

CONSUMER-CENTRIC INNOVATION FOR MOBILE APPS EMPOWERED BY SOCIAL MEDIA ANALYTICS

Zhilei Qiao

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in
partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Business Information Technology

G. Alan Wang, Co-Chair

Tabitha L. James, Co-Chair

Alan S. Abrahams

Weiguo Fan

Loren P. Rees

Wenqi Shen

Christopher W. Zobel

May 8th, 2018
Blacksburg, VA

Keywords: innovation, mobile apps, social media, online reviews, competition, deep learning

Copyright 2018, Zhilei Qiao

CONSUMER-CENTRIC INNOVATION FOR MOBILE APPS EMPOWERED BY SOCIAL MEDIA ANALYTICS

Zhilei Qiao

Abstract

Due to the rapid development of Internet communication technologies (ICTs), an increasing number of social media platforms exist where consumers can exchange comments online about products and services that businesses offer. The existing literature has demonstrated that online user-generated content can significantly influence consumer behavior and increase sales. However, its impact on organizational operations has been primarily focused on marketing, with other areas understudied. Hence, there is a pressing need to design a research framework that explores the impact of online user-generated content on important organizational operations such as product innovation, customer relationship management, and operations management. Research efforts in this dissertation center on exploring the co-creation value of online consumer reviews, where consumers' demands influence firms' decision-making. The dissertation is composed of three studies. The first study finds empirical evidence that quality signals in online product reviews are predictors of the timing of firms' incremental innovation. Guided by the product differentiation theory, the second study examines how companies' innovation and marketing differentiation strategies influence app performance. The last study proposes a novel text analytics framework to discover different information types from user reviews. The research contributes theoretical and practical insights to consumer-centric innovation and social media analytics literature.

CONSUMER-CENTRIC INNOVATION FOR MOBILE APPS EMPOWERED BY SOCIAL MEDIA ANALYTICS

Zhilei Qiao

General Audience Abstract

The IT industry, and especially the mobile application (app) market, is intensively competitive and propelled by rapid innovation. The number of apps downloaded worldwide is 102,062 million, generating \$88.3 billion in revenue, and projections suggest this will rise to \$189 billion in 2020. Hence, there is an impetus to examine competition strategies of app makers to better understand how this important market functions. The app update is an important competitive strategy. The first study investigates what types of public information from both customers and app makers can be used to predict app makers' updating decisions. The findings indicate customer provided information impacts app makers' updating decisions. Hence, the study provides insights into the importance of customer-centric strategy to market players. In the second study, it explores the impacts of product differentiation strategies on app product performance in the mobile app marketplace. The results indicate that product updates, which the first study showed are influenced by consumer feedback, are a vertical product differentiation strategy that impacts app performance. Therefore, the results from the two studies illustrate the importance of integrating online customer feedback into companies' technology strategy. Finally, the third study proposes a novel framework that applies a domain-adapted deep learning approach to categorizing and summarizing two types of innovation opportunities (i.e., feature requests) embedded in app reviews. The results show that the proposed classification approach outperforms traditional algorithms.

Dedication

To my heavenly father, for his everlasting love and limitless wisdom!

Acknowledgements

First, I would like to acknowledge and thank my co-advisors, Dr. G. Alan Wang and Dr. Tabitha L. James, for their guidance and support throughout my time in the Ph.D. program. Dr. Wang set a great example for me of how to be an excellent teacher and researcher. He devoted himself to offering great courses and conducting rigorous research that showed me how to excel in these areas. His support, patience, and trust helped me prevail throughout my time in the Ph.D. program. He is not only an advisor, but a life-long mentor. Dr. James is always open and willing to answer my questions. I am grateful for her strong support and willingness to help me develop my research skills. Her many hours spent mentoring, developing, and investing in a junior scholar are greatly appreciated.

I also owe my sincere gratitude to my committee members. Whenever I go to Dr. Weiguo Fan with a question, he always provides me with constructive suggestions and research directions. He is a fount of knowledge and wisdom. I would like to thank Dr. Alan Abrahams for his assistance in helping me secure assistance with tagging data for Chapter 3 of my dissertation. I would thank Dr. Loren Rees for helping me focus my efforts on my dissertation. My gratitude goes to Dr. Wenqi Shen for the research guidance and direction she provided. I would also like to thank Dr. Christopher Zobel because I have benefited from his experience and encouragement.

I have also greatly benefited from the support of the department of Business Information Technology. I would like to thank Dr. Cliff Ragsdale for accepting me into this wonderful Ph.D. program and supporting me during my job search. I would also like to thank Dr. Roberta Russell, as well as all the other faculty and staff in the Department of Business Information Technology, for all their support throughout my time in the Ph.D. program.

I had the pleasure of sharing this experience with an excellent group of colleagues at Virginia Tech. I would like to thank Qianzhou Du, Long Xia, Yufeng Ma, Milad Baghersad, Nohel Zaman, Rich Gruss, Josey Chacko, Yuhong Li, David Goldberg, Jamie Zhou, Shihao Zhou, and Xuan Zhang for their friendship throughout this journey.

Last but not least, I would like to thank my precious family. I owe a huge debt of gratitude to my wife, Dr. Liang Shan, for her love, assistance, encouragement, and companionship. I would also like to thank my parents, siblings, and parents-in-law for their sacrifice and everlasting love. I will be forever grateful to my kids John and Abigail, for adding so much joy to my Ph.D. life.

Table of Contents

Dedication	iv
Acknowledgements	v
List of Tables	x
List of Figures	xi
1 Introduction	1
1.1 Goals of the Dissertation	2
1.2 Organization of the Dissertation	4
1.3 The Framework of the Dissertation	7
1.4 Research Methodologies of the Dissertation	7
1.5 Scope and Limitations	8
1.6 Contributions of the Dissertation	8
2 Examining Incremental Innovation in the Mobile Application Marketplace: A Signaling Theory Perspective	10
2.1 Introduction	10
2.2 Theoretical Development	13
2.2.1 <i>Incremental Innovation and App Updating</i>	13
2.2.2 <i>Signaling Theory</i>	18
2.3 Method	31
2.3.1 <i>Model Development</i>	31
2.3.2 <i>Research Setting</i>	35
2.3.3 <i>Data Collection and Sample</i>	37
2.3.4 <i>Descriptive Statistics</i>	41
2.3.5 <i>Results</i>	41
2.4 Discussion	42

2.4.1	<i>Contributions to Research and Theory</i>	47
2.4.2	<i>Implications for Society and Practice</i>	48
2.5	Limitations and Directions for Future Research	49
2.6	Conclusions	50
2.7	References	52
3	The Impact of Product Differentiation Strategies on Mobile Application Performance: An Empirical Study	58
3.1	Introduction	58
3.2	The Mobile App Marketplace	62
3.3	Product Differentiation.....	64
3.4	Research Model.....	66
3.4.1	<i>Product Innovation Differentiation</i>	67
3.4.2	<i>Marketing Differentiation</i>	70
3.5	Research Methodology.....	73
3.5.1	<i>Data</i>	73
3.5.2	<i>Measures</i>	74
3.5.3	<i>Empirical Model</i>	77
3.5.4	<i>Results</i>	80
3.6	Conclusions and Discussions	83
3.6.1	<i>Theoretical Contributions</i>	84
3.6.2	<i>Practical Implications</i>	85
3.7	References	87
4	Deep Learning-Based USER FEEDBACK Classification in Mobile App Reviews	96
4.1	Introduction	96

4.2	Related Work.....	99
4.2.1	<i>Customer Value Co-Creation</i>	99
4.2.2	<i>Information Types in App Reviews</i>	100
4.2.3	<i>Sentence-Level User Feedback Classification</i>	101
4.2.4	<i>Word Embedding</i>	104
4.2.5	<i>Deep Learning for Text Classification</i>	105
4.3	Research Objectives	106
4.4	Research Design.....	107
4.4.1	<i>A Deep Learning-Based, Sentence-Level User Feedback Classification Framework</i>	107
4.5	Experiments.....	114
4.5.1	<i>Data Description</i>	114
4.5.2	<i>Performance Metrics</i>	115
4.5.3	<i>Experiment</i>	116
4.5.4	<i>Experimental Results</i>	119
4.6	Conclusion and Discussion	122
4.7	References	124
5	Conclusions	133
6	Bibliography.....	136

List of Tables

Table 2-1 Descriptions of Independent Variables	39
Table 2-2 Descriptive Statistics and Bivariate Correlations (n = 7,208)	43
Table 2-3 Main Effects of App Maker Provided and Consumer Provided Quality Signals and App and App Maker Credibility Signals	44
Table 2-4 Explanation of Statistical Results	45
Table 3-1 Game App Categories.....	74
Table 3-2 Types of Update Based on Keyword Matching	78
Table 3-3 Description of Variables.....	79
Table 3-4 Descriptive Statistics	81
Table 3-5 Fixed Effects Estimation on Product Performance (-Log [ProductRankit]).....	82
Table 3-6 Explanations of Statistical Analysis	83
Table 4-1 Examples of Different Types of App Reviews.....	101
Table 4-2 Literature Reviews of Sentence-level Objective Information Classification	103
Table 4-3 Notations in the Proposed Framework	112
Table 4-4 Word Embedding for Capturing Semantic Relationships between Words	119
Table 4-5 Word Embedding for Capturing Syntactic Relationships between Words	120
Table 4-6 Performance of the Proposed Framework vs. Baseline Methods.....	121

List of Figures

Figure 1-1 Research Framework.....	7
Figure 2-1 Research Model.....	32
Figure 3-1 Conceptual Model	67
Figure 3-2 Research Model.....	73
Figure 4-1 Word Embedding for Predicting Contextual Information	108
Figure 4-2 CNN for Capturing N-Grams Contextual Information	110
Figure 4-3 RNN for Capturing Word Sequence-based Contextual Information	111
Figure 4-4 Deep Learning Based User Feedback Classification Framework.....	113
Figure 4-5 Training Accuracy with Different Window Sizes (word count=5 and dimension=300)	117
Figure 4-6 Training Accuracy with Different Word Counts (window size=4 and dimension=300)	117
Figure 4-7 Training Accuracy with Different Dimensions (window size=4 and word count= 5)	118

1 INTRODUCTION

With the proliferation of ICTs, online review platforms have exploded in popularity among customers (Luo 2009). Online customer reviews represent the “wisdom of the crowd.” Existing literature shows that online customer reviews reveal information pertinent to actual customer needs that is unobtainable from traditional media (Luo et al. 2013). More importantly, user-generated content in online review platforms is updated rapidly and spreads virally at an unprecedented speed, providing first-hand information ahead of other sources. Therefore, firms now have the opportunity to utilize online user-generated content in developing their operating and decision-making strategies (Ghose and Ipeirotis 2011).

In particular, the dissertation focuses on knowledge discovery from user-generated content and its impact on organizational operations and consumer behavior. Due to the rapid development of Internet and mobile phone technologies, an increasing number of social media platforms exist where consumers can exchange comments online about the products and services businesses offer. The existing literature has demonstrated that online user-generated content can significantly influence consumer behavior and increase sales (Duan et al. 2008; Mudambi and Schuff 2010; Shen et al. 2015). However, its impact on organizational operations has been primarily focused on marketing, while other areas such as product innovation remain understudied. The dynamic environment of the mobile app marketplace provides a unique context in which to study the impact of user-generated content. Therefore, a research framework is developed to explore how online user-generated content impacts organizational operations, such as product innovation, customer relationship management, and operations management, as well as consumer behavior (e.g., product adoption).

1.1 Goals of the Dissertation

There are many potential benefits available to firms from incorporating knowledge from online consumer reviews into their operational processes. Additionally, the opportunities for processing large amounts of data provided by the rapid development of cutting-edge data mining technologies (e.g., natural language processing, machine learning, and deep learning) make it possible to more easily obtain useful information from online consumer reviews. In consideration of these two factors, I investigate how online consumer reviews influence app makers' innovation choices and design a novel text analytics framework to extract valuable information from online reviews. In this dissertation, I examine the following three research questions:

RQ1: How are online consumer reviews and app characteristics associated with the likelihood of incremental product innovation (i.e., app updates) in the mobile app marketplace?

Product innovation is important for firms to gain competitive advantage in a dynamic business environment. With the increased use of the Internet and mobile technology, online users are enabled and motivated to provide reviews and discussions about product features and use experiences. User-generated product reviews have been identified as an important component of a company's marketing mix and shown to impact sales. However, their impact on the timing of product innovation cycles has not been well studied. Given the competitive and dynamic environment of the mobile app marketplace, I suggest that the in-depth analysis of consumers' reviews and app firm-provided information could help app firms or their competitors' determine their products' updating strategies.

RQ2: How can product innovation differentiation strategies together with marketing differentiation strategies influence product performance in the competitive mobile app industry?

Product differentiation strategies have been widely considered to be important methods for firms to survive in a highly competitive industry. Empirical studies in the management and market literature have shown that product innovation differentiation strategies, also referred to as “vertical product differentiation” (Animesh et al. 2011; Beal 2000), are generally more important and useful in dynamic and competitive environments – for example, the software industry – in which products and practices change quickly. Marketing differentiation, which attempts to create a unique product image via product price, product rank and product rating, is another strategy firms use to differentiate their products. Although the impact of both product differentiation strategies in the traditional market is well known, much less is known about their impact on product performance in the mobile app marketplace.

RQ3: Can the deep-learning based approach achieve better performance in terms of accuracy than previous traditional supervised methods?

Discovering and extracting relevant information from the large amount of online reviews for a firm plays an increasingly important role in various business contexts. Although existing supervised and unsupervised machine learning algorithms can achieve good performance on some text mining tasks, traditional supervised methods primarily rely on bag-of-words and lexical features and ignore other important features (e.g., syntactic and semantic). Moreover, many traditional methods are not designed to dynamically capture new user needs and cannot be used to monitor user needs in real time. I explore the development of a deep-learning based algorithm for text mining product features from user reviews.

1.2 Organization of the Dissertation

The dissertation is composed of three studies. The first study finds empirical evidence that quality signals in online product reviews are one predictor of the timing of firms' incremental innovation. In the second study, I examine how innovation and marketing-based product differentiation strategies influence product performance. In the third study, I propose a novel text mining framework that effectively discovers and quantifies useful information from online consumer reviews. The research contributes theoretical and practical insight into customer-centric innovation and social media analytics. The major findings and contributions of the three studies are described below.

The first study, "Exploring Incremental Innovation in the Mobile Application Marketplace: A Signaling Theory Perspective," empirically examines how online consumer reviews and app characteristics are associated with the likelihood of incremental product innovation (i.e., app updates) in the mobile app marketplace. Incremental innovation plays a significant role in maintaining the competitiveness of mobile apps as a key component of the app maker's technology strategy. Building on signaling theory, I examined whether and how quality signals from both the app makers and app consumers affected app updating. The findings showed that the app updating rate was faster when the novelty of the app decreased and when customer feedback was negative or diverse. I also found that the app updating rate was slower as the app and app maker matured and when consumer feedback indicated the app was popular and of sufficiently high quality. The chance of an update was also diminished when consumer feedback was addressed in the last update. These findings indicate that marketplace signals can indeed help predict app updates and thus provide an indication of the product update strategies of market competitors.

