierarchical structure in the model. I proxy the growth rate as the number of days since the launch date to the date on which I collected the data. Table IV-1 provides descriptive statistics of measures of both customer acquisition and growth rate.

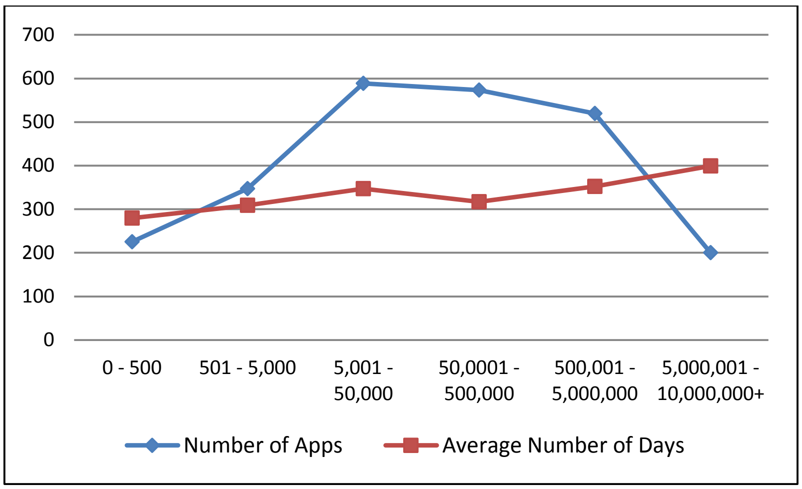
Table IV-1 Descriptive Statistics of Customer Acquisition and Growth

Level	Number of Users	Number of Apps	% of Sample	Average Number of Days to Reach the Tier
1	0 - 500	225	9.20%	279
2	501 - 5,000	347	14.10%	308
3	5,001 - 50,000	589	24.00%	347
4	50,0001 - 500,000	573	23.30%	317
5	500,001 - 5,000,000	520	21.20%	352
6	5,000,001 - 10,000,000+	200	8.10%	399

Interestingly, the tiers are not equal in size. Lower tiers have a much smaller size than the higher ones. This, perhaps, reflects the realization in the industry as well that early growth is much slower than later growth. Second, the distribution of the number of apps at each level of customer acquisition follows a normal distribution. Fewer mobile apps in the store have extremely low or high levels of customer acquisition, i.e., downloads of 0 to 500 and downloads of more than 5 million. The number of users of sixty eight percent of mobile apps in the sample lies between 5 thousand and 5 million.

Third, the number of days indicates the growth rate of customer acquisition. In the initial release stage, an app may gain a small user group, i.e., from 0 to 500, in a relatively short time period. However, the growth rate of user base decreases after the initial stage. It takes longer time for an app to reach a higher level of customer acquisition. Interestingly, after the customer reached the third level, i.e., from 5,001 to 50,000, the growth increased. After this stage, however, the growth rate decreased again. Both the last column of Table IV-1 and the Figure 1 illustrate the pattern.

Figure 1: Growth across Different Levels of Customer Acquisition



Independent Variables:

Mobile apps are heterogeneous in terms of their themes (e.g., arcade, brain, etc.), and the number and complexity of attributes (Wall Street Journal 2013). Such differences can affect how consumers choose, use, and assess products (Bettaman, Luce and Payne 1998) based on weighted additive-utility model. I, therefore, use multiple variables to capture the effects of product attributes in customer acquisition and growth rate. Specifically, I include the following variables: (1) App price (*PRICE*) – the app price ranges from $0 to

$17.41. I recode the price variable into a dummy variable with 0 represents “free apps” and 1 represents “paid apps”. There are two reasons. First, the price dispersion of application is small. The most expensive applications in the dataset is $17.41. Second, more than half of applications are offered for free (63.74%). Therefore, it’s not appropriate to treat the variable as continuous variable (2) The promotional video (*VIDEO*) – about one third of mobile apps in the sample has promotional video – may increase visibility of products and thus serving as an advertising vehicle (3) The Product Category - include *Arcade*, *Brain*, *Cards*, *Casual*, *Sports*, and *Racing*. I recode it to six dummy variables with *Sports* serving as the reference category and include the dummies into model to control for the differences in product categories (4) Average Rating (*USR_RATING*) – represents the perceived quality from users who actually installed the products. The average rating is very positive with mean average rating of 4.005. Therefore, the perceived quality of most mobile apps in the store is very high (5) the File Size in megabytes (*FSIZE*) – represent product complexity and therefore affect customer acquisition and growth rate. Table IV-2 provide the descriptive statistics of products attributes illustrated above.

Table IV-2 Descriptive Statistics of Product Attributes

Variables	Categories	Number of Apps in the Category	Percentages
	Free	1872	63.74%
PRICE	Paid	1065	36.26%
	No	1975	67.25%
VIDEO	Yes	962	32.75%
	arcade	589	20.05%
	brain	407	13.86%
	cards	455	15.49%
	casual	805	27.41%
CATEGORY	sports	340	11.58%

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
racing	341		11.61%			
USR_RATING	0	3.8	4.2	4.005	4.5	5
FSIZE	0.001	2.6	7	21.175	17	1024

For every app, Google Play also presents information on the other apps "*viewed*" or "*installed*" by the user. I constructed four variables from this information to operationalize context effects: (1) I compute the average price (*VIEW_PRICE*) and average volume of ratings (*VIEW_NUSERS*) of mobile apps featured as "viewed by the same user" on the website. (2) I also compute the average price (*INSTL_PRICE*) and average volume of ratings (*INSTL_NUSERS*) of mobile apps featured as "installed by the same user" on the website. Table IV-3 provides descriptive statistics of the four variables representing context effects. The mean average price of the apps "viewed by the same user" (1.078) is higher than that of the apps "installed by the same user"(0.775). In addition, the mean average volume of the apps "installed by the same user" (30146) is much higher than that of the apps "installed by the same user" (808).