The second study, “The Impact of Product Differentiation Strategies on Mobile App Performance: An Empirical Study,” extends the previous research and examines the effect of product differentiation strategies based on innovation and marketing information on product performance in the mobile app industry. Although the impact of product differentiation strategies in the traditional market has been well studied, much less is known about firm’s differentiation strategies in the competitive mobile app market. Mobile apps, like information goods, have negligible reproduction costs and high development and maintenance costs, and thus the mobile apps industry is highly competitive. This paper primarily focuses on whether and how the product differentiation strategies influence mobile app performance. Specifically, I employ a panel data regression model to examine the impact of the two differentiation strategies on app performance. Utilizing a data set of mobile apps from the Google Play store, I find that the quantity measure of adding new feature requests relative to its competitors has a positive effect on app performance, but the relative quantity measure of fixing bugs has no significant impact. Moreover, price and reputation differentiation strategies have a negative effect on app performance. The findings regarding the product differentiation strategies imply the need to study the impact of the operational environment on app performance together with market power and prices in one integrated framework.

The third study, “Deep Learning Based User Feedback Classification in Mobile App Reviews,” is a design science study. Guided by the seven guidelines for conducting design science research (Hevner et al. 2004), this study proposes a novel framework that employs a domain-specific deep learning approach to categorize and summarize two types of latent innovation opportunities embedded in online product reviews, namely, bug reports and feature requests. Because online users are interacting with many mobile apps under different usage contexts, user needs have

become a critical consideration in the app design process. Existing studies indicate timely and constructive online reviews from users create innovation opportunities and have become extremely crucial to help developers understand user needs. However, discovering and quantifying potential user needs from large amounts of unstructured text is a nontrivial task. In this study, I propose a *domain-oriented deep learning framework* that can discover the most critical user needs (i.e., new features of apps and bug reports) from online product reviews. I conduct a systematic evaluation including quantitative evaluations to ensure the quality of discovered information. Experimental results demonstrate that the proposed framework outperforms the baseline models and can accurately find more valuable information. The research has significant managerial implications for app developers, app customers, and app platform providers.

The research framework is summarized and illustrated in Figure 1-1. In summary, the first study can potentially help app makers develop informed updating strategies. The second study explores how app makers' product differentiation impacts product performance, which has important theoretical and managerial implications for researchers and executives. Finally, I develop a novel deep-learning based algorithm to summarize and categorize objective information in online reviews.

1.3 The Framework of the Dissertation

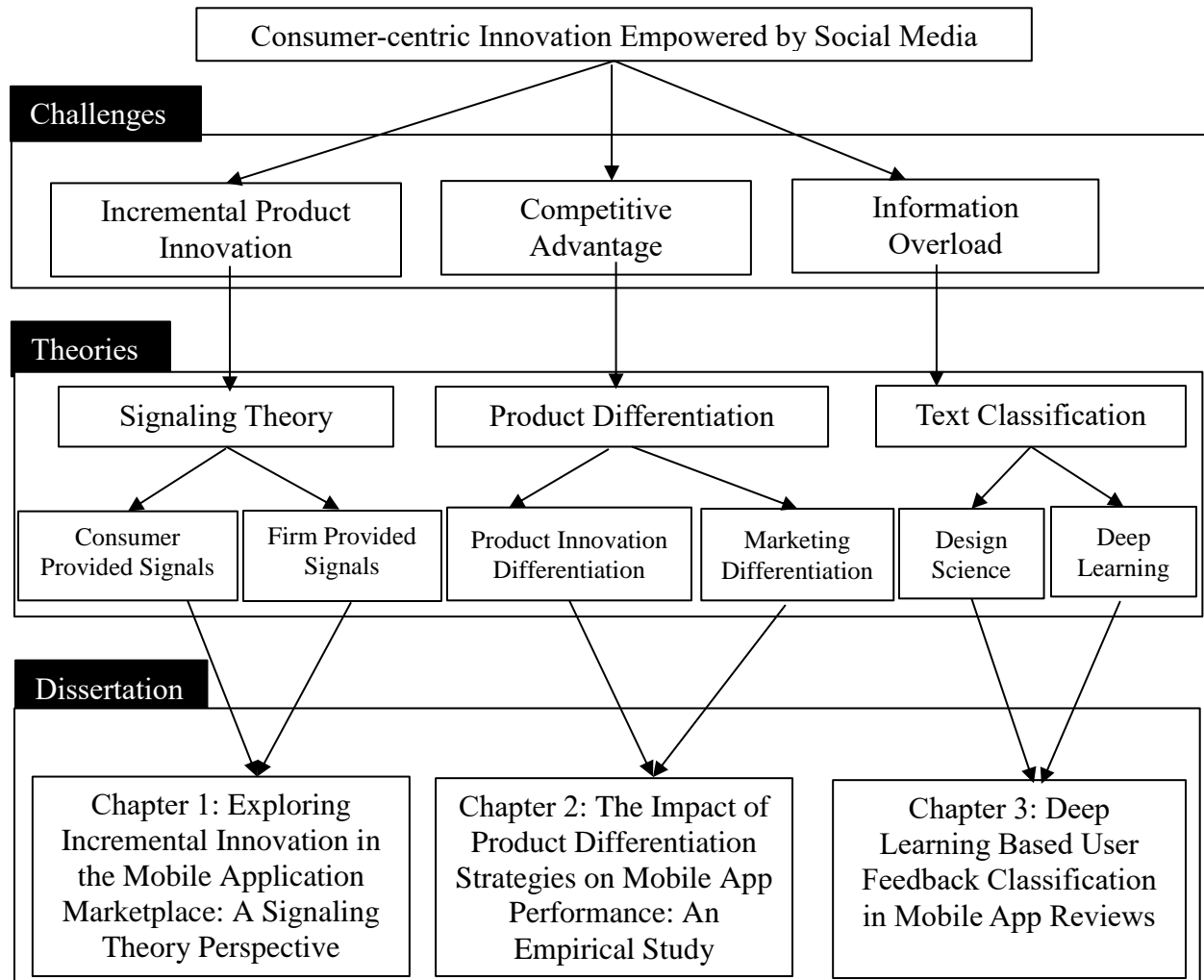


Figure 1-1 Research Framework

1.4 Research Methodologies of the Dissertation

Both empirical studies and design science research are conducted in the dissertation. The empirical studies in the dissertation include the application of survival analysis to understand how online reviews and other signals influence the timing of app updates. Survival analysis is used when an event of interest occurs more than once over the observation period. In the data set, an app update

is defined to be the event of interest and app updates occur multiple times for each app, and therefore survival analysis is used. The second study employs a panel regression model to analyze the impact of different relative (i.e., between focal products and competing products) factors on product performance. Finally, a design science research study is conducted. In particular, a novel text analytics framework is developed. It uses extensive text mining and social media analysis techniques to understand consumer demands embedded in online user-generated content.

1.5 Scope and Limitations

Product innovation has been extensively studied in the literature. To keep the scope of this study manageable, I focus on incremental product innovation in the mobile app industry. As such, the unit of analysis is at the product (app) level. The study does not examine innovation decision making at a firm level. Moreover, the dataset does not include all types of mobile apps, but is limited to primarily paid-game apps. The data set used in Chapter 2 includes only paid games in the top chart of the Google Play app store. Therefore, the findings of the study cannot be generalized to free apps or other app stores. While the deep learning model developed can capture general semantic and syntactic relations, they could be enriched in future studies to capture more micro-level analysis such as aspect or entity level analysis.

1.6 Contributions of the Dissertation

The dissertation research provides several contributions to research and theory.

- It provides insight into the strategic value of online consumer reviews on app firms' operational behaviors, with a specific focus on app updates (i.e., incremental product innovation). By focusing on consumer feedback and operational decision making in the mobile app context, it also fills a gap in the literature of app marketplaces.

- It extends the application of signaling theory in IS by examining signals from app consumers, app platforms, and app developers simultaneously.
- It extends the product differentiation perspective to examine the relationships between innovation and marketing differentiation strategies and product performance in the mobile app industry.
- It provides a novel deep learning-based user feedback classification framework to efficiently and effectively classify online reviews to possible sets of user needs for app developers.

2 EXAMINING INCREMENTAL INNOVATION IN THE MOBILE APPLICATION MARKETPLACE: A SIGNALING THEORY PERSPECTIVE

2.1 Introduction

The IT industry, and especially the mobile application (app) market, is intensively competitive and propelled by rapid innovation (Comino et al. 2016; Harter et al. 2000). The number of apps on the Apple App Store as of March 2017 was approximately 2.2 million, and in the Google Play Store approximately 2.8 million¹. These two digital app marketplaces (Ghazawneh and Henfridsson 2015) serve the primary smartphone types used today, the iOS-based iPhones (Apple App Store) and the Android-based smartphones (Google Play). The number of apps downloaded worldwide is 102,062 million generating \$88.3 billion, and projections suggest this rose to 268,692 million in 2017². For app makers competing in such a thriving marketplace, it is important for their app to remain visible to and popular amongst the consumers. In app marketplaces this is often achieved using software updating strategies that are conducted to improve quality through incremental innovation to retain old, as well as attract new, consumers in order to stay in the “top charts” (i.e., remain visible and popular). In fact, research has argued that it is vital to app makers that tend to be smaller enterprises to make the “right release decisions” (Nayebi et al. 2016, p. 552). Hence, there is an impetus to determine updating strategies of other app makers to outstrip the competition.

¹ <https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/>

² <https://www.statista.com/topics/1002/mobile-app-usage/>

Researchers exploring the app marketplaces have commented on several important features. First, app makers have only one channel (i.e., marketplace) on which to sell the apps they develop for a particular operating system, and operating on the channel is uniform for all app makers (Lee and Raghu 2014). For example, if an app maker produces a game for the iPhone, the only option it has for selling that game is through the Apple App Store and the terms and conditions for doing so are the same for every app maker. Second, app makers “have the opportunity to change not only price but also the features and characteristics of the app based on user feedback and reviews” and “mobile apps offer a greater range of flexibility to sellers in versioning strategies (e.g., feature-based or price-based differentiation, in-app purchase, subscription length, etc.)” (Lee and Raghu 2014 p. 134). This suggests that app developers are constantly innovating to better position their apps in a highly competitive single marketplace. Third, research indicates one competitive advantage in app marketplaces is to stay in the “top charts” whose apps are downloaded much more often than those not in the “top charts” (Garg and Telang 2012). One strategy to maintain status in the “top charts” is to strategically update the app in a manner that will stimulate demand at crucial moments (Comino et al. 2016).

The previous discussion illustrates that it is useful for firms in the hypercompetitive app marketplaces to utilize available information to help develop their updating strategies in order to maintain a competitive advantage. Prior research suggests that as concentration increases in competitive markets such as software, firms increasingly respond to innovation activity from external actors (Turner et al. 2010). The uniformity of the app marketplace creates a standardized and rich information environment that both the app makers and the consumers can contribute to and access. Specifically, the app maker contributes information about the app on its app information page when it is released or updated (e.g., price, app description) that signals quality to

others on the app marketplace. The consumers contribute information through assigning star ratings and writing comments regarding their perceptions of the app, both of which also signal app quality to others on the marketplace. All of this information is equally available to the app maker, its competitors, and the consumers. I argue that this information provides signals to the market regarding an app makers' updating strategies, and suggest that information obtained from observing competitor's marketplace signals can be used to inform an app maker's own incremental innovation strategy (Rose et al. 2016; Turner et al. 2010; Wen and Zhu 2016). However, while this information provides clues about when an app may update, only individuals internal to the app maker are privy to perfect information regarding the updating strategy, indicative of information asymmetry in the app marketplace. The study seeks to determine if, and which, app marketplace signals may predict app updates.

Researchers suggest that “the growth of the app market provides a great opportunity to examine important questions around software innovation, firm entry and exit strategy, software product pricing and promotion, platform leadership, and externality” (Garg and Telang 2012, p. 1254) and the study contributes to this conversation by focusing on the relationship between quality signaling and incremental software innovation. Building on signaling theory, I consider the use of quality signals to predict incremental app innovation (i.e., updating). Specifically, I explore the following research questions:

1. How are app maker quality signals (i.e., app characteristics) and credibility signals associated with incremental innovation (i.e. updating) rates in mobile apps?
2. How are consumer quality signals (i.e., online user feedback) associated with incremental innovation (i.e. updating) rates in mobile apps?

The study provides several contributions to literature and practice. First, I build on signaling theory to explore how publically available information in a highly competitive online marketplace can indicate app makers' updating strategies. Second, I extend the treatment of signals by incorporating several text-based variables into the analysis. I use text-mining techniques to develop variables that allow me to examine the characteristics of the textual quality signals from both the app makers and the consumers. This provides insight into whether textual marketplace commentary is an indicator of app makers' incremental innovation strategies. Finally, I contribute to the burgeoning text mining literature by incorporating a broad palette of text-based variables into the model. The results show how various signals predict updating strategies in one popular app marketplace. This information may be valuable to companies considering updating strategies to better position their apps in a highly competitive marketplace.

2.2 Theoretical Development

2.2.1 Incremental Innovation and App Updating

Innovation has been defined as “an idea, practice or a material artifact perceived to be new by the relevant unit of adoption” (Zaltman et al. 1973 p. 10). Different forms of innovation have been proposed (Bhoovaraghavan et al. 1996; Dewar and Dutton 1986), including the distinction between radical and incremental innovation suggested to “pertain to distinctions along a theoretical continuum of the level of new knowledge embedded in an innovation” (Dewar and Dutton 1986 p. 1423). While there is agreement on the fact that different forms of innovation exist, there has been trouble concretely defining the forms (Bhoovaraghavan et al. 1996). I adopt the definition of incremental product innovation provided in Banbury and Mitchell (1995, p. 161) as “refinements and extensions of established designs that results in substantial price or functional benefits to users,”

where they suggest that “incremental product innovation is a critically important competitive factor.” In the current paper, the dependent variable of interest is the hazard rate for app updates. The hazard rate is defined as the rate of an interest of event occurring at some time (Kleinbaum 1998). App updates equate to incremental innovation. That is, app updates reflect refinements (e.g., bug fixes) and extensions (e.g., new features) to an established app that benefit the users.

The innovation process for apps has several key properties that should be considered (Boudreau 2012). First, apps are bound to a technology platform (Ghazawneh and Henfridsson 2015; Yoo et al. 2012), which presents the possibility of network effects, where more users are drawn to it resulting in more customers for app makers as more apps are created for the platform. Second, there are vast possibilities to create novel apps. Boudreau (2012, p. 1411), citing Stoneman (2010), states that apps are “‘soft’ innovations, a class apart from the output of more usual industrial innovation,” and citing Zittrain (2006) describes “innovation in digital outputs as possessing ‘generativity,’ a tendency to expansive novelty” (Boudreau 2012, p. 1411). Third, app development benefits from code reuse and recombination that allows for small changes to be implemented with relative ease that have a large effect on the product. Software development has been described as being “characterized by rapid sequential innovation, reuse, and recombination” (Cohen and Lemley 2001, p. 3). Furthermore, researchers have suggested that software platforms such as iOS encourage generativity by “enabling continual reinterpretations, expansions and refinements of products and services” (Yoo et al. 2010, p. 14). Fourth, Boudreau (2012, p. 1412) argues the app market is characterized by uncertainty that may lead to “many failed ‘experiments’ and only isolated successes.” Fifth, platform development for apps is “explicitly intended to reduce development costs, speed product creation, reduce coordination costs, and facilitate experimentation by enabling reuse of core design elements” (Boudreau 2012, p. 1412), which

means that apps can be rather cheaply and easily built. Finally, the low cost of entry and the reduction in skill needed lowers the barriers to entry into the app market. Hence, incremental innovation in the app market is crucial and differs in many key respects from innovation in other products and services.