Table IV-3 Descriptive Statistics of Variables Representing Context Effects

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
VIEW_PRICE	0	0.248	0.743	1.078	1.495	12.598
VIEW_NUSERS	0	1924	7364	30146	30905	697185
INSTL_PRICE	0	0	0	0.775	1.490	9.168
INSTL_NUSERS	0	31	70	808	230	107572

Finally, Google Play also gives users the opportunity to learn more about the app developers' features. For every app in the sample, Google Play provides a product assortment created by the developer and a hyperlink to a more detailed summary screen which includes all apps created by the developer. I use the information to construct three variables to represent the influence of framing effects: (1) The total number of mobile apps created by the developer (*NAPPS*) – represents the size of product assortment; (2) The average volume of ratings of mobile apps created by the developer (*NUSERS*) – represent the general popularity of the products created by the developer; (3) The average price of mobile apps created by the developer (*AVEPRICE*), which may signal the average product quality of the products created by the developer. Table IV-4 provides descriptive statistics of the three variables representing framing effects. The developers exhibit substantial heterogeneity in the size of the product assortments, the average volume of ratings, and the average price of product assortments. I relied on these three variables to control developers' fixed effects in the empirical investigation. The control of heterogeneity reduces the likelihood of systematic correlation between app attributes and developers' experience.

Table IV-4 Descriptive Statistics of Variables Representing Framing Effects

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
NAPPS	5	7	12	17.9	22	67
NUSERS	1	95	1037	16791	14880	381976
AVEPRICE	0	0	0.398	0.830	1.273	10.928

4.4 Modeling Approach

Since the goal is to investigate the factors that affect the time taken by app i (t_{ik},) to reach category k (k_i) since its launch, I decide to implement the issue using a joint model (Henderson, Diggle, and Dobson 2000) with survival model on growth rate and ordinal logit model on levels of customer acquisition. The probability for each app i taking time t to reach level k of customer acquisition, therefore, is:

$$L_{ik} = p(k_i)p(t_{ik})$$

Where,

$\mathrm{p}(k_i)$ is the density function of the ordinal logit model;

$\mathrm{p}(t_{ik})$ is the density function of the survival model;

Model for Growth: Survival Model

Since the number of days in release is the proxy for growth rate, I take a survival analysis approach to investigate growth rate. For each app i, I assume that the baseline density of the event time follow a Weibull distribution, where μ_{1i} is the scale parameter and λ is the shape parameter.

$$t_{ik}|\lambda, \mu_{1i} \sim weibull(\lambda, \mu_{1i})$$

$$f(t_{ik}|\lambda, \mu_{1i}) = \mu_{1i}\lambda {t_{ik}}^{\lambda-1} e^{-\mu_{1i}{t_{ik}}^{\lambda}}$$

$$h(t_{ik}|\lambda, \mu_{1i}) = \mu_{1i}\lambda {t_{ik}}^{\lambda-1}$$

The hazard rate $h(t_{ik}|\lambda, \mu_{1i})$ indicates the probability of app i will reach a higher level of customer acquisition given that they have reached the current level.

The likelihood function of reaching a higher level given the current acquired customer base, therefore, is:

$$\mathrm{p}(t_{ik}|\beta, \lambda) = \prod_{i=1}^{n} \left\{\mu_{1i}\lambda {t_{ik}}^{\lambda-1} e^{-\mu_{1i}{t_{ik}}^{\lambda}}\right\}, i = 1,2, \dots, 2542$$

The proportional term is: $\mu_{1ik} = \exp(\overline{x_1'}.\overline{\beta_1} + \sum_{h=1,h\neq k}^{k-1} \gamma_k d_k)$. The covariates x_1' includes product attributes, variables representing context effects, and variables representing framing effects. Additionally, figure 1 suggests that growth rate is affected by level of the customer acquisition reached. It is likely that, in turn, mobile apps that are growing at a faster rate attract more users in later stages. I, therefore, allow for potential simultaneity between the two outcomes. Operationally, I includes the tiers reached as indicators (d_{kij}) in the growth model. Therefore, in survival model, $\gamma_k$captures the effect of passing the levels lower than k on the time taken to reach k.

$$
\begin{aligned}
log(\mu_{1,ik}) = {} & \beta_{1,1} \times USR\_RATING_{ij} + \beta_{1,2} \times VIDEO_{ij} + \beta_{1,3} \times PRICE_{ij} + \beta_{1,4} \times FSIZE_{ij} \\
& + \beta_{1,5} \times VIEW\_NUSERS_{ij} + \beta_{1,6} \times VIEW\_PRICE_{ij} + \beta_{1,7} \\
& \times INSTL\_NUSERS_{ij} + \beta_{1,8} \times INSTL\_PRICE_{ij} + \beta_{1,9} \times ARCADE_{ij} + \beta_{1,10} \\
& \times BRAIN_{ij} + \beta_{1,11} \times CARDS_{ij} + \beta_{1,12} \times CASUAL_{ij} + \beta_{1,13} \times RACING_{ij} + \gamma_1 \\
& \times d_{1ijk} + \gamma_2 \times d_{2ijk} + \gamma_3 \times d_{3ijk} + \gamma_4 \times d_{4ijk} + \gamma_5 \times d_{5ijk} + u_{1,j}
\end{aligned}
$$

$d_k: dummies\ representing\ if\ the\ download\ of\ the\ app\ jumped\ to\ the\ higher\ level.\ k = 1,2,3,4,5$

Table IV-5 Dummy Coding of Indicators d_k,k=1,2,3,4,5

Indicator d_k	Dummy Coding
d_1	"1" if reached level 2 of customer acquisition, "0" otherwise.
d_2	"1" if reached level 3 of customer acquisition, "0" otherwise.
d_3	"1" if reached level 4 of customer acquisition, "0" otherwise.
d_4	"1" if reached level 5 of customer acquisition, "0" otherwise.
d_5	"1" if reached level 6 of customer acquisition, "0" otherwise.