The app market is highly competitive and fast-paced, and hence, of utmost importance to companies producing apps is the determination of how to attract new customers and retain existing ones. One approach is to strategically update the app (i.e., incrementally innovate) in a manner that will stimulate demand at the crucial moments to keep the app in the “top charts” (i.e., maintain the app’s visibility). For example, one author states “practitioners and developers are well aware that managing app updates (i.e., releasing new versions of an existing app) is critical to increase app visibility and to keep users engaged, disguising a hidden strategy to stimulate downloads” (Comino et al. 2016, p. 1). Further reinforcing this pattern, another study found that an app’s survival in the top 300 chart of the Apple App Store was improved threefold by app updates (Lee and Raghu 2014). High numbers of downloads will keep an app in the “top charts,” which contain the most visible apps in the marketplace. In fact, research has suggested that the top ranked, paid iPhone app may garner more than 150 times more downloads than the paid iPhone app ranked at 200 (Garg and Telang 2012). Even without the stimulus of remaining in the “top charts”, pacing the product updates of high-technology products is an important consideration, with researchers suggesting that introducing new generations quickly may lead to premature cannibalization of the old product and high development costs, whereas a slow pace may result a failure to “capitalize on customer willingness-to-pay for improved technology” (Druehl et al. 2009, p. 621).

While strategically managing app updates may be used as a tactic to improve visibility, which may lead to more downloads and improved rankings, what is not very well known is whether

updating strategies can be determined from the information publicly available on the app marketplace. Being able to use such information to determine the updating strategies of the competition may be useful to app makers trying to stay in the “top charts.” Hence, in this paper I explore factors that may predict app makers’ incremental innovation (i.e., updating). I group these factors into two categories: 1) app maker provided quality signals – the app maker releases information about its app to the marketplace (e.g., screenshots, app description) that may, intentionally or not, signal the app maker’s updating strategy, and 2) consumer provided quality signals – the consumers provide quality information in their online feedback that may be unintentionally reflective of the app maker’s updating strategy, or suggestive of the consumers’ desired updating strategy that the app maker may be pressured to adopt.

The literature has described the cyclic development process in the mobile app market and the importance of the user to it:

The App Store marketplace can be thought of as a highly user-participatory cyclic development model that partly involves ‘requirements for the masses; requirements from the masses’. In this adaptive development space, users express their needs and desires by voting apps up and down and contributing product reviews. Users also tacitly express support for a feature by downloading apps that offer it. Developers may observe this behavior and respond accordingly by adding popular features, where appropriate, to their own products. In this way, the developers triage the perceived desires of their users and make strategic decision as to which features to adopt. (Sarro et al. 2015, p. 76)

Going even further, another team of researchers notes that “many software requirements are identified only after a product is deployed, once users have had a chance to try the software and provide feedback” (Ko et al. 2011, p. 1).

Hence, the app's development lifecycle could be interpreted as incremental innovation described in the following set of steps:

1. The firm develops the initial (or updated) app and releases it on the app marketplace with several characteristics listed on the app's information page that signal to the marketplace the quality of the app (e.g., price, description, rating, supported devices, screenshots).
2. After release, the consumers provide feedback regarding their perceptions of the app, which also signals to the marketplace the quality of the app. This feedback in the app marketplace comes in the form of star ratings, entry or exit into or from the top charts, and textual feedback describing the consumers' perceptions of the app.
3. The app is updated creating a new version to release to the app marketplace. The app maker's and consumers' quality signals may indicate the likelihood of an update to an app being released. Characteristics of the app and app maker (e.g., firm age and app portfolio diversity) that signal credibility, as well as changes to the supporting platform (i.e., iOS updates), may also indicate the app maker's updating strategy.
4. The process cycles back to Step 1 for the life of the app.

Hence, incremental innovation is crucial to the thriving app marketplace, and improving the understanding of it has important implications for both competitors and consumers of apps. First, as previously mentioned, being able to gauge the updating strategy of other app makers may provide information that competitors can use in developing their own updating strategy. Factoring such knowledge into their app updating strategies can help app makers better serve their customer base, retain visibility of their apps, and remain competitive. Second, app consumers may benefit from an improved understanding of which apps may be more aggressively updated. Such knowledge may enable the user to select apps that are more innovative, competitive, and

responsive to market desires and trends. Drawing from the properties of the innovation process for mobile apps and building on signaling theory, I develop a model to predict incremental app innovation (hazard rate for app updates) from app maker and consumer quality signals.

2.2.2 Signaling Theory

The concept of a market signal was developed in research by Michael Spence (1974) and can be described as “a way to pass information to other individuals in the marketplace” (Hoxmeier 2000, p. 119). Porter (1980, p. 75) defined a market signal as “any action by a competitor that provides a direct or indirect indication of its intentions, motives, goals, or internal situation.” Yet another definition suggests “a firm or individual credibly communicates the level of some unobservable element in a transaction by providing an observable signal,” and suggests that this may be applied to the conveyance of quality information (Kirmani and Rao 2000, p. 66). As a broad concept, market signals are suggestive of information asymmetry that informational cues can help resolve (Connelly et al. 2011). In fact, market signals have been broadly applied to judge employee potential in the hiring process (Spence 1974), as announcements or previews that signal firm actions to competitors (Heil and Robertson 1991), and as a means of conveying quality information about a product to consumers (Dawar and Parker 1994; Ippolito 1990; Kihlstrom and Riordan 1984; Kirmani and Rao 2000; Milgrom and Roberts 1986). In the information technology (IT) literature, IT features have been examined as quality signals that consumers use to make trust and participation decisions online (Benlian and Hess 2011; Mavlanova et al. 2012; Mavlanova et al. 2016). Ou and Chan (2014) examined seller and product characteristics as quality signaling mechanisms and found them to impact sales volume. Signaling theory has also been applied to consumer market moves providing cues to online purchasing intention (Cheung et al. 2014). Thus,

IT research has examined signals emanating from both the firm and the consumer. Following this literature, I explore both firm and consumer information signals on the app marketplace as predictors of incremental app innovation.

I suggest that quality information published to the app marketplace is signaling (intentionally or not) the app makers' updating strategy, or in other words, the quality signals can indicate when an app maker is likely to provide an update to the software. The context is unusual because the app makers have only a single online marketplace on which to offer their apps for a particular platform (iOS or Android). This online marketplace maintains publicly available quality information in a standardized format (app information page) from both the app makers and the consumers. Furthermore, there is an incentive for app makers to strategically release new versions of their apps at crucial times to stimulate demand to keep the app in the "top charts" where it is most visible to consumers. Hence, there is incomplete information on the marketplace (i.e., when an app maker will update their app), and an incentive for competitors to use marketplace signals to determine the updating rate in order to best position their own apps.

In the current study, I build on signaling theory to consider that two groups of signals from two different senders may indicate the updating rate to marketplace observers (e.g., app competitors or consumers). I use the app marketplace information categorization described in Lee and Raghu (2014). First, the app firm has a set of product characteristics (e.g., pricing, description, app update notes) that it must decide prior to the release of the initial or updated app that are intended to advertise the app's quality to consumers. I add to this framework credibility signals from the electronic auction literature that provide information on the maturity, success, and depth of experience of the app maker. Second, the consumers of the app respond to the quality of the app by providing their own quality perceptions through online ratings, which drive rankings, and

comments. Such feedback may also inadvertently signal the app maker's updating strategy, or may prompt the app maker to update earlier. In the following subsections, I concentrate on how these two categories of signals may predict incremental app innovation.

2.2.2.1 Predicting Incremental Innovation Using App Maker App Quality and Credibility

Signals

Benlian and Hess (2011, p. 11) state that “signaling theory applied in consumer research argues that when facing such situations of information asymmetry and difficult decisions about quality, consumers attend to particular kinds of informational cues that are actions or artifacts of businesses that credibly relay information about unobservable product (or company) quality to the consumer.” In their study, they argue that IT features act as signals to consumers of the quality of online communities and explore the relationship between the IT feature signals and online community trust and participation. Extending this logic, such IT features signal quality information to anyone using the marketplace (e.g., consumers and competitors).

Another study used signaling theory to explore quality of online commerce websites and revealed that high-quality retailers display more signals (Mavlanova et al. 2012). They classified signals as high cost (e.g., third-party seals, live chat, regulatory compliance, coupon redemption, consumer feedback) and low cost (e.g., contact information, shipping methods, privacy, return, security policies). They also differentiated between easy-to-verify (e.g., contact information, email confirmation) and difficult-to-verify signals (e.g., privacy, return security policies, coupon redemption, consumer feedback), and explored pre-purchase, purchasing, and post-purchase signals. The difficult-to-verify signals in the study refer to a number of policies that show effort on the part of the website, even if their ultimate ease-of-use may not be easy to verify until after the consumer makes a purchase. The analysis conducted by Mavlanova et al. (2012) indicates that

high-quality websites not only display more signals, they also use high cost signals and signals classified as difficult-to-verify. Following similar logic, I suggest that the information provided by the app maker on the app information page signals the app quality. Building on Mavlanova et al. (2012), I also suggest that higher quality apps will use more, and more robust, signals. While the fields on the App Store page are standardized, the app maker can choose how much information they input into the app information page and how detailed that information is. For example, the app maker can choose to have more screenshots of the app, a longer app description, more verbose update notes, and which subcategories its app will compete in.

Quality signals in the online auction environment (i.e., eBay) have been classified as direct quality indicators (e.g., picture postings, money-back guarantee, certification of product, product description), indirect quality indicators (e.g., minimum bid, hidden reserve price, hidden reserve price, BIN option), and seller credibility options (e.g., seller's rating, third-part payment, certification of seller, credit card payment, escrow service) (Li et al. 2009). Similar to the direct quality indicators in the online auction context, the app makers in the context have the opportunity to post multiple pictures (i.e., screenshots) of their app, write a description of their app, post update notes describing what was updated in a new version, and list the iOS versions and languages supported. Following Li et al. (2009), I suggest that these are direct quality indicators. One key difference between the context and the online auction context is price. In the App Store marketplace, the price is set upfront and non-negotiable, and thus it is a direct quality indicator in the context because there is no need for the consumer to speculate or adjust their valuation like there is in an online auction. Hence, I suggest that app price, number of screenshots provided, description length, content type, number of supported devices, number of hyper-competitive subcategories the app is placed in, and the update topic diversity are all direct quality indicators

because this information is purposefully considered, developed, and released by the app maker with the initial release of the app or an app update with the intent of projecting its quality in order to encourage downloads. Such informational cues, e.g., the app description, company policy, services and protections offered have been shown to influence sales, which suggests that consumers do pay attention to such signals to evaluate the app (Lee et al. 2015; Li et al. 2015; Ou and Chan 2014; Svedic 2015), and thus they serve as marketplace quality signals.

In the Apple App Store there are more than 2 million apps available as of March 2017³. The properties of the app marketplace create an environment where barriers to entry are lowered (Boudreau 2012). That is, the platform environment means that app development is purposefully made as easy as possible and the cost of developing an app is often rather negligible compared to other products and services. Both of these characteristics encourage new entries to the marketplace, and likely contributed to the uncertainty and skewed outcomes described by Boudreau (2012, p. 1412) of “many failed ‘experiments’ and only isolated successes.” This trend can be seen in the number of abandoned apps that have not been updated for long periods of time and no longer function under the latest iOS versions. This phenomenon has become such a problem in the App Store that Apple sent a letter to developers informing them that abandoned apps would be removed from the marketplace⁴. In fact, it has been reported that more than half of the apps on Apple’s App Store have not been updated in a year or more, with the games category accounting for 19% of the apps that had been abandoned⁵.

³ <https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/>

⁴ <https://techcrunch.com/2016/09/01/apple-is-going-to-remove-abandoned-apps-from-the-app-store/>

⁵ <https://sensortower.com/blog/abandoned-apps>

Research has suggested that an app maker's release strategy "is a factor that affects the ongoing success of mobile apps" and that "the recent practice of continuous delivery is a parallel attempt to simplify release planning by releasing more frequently and reducing the scope of releases" (Nayebi et al. 2016, p. 552). Furthermore, it has been argued that consumers select apps that are recently updated (Nayebi et al. 2016) and may reward apps for updating in the rankings (Martin et al. 2016; McIlroy et al. 2016). Nayebi et al. (2016) found that frequently updated apps were not perceived to include apps of low quality, which suggests that high quality apps may update more frequently. Hence, I argue that apps that make the effort to directly signal high quality (e.g., longer descriptions, more screenshots) on their app page will also take care with the maintenance of their app and release more timely and frequent updates indicative of a well-thought out release strategy. Conversely, apps that send signals of lower app quality are more likely to be apps that are abandoned. This logic suggests the following hypothesis.

H1: App makers who send signals of high app quality will update their app at a faster rate than app makers who send low direct quality signals.

Boudreau (2012) used producer-level (i.e., firm-level) scale and scope characteristics to explore innovation incentives on an early third-party app platform. The current study differs in that I am examining app-level variables. The analysis performed by Boudreau (2012, p. 1409) indicated that "adding large numbers of producers led innovation to become more dependent on population-level diversity, variation, and experimentation—while drawing less on the heroic efforts of any one individual innovator". Employing a text-mining method, I was able to calculate a similarity measure of an app's description relative to the other app's descriptions in the same subcategory. This measure allows me to explore population-level diversity and variation between apps of a subcategory and determine its impact on incremental innovation.

In describing the competitive environment for apps, Boudreau (2012, p. 1414) states that while some protections, such as patents, are available to protect intellectual property, these “protections do not necessarily extend to the protection of visual aspects, functionality, and sequences in a program’s use.” He goes on to argue that this characteristic in combination with common development tools that support quick and easy software development could lead to a situation where competitors simply copy features or concepts once they have proven valuable in the marketplace, a scenario which has also been described by Wang et al. (2015) and Li et al. (2014). Such characteristics may change the nature of competition on app platforms and make it difficult for an app to maintain product differentiation (Porter 1980) from its competition. In fact, research has explored “migratory features” and their movement between categories, supporting the idea that popular features move between apps that are functionally similar (Sarro et al. 2015). I argue that as the similarity of an app to the other apps increases, the app maker will update at a faster rate in order to maintain the competitive advantage provided by product differentiation (Porter 1980). This logic suggests the following hypothesis.

H2: Apps that are more similar to the other apps in their subcategory will update at a faster rate than apps that are less similar to the other apps in their subcategory.

Li et al. (2009) categorize one set of information available on electronic auction websites (e.g., eBay) as “seller credibility indicators.” They suggest seller credibility indicators include longer term indications of performance such as seller ratings, seller certifications, and third-party payment methods. I examine seller ratings in the current study, but I follow Lee and Raghu (2014) in classifying these as consumer responses to the app makers release decisions and explore them in the next section. However, I acknowledge that consumer feedback may also enhance the app maker’s credibility. In the context, all payment is handled through a single method provided by

Apple, so there is no variation between app makers. Yet, I suggest that there is some information in the online marketplace that indicates the maturity and experience of the app maker, and thus may be used as proxies for credibility, such as the length of time the app has been in the App Store, the length of time the app maker has had any app in the App Store, the number of apps the app maker has in the App Store, and diversity of the app maker's app portfolio (i.e., the number of different subcategories the app maker has an offering in).