We assume non-informative priors on all the parameters:

$$\beta_1 \sim dnorm(0,0.001)$$

$$\lambda \sim dgamma(0.001,0.001)$$

$$d_k \sim dnorm(0,0.001)$$

$$\gamma \sim dnorm(0,0.001)$$

Model of Customer Acquisition: Ordinal Logit Model

In the model of customer acquisition, I assume that the utility of application i is represented by an unobservable latent variable U_i. The downloads of application i will jump to a certain level on the basis of U_i.

$$U_i = \mu_{2,i} + \varepsilon$$

The threshold parameters obey the ordering constraint: $\theta_1 < \theta_2 < \theta_3 < \theta_4 < \theta_5$

$$\begin{gathered}
downloads_i = 1 \xrightarrow{yield} U_i < \theta_1 \\
downloads_i = 2 \xrightarrow{yield} \theta_1 < U_i < \theta_2 \\
downloads_i = 3 \xrightarrow{yield} \theta_2 < U_i < \theta_3 \\
downloads_i = 4 \xrightarrow{yield} \theta_3 < U_i < \theta_4 \\
downloads_i = 5 \xrightarrow{yield} \theta_4 < U_i < \theta_5 \\
downloads_i = 6 \xrightarrow{yield} U_i > \theta_5
\end{gathered}$$

Assume that ε follows a logistic distribution, which means the cumulative distribution of ε is $F(\varepsilon) = \exp(\varepsilon)/(1 + \exp(\varepsilon))$.

Specifically,

$$\begin{gathered}
p_{i1} = Prob(U_i < \theta_1) = Prob(\mu_{2,i} + \varepsilon < \theta_1) = Prob(\varepsilon < \theta_1 - \mu_{2,i}) \\
p_{i2} = Prob(\theta_1 < U_i < \theta_2) = Prob(\theta_1 < \mu_{2,i} + \varepsilon < \theta_2) = Prob(\theta_1 - \mu_{2,i} < \varepsilon < \theta_2 - \mu_{2,i}) \\
p_{i3} = Prob(\theta_2 < U_i < \theta_3) = Prob(\theta_2 < \mu_{2,i} + \varepsilon < \theta_3) = Prob(\theta_2 - \mu_{2,i} < \varepsilon < \theta_3 - \mu_{2,i}) \\
p_{i4} = Prob(\theta_3 < U_i < \theta_4) = Prob(\theta_3 < \mu_{2,i} + \varepsilon < \theta_4) = Prob(\theta_3 - \mu_{2,i} < \varepsilon < \theta_4 - \mu_{2,i}) \\
p_{i5} = Prob(\theta_4 < U_i < \theta_5) = Prob(\theta_4 < \mu_{2,i} + \varepsilon < \theta_5) = Prob(\theta_4 - \mu_{2,i} < \varepsilon < \theta_5 - \mu_{2,i}) \\
p_{i6} = Prob(U_i > \theta_5) = Prob(\mu_{2,i} + \varepsilon > \theta_5) = Prob(\varepsilon > \theta_5 - \mu_{2,i})
\end{gathered}$$

In the ordinal logit model, similarly, the three sets of covariates representing the product attributes, context effects, and framing effects were included in the regression function. Additionally, the number of days since launch is included as a predictor in the customer acquisition in the model to control for simultaneity between growth rate and customer acquisition.

$$\begin{aligned}
\mu_{2,i} = \beta_{2,1} \times USR_RATING_{ij} + \beta_{2,2} \times VIDEO_{ij} + \beta_{2,3} \times PRICE_{ij} + \beta_{2,4} \times FSIZE_{ij} + \beta_{2,5} \\
\times VIEW_PRICE_{ij} + \beta_{2,6} \times VIEW_NUSERS_{ij} + \beta_{2,7} \times INSTL_NUSERS_{ij} \\
+ \beta_{2,8} \times INSTL_PRICE_{ij} + \beta_{2,9} \times ARCADE_{ij} + \beta_{2,10} \times BRAIN_{ij} + \beta_{2,11} \\
\times CARDS_{ij} + \beta_{2,12} \times CASUAL_{ij} + \beta_{2,13} \times RACING_{ij} + \beta_{2,14} \times days_{ij} + u_{2,j}
\end{aligned}$$

I again assume non-informative priors on all parameters:

$$\beta_2 \sim Normal(0,0.001)$$

The sample also suggests that the growth rate and customer acquisition can be related to two groups of observable variables: the application-level variables and the developer-level variables. For instance, a lower price of a mobile app may attract more users. However, products created by developers with good reputation and more resources might be more popular in the market. Therefore, I allow for a frailty in survival model and the random effects in the ordinal logit model to control for the heterogeneity of developers in the market. Operationally, I assume that they follow a bivariate normal distribution. μ_{uj} capture the fixed effects of developers by regressing on three variables representing framing effects created by developers: (1) the total number of products (2) the average volume of ratings (3) and the average price. Σ_u captures both the variances and covariance of growth rate and customer acquisition. Specifically, the variance of u_{1j} captures the random effects of developers on growth rate, whereas the variance of u_{2j} captures the random effects of developers on customer acquisition. Finally, the covariance captures the random effects of simultaneity between growth rate and customer acquisition. I assume non-informative priors and hyper priors on all the parameters.

$$\binom{u_{1j}}{u_{2j}} \sim bivariate\ normal(\mu_{uj}, \Sigma_u)$$

Priors:

$$\mu_{uj} \sim bivariate\ normal(mn_{uj}, B)$$
$$\Sigma_u \sim wishart(R, nu)$$

$$mn_{1,j} = \alpha_{1,1} \times NAPPS_j + \alpha_{1,2} \times NUSERS_j + \alpha_{1,3} \times AVEPRICE_j$$
$$mn_{2,j} = \alpha_{2,1} \times NAPPS_j + \alpha_{2,2} \times NUSERS_j + \alpha_{2,3} \times AVEPRICE_j$$

Hyper priors:

$$\alpha_{q,r} \sim Normal(0,0.001)$$

I calibrate the model using MCMC methods in Bayesian framework. Prior to model calibration, I log-transform and standardize all continuous variables in the sample due to skewness.

4.5 Empirical Results

4.5.1 Empirical Results of Growth Model

Estimates of the parameters of the survival model are presented in Table IV-5 to Table IV-8. The research question is on what factors affect the growth rate.

Effects of the Weighted Additive Utility Model:

First, as what we expected, price has a significant negative effect on hazard rate due to its allocative role ($\beta_{1,3} = -1.18$). Thus, in the mobile app market, high price of products will reduce growth rate. In other words, it takes longer time for a high-priced item to reach a higher level. More importantly, however, product complexity has a positive effect thus increases hazard rate ($\beta_{1,4} = 0.151$). Thus, a product with bigger file size will take shorter time to reach a higher level. Interestingly, both promotional videos and valence of customer ratings do not have any significantly effect on growth rate. As for the product categories, compared with Sports mobile apps, Arcade, Racing, and Casual mobile apps can increase hazard rate, whereas Cards will decrease hazard rate. Therefore, a card game will take a longer time to reach a higher customer acquisition level compared to games of other categories.

Table IV-6 Empirical Results of Price, Promotion and Product Attributes on Growth

		Posterior Mean	2.5%	97.5%
USR_RATING	$\beta_{1,1}$	-0.0411	-0.0949	0.0124
VIDEO	$\beta_{1,2}$	0.1201	-0.0105	0.2479
PRICE	$\beta_{1,3}$	-1.1800**	-1.4910	-0.8604
FILE SIZE	$\beta_{1,4}$	0.1510**	0.0797	0.2259
ARCADE	$\beta_{1,9}$	0.4115**	0.2412	0.5819
BRAIN	$\beta_{1,10}$	0.1368	-0.0581	0.3354
CARDS	$\beta_{1,11}$	-0.3422**	-0.5957	-0.0977
CASUAL	$\beta_{1,12}$	0.2751**	0.0965	0.4516
RACING	$\beta_{1,13}$	0.3675**	0.1790	0.5569

Empirical Results of Context Effects:

Interestingly, two of the variables representing part of the effect of the context of mobile apps “viewed by similar users” do not have any significant effects on hazard rate. Mobile apps “installed by similar users”, however, do have significantly negative effects on hazard rate. Specifically, as the price of mobile apps “installed by similar users” increases, hazard rate decreases and thus the product being acquired will take longer time to reach a higher customer acquisition level, which is opposite to our expectation. The possible explanation is that customers may have a budget constraints on the apps that one can consume. Therefore, as the price the apps installed on their device increases, they are less likely to acquire a similar one. Similarly, as the number of users who installed similar apps increases, hazard rate decreases and thus it takes longer time for the product to reach a higher customer acquisition level. The possible explanation is that the competitive apps may have taken a considerable shares and thus decrease the rate to acquire the product. Therefore, adding a high-priced and widely-discussed similar product to the context may hurt the growth rate, providing negative contexts for the product.

Table IV-7 Empirical Results of Context Effects on Growth

		Posterior mean	2.5%	97.5%
VIEW_PRICE	$\beta_{1,5}$	0.0231	-0.0286	0.0765
VIEW_NUSERS	$\beta_{1,6}$	0.0202	-0.0333	0.0752
INSTL_PRICE	$\beta_{1,7}$	-0.1359**	-0.2612	-0.0081
INSTL_NUSERS	$\beta_{1,8}$	-0.2304**	-0.3022	-0.1584

Empirical Results of Framing Effects:

Finally, the framing effects represented by the three variables of developer's own context do not have any significant effects on growth rate. Specifically, increasing the size of product assortment will reduce the growth rate but the effect is not significant. Interestingly, knowing that the other apps having more users would increase the growth rate and thus decrease the duration, this can probably attributed to the enhanced reputation of the developer. With more users using other products, a develop is more likely to be recognized by customers in the market and thus the reputation would be enhanced. Additionally, adding a high-priced item into the product assortment will reduce the hazard rate and thus increase the duration, which is opposite to our expectation. Other products with high prices may require more resources and thus reduce the growth rate of the product being evaluated.