Mavlanova et al. (2016) explored the role of internal and external quality signals offered by the e-commerce seller on consumers perceptions of trust, seller quality, and deception, leading to perceived product quality and finally to purchase intention. They found both types of signals, internal (e.g., posted firm policies) and external (e.g., third-party trust seals) to affect consumer perceptions, suggesting that credibility signals do have market effects. Previous research has also found that "more successful producers tend to take longer or wait longer to release new versions" (Boudreau 2012, p. 1419). Another study that surveyed both users and developers suggests that more experienced developers believe having an app updating strategy can impact user feedback (Nayebi et al. 2016). Their study found that users prefer to download apps that have been updated recently, but the findings were mixed regarding consumer preferences towards frequently updated apps. Hence, previous literature suggests that experienced and successful app makers may carefully consider their app updating strategy and plan to release updates more slowly, less frequently, and at times where they may be the most beneficial. This logic leads me to the following hypothesis.

H3: Apps with app maker credibility signals that suggest maturity, experience, and success will update at a slower rate than apps whose app maker credibility signals suggest immaturity, inexperience, and lack of success.

2.2.2.2 *Predicting Incremental Innovation Using Consumer Quality Signals*

The involvement of the consumer has often been cited as crucial to innovation (Arakji and Lang 2007; Chatterji and Fabrizio 2014; Dahlander and Frederiksen 2012; Di Gangi and Wasko 2009; Kankanhalli et al. 2015; Nambisan and Baron 2007; Rose et al. 2016; von Hippel and Katz 2002; von Hippel and von Krogh 2003), new product development (NPD) (Chan et al. 2015; Fuchs and Schreier 2011; Füller et al. 2006; Hoyer et al. 2010; Nambisan 2002), marketing (Chen and Xie 2008; Cheung et al. 2003; Zhu and Zhang 2010), and quality assurance, process control, and improvement (Abrahams et al. 2015; Dellarocas 2003; Griffin and Hauser 1993; Yeung et al. 2005). The role of the consumer has garnered even more interest with the emergence of public online forums where consumers can easily relate their opinions to firms and other consumers. For example, Fuchs and Schreier (2011, p. 17) state that “many have advocated the idea of democratizing innovation by empowering customers to take a much more active stake in corporate NPD. This has become feasible because the Internet now allows companies to build strong online communities through which they can listen to and integrate thousands of customers from all over the world.”

Using signaling theory, Cheung et al. (2014) explored the role of both consumer action signals, operationalized as peer consumer purchases, and consumer opinion signals, operationalized as peer consumer reviews, in influencing purchase decisions. In the model, I similarly consider consumer action signals because ranking information is based on downloads of the app. I also consider consumer opinion signals in the form of both the star ratings assigned to an app and the opinions (i.e., textual feedback) regarding the app left by the consumers. However, rather than purchase decisions as the dependent variables, I am interested in how such consumer feedback may predict incremental innovation in the app market.

Many studies have examined the relationships between consumer ratings, consumer commentary, and product sales (Chevalier and Mayzlin 2006; Chintagunta et al. 2010; Duan et al. 2008; Ghose and Ipeirotis 2011; Gu et al. 2012; Li et al. 2015; Ou and Chan 2014). This line of work has found that better consumer reviews can improve the sales of some products, such as books sold online, and that negative ratings may be more influential than positive ones (Chevalier and Mayzlin 2006), but that the influence of reviews may change or diminish over time (Hu et al. 2008; Li and Hitt 2008). Research has also explored the impact of volume and valence (ratings) on, for example, box office performance of movies (Chintagunta et al. 2010) and found both to matter depending on the market examined, while others have found volume but not valence to have an impact (Duan et al. 2008). Studies have also explored the impact of consumer reviews on competition (Kwark et al. 2014). Prior research has also shown that adapting strategy, such as pricing or offering frills, can impact consumer ratings and profits (Kuksov and Xie 2010), which suggests that it may be valuable to an app maker to consider consumer feedback in developing their updating strategy.

While consumer feedback has been shown to influence other consumers' purchase decisions, less attention has been paid to how it may influence the decisions made by the firm. Li and Hitt (2008, p. 456) do suggest that "firms could benefit from altering their marketing strategies such as pricing, advertising, or product design to encourage consumers likely to yield positive reports to self-select into the market early and generate positive word-of-mouth for new products." Similarly, updating an app could be stimulated by undesirable consumer feedback in hopes of correcting future consumer feedback. Decker and Trusov (2010) suggest that consumer feedback could be aggregated to determine consumer preferences that could be used to support product development and product improvement processes. One study stated that for iTunes apps, "worsening of app

performance increases the likelihood of releasing minor updates only” and interpreted this finding as suggesting “that minor updates are better suited to be used as a strategic tool to counter poor past performances” (Comino et al. 2016, p. 6). This suggests that app makers whose consumer feedback indicates popularity and happiness with the quality of the app may release more extensive updates that add new product functionality and features, and therefore would take longer to develop. Whereas, companies that are experiencing waning popularity and feedback expressing quality concerns may release minor updates that rectify complaints or show that the company is trying to address concerns. This logic leads me to the following hypothesis.

H4: Apps whose consumer feedback signals popularity and high quality will be updated at a slower rate than apps whose consumer feedback signals waning popularity and low quality.

Yeung et al. (2005) explored quality management in a high-technology industry (electronics) and found that a customer focus led to time-based efficiency, which was important to customer satisfaction. Their items to measure customer focus dealt with “acquiring customer information”, “analyzing customers’ feedback”, and “working with customers in product design” (Yeung et al. 2005, p. 195). They argued that it is necessary for volatile technologies in highly competitive marketplaces with limited negotiation power to be customer-oriented, and thus, time-based efficiencies such as delivery timeliness and manufacturing lead-times become paramount (Yeung et al. 2005). These findings suggest that consumer feedback in a high-technology market may result in shorter updating intervals in order to quickly address quality issues. In fact, updating on the app marketplace is quite frequent, with studies finding apps being updated on the Apple App Store on average every 58 days and on Google Play every 13 days on average (Comino et al. 2016). This study also found performance gains for apps that introduced more robust updates that included new features rather than just bug fixes (Comino et al. 2016). I argue that an app whose

consumer feedback contains a robust discussion of different topics suggests that consumers are interested in the app and that the app has ample opportunities to produce updates that add features or rectify complaints that will please their customers. Thus, apps with diverse feedback may update faster in order to incrementally address consumers desires and complaints in order to encourage positive consumer sentiment and maintain consumer interest. Whereas, consumer feedback with less diversity of topics being discussed may indicate stable apps or apps that are not as interesting to consumers, and such apps may be updated more slowly. This logic leads to the following hypothesis.

H5: Apps whose consumer feedback contains discussion of a diverse the array of topics will be updated at a faster rate than apps that have less topics discussed in their consumer feedback.

There is some agreement that user needs are an important factor to consider in the innovation process (Abrell et al. 2016; Arakji and Lang 2007; Bhoovaraghavan et al. 1996; Di Stefano et al. 2012; Griffin and Hauser 1993). For example, Bhoovaraghavan et al. (1996 p. 234) state that “an interesting perspective stems from the product life-cycle theory, wherein innovative products in the early stages of the product life cycle are in many ways not well adapted to the needs of the user. As one progresses through the product life cycle, the awareness of the product becomes greater and, consequently, incremental changes are the catalysts for achieving higher sales and returns.” Similarly, Quintas (1994), drawing on Rosenberg (1972) explicates that “diffusion of innovation is accompanied by continuous process improvements and modifications which tailor the innovation to specific needs or improve the performance of the original. Thus the diffusion process generally leads to further innovation as technologies are extended to meet new applications.”

Prior research has indicated the importance of being considerate of consumers requests and concerns in the technology industry (Yeung et al. 2005) and in the app market specifically (Lee and Raghu 2016; Nayebi et al. 2016). Nayebi et al. (2016) surveyed app developers and found that slightly over 50% of their respondents had strategies for app updating and believed it to be important to consumers. Other research has indicated the depth of information that may be gleaned from consumer feedback (Abrahams et al. 2015), and in particular from consumer feedback for apps (Khalid et al. 2015). Additionally, Hoxmeier (2000) determined that producing software that provided promised functionality was more important than timely delivery. Hence, it may be more important to provide requested features, functionality, and fixes than it is to provide fast but minor updates. Thus, these areas of research suggest that some app makers may factor consumer comments into their updating strategies, and perhaps do so in an attempt to earn goodwill from their customers by complying with their requests and responding to their complaints. I argue that this suggests that app firms that carefully fulfill consumer demands may update more slowly, and their carefully considered updates may provide them with quality improvements that provide a buffer before new updates need to be addressed. Whereas app makers that are less responsive to the concerns and desires of their consumers may be releasing faster minor updates that are less carefully planned. This logic leads to the following hypothesis.

H6: Apps whose app makers are less responsive to the consumers' feedback will be updated at a faster rate than those apps whose app makers are more responsive to consumers' feedback.

2.3 Method

The relationships between the quality signals and incremental app innovation proposed in the hypotheses presented above are summarized in Figure 2-1. In the sections that follow, I provide details of the model development, research setting, data collection, sample, and analysis.

2.3.1 Model Development

To model the effect of the theoretical factors on the rate of app updates, I denote the beginning time point of the observation as *time 0* and employ a survival model that accounts for (1) right censored observations (Kleinbaum 1998), (2) the influence of covariates (most of them are time variant) and control variables (Mallapragada et al. 2012), (3) unobserved heterogeneity across product clusters (Aalen 1987), (4) unobserved heterogeneity for recurrent update events (Box-Steffensmeier and De Boef 2006), and (5) event dependence among recurrent update events (Box-Steffensmeier and De Boef 2006). In the model, the random variable T represents the time to the next app update when the hazard rate (i.e., probability of app update at time T), if the first update in the sampling period has not occurred, is $h(T) = \frac{f(T)}{S(T)}$, where $f(T)$ is the probability density function and $S(T)$ is the survival function for T , $S(T) = 1 - F(t)$, and $f(T) = \frac{d(F(T))}{dt}$ with $F(T)$ as the cumulative distribution function (Kleinbaum 1998).

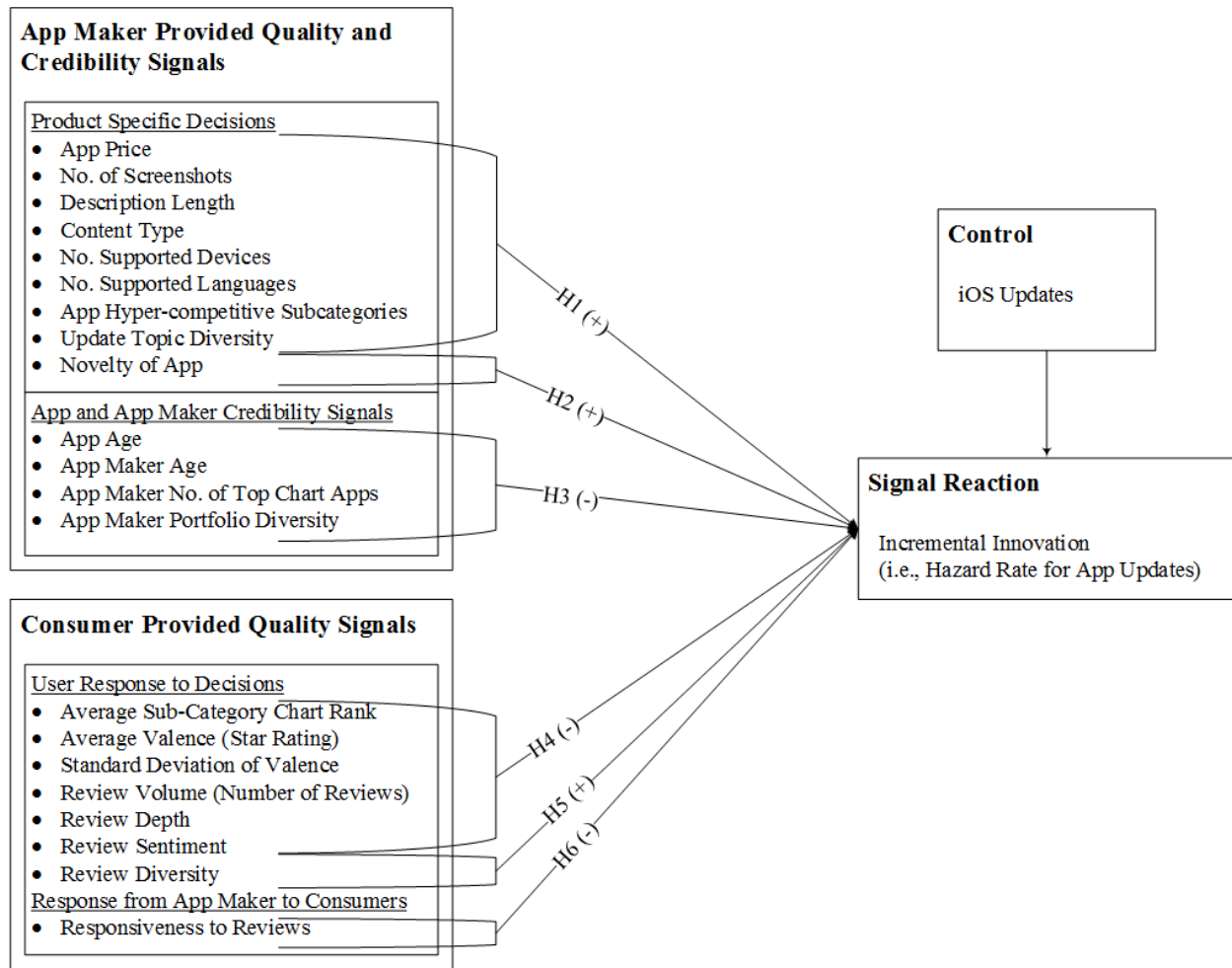


Figure 2-1 Research Model

2.3.1.1 Right censored observations

Of concern in the model is the possible inclusion of censored observations (i.e., observations in which the exact time between updates for some apps is not observed due to when the data collection ends). The existing literature shows three types of censored observations: right censored, left censored, and interval censored. However, here I am only concerned with right censored observations because I can observe the exact time between events if the event occurred before the end of the observation period. To deal with the right censored observations and use the minimum assumptions, I employ a Cox Proportional Hazard model, which is commonly used in management science and economics (Arora et al. 2010). In addition, the Cox model is preferred to a logistic

model because it can use more information including survival times and censoring data, whereas the logistic model ignores survival times and censoring data and only considers a dichotomous dependent variable (0, 1).

Let $h_i(t)$ indicate the hazard rate of an update of app i at time t . Then, the Cox Proportional Hazard model is

$$h_i(t, X_i, \beta) = h_0(t) \exp(X_i, \beta), \quad (1)$$

where $h_0(t)$ is the baseline hazard rate at time t , β is the vector of coefficients that will be estimated, and X_i is a covariate vector.