Table IV-8 Empirical Results of Framing Effects on Growth

		Posterior mean	2.5%	97.5%
NAPPS	$\alpha_{1,1}$	-0.0031	-0.8303	0.7680
NUSERS	$\alpha_{1,2}$	0.1336	-0.7651	1.0550
AVEPRICE	$\alpha_{1,3}$	-0.1655	-1.0630	0.6663

Empirical Results of Information Cascade Effects:

We included five indicators in the survival model to represent the tiers crossed by the app with the first tier serving as the base. Since each higher tier does take more days to be reached, hazard rate decreases as it proceeds through the tiers. But, consistent with the information cascade theory, the estimates suggest that the reduction in hazard rate slows down, i.e., the growth rate picks up, as the customer acquisition of the product moves towards the higher tiers. Additionally, as we expected, the growth rate slows down after reaching the third level of customer acquisition, perhaps because the customer acquisition reached its early mature stage, which is also consistent with our expectation.

Table IV-9 Empirical Results of Cascade Effects on Growth

		Posterior mean	**2.5%**	**97.5%**
d1	γ_1	-0.8743**	-1.1150	-0.6243
d2	γ_2	-0.5469**	-0.7056	-0.3721
d3	γ_3	-0.2575**	-0.4182	-0.0863
d4	γ_4	-0.4927**	-0.6691	-0.3091
d5	γ_5	-0.2025**	-0.3925	-0.0149

4.5.2 Empirical Results of Customer Acquisition Model

Estimates of the parameters of the ordinal logit model are presented in Table IV-9 to Table IV-12. The research question was on what factors affect the customer acquisition.

Effects of the Weighted Additive Utility Model

First, as expected, increasing product visibility by providing promotional videos can significantly increase customer acquisition ($\beta_{2,2} = 0.6436$). Second, increases in product complexity suppresses acquisition due to information overload ($\beta_{2,4} = -0.1205$), which is also consistent with our expectation. Lastly, price has a significantly negative effect suggesting that apps have to be priced competitively to attract more users ($\beta_{2,3} =$

-6.195) since it plays a strong allocative role in extremely competitive markets. Additionally, the decision to allow for simultaneity is also supported – the number of days since launch has a significant positive effect on acquisitions ($\beta_{2,14} = 0.7795$). However, a display of product category wouldn't help to increase customer acquisition in the case of extremely competitive markets.

Table IV-10 Empirical Results of Effects of Price, Promotion, and Product Attributes on Customer Acquisition

		Posterior mean	val2.5pc	val97.5pc
USR_RATING	$\beta_{2,1}$	0.5369**	0.3699	0.7180
VIDEO	$\beta_{2,2}$	0.6436**	0.4033	0.8873
PRICE	$\beta_{2,3}$	-6.1950**	-6.8120	-5.6440
FSIZE	$\beta_{2,4}$	-0.1205**	-0.2327	-0.0072
ARCADE	$\beta_{2,9}$	-0.0054	-1.5070	1.4720
BRAIN	$\beta_{2,10}$	0.0075	-1.5210	1.5740
CARDS	$\beta_{2,11}$	-0.0095	-1.5620	1.4960
CASUAL	$\beta_{2,12}$	-0.0016	-1.5000	1.4660
RACING	$\beta_{2,13}$	-0.0038	-1.5020	1.5080
DAYS	$\beta_{2,14}$	0.7795**	0.6308	0.9322

Empirical Results of Context Effects:

The average volume of customer ratings for the mobile apps "viewed by the similar users" has a significantly positive effect on customer acquisition ($\beta_{2,6} = 0.1293$) and so does the average price of the mobile apps "viewed by similar users" ($\beta_{2,5} = 0.2372$). This is a positive context effect. Therefore, the competitive products with high price will drive customers away, compromise effect occurs in extremely competitive markets. In other words, customers are more likely to select a middle of the road product rather than a product with high price. The significantly positive effect of the average volume of customer ratings for the competitive products indicates that customers tend to choose product which are similar to the ones popular in the market but not the most popular ones.

On the other hand, as the average price of similar products installed by others increases, the number of customer acquisitions comes down ($\beta_{2,7} = -0.3796$), which serves as a negative context. This, again, suggests that customers have a budget constraints on the consumptions of mobile apps. As they installed one or more mobile apps on the device with relatively high price, they are less likely to install a similar one. As the average volume of customer ratings of the similar products installed by others increase, however, the number of customer acquisition goes up ($\beta_{2,8} = 0.8319$), providing a positive context. Therefore, users tend to download a mobile app when the similar ones installed on their device have higher average volume of customer ratings.

Table IV-11 Empirical Results of Context Effects on Customer Acquisition

		Posterior mean	**2.5%**	**97.5%**
VIEW_PRICE	$\beta_{2,5}$	0.2372**	0.1330	0.3420
VIEW_NUSERS	$\beta_{2,6}$	0.1293**	0.0188	0.2382
INSTL_PRICE	$\beta_{2,7}$	-0.3796**	-0.6372	-0.1131
INSTL_NUSERS	$\beta_{2,8}$	0.8319**	0.6832	0.9833

Empirical Results of Framing Effects:

Turning to the framing effects, the average volume of customer ratings of other apps from the same developer has a significantly positive effect on customer acquisition ($\alpha_{2,2} = 1.5480$), which is consistent with our expectation. Therefore, improving the reputation of other products created by the developer can provide a positive frame for the product being evaluated and thus help to attract more users. Additionally, adding high-priced items to the product assortment will increase the attractiveness of the product being evaluated and thus form a positive frame as we expected, though the effect is not significant. Interestingly, increase the size of the product assortment will decrease the customer acquisition of the product and thus form a negative frame. A possible explanation is that providing more products for customers may reduce the resources allocated to each product and thus lower the product quality, which leads to less customer acquisitions.

Table IV-12 Empirical Results of Framing Effects on Customer Acquisition

		Posterior mean	val2.5pc	val97.5pc
NAPPS	$\alpha_{2,1}$	-0.1096	-0.6693	0.5034
NUSERS	$\alpha_{2,2}$	1.5480**	0.8543	2.2200
AVEPRICE	$\alpha_{2,3}$	0.1543	-0.4601	0.8025

Empirical Results of Frailty and Random Effects:

Finally, the elements in the precision matrix are significantly positive. Therefore, the variance across different developers affects both growth rate and customer acquisition. Additionally, the significant covariance between growth rate and customer acquisition indicates the two are conditional dependent. In other words, the growth rate is related to customer acquisition. The result was also confirmed by the estimates of indicators in the survival model and the growth rate in the ordinal logit model. Specifically, the growth rate will be accelerated as it gained more customers. In the meanwhile, the increased growth rate helps to gain more users.

Table IV-13 Empirical Results of Frailty and Random Effects

Precision	Posterior mean	val2.5pc	val97.5pc
τ_{11}	0.5288	0.2842	1.0350
τ_{12}	0.7259	0.3839	1.4440
τ_{21}	0.7259	0.3839	1.4440

τ_{22}	1.0120	0.5322	2.0350

4.6 Discussion and Future Research Directions

In extremely competitive markets, some products can attract a large number of customers in a very short time period but others are struggling to win customers over for a very long time. However, there is little research that provides insights regarding how firms can succeed in these categories in this regard. This is the issue that I address in this research. Specifically, I investigate the factors that affect the time taken by an app to reach a certain download category since its launch.

I draw on an extensive research on decision making in economics and marketing and develop a conceptual model based on four components in consumers' evaluation of products: a weighted additive utility component (Bettman et al 1998), a component based on context effects (Iyengar et al 2010), a component based on framing effects (Tversky and Kahneman 1981; Kahneman and Tversky 1986; Levin et al 1998) component, and a component of information cascades (Bikhchandani, Hirshleifer and Welch 1992, 1998).

To empirically test the proposed model, I collect a sample of mobile app data from Google Play Store and the website of App Annie. I reply on a simultaneous equation model with joint random effects to explore the effects of weighted additive utility model, contexts, frames, and information cascades on customer acquisition and growth rate. Specifically, on customer acquisition, I use an ordinal logit model with duration added as a control variable. On growth rate, I utilize a survival model with indicators of download ranges included serving as the proxy of the effect of information cascades.

The empirical results provide important managerial implications for the marketers in extremely competitive categories. First, since customers may rely on weighted additive utility model to evaluate products, to acquire more customers and increase growth rate, developers should create less complex products offering at lower price. It is also necessary to provide promotional video to increase visibility of the products. Maintaining higher valence of customer ratings will help to increase customer acquisition as well. Second, developers should carefully position its product against competitors in the market to provide a positive context for their products. Specifically, adding a group of high-priced and widely discussed competitive products will increase the chances of the product being selected and thus provides a positive frame for the product. However, marketers should be aware that customers have a budget constraints on the consumption of mobile app products. Adding a high-priced similar product may decrease the customer acquisition and thus growth rate. But adding a more popular product may increase customer acquisition but decrease the growth rate of product being evaluated.

Overall, to increase customer acquisition and growth rate, the empirical results suggest that marketers should (1) provide a product line (2) encourage users to spread word of mouth (3) offer lower prices and (4) carefully select competitors to position against.

One limitation of the investigation is that our data is not longitudinal and thus we cannot explore the effects of previous downloads on subsequent customer acquisitions directly, though we believe the information cascade effect on the growth rate reflects that a “snowball effect” would occur in an indirect way, i.e., more customers would acquire the product if it has already had a large group of user base. Another limitation is that our data prevent us from exploring the effects of developers' reputation and ability on customer acquisition and growth rate, such as the ability of developing outstanding products and implement effective promotional or advertising strategies. The effects of average price, average volume of customer ratings, and total number of products do not have significantly effects on the growth rate indicates that providing low-priced items, gaining more word of mouth, or expand the product assortment won't help to increase the growth rate. It will therefore be useful for future research to empirically investigate the issue.

4.7 References of Essay III

Alevy, Jonathan E., Michael S. Haigh, and John A. List (2006). "Information Cascades: Evidence from a Field Experiment with Financial Market Professionals." *Journal of Finance*

Anderson, L.R. (2001), "Payoff Effects in Information Cascade Experiments," *Economic Inquiry* (39:4), pp. 609-615.

Anderson, L.R., and Holt, C.A. (1996), "Classroom Games: Information Cascades," *The Journal of Economic Perspectives* (10:4), 1996, pp. 187-193.

Anderson, L.R., and Holt, C.A. (1997), "Information Cascades in the Laboratory," *The American Economic Review* (87:5), 1997, pp. 847- 862.

Bertini, Marco, Elie Ofek, and Dan Ariely. (2009), "The Impact of Add-On Features on Consumer Product Evaluations." *Journal of Consumer Research* 36.1 (2009): 17-28.

Bertini, Marco, Luc Wathieu and Sheena S. Iyengar (2012), "The Discriminating Consumer: Product Proliferation and Willingness to Pay for Quality," *Journal of Marketing Research*: February 2012, Vol. 49, No. 1, pp. 39-49.

Bettman, James R., Mary Frances Luce, and John W. Payne. (1998), "Constructive consumer choice processes." *Journal of consumer research* 25.3 (1998): 187-217.