2.3.1.2 Model Covariates and Controls

The covariate vector is denoted as $X_i = (X_i^F, X_i^{CUS}, X_i^C)$, where X_i^F represents the app maker provided quality and credibility signals, X_i^{CUS} represents the customer provided quality signals, and X_i^C is the market characteristic, i.e., control variable. The covariates may change from one interval to the next but are treated as constant covariates within an update interval. I examine the effects of app maker provided quality and credibility signals (X_i^F), customer provided quality signals (X_i^{CUS}), and the control (iOS updates) (X_i^C) on the hazard rate. I incorporate these effects by modifying the hazard model function as follows:

$$h_i(t, X_i, \beta', \gamma', v') = h_0(t) \exp(\beta' X_i^F + \gamma' X_i^{CUS} + v' X_i^C), \quad (2)$$

where β' captures the effect of app maker provided quality and credibility signals, γ' is the main effect of customer provided quality signals, and v' is the effect of the control variable X_i^C .

2.3.1.3 Unobserved Heterogeneity across Products

To control for unobserved heterogeneity across different apps, I also run a frailty hazard model. Some mobile apps are intrinsically more or less prone to updating than are others, and the distribution of these effects can be estimated. The frailty hazard model function is specified as

$$\begin{aligned} h_i(t, X_i, \beta', \gamma', v', \omega_i) &= h_0(t) \exp(\beta' X_i^F + \gamma' X_i^{CUS} + v' X_i^C + \omega_i) \\ &= h_0(t) e_i \exp(\beta' X_i^F + \gamma' X_i^{CUS} + v' X_i^C), \end{aligned} \quad (3)$$

where $e_i = \exp(\omega_i)$ is the unobserved frailty and is conventionally assumed to follow a gamma distribution with a mean of 1 and an unknown variance σ^2 that needs to be estimated. The frailty model is similar to a “random effects” model, where a random variable e is fixed within the product of interest but varies across different products in terms of the distribution. To control for product specific unobserved heterogeneity, I let e be fixed within apps. The idea is that that app makers may have specific policies or strategies that will influence that app makers’ decisions that cannot be observed.

2.3.1.4 Unobserved Heterogeneity for Recurrent Update Events

Because each app may have recurrent update events, I have to capture the random effects across a more general unit level heterogeneity among the events within an app. In this study, I have K observations for I apps. The hazard rate for the k th event for the i th app is

$$\begin{aligned} h_{ik}(t, X_{ik}, \beta', \gamma', v', \omega'_i) &= h_0(t) \exp(\beta' X_{ik}^F + \gamma' X_{ik}^{CUS} + v' X_{ik}^C + \omega'_i) \\ &= h_0(t) e'_i \exp(\beta' X_{ik}^F + \gamma' X_{ik}^{CUS} + v' X_{ik}^C) \end{aligned} \quad (4)$$

here $e'_i = \exp(\omega'_i)$ is the random effect across recurrent update events of the i th app, and they are constant over time. There is only one single update event for each value of the random effect, so that it is shared over time by a single update event, rather than shared across apps. This introduces

heterogeneity across recurrent update events and produces within product correlation in the occurrence and time of recurrent events within a given app.

2.3.1.5 Event Dependence among Recurrent Update Events

The occurrence of one app update event may influence the probability of further events' occurrences. The dependence among events may be caused by endogenous positive or negative effects. Either of the scenarios indicates that the risk for an app update event is a function of the occurrence of a previous event. Therefore, this dependence relationship also generates some within subject correlation. To characterize the relationship, existing studies show that event based stratification (varying baseline hazards) can be used.

$$\begin{aligned} h_{ik}(t, X_{ik}, \beta', \gamma', v', \omega'_i) &= h_{0k}(t) \exp(\beta' X_{ik}^F + \gamma' X_{ik}^{CUS} + v' X_{ik}^C + \omega'_i) \\ &= h_{0k}(t) e_i' \exp(\beta' X_{ik}^F + \gamma' X_{ik}^{CUS} + v' X_{ik}^C) \end{aligned} \quad (5)$$

here $h_{0k}(t)$ is the baseline hazard rate and differs by event number.

2.3.2 Research Setting

In this study, I explore incremental innovation (i.e., app updating) in the mobile app marketplace, specifically the Apple App Store that is used to distribute mobile apps compatible with the iOS platform. As one of the two major app marketplaces, it is an appropriate context for the study. Game apps account for a large majority of revenue on app marketplaces, with some statistics indicating that games account for more than 90% of the revenue on the Google Play marketplace⁶. Mobile phone users have also been reported to spend the majority of their app time in games –

⁶ <http://www.androidauthority.com/2016-recap-90-percent-google-play-revenue-gaming-fun-stats-743626/>

32% of the time spent in apps⁷. Hence, I chose to focus on the Game category and its subcategories that garner the most user interest.

At the time of the data collection, the Apple App Store published ranking lists for several major categories of apps (e.g., Games), as well as ranking lists for subcategories of these major category headings (example subcategories of the Games category include “Action” and “Family,” “Arcade,” “Card,” “Puzzle”). These lists were in addition to the “Top Free” and “Top Paid” app ranking lists. The “Top Free” and “Top Paid” ranking lists include the top 100 apps in the free and paid categories based on the number of downloads, whereas the ranking list for a major category (e.g., Games) or a subcategory includes the top 240 apps, based on the number of downloads, classified under that category or subcategory.

The Apple App Store offers an app information page for every app. The app information page contains a variety of fields that the app maker fills out when it places its new or updated app in the marketplace (e.g., app description, app subcategories, screenshots of the app, update notes). Consumer can post reviews for any app and these reviews consist of two components: 1) a star rating on a scale of 1 to 5 with 5 being the best, and 2) textual feedback (i.e., comments) regarding their perceptions of the app. The consumer can leave a star rating without any textual feedback. The app information page contains all of the individual consumer star rating assignments along with the textual feedback (i.e., comments) provided by the consumer, as well as an average of all the consumer star ratings.

⁷ <http://www.adweek.com/digital/mobile-users-spend-86-time-apps-32-gaming-17-facebook/>

2.3.3 Data Collection and Sample

Because the ranking lists are updated daily, I crawled the “Top Free,” “Top Paid,” “Top Game,” and the top game subcategory lists, as well as their app information page on the Apple App Store, each day over the timespan of the data collection. However, while I crawled the “Top Free” and “Top Paid” lists daily to retain the ranking information for the games in these lists, the analysis included only apps in the “Games” category.

The data collection spanned 298 days, from February 4th, 2015 to November 29th, 2015. The data from the 18 game subcategories contained app information page, chart ranks, and update data from 4,215 game apps that appeared in at least one of the charts (i.e., “Top Free,” “Top Paid,” “Top Game,” and top subcategory lists). Not all apps provided information for all fields in the app information page, which excluded 796 apps due to left censoring issues arising from the missing data fields. Furthermore, 1,515 of the apps did not have an update during the 298-day observation period and had to be excluded to meet the model parameters. Another 230 apps did not have consumer reviews and were excluded. Finally, 97 apps had been removed from the marketplace before the end of the observation period and were also removed from the sample. The final dataset included data from 1,577 apps from 844 unique app makers. These 1,577 apps had a total of 2,397,564 consumer reviews from approximately 2 million unique reviewers.

The data is collected over time, and therefore, one app may have multiple update events over the course of the data collection. Thus, the unit of observation is one update interval for one app. The final dataset of 1,577 apps had 7,208 usable update events during the observation period. I capture data for all the variables at the start of each updating period for an app. Hence, I have the app information at the beginning of each update interval and at the end of each update interval (the beginning of the subsequent update interval).

The primary dependent variable in the study is the hazard rate for an app update. The hazard rate integrates information on whether the update event happened and the update interval, measured by the number of days from one update of the app to the next update of the app. In the observation period, 72.12% of the apps in the final data set released more than one update. The hazard rate function (Hoang and Rothaermel 2010) is given by

$$\begin{aligned}
 h(t) &= \frac{f(t)}{S(t)} = \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t \cdot S(t)} = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t \leq T \leq t + \Delta t)}{\Delta t \cdot \Pr(T > t)} \\
 &= \lim_{\Delta t \rightarrow 0} \frac{\Pr(t \leq T \leq t + \Delta t | T > t)}{\Delta t} \tag{6}
 \end{aligned}$$

where T is the time spent at risk of an update event occurring, and the probability Pr indicates the possibility of experiencing an update event during the time interval from t to $t + \Delta t$, conditional on the app being at risk of having an update event happen at time t .

The independent variables derived from the data collection are defined in Table 2-1 and organized by their classification in Figure 2-1. I used text mining tools to calculate several of the independent variables in Table 2-1. To calculate similarity for the variable *Novelty of App*, I collected all product descriptions for each subcategory and construct a binary vector (V) with a length of the count of unique words used in all product descriptions for the subcategory. For a given app i , I build a new vector P_i in which a new element is 1 if it occurs in the vector (V) and zero otherwise. According to the literature, I restrict my words to nouns, verbs, and adjectives because they best represent the apps functionality and features (Ghose and Ipeirotis 2011). Thus, I can define the normalized frequency vector F_i to be

$$F_i = \frac{P_i}{\sqrt{P_i \cdot P_i}} \tag{7}$$

The measure of similarity between game subcategory j and app i is calculated by the dot product of their normalized vectors:

$$Product\ Similarity_{i,j} = (F_i \cdot F_j) \quad (8)$$

The values for the *Responsiveness to Reviews* variable are similarly constructed, but in this case the similarity is calculated between all the consumer reviews from the start of the app update interval to the end and the update notes the app maker releases at the end of the app update interval (i.e., start of the next update interval). To construct the *Update Topic Diversity* and *Review Diversity* variables, I count the number of unique noun or noun phrases to represent the number of different topics (Wei et al. 2008). To extract the noun and noun phrases from the reviews and update notes, I use the Stanford NLP⁸ tool to conduct Part-of-Speech tagging analysis. To construct the *Review Sentiment* variable, I use SentiStrength, which is a lexicon-based classifier that uses supplementary (non-lexical) linguistic information and rules to identify sentiment strength in short informal English text. For each comment, the SentiStrength output has two integer values ranging from 1 to 5, one for positive sentiment strength and the other for negative sentiment strength (Thelwall et al. 2010). I calculate the sentiment variable by subtracting the negative sentiment strength from the positive sentiment strength.

Table 2-1 Descriptions of Independent Variables

Variable	Definition
App Maker Provided Quality Signals	
App Price	A binary variable indicating if the app is free or paid at time t .
No. of Screenshots	A count of the number of screenshots of the app provided by the app maker on the app's page at time t .
Description Length	A count of the number of words in the description of the app provided by the app maker on the app's page at time t .

⁸ <https://nlp.stanford.edu/software/>

Content Type	A categorical variable indicating the age restriction on the app (“4+”, “9+”, “12+”, and “17+”) at time t .
No. of Supported Devices	A count of the number of devices (e.g., iPhone, iPad) the app supports at time t .
No. of Supported Languages	A count of the number of languages (e.g., English, Spanish) the app supports at time t .
App Hyper-competitive Subcategories	Count of the number of the hyper-competitive subcategories (6 subcategories that represent 70% of the game downloads) to which the app belongs at time t .
Update Topic Diversity	A count of the number of unique nouns or noun phrases present at time t in the update notes for the previous update of the app.
Novelty of App	A measure of similarity of the unique words in the descriptions of an app and all the other apps in its subcategory (where a larger similarity value means an app is more similar to all the other apps in the subcategory) at time t .
App and App Maker Credibility Signals	
App Maker Age	An integer variable representing the number of months the app maker has had an app in the App Store at time t .
App Maker No. of Top Chart Apps	How many apps an app maker has in the “top charts” of the App Store at time t .
App Maker Portfolio Diversity	A count of the number of game subcategories the app maker has at least one app in at time t .
Consumer Provided Quality Signals	
Average Subcategory Chart Rank	The average rank over the update interval of the app in its primary subcategory top chart.
Average Valence (Star Rating)	The average star rating for the app during the update interval.
Std. Dev. of Valence	The standard deviation of the star rating for the app during the update interval.
Review Volume	A count of the number of reviews for an app during the update interval.
Review Depth	The average length, in number of words, of the reviews from the current update interval.
Review Sentiment	A calculated variable in the range of 1 to 5, where 1 indicates positive sentiment and 5 indicates negative sentiment.
Review Diversity	A count of the number of unique nouns or noun phrases present in the reviews of the app during the current update interval.
Responsiveness to Reviews	A measure of the similarity between the reviews for an app and the update notes for the previous update interval (where a larger similarity value means the reviews are more similar to the update notes).

iOS Updates	A count of the number of iOS updates at time t since the last app update.
-------------	---

2.3.4 Descriptive Statistics

Table 2-2 presents descriptive statistics and the correlation matrix for all the variables in the model. The mean of the app update hazard rate is 0.02 and the average app update interval is around 50 days. The iOS was updated eight times during the observation period. The mean of iOS update is 0.58, which means an update of iOS occurs during 58% of the app update intervals in the dataset. A mean app maker age is 43.28 months, and the average app age is 26 months. The mean review valence (star rating) is 3.82 and most app reviews are positive. The average review volume is 333 reviews for an app update interval. The correlation among independent variables is low in almost all cases. However, the variation inflation factor (VIF) of the independent variables' regression result even for those with higher correlations is less than five, suggesting that multicollinearity is not an issue (Larose and Larose 2015).

2.3.5 Results

To estimate the models in Equations (1), (2), (3), (4) and (5), I characterize the baseline function $h_0(t)$ in the nonparametric form suggested in the Cox proportional hazard model (Cox 1975). Nonparametric forms require no assumptions about the baseline hazard function and have very few restrictions. I report estimates for all models by the maximum likelihood method in Table 3. Estimates for the shared frailty and without frailty models are also provided in Table 2-3. The shared frailty parameter (θ) is significant, denoting that controlling for unobserved heterogeneity across recurrent update events within an app is necessary and effective (Box-Steffensmeier and De Boef 2006; Mallapragada et al. 2012). In both models, the parameters estimates are similar. I

report hazard ratios, which are easier to interpret (the hazard ratio is $\exp(\beta')$). If a hazard ratio is greater than one, the covariate (X_i) will increase the hazard (instantaneous probability) of an app update; while a hazard ratio is less than one, the covariate (X_i) will decrease the hazard (instantaneous probability) of an app update.

2.4 Discussion

An explanation of each result in relation to the hypotheses is provided in Table 2-4. The statistical results generally support the hypotheses. I hypothesized that apps with app maker provided information that signaled high quality would have an increased likelihood of an app update, which was found to be true in all but two cases. The *No. of Supported Devices* had a negative relationship with the hazard rate of app updates, suggesting that as the number of supported devices increased, the likelihood of an app update decreased. One explanation for this finding may simply be in the effort required to develop and test an app release for multiple devices. The *App Price* was found to have no relationship with the hazard rate of app updates, which may be explained by the small variation in price of apps in general.