Bowden, Mark P. "Social learning, herd behavior and information cascades (2013), A review of the recent developments in relation to their criticisms." *JOURNAL OF ECONOMIC RESEARCH* 18.3 (2013): 205-440.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992), "A theory of fads, fashion, custom and cultural change as informational cascades," *Journal of Political Economy* (100:5), 1992, pp. 992-1026.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998), "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades," *The Journal of Economic Perspectives* (12:3), 1998, pp. 151-170.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. (2005), "Information Cascades and Observational Learning". Charles A. Dice Center for Research in Financial Economics, Fisher College of Business, Ohio State University, 2005.

Çelen, Boğaçhan, and Shachar Kariv. (2004) "Distinguishing informational cascades from herd behavior in the laboratory." *American Economic Review* (2004): 484-498.

Chakravarti, Dipankar, and John G. Lynch Jr. (1983), "A FRAMEWORK FOR EXPLORING CONTEXT EFFECTS ON CONSUMER JUDGMENT AND CHOICE." *Advances in Consumer Research* 10.1 (1983).

Chen, Yubo, Qi Wang, and Jinhong Xie (2011), "Online social interactions: A natural experiment on word of mouth versus observational learning." *Journal of marketing research* 48.2 (2011): 238-254.

Dellarocas, Chrysanthos, Neveen Awad, and M. Zhang (2005), "Using online ratings as a proxy of word-of-mouth in motion picture revenue forecasting". Working paper, 2005.

Dhar, Ravi, Stephen M. Nowlis, and Steven J. Sherman (2000), "Trying hard or hardly trying: An analysis of context effects in choice." *Journal of Consumer Psychology* 9.4 (2000): 189-200.

Dodds, William B., Kent B. Monroe, and Dhruv Grewal (1991), "Effects of price, brand, and store information on buyers' product evaluations." *Journal of marketing research* (1991): 307-319.

Duan, Wenjing, Bin Gu, and Andrew B. Whinston (2008), "The dynamics of online word-of-mouth and product sales - an empirical investigation of the movie industry." *Journal of retailing* 84.2 (2008): 233-242.

Erdem, Tülin, Michael P. Keane, and Baohong Sun (2008), "A dynamic model of brand choice when price and advertising signal product quality." *Marketing Science*27.6 (2008): 1111-1125.

Erat, Sanjiv, and Sreekumar R. Bhaskaran (2012), "Consumer mental accounts and implications to selling base products and add-ons." *Marketing Science* 31.5 (2012): 801-818.

Forbes (2013), “How Much Do Average Apps Make?”, Aug 10th, http://www.forbes.com/

Godes, David, and Dina Mayzlin (2004). "Using online conversations to study word-of-mouth communication." *Marketing science* 23.4 (2004): 545-560.

Godes, David, and José C. Silva (2012) "Sequential and temporal dynamics of online opinion." *Marketing Science* 31.3 (2012): 448-473.

Golder, Peter N., and Gerard J. Tellis (2004), "Growing, growing, gone: Cascades, diffusion, and turning points in the product life cycle." *Marketing Science* 23.2 (2004): 207-218.

Golder, Peter N., and Gerard J. Tellis (1997), "Will it ever fly? Modeling the takeoff of really new consumer durables." *Marketing Science* 16.3 (1997): 256-270.

Hao, Lin, Tan and Xu (2011), "The economic value of ratings in app market." Available at *SSRN* 1892584 (2011).

Hung, Angela A., and Charles R. Plott (2001), "Information cascades: Replication and an extension to majority rule and conformity-rewarding institutions." *American Economic Review* (2001): 1508-1520.

Iyengar, Sheena S., and Mark R. Lepper (2000), "When choice is demotivating: Can one desire too much of a good thing?" *Journal of personality and social psychology* 79.6 (2000): 995.

Iyengar, Sheena S., and Emir Kamenica (2010), "Choice proliferation, simplicity seeking, and asset allocation." *Journal of Public Economics* 94.7 (2010): 530-539.

Jacoby, Jacob (1977), "Information load and decision quality: Some contested issues." *Journal of Marketing Research* (1977): 569-573.

Janiszewski, Chris, Tim Silk, and Alan DJ Cooke (2003), "Different Scales for Different Frames: The Role of Subjective Scales and Experience in Explaining Attribute-Framing Effects." *Journal of Consumer Research* 30.3 (2003): 311-325.

Kivetz, Ran, Oded Netzer, and V. Srinivasan (2004), "Alternative models for capturing the compromise effect." *Journal of marketing research* 41.3 (2004): 237-257.

Mela, Carl F., Sunil Gupta, and Donald R. Lehmann (1997), "The long-term impact of promotion and advertising on consumer brand choice." *Journal of Marketing Research* (1997): 248-261.

Noble, Peter M. and Thomas S. Gruca (1999), “Industrial Pricing: Theory and Managerial Practice,” *Marketing Science*, 18, 435-454.

Lalwani, Ashok K., and Sharon Shavitt (2013), "You get what you pay for? Self-construal influences price-quality judgments." *Journal of Consumer Research* 40.2 (2013): 255-267.

Lee, Young Jin, Yong Tan and Kartik Hosanagar (2009), “Do I Follow My Friends or the Crowd? Informational Cascades in Online Movie Rating”, *Proceedings of the 19th Workshop on Information Technology and Systems* (WITS) 2009.

Levin, Irwin P., Sandra L. Schneider, and Gary J. Gaeth (1998), "All frames are not created equal: A typology and critical analysis of framing effects." *Organizational behavior and human decision processes* 76.2 (1998): 149-188.