Table 2-2 Descriptive Statistics and Bivariate Correlations (n = 7,208)

	Mean	SD	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
1.App Update	46.71	57.36	0.00	298.00	1																							
2.App Price	0.04	0.20	0.00	1.00	0.06	1																						
3.App Hyper-competitive Subcategories	0.67	0.72	0.00	2.00	-0.17	0.11	1																					
4.iOS Updates	0.58	0.49	0.00	1.00	0.73	0.00	-0.13	1																				
5.Content Type	7.74	4.18	4.00	17.00	-0.38	-0.19	-0.25	-0.27	1																			
6.Description Length	272.08	136.29	0.00	759.00	-0.38	-0.01	0.03	-0.28	0.13	1																		
7.Novelty of App	0.00	0.00	0.00	0.01	-0.45	-0.07	0.00	-0.34	0.29	0.58	1																	
8.Update Topic Diversity	11.24	13.88	0.00	141.00	-0.09	-0.05	-0.53	-0.06	0.01	0.20	-0.01	1																
9.No. of Supported Devices	18.75	1.92	3.36	25.00	-0.37	-0.15	-0.01	-0.24	0.20	0.32	0.23	0.11	1															
10.No. of Supported Languages	6.22	7.14	0.00	40.00	0.34	-0.05	-0.26	0.22	-0.11	-0.31	-0.28	0.12	-0.2	1														
11.No. of Screenshots	4.80	0.58	0.00	5.00	-0.30	-0.10	0.12	-0.22	-0.02	0.13	0.04	-0.18	0.15	-0.34	1													
12.App Maker No. of Top Chart Apps	9.84	13.85	1.00	80.00	-0.37	-0.14	0.07	-0.27	0.09	0.38	0.27	-0.01	0.18	-0.06	0.16	1												
13.App Maker Age	43.28	20.75	0.00	93.00	0.24	0.13	-0.02	0.11	-0.20	-0.02	-0.16	-0.04	-0.16	-0.02	-0.08	-0.15	1											
14.App Maker Portfolio Diversity	2.52	2.15	1.00	12.00	-0.15	0.01	0.23	-0.12	-0.12	0.31	0.00	-0.10	0.06	-0.35	0.19	0.09	0.29	1										
15.Average Subcategory Chart Rank	101.98	70.18	1.00	240.00	-0.11	0.01	0.25	-0.08	-0.14	0.28	-0.03	-0.12	0.03	-0.33	0.20	0.07	0.25	0.98	1									
16.App Age	26.00	16.79	0.00	90.00	0.11	0.06	0.02	0.10	0.02	-0.18	-0.08	-0.06	-0.14	-0.07	-0.16	-0.17	-0.30	-0.36	-0.32	1								
17.Average Valence (Star Rating)	3.82	0.84	1.00	5.00	-0.09	-0.18	0.08	-0.03	-0.10	-0.17	-0.05	-0.11	-0.04	-0.08	0.00	0.03	-0.13	-0.16	-0.21	0.03	1							
18.Std. Dev. of Valence	1.20	0.44	0.00	2.00	-0.08	0.06	0.04	-0.10	0.17	0.18	0.12	-0.09	0.13	-0.05	0.07	0.07	-0.03	0.19	0.23	-0.21	-0.76	1						
19.Review Volume	332.63	1618.61	1.00	83471.00	0.03	-0.04	0.05	-0.01	-0.12	-0.14	-0.12	-0.10	0.00	-0.04	-0.05	-0.08	-0.18	-0.02	-0.03	-0.22	0.19	-0.21	1					
20.Review Depth	26.55	13.68	1.00	216.60	-0.01	0.00	0.07	-0.04	-0.10	-0.08	-0.09	-0.09	0.06	-0.08	-0.02	-0.05	-0.18	0.05	0.03	-0.29	0.13	-0.13	0.98	1				
21.Review Sentiment	0.87	0.55	3.00	3.00	-0.07	-0.19	-0.02	-0.02	-0.10	-0.17	-0.05	-0.06	-0.09	-0.04	-0.05	0.03	-0.10	-0.17	-0.22	0.04	0.94	-0.72	0.11	0.05	1			
22.Review Diversity	5.80	3.00	0.00	53.50	-0.22	0.12	-0.06	-0.18	0.23	0.28	0.18	0.15	0.26	-0.09	0.03	0.10	-0.01	0.11	0.14	-0.07	-0.79	0.67	-0.24	-0.14	-0.74	1		
23.Responsiveness to Reviews	0.05	0.10	0.00	1.00	0.20	0.01	-0.17	0.08	-0.14	-0.27	-0.20	0.01	-0.51	0.13	-0.13	-0.18	0.03	-0.14	-0.12	-0.03	-0.12	0.04	-0.05	-0.08	-0.08	-0.05	1	

Table 2-3 Main Effects of App Maker Provided and Consumer Provided Quality Signals and App and App Maker Credibility Signals

	(1) Without Frailty		(2) With Heterogeneity		(3) Shared Frailty		(4) Event Dependence		(5) Event Dependence With Shared Frailty	
	Hazard ratio	Std. error	Hazard ratio	Std. error	Hazard ratio	Std. error	Hazard ratio	Std. error	Hazard ratio	Std. error
App Maker Provided Quality Signals										
<i>App Price</i>	1.037	0.076	1.037	0.076	1.130	0.089	1.089	0.077	1.149	0.083
<i>No. of Screenshots</i>	1.135***	0.029	1.135**	0.029	1.187***	0.035	1.058+	0.030	1.051	0.034
<i>Description Length</i>	1.000*	0.000	1.000	0.000	1.001***	0.000	1.000	0.000	1.000*	0.000
<i>Content Type</i>	1.029***	0.003	1.029***	0.003	1.050***	0.004	1.02***	0.004	1.032***	0.004
<i>No. of Supported Devices</i>	0.952***	0.008	0.952***	0.008	0.900***	0.010	0.966***	0.008	0.937***	0.010
<i>No. of Supported Languages</i>	1.013***	0.002	1.013***	0.002	1.012***	0.003	1.001	0.002	1.002	0.002
<i>App Hyper-competitive Subcategories</i>	1.183***	0.021	1.183***	0.021	1.348***	0.026	1.105***	0.022	1.213***	0.025
<i>Update Topic Diversity</i>	1.001	0.001	1.001	0.001	0.993***	0.002	1.001	0.001	0.997+	0.002
<i>Novelty of App</i>	1.054***	0.014	1.054*	0.014	1.087**	0.018	1.028+	0.015	1.058*	0.017
App and App Makers Credibility Signals										
<i>App Age</i>	0.981***	0.001	0.981***	0.001	0.969***	0.001	0.985***	0.001	0.974***	0.001
<i>App Maker Age</i>	0.997**	0.001	0.997*	0.001	0.996**	0.001	0.997***	0.001	0.995***	0.001
<i>App Maker No. of Top Chart Apps</i>	0.980***	0.002	0.980***	0.002	0.967***	0.003	0.990***	0.002	0.980***	0.002
<i>App Maker Portfolio Diversity</i>	1.085***	0.013	1.085***	0.013	1.174***	0.016	1.041**	0.014	1.099***	0.015
Consumer Provided Quality Signals										
<i>Average Valence (Star Rating)</i>	0.920**	0.027	0.920**	0.027	0.748***	0.031	0.862***	0.027	0.764***	0.030
<i>Std. Dev. of Valence</i>	0.825***	0.039	0.825***	0.039	0.654***	0.046	0.847***	0.039	0.729***	0.044
<i>Review Volume</i>	0.999***	0.000	0.999***	0.000	0.999***	0.000	0.999***	0.000	0.999***	0.000
<i>Review Depth</i>	1.089*	0.038	1.089+	0.038	1.165**	0.005	0.986**	0.004	0.984**	0.005
<i>Review Sentiment</i>	0.982***	0.005	0.982***	0.005	0.985**	0.045	1.104*	0.039	1.143**	0.044
<i>Review Diversity</i>	1.103***	0.020	1.103***	0.020	1.069**	0.024	1.065**	0.020	1.064**	0.023
<i>Responsiveness to Reviews</i>	0.542***	0.152	0.542**	0.152	0.528***	0.176	0.751+	0.156	0.678*	0.175
Control										
<i>iOS Updates</i>	0.296***	0.034	0.296***	0.034	0.381***	0.036	0.3063***	0.034	0.345***	0.036
σ (Frailty Parameter)					0.738***	0.197			0.420***	0.221
Likelihood Ratio Test		-39640.24		-39640.24		-37971.72		-26298.79		-25300.61

+p≤0.1, *p <0.05, **p <0.01, ***p <0.001

Table 2-4 Explanation of Statistical Results

Variable	Hypothesis	Explanation
App Price	Not Supported	Whether the app is provided free of charge or for a fee does not impact the likelihood of an app update.
No. of Screenshots	(+) Supported	As the number of screenshots on the app page increases, the likelihood of an app update increases.
Description Length	(+) Supported	As the length of the description increases, the likelihood of an app update increases.
Content Type	(+) Supported	As the age restriction on the app increases, the likelihood of an app update increases.
No. of Supported Devices	(-) Significant in Opposite Direction	As the number of devices the app supports increases, the likelihood of an app update decreases.
No. of Supported Languages	(+) Supported	As the number of languages the app supports increases, so does the likelihood of an app update.
App Hyper-competitive Subcategories	(+) Supported	As the number of hyper-competitive subcategories an app belongs to increases, the likelihood of an app update increases.
Update Topic Diversity	(+) Supported	As the count of unique nouns or noun phrases in the update notes of the last period increases, so does the likelihood of an app update.
Novelty of App	(+) Supported	As the similarity of an app to the other apps in its subcategory increases, the likelihood of an app update increases.
App Age	(-) Supported	As the age of the app increases, the likelihood of an app update decreases.
App Maker Age	(-) Supported	As the age of the firm increases, the likelihood of an app update decreases.
App Maker No. of Top Chart Apps	(-) Supported	As the number of apps the app maker has in the top charts increases, the likelihood of an app update decreases.
App Maker Portfolio Diversity	(+) Supported	As the number of game subcategories an app maker has an app in increases, so does the likelihood of an app update.
Average Subcategory Chart Rank	(-) Supported	As the average subcategory chart rank increases, the likelihood of an app update decreases.
Average Valence (Star Rating)	(-) Supported	As the average star rating of the app during an update interval increases, the likelihood of an app update decreases.

Std. Dev. of Valence	(-) Supported	As the standard deviation of the star ratings of the app during an update interval increases, the likelihood of an app update decreases.
Review Volume	(-) Supported	As the number of reviews for the app during an update interval increases, the likelihood of an app update decreases.
Review Depth	(-) Supported	As the average length of the reviews for the app during an update interval increases, the likelihood of an app update decreases.
Review Sentiment ⁹	(+) Supported	As the sentiment of the customer reviews becomes more negative, the likelihood of an update increases.
Review Diversity	(+) Supported	As the count of unique nouns or noun phrases in the customer reviews from the last period increases, so does the likelihood of an app update.
Responsiveness to Reviews	(-) Supported	As the similarity between the reviews and the update notes from the last period increases, the likelihood of an app update decreases.
iOS Updates	(-) Supported	As the number of iOS updates increases, the likelihood of an app update decreases.

The findings revealed that the less novel an app was (i.e., the more similar it was to the other apps in its subcategory), the more likely it was to update, supporting the second hypothesis. The more mature the app and the app maker, the more successful the app maker, and the more diverse the app maker's portfolio of apps was, the less likely it was to update. These findings support the third hypothesis. Concerning consumer provided quality signals, I hypothesized that apps whose consumer feedback signaled popularity and high quality would be updated at a slower rate, which the results supported. Thus, the fourth hypothesis was supported. Finally, I hypothesized that apps whose reviews had a large variety of topics would update at a faster rate, and apps whose app updates were more responsive to the consumers' reviews would update at a slower rate. The results confirm the last two hypotheses.

⁹ Note that because of the scaling for the sentiment variable the hazard rate is negative, but the logic is still that as the reviews get better (more positive), the updating rate will slow.

2.4.1 Contributions to Research and Theory

The study makes several contributions to research and theory. First, I use marketplace data to explore incremental app innovation, which helps fill a gap in the literature of app marketplaces. Previous researchers have stated that “the growth of the app market provides a great opportunity to examine important questions around software innovation, firm entry and exit strategy, software product pricing and promotion, platform leadership, and externality” (Garg and Telang 2012, p. 1254), and the study contributes to this call to action. Second, I extend previous applications of signaling theory in IT by examining signals from two sources in the same model. Taking this approach allowed me to examine how the different sets of signals complemented each other to uncover app incremental innovation strategy. Third, I contribute to the body of research that has investigated innovation in the app marketplace (e.g., Boudreau 2012; Li et al. 2014; Nayebi et al. 2016), by examining incremental app innovation (i.e., app updating) 1) at an app-level, 2) with a focus on text-based variables, and 3) using both app maker and consumer provided quality signals. I find that publicly available information on the app marketplace can predict app updating rates, and such knowledge may be used to help inform strategy decisions. Finally, I incorporate several novel text-based variables that are not typically used in research (i.e., extend past valence and volume) that allow me to contribute richly to the burgeoning text-mining literature. From a theoretical perspective, these novel text-based variables let me explore app similarity, topic diversity, and gauge the responsiveness of the app maker to the consumers’ feedback. These variables contribute knowledge to the innovation literature that has been less widely examined.

2.4.2 Implications for Society and Practice

One implication of the findings for app makers is that they would be well advised to strategically plan their app updates factoring in information about their competitors that can be obtained on the marketplace. Such knowledge can help app makers strategically plan their updates to maintain or accrue positions in the “top charts” and respond agilely to consumers’ market-wide desires and complaints (Harter et al. 2000; Roberts and Grover 2012), which may help app firms pursue competitive advantages by finding and releasing new features quickly. Research has suggested that consumers prefer to download recently updated apps, and may not penalize frequent updating (McIlroy et al. 2016; Nayebi et al. 2016). In fact, well-managed updating may be rewarded by consumers (Martin et al. 2016; McIlroy et al. 2016; Nayebi et al. 2016), further reinforcing the need for app makers to strategically plan updates. Moreover, research suggests that features do migrate through the app marketplaces and that it may not be easy to maintain product differentiation with other apps because of the ease with which preferred features can be implemented into competitors’ apps (Boudreau 2012; Li et al. 2014; Sarro et al. 2015; Wang et al. 2015). Prior research also suggests that firms who first introduce an incremental innovation enjoy the greatest competitive advantage, and that while followers enjoyed some increased market share, the more followers the more advantage the first firm enjoys (Banbury and Mitchell 1995). This suggests that there is an advantage to being the first to introduce new features. The results show that the updating rate of a competitor can be predicted using market signals, and this information can be used in app makers’ strategic planning of their own app updates.

The findings also have practical implications for consumers. That negative and diverse consumer feedback predicts a faster updating rate suggests that app makers are sensitive to the requests and complaints of their users. This means that consumers should be encouraged to provide

constructive feedback to app makers as it may be crucial to the innovative processes on app marketplaces where competition is fierce and innovation fast-paced (Comino et al. 2016; Harter et al. 2000). However, high volume, positive commentary and rankings appear to slow updating rates, which may indicate either stable app makers with well-managed update regimens or that app makers in comfortable positions feel less motivation to quickly provide new features. Prior research has suggested that consumers like to comment on products that many others have already commented upon, but that they also prefer to post on less successful and available products (Dellarocas et al. 2010). The results suggest the latter may be most instrumental to stimulating change in apps. Overall, the findings suggest that consumer feedback is indeed an integral factor to incremental app innovation, and that both competitors and consumers can extract useful information signaling the future of the app from it.

2.5 Limitations and Directions for Future Research

The study is limited by the scope of its data that limits its generalizability. I examined data from only the games category of the Apple App Store. Future research may want to explore the generalizability of the findings across different categories of apps or in both the Apple App Store and Google Play. The two major app marketplaces have been found to have some distinct characteristics that may differentiate the two (Comino et al. 2016; Ghazawneh and Henfridsson 2015; Roma et al. 2016).

The study heavily relied on text mining methods to construct variables that explored the similarity of apps within subcategories, the diversity of the text for both the app descriptions and the consumer comments, the similarity between the update notes and the consumer feedback, and the sentiment of the consumer feedback. However, future studies could focus on more granular

analysis of the features requested or the complaints made by the consumers (Chan et al. 2015; Decker and Trusov 2010), and the relationship of these to incremental app innovation. Prior studies have found that consumers tend to complain about, for example, errors, privacy, hidden costs, interface design, feature removal or request features to be added (Khalid et al. 2015). Such information could be combined with knowledge of app performance and usage patterns from an app's own customer base (Chen et al. 2017; Gómez et al. 2017) and knowledge regarding app feature preferences among consumers (Ghose and Han 2014) to make decisions regarding what features to add, update, or remove in order to remain competitive in the app marketplace. While the results indicate that certain consumer feedback may spur app updates, and thus be important for app makers to consider in determining competitive updating strategies, interesting future research may want to consider the impact of particular feature requests or complaints.

2.6 Conclusions

I explored incremental app innovation using signaling theory. The findings reveal what marketplace information may provide insight into competitors' app updating strategies. The results indicate that app makers signaling high quality in the information provided to their app information pages will update at a faster rate than those that do not, and therefore these apps should be watched the most closely by competitors in timing their own updating strategies in hopes of maintaining ranking position or experiencing a first mover advantage. The results also confirm that maintaining product differentiation is a predictor of app updating. In other words, apps that are more similar to others in their category will update at a faster rate. However, app and app maker experience, maturity, and popularity all signal apps that will update at a slower rate. This suggests that older apps, experienced app makers with more "top chart" apps and a bigger portfolio, as well as popular

apps in the rankings with good and numerous reviews may feel less rushed to provide updates or already have long-term, repetitive updating strategies that are managed and slower (Nayebi et al. 2016). However, if the consumer feedback is negative or the topics discussed in the feedback numerous, the apps will update at a faster rate. A high number of topics discussed in the consumer feedback may suggest either a plethora of complaints or requests, either one of which it is likely in the best interest of the app maker to capitalize upon. The results also illustrate that apps whose makers are responsive to the feedback received from the consumers (i.e., their update notes reflect the consumers' requests) also update at a slower rate, which suggests that being responsive to consumers may pay off in providing some time before there is pressure to update again. Hence, the findings suggest that app makers may be considerate of consumer feedback and realize its importance to their innovativeness, competitiveness, and survival, a suggestion that has also been made in prior research (Chan et al. 2015; Gerstheimer and Lupp 2004). Savvy competitors may therefore benefit not only from using the textual feedback to determine their competitors' updating strategies, but may also be able to determine the features the competitors may add by mining consumer comments. Future research should continue to investigate in more detail the benefits of being responsive to consumer feedback, especially with respect to innovation.

2.7 References

- Aalen, O. O. 1987. "Two Examples of Modelling Heterogeneity in Survival Analysis," *Scandinavian Journal of Statistics* (14:1), pp. 19-25.
- Abrahams, A. S., Fan, W., Wang, G. A., Zhang, Z. J., and Jiao, J. 2015. "An Integrated Text Analytic Framework for Product Defect Discovery," *Production and Operations Management* (24:6), pp. 975-990.
- Abrell, T., Pihlajamaa, M., Kanto, L., vom Brocke, J., and Uebernickel, F. 2016. "The Role of Users and Customers in Digital Innovation: Insights from B2b Manufacturing Firms," *Information & Management* (53:3), pp. 324-335.
- Arakji, R. Y., and Lang, K. R. 2007. "Digital Consumer Networks and Producer-Consumer Collaboration: Innovation and Product Development in the Video Game Industry," *Journal of Management Information Systems* (24:2), pp. 195-219.
- Arora, A., Krishnan, R., Telang, R., and Yang, Y. 2010. "An Empirical Analysis of Software Vendors' Patch Release Behavior: Impact of Vulnerability Disclosure," *Information Systems Research* (21:1), pp. 115-132.
- Banbury, C. M., and Mitchell, W. 1995. "The Effect of Introducing Important Incremental Innovations on Market Share and Business Survival," *Strategic Management Journal* (16:SI), pp. 161-182.
- Benlian, A., and Hess, T. 2011. "The Signaling Role of It Features in Influencing Trust and Participation in Online Communities," *International Journal of Electronic Commerce* (15:4), pp. 7-56.
- Bhoovaraghavan, S., Vasudevan, A., and Chandran, R. 1996. "Resolving the Process Vs. Product Innovation Dilemma: A Consumer Choice Theoretic Approach," *Management Science* (42:2), pp. 232-246.
- Boudreau, K. J. 2012. "Let a Thousand Flowers Bloom? An Early Look at Large Numbers of Software App Developers and Patterns of Innovation," *Organization Science* (23:5), pp. 1409-1427.
- Box-Steffensmeier, J. M., and De Boef, S. 2006. "Repeated Events Survival Models: The Conditional Frailty Model," *Statistics in Medicine* (25:20), pp. 3518-3533.
- Chan, H. K., Wang, X., Lacka, E., and Zhang, M. 2015. "A Mixed - Method Approach to Extracting the Value of Social Media Data," *Production and Operations Management* (25:3), pp. 568-583.
- Chatterji, A. K., and Fabrizio, K. R. 2014. "Using Users: When Does External Knowledge Enhance Corporate Product Innovation?," *Strategic Management Journal* (35:10), pp. 1427-1445.
- Chen, Q., Zhang, M., and Zhao, X. 2017. "Analysing Customer Behaviour in Mobile App Usage," *Industrial Management & Data Systems* (117:2).
- 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. 477-491.
- Cheung, C. M., Xiao, B. S., and Liu, I. L. 2014. "Do Actions Speak Louder Than Voices? The Signaling Role of Social Information Cues in Influencing Consumer Purchase Decisions," *Decision Support Systems* (65), pp. 50-58.
- Cheung, K.-W., Kwok, J. T., Law, M. H., and Tsui, K.-C. 2003. "Mining Customer Product Ratings for Personalized Marketing," *Decision Support Systems* (35:2), pp. 231-243.

- Chevalier, J. A., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research* (43:3), pp. 345-354.
- Chintagunta, P. K., Gopinath, S., and Venkataraman, S. 2010. "The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation across Local Markets," *Marketing Science* (29:5), pp. 944-957.
- Cohen, J. E., and Lemley, M. A. 2001. "Patent Scope and Innovation in the Software Industry," *California Law Review* (89:1), pp. 1-57.
- Comino, S., Manenti, F. M., and Mariuzzo, F. 2016. "To Upgrade or Not to Upgrade? The Release of New Versions to Survive in the Hypercompetitive App Market," *International Workshop on App Market Analytics (WAMA 2016)*.
- Connelly, B. L., Certo, S. T., Ireland, R. D., and Reutzel, C. R. 2011. "Signaling Theory: A Review and Assessment," *Journal of Management* (37:1), pp. 39-67.
- Cox, D. R. 1975. "Partial Likelihood," *Biometrika* (62:2), pp. 269-276.
- Dahlander, L., and Frederiksen, L. 2012. "The Core and Cosmopolitans: A Relational View of Innovation in User Communities," *Organization Science* (23:4), pp. 988-1007.
- Dawar, N., and Parker, P. 1994. "Marketing Universals: Consumers' Use of Brand Name, Price, Physical Appearance, and Retailer Reputation as Signals of Product Quality," *The Journal of Marketing* (58:2), pp. 81-95.
- Decker, R., and Trusov, M. 2010. "Estimating Aggregate Consumer Preferences from Online Product Reviews," *International Journal of Research in Marketing* (27:4), pp. 293-307.
- Dellarocas, C. 2003. "The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms," *Management Science* (49:10), pp. 1407-1424.
- Dellarocas, C., Gao, G., and Narayan, R. 2010. "Are Consumers More Likely to Contribute Online Reviews for Hit or Niche Products?," *Journal of Management Information Systems* (27:2), pp. 127-158.
- Dewar, R. D., and Dutton, J. E. 1986. "The Adoption of Radical and Incremental Innovations: An Empirical Analysis," *Management Science* (32:11), pp. 1422-1433.
- Di Gangi, P. M., and Wasko, M. 2009. "Steal My Idea! Organizational Adoption of User Innovations from a User Innovation Community: A Case Study of Dell Ideastorm," *Decision Support Systems* (48:1), pp. 303-312.
- Di Stefano, G., Gambardella, A., and Verona, G. 2012. "Technology Push and Demand Pull Perspectives in Innovation Studies: Current Findings and Future Research Directions," *Research Policy* (41:8), pp. 1283-1295.
- Druehl, C. T., Schmidt, G. M., and Souza, G. C. 2009. "The Optimal Pace of Product Updates," *European Journal of Operational Research* (192:2), pp. 621-633.
- Duan, W., Gu, B., and Whinston, A. B. 2008. "Do Online Reviews Matter?—an Empirical Investigation of Panel Data," *Decision Support Systems* (45:4), pp. 1007-1016.
- Fuchs, C., and Schreier, M. 2011. "Customer Empowerment in New Product Development," *Journal of Product Innovation Management* (28:1), pp. 17-32.
- Füller, J., Bartl, M., Ernst, H., and Mühlbacher, H. 2006. "Community Based Innovation: How to Integrate Members of Virtual Communities into New Product Development," *Electronic Commerce Research* (6:1), pp. 57-73.
- Garg, R., and Telang, R. 2012. "Inferring App Demand from Publicly Available Data," *MIS Quarterly* (34:4), pp. 1253-1264.

- Gerstheimer, O., and Lupp, C. 2004. "Needs Versus Technology—the Challenge to Design Third-Generation Mobile Applications," *Journal of Business Research* (57:12), pp. 1409-1415.
- Ghazawneh, A., and Henfridsson, O. 2015. "A Paradigmatic Analysis of Digital Application Marketplaces," *Journal of Information Technology* (30:3), pp. 198-208.
- Ghose, A., and Han, S. P. 2014. "Estimating Demand for Mobile Applications in the New Economy," *Management Science* (60:6), pp. 1470-1488.
- Ghose, A., and Ipeirotis, P. G. 2011. "Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics," *IEEE Transactions on Knowledge and Data Engineering* (23:10), pp. 1498-1512.
- Gómez, M., Adams, B., Maalej, W., Monperrus, M., and Rouvoy, R. 2017. "App Store 2.0: From Crowdsourced Information to Actionable Feedback in Mobile Ecosystems," *IEEE Software* (34:2), pp. 81-89.
- Griffin, A., and Hauser, J. R. 1993. "The Voice of the Customer," *Marketing Science* (12:1), pp. 1-27.
- Gu, B., Park, J., and Konana, P. 2012. "Research Note—the Impact of External Word-of-Mouth Sources on Retailer Sales of High-Involvement Products," *Information Systems Research* (23:1), pp. 182-196.
- Harter, D. E., Krishnan, M. S., and Slaughter, S. A. 2000. "Effects of Process Maturity on Quality, Cycle Time, and Effort in Software Product Development," *Management Science* (46:4), pp. 451-466.
- Heil, O., and Robertson, T. S. 1991. "Toward a Theory of Competitive Market Signaling: A Research Agenda," *Strategic Management Journal* (12:6), pp. 403-418.
- Hoang, H., and Rothaermel, F. T. 2010. "Leveraging Internal and External Experience: Exploration, Exploitation, and R&D Project Performance," *Strategic Management Journal* (31:7), pp. 734-758.
- Hoxmeier, J. A. 2000. "Software Preannouncements and Their Impact on Customers' Perceptions and Vendor Reputation," *Journal of Management Information Systems* (17:1), pp. 115-139.
- Hoyer, W. D., Chandy, R., Dorotic, M., Krafft, M., and Singh, S. S. 2010. "Consumer Cocreation in New Product Development," *Journal of Service Research* (13:3), pp. 283-296.
- Hu, N., Liu, L., and Zhang, J. J. 2008. "Do Online Reviews Affect Product Sales? The Role of Reviewer Characteristics and Temporal Effects," *Information Technology and Management* (9:3), pp. 201-214.
- Ippolito, P. M. 1990. "Bonding and Nonbonding Signals of Product Quality," *Journal of Business* (63:1), pp. 41-60.
- Kankanhalli, A., Ye, H., and Teo, H. H. 2015. "Comparing Potential and Actual Innovators: An Empirical Study of Mobile Data Services Innovation," *MIS Quarterly* (39:3), pp. 667-682.
- Khalid, H., Shihab, E., Nagappan, M., and Hassan, A. E. 2015. "What Do Mobile App Users Complain About?," *IEEE Software* (32:3), pp. 70-77.
- Kihlstrom, R. E., and Riordan, M. H. 1984. "Advertising as a Signal," *Journal of Political Economy* (92:3), pp. 427-450.
- 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.
- Kleinbaum, D. G. 1998. "Survival Analysis, a Self - Learning Text," *Biometrical Journal* (40:1), pp. 107-108.

- Ko, A. J., Lee, M. J., Ferrari, V., Ip, S., and Tran, C. 2011. "A Case Study of Post-Deployment User Feedback Triage," *ACM, Proceedings of the 4th International Workshop on Cooperative and Human Aspects of Software Engineering*, Waikiki, Honolulu, HI, USA, pp. 1-8.
- Kuksov, D., and Xie, Y. 2010. "Pricing, Frills, and Customer Ratings," *Marketing Science* (29:5), pp. 925-943.
- Kwark, Y., Chen, J., and Raghunathan, S. 2014. "Online Product Reviews: Implications for Retailers and Competing Manufacturers," *Information Systems Research* (25:1), pp. 93-110.
- Larose, D. T., and Larose, C. D. 2015. *Data Mining and Predictive Analytics*: John Wiley & Sons.
- Lee, G., and Raghu, T. 2014. "Determinants of Mobile Apps' Success: Evidence from the App Store Market," *Journal of Management Information Systems* (31:2), pp. 133-170.
- Lee, G., Raghu, T., and Park, S. 2015. "Do App Descriptions Matter? Evidence from Mobile App Product Descriptions."
- Lee, G.-w., and Raghu, T. 2016. "The Role of Quality in Mobile App Markets," *37th International Conference on Information Systems*, Dublin.
- Li, B., Singh, P., and Wang, Q. 2014. "Zoom in Ios Clones: Examining the Antecedents and Consequences of Mobile App Copycats," *International Conference on Information Systems*, Auckland, New Zealand.
- Li, H., Fang, Y., Wang, Y., Lim, K. H., and Liang, L. 2015. "Are All Signals Equal? Investigating the Differential Effects of Online Signals on the Sales Performance of E-Marketplace Sellers," *Information Technology & People* (28:3), pp. 699-723.
- Li, S., Srinivasan, K., and Sun, B. 2009. "Internet Auction Features as Quality Signals," *Journal of Marketing* (73:1), pp. 75-92.
- Li, X., and Hitt, L. M. 2008. "Self-Selection and Information Role of Online Product Reviews," *Information Systems Research* (19:4), pp. 456-474.
- Mallapragada, G., Grewal, R., and Lilien, G. 2012. "User-Generated Open Source Products: Founder's Social Capital and Time to Product Release," *Marketing Science* (31:3), pp. 474-492.
- Martin, W., Sarro, F., and Harman, M. 2016. "Causal Impact Analysis for App Releases in Google Play," *ACM, Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering*, Seattle, WA, pp. 435-446.
- Mavlanova, T., Benbunan-Fich, R., and Koufaris, M. 2012. "Signaling Theory and Information Asymmetry in Online Commerce," *Information & Management* (49:5), pp. 240-247.
- Mavlanova, T., Benbunan-Fich, R., and Lang, G. 2016. "The Role of External and Internal Signals in E-Commerce," *Decision Support Systems* (87), pp. 59-68.
- McIlroy, S., Ali, N., and Hassan, A. E. 2016. "Fresh Apps: An Empirical Study of Frequently-Updated Mobile Apps in the Google Play Store," *Empirical Software Engineering* (21:3), pp. 1346-1370.
- Milgrom, P., and Roberts, J. 1986. "Price and Advertising Signals of Product Quality," *Journal of Political Economy* (94:4), pp. 796-821.
- Nambisan, S. 2002. "Designing Virtual Customer Environments for New Product Development: Toward a Theory," *Academy of Management Review* (27:3), pp. 392-413.