Liu, Yong (2006), "Word of mouth for movies: Its dynamics and impact on box office revenue." *Journal of marketing* 70.3 (2006): 74-89.

Park, C. Whan, and V. Parker Lessig (1981), "Familiarity and its impact on consumer decision biases and heuristics." *Journal of consumer research* (1981): 223-231.

Payne, John W (1982), "Contingent decision behavior." *Psychological Bulletin* 92.2 (1982): 382.

Prelec, Drazen, Birger Wernerfelt, and Florian Zettelmeyer (1997), "The role of inference in context effects: Inferring what you want from what is available." *Journal of Consumer research* 24.1 (1997): 118-126.

Ratneshwar, Srinivasan, Allan D. Shocker, and David W. Stewart (1987), "Toward understanding the attraction effect: The implications of product stimulus meaningfulness and familiarity." *Journal of Consumer Research* (1987): 520-533.

Sela, Aner, Jonah Berger, and Wendy Liu (2009), "Variety, vice, and virtue: How assortment size influences option choice." *Journal of Consumer Research* 35.6 (2009): 941-951.

Simonson, Itamar (2008), "Regarding inherent preferences." *Journal of Consumer Psychology* 18.3 (2008): 191-196.

Swaminathan, Vanitha (2003), "The impact of recommendation agents on consumer evaluation and choice: the moderating role of category risk, product complexity, and consumer knowledge." *Journal of Consumer Psychology* 13.1 (2003): 93-101.

Tellis, Gerard J., Stefan Stremersch, and Eden Yin (2003), "The international takeoff of new products: The role of economics, culture, and country innovativeness." *Marketing Science* 22.2 (2003): 188-208.

Tellis, Gerard J (1988), "Advertising exposure, loyalty, and brand purchase: A two-stage model of choice." *Journal of marketing research* (1988): 134-144.

Tripsas, Mary and Giovanni Gavetti (2000), “Capabilities, Cognition, and Inertia: Evidence From Digital Imaging,” *Strategic Management Journal*, 21, 1147-1161.

Tversky, Amos, and Daniel Kahneman (1981), "The framing of decisions and the psychology of choice." *Science* 211.4481 (1981): 453-458.

Tversky, Amos, and Daniel Kahneman (1986), "Rational choice and the framing of decisions." *Journal of business* (1986): S251-S278.

Tversky, Amos, and Itamar Simonson (1993), "Context-dependent preferences." *Management science* 39.10 (1993): 1179-1189.

Walden, Eric, and Glenn Browne (2002), "Information cascades in the adoption of new technology." *ICIS 2002 Proceedings* (2002): 40.

Watts, Duncan J., and Peter Sheridan Dodds (2007). "Influentials, networks, and public opinion formation." *Journal of consumer research* 34.4 (2007): 441-458.

Wathieu, Luc and Marco Bertini (2007), “Price as a Stimulus to Think: The Case for Willful Overpricing,” *Marketing Science*, 26(1), 118-129.

Wernerfelt, Birger (1995), "A rational reconstruction of the compromise effect: Using market data to infer utilities." *Journal of Consumer Research* (1995): 627-633.

Zhang, Juanjuan (2010), "The sound of silence: Observational learning in the US kidney market." *Marketing Science* 29.2 (2010): 315-335.

Zhu, Feng, and Xiaoquan Zhang (2010), "Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics." *Journal of marketing* 74.2 (2010): 133-148.

CURRICULUM VITAE

Ruijiao Guo

Place of birth: Hunan, China

Education

- B.B.A., Shaanxi University of Science and Technology, July 2005
  Major: Management
- M.A., Nanjing University, July 2008
  Major: Management
- Ph.D. Candidate, University of Wisconsin– Milwaukee, 2009-2015
  Major: Quantitative Marketing;
  Minor: Advanced Marketing Models/Econometrics/Bayesian Statistics
  Dissertation: "Succeeding in Extremely Competitive Markets: Insights from the Mobile App Market"

Research Interests:

- Social Media, Mobile App Market, Motion Picture Industry
- Quantitative Marketing, Marketing Analytics, Bayesian Statistics
- Online Retailing, Online Pricing, Online Competition, Online Word of Mouth

Teaching Experience:

- Lecturer at Sheldon B. Lubar School of Business, UW-Milwaukee: Teach courses such as Marketing Research, Consumer Behavior, Marketing Management, and Principles of Marketing, Jan. 2013 – May, 2015.
- Teaching Assistants at Sheldon B. Lubar School of Business, UW-Milwaukee: Assist to teach courses such as Introduction to Management Statistics, Principles of Marketing, August 2009 – December 2012

Academic Honors

- Chancellor's Graduate Student Award, UW-Milwaukee, 2013, 2014, 2015
- Sheldon B. Lubar Doctoral Scholarship, UW- Milwaukee
- Graduate Student Travel Award, Graduate School, UW-Milwaukee, 2011, 2012, 2013
- Outstanding Graduate Award, Nanjing University, 2008
- Outstanding Graduate Award, Shaanxi University of Science and Technology, Top 1, 2005; Outstanding student, People's Scholarship, Shaanxi University of Science and Technology, 2002, 2003, 2004

Volunteer Work:

- Volunteer Assistant for Classification Society 2013 Annual Meetings at UW-Milwaukee, 2013
- Secretary of UW Milwaukee Doctoral Student Marketing Research Society, 2013, 2012, 2011, 2010
- Volunteer teacher at Yili Radio and Television University, Feb – Jun, 2006, Xinjiang, China

Professional Affiliations:

- Institute for Operations Research and the Management Sciences (INFORMS)
- American Marketing Association (AMA)