- Nambisan, S., and Baron, R. A. 2007. "Interactions in Virtual Customer Environments: Implications for Product Support and Customer Relationship Management," *Journal of Interactive Marketing* (21:2), pp. 42-62.
- Nayebi, M., Adams, B., and Ruhe, G. 2016. "Release Practices for Mobile Apps--What Do Users and Developers Think?," IEEE, *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, Osaka, Japanpp. 552-562.
- Ou, C. X., and Chan, K. C. 2014. "Developing a Competitive Edge in Electronic Markets Via Institutional and Social Based Quality Signaling Mechanisms," *Information & Management* (51:5), pp. 532-540.
- Porter, M. E. 1980. *Competitive Strategy: Techniques for Analyzing Industries and Competitors*, New York, NY: The Free Press.
- Quintas, P. 1994. "A Product-Process Model of Innovation in Software Development," *Journal of Information Technology* (9:1), pp. 3-17.
- Roberts, N., and Grover, V. 2012. "Leveraging Information Technology Infrastructure to Facilitate a Firm's Customer Agility and Competitive Activity: An Empirical Investigation," *Journal of Management Information Systems* (28:4), pp. 231-270.
- Roma, P., Zambuto, F., and Perrone, G. 2016. "The Role of the Distribution Platform in Price Formation of Paid Apps," *Decision Support Systems* (91), pp. 13-24.
- Rose, J., Jones, M., and Furneaux, B. 2016. "An Integrated Model of Innovation Drivers for Smaller Software Firms," *Information & Management* (53:3), pp. 307-323.
- Rosenberg, N. 1972. "Factors Affecting the Diffusion of Technology," *Explorations in Economic History* (10:1), pp. 3-33.
- Sarro, F., Al-Subaih, A. A., Harman, M., Jia, Y., Martin, W., and Zhang, Y. 2015. "Feature Lifecycles as They Spread, Migrate, Remain, and Die in App Stores," IEEE, *2015 IEEE 23rd International Requirements Engineering Conference (RE)*, Ottawa, ON, Canada, pp. 76-85.
- Spence, A. M. 1974. *Market Signaling: Informational Transfer in Hiring and Related Screening Processes*, Cambridge, MA: Harvard University Press.
- Stoneman, P. 2010. "Soft Innovation: Economics, Design, and the Creative Industries," Oxford University Press, Oxford, UK.
- Svedic, Z. 2015. "The Effect of Informational Signals on Mobile Apps Sales Ranks across the Globe," Unpublished dissertation: *Beedie School of Business*, Simon Fraser University, Canada.
- Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., and Kappas, A. 2010. "Sentiment Strength Detection in Short Informal Text," *Journal of the American Society for Information Science and Technology* (61:12), pp. 2544-2558.
- Turner, S. F., Mitchell, W., and Bettis, R. A. 2010. "Responding to Rivals and Complements: How Market Concentration Shapes Generational Product Innovation Strategy," *Organization Science* (21:4), pp. 854-872.
- von Hippel, E., and Katz, R. 2002. "Shifting Innovation to Users Via Toolkits," *Management Science* (48:7), pp. 821-833.
- von Hippel, E., and von Krogh, G. 2003. "Open Source Software and the "Private-Collective" Innovation Model: Issues for Organization Science," *Organization Science* (14:2), pp. 209-223.

- Wang, Q., Li, B., and Singh, P. V. 2015. "The Impact of Copycats on an Original Mobile App's Demand: Empirical Analysis and a Method for Detecting Copycat Apps," Carnegie Mellon University, Pittsburg, PA.
- Wei, C.-P., Yang, C. C., and Lin, C.-M. 2008. "A Latent Semantic Indexing-Based Approach to Multilingual Document Clustering," *Decision Support Systems* (45:3), pp. 606-620.
- Wen, W., and Zhu, F. 2016. "How Do Complementors Respond to the Threat of Platform Owner Entry? Evidence from the Mobile App Market," University of Texas Austin, Available at SSRN: <https://ssrn.com/abstract=2848533>.
- Yeung, A. C., Cheng, T., and Lai, K. h. 2005. "An Empirical Model for Managing Quality in the Electronics Industry," *Production and Operations Management* (14:2), pp. 189-204.
- Yoo, Y., Boland Jr, R. J., Lyytinen, K., and Majchrzak, A. 2012. "Organizing for Innovation in the Digitized World," *Organization Science* (23:5), pp. 1398-1408.
- Yoo, Y., Lyytinen, K., Boland, R., Berente, N., Gaskin, J., Schutz, D., and Srinivasan, N. 2010. "The Next Wave of Digital Innovation: Opportunities and Challenges: A Report on the Research Workshop: "Digital Challenges in Innovation Research"," Temple University, Philadelphia, PA.
- Zaltman, G., Duncan, R., and Holbek, J. 1973. *Innovations and Organizations*, New York, NY: John Wiley & Sons.
- Zhu, F., and Zhang, X. 2010. "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing* (74:2), pp. 133-148.
- Zittrain, J. L. 2006. "The Generative Internet," *Harvard Law Review* (119:7), pp. 1974-2040.

3 THE IMPACT OF PRODUCT DIFFERENTIATION STRATEGIES ON MOBILE APPLICATION PERFORMANCE: AN EMPIRICAL STUDY

3.1 Introduction

Mobile apps are digital goods that are produced, distributed, and consumed through information technology (Bakos and Brynjolfsson 1999). Digital goods often benefit from significant cost savings through eliminating and replacing traditional value-added processes (Amit and Zott 2001). Moreover, the mobile platform providers (MPP), such as the Apple App Store and Google Play Store, offer many value-adding activities by providing developers with great cost savings and ease of marketing and distribution for mobile application developers (Basole and Karla 2011; MacMillan et al. 2009). As a result, the mobile app marketplaces have attracted many developers, with small companies and individual developers being the majority (Basole and Karla 2012). The mobile app industry has experienced tremendous growth during the last few years (Comino et al. 2016). According to Statista (2017)¹⁰, the number of apps in the Apple App Store has increased from 0.15 million in March 2013 to approximately 2.2 million in 2017. A similar trend is observed in the Google Play Store where the number of apps increased from 1 million in July 2013 to 3.3 million in 2017. App-generated revenue from end users is forecasted to increase from \$88 billion in 2016 to \$189 billion in 2020. As a result, the mobile app marketplace has become economically important and intensively competitive.

¹⁰<https://www.statista.com/statistics/269025/worldwide-mobile-app-revenue-forecast/>

In a highly competitive industry, product differentiation strategies are key to firms' survival and success (Bhargava and Choudhary 2001). Chamberlin (2010) defines product differentiation as “a general class of products which differentiates from any other types of products based on any substantial characteristics and leads to a preference for one variety of the product over another.” Existing marketing literature focuses on two types of differentiation strategies: product innovation based and marketing based differentiation (Miller 1986), which have been empirically supported (Lee and Miller 1999). A product innovation-based strategy strives to create the most up-to-date and attractive products through improvements in quality, efficiency, design innovations, and style. The strategy can differentiate from competitors by innovation complexity (Gnyawali et al. 2010), quantities (Dougherty and Dunne 2011), and timeliness (Chen et al. 1992). Firms taking this type of differentiation strategy lead their competitors in innovation. For example, mobile app firms can achieve product innovation differentiation through incremental product innovation, also known as “vertical product differentiation” (Animesh et al. 2011; Beal 2000). The positive “*leapfrog*” effect of incremental innovation on firms’ competitive advantage has been documented in Xu’s study (Xu 2015).

The other differentiation strategy - marketing-based differentiation strategy comes from the literature on consumer information and transaction costs (Caves and Williamson 1985; Miller 1986). The consumers, constrained by information costs and bounded rationality in making the best decision, are influenced by various types of information up to the point where the expected marginal benefits of making a purchase decision equal the marginal information costs. In the context of mobile apps, consumers can be influenced by various types of information available in the mobile app marketplace, including price, app description, update release notes, peer consumer reviews and ratings, and technical details (e.g., supported devices). The app developers can release

or revise information influencing consumers' perceived product quality and uncertainty, even if all consumers had identical preferences about the primary product attributes (Cusumano et al. 2008). For example, an optimal price can help firms differentiate from others in a competitive market and gain as much profit as possible (Guo et al. 2013). Although app price is an important determinant of consumers' purchase decisions, product quality is also an important factor in determining consumer demands. Online product ratings and reviews provide an information channel that may reflect consumers' perceived product quality (Ghose and Ipeiritis 2011). Existing studies find that consumers' perceived product quality will influence product sales (Duan et al. 2008; Zhu and Zhang 2010). In addition, online product ranking charts provide an information environment where consumers tend to search through the listings sequentially in the order they are listed (Lee and Raghu 2014). The ranking creates a directional market where products listed at the top have an advantage over those appearing lower in the ranking (Animesh et al. 2011). Therefore, marketing-based product differentiation can effectively complement product innovation based differentiation to sustain a firm's competitive advantage, especially in the mobile app marketplace where developers and consumers can easily share information and access value-added information services provided by the platforms.

While the mobile app industry has attracted a lot of research interest, very few existing studies have focused on the differentiation strategies of mobile apps. Given the highly competitive nature of the mobile app marketplace, it is not a trivial task to study the impact of product differentiation strategy on product performance and sustainability. Product innovation based differentiation powered by incremental innovation is important but constrained by the limited development and maintenance resources of app developers (Animesh et al. 2011). With the information services provided by the mobile platform providers, marketing-based differentiation costs much less than

incremental product innovation. However, the impact of different information types on the product success of mobile apps has not been well explored. Drawing on the characteristics of the mobile app industry and aforementioned product differentiation alternatives (Dickson and Ginter 1987; Greenstein and Markovich 2012; Systems 2017), the study extends the literature on the mobile app marketplace by empirically investigating how different differentiation strategies affect the product performance of mobile apps. Specifically, the study explores the following research questions:

1. How does a mobile app firm differentiate its app relative to its competitors' apps through product innovation based and marketing based differentiation strategies?
2. How do these differentiation strategies influence product performance?

This study contributes to the literature in several ways. First, this study collectively considers product innovation differentiation and marketing differentiation strategies in the context of mobile apps. Taking this approach allows me to examine how the different differentiation strategies complement each other to influence product performance. Past empirical studies on differentiation strategies mostly focus on one of the two differentiation strategies. However, the mobile app marketplace is highly competitive and has multiple information channels including focal firms' public information, such as incremental innovation profile, and marketing information, such as top charts, user reviews and ratings, and app information, to help the users make purchasing decisions. It is likely that to stay competitive in such a competitive marketplace an app developer may need to apply both differentiation strategies. Second, the study is also one of the first to study how the two types of software innovation activities, namely bug fixes and new features, influence app performance. Such knowledge can be used to help app firms make informed innovation decisions to help sustain competitive advantage. Finally, this study contributes to the traditional differentiation literature by taking advantage of new information channels available in the app

store, such as the top chart rankings and user ratings, to infer perceived product reputation and quality.

In the remainder of this chapter, the theoretical background and hypotheses development are discussed in Section 2. In Section 3, the data collection process and experimental design are described. The empirical model is explained in Section 4. The findings are presented in Section 5. The last section of the chapter ends by presenting conclusions and limitations of the study.

3.2 The Mobile App Marketplace

Compared to traditional software products and other digital goods, mobile apps have some unique characteristics that may have an impact on product performance in this marketplace. First, the mobile app marketplace is propelled by app developers' rapid incremental innovation through frequent updates (Zhou et al. 2018). In the traditional on-premises software industry, product patching management or updating is costly because firms need to allocate the necessary resources, both human and technology, to release or distribute patches into diverse and complex client-side environments. By contrast, mobile app makers have only one channel to release or update their products via the Internet. For example, a mobile app developed for iOS can only be released or updated in the Apple App Store. The uniform distribution channel can help app developers reduce costs and ease the product managers' burden to update products in a complex client-side environment. Once an app update is released, the app users will acquire the update as easy as a single tap. It has been an established practice that updates are made available free of charge to anyone who has previously downloaded the app (Comino et al. 2015). Moreover, mobile app developers mostly rely on the updating mechanism rather than the liability mechanism to achieve customer satisfaction. Kim et al. (2011) find that the updating mechanism is more effective than

the liability mechanism when users' and the vendor's updating costs are low. While in many “long-tail markets” such as books, music, and movies, product versioning is constrained by an often-fixed time-to-release period (e.g., once a year), app developers are not. They have the flexibility to quickly respond to customer feedback and requests through a versioning or product updating strategy. Given that customers may have a dynamic range of needs, Nayebi et al. (2016) contend that app makers should make the "right release decisions" to remain visible and popular among consumers. Hence, the product updating strategy has an important impact on the competitive advantage of mobile apps.

Second, a large amount of customer reviews plays an important role in the success of mobile apps. In contrast to search goods whose quality assessment can be made before purchase, mobile apps are experience goods whose true quality can only be learned after consumption (Mudambi and Schuff 2010). The review systems provided by the mobile app marketplaces allow the users to share their user experiences, quality assessment, and feature requests after they have downloaded and installed the mobile app. Compared to other online stores such as Amazon, the "verified-purchase" requirement is strongly enforced in the mobile app marketplace. Mobile app developers, therefore, can improve customer satisfaction and product success by strategically responding to customer feedback and needs through updates.

Third, the mobile app marketplace is highly competitive due to the sheer number of apps available in the market. It is considered to be a “hyper-competitive” marketplace where developers struggle to attract consumers (Datta and Kajanan 2013). This brings out the issue of how to draw the attention of consumers so that a mobile app product can emerge from millions of apps. The appearance and ranking in the top charts can significantly influence product discoverability, downloads, and revenue (Ragaglia and Roma 2014). Therefore, mobile app developers have to

carefully determine their competition and product differentiation strategies to succeed in this highly competitive marketplace.

3.3 Product Differentiation

Product differentiation is one of the two competition strategies identified by Porter (2010) for highly competitive industries besides cost leadership. A cost leadership strategy emphasizes a relatively low cost compared to the competitors. It aims to provide the lowest price for the consumers. However, in the mobile app marketplace, the acquisition cost of 94% mobile app products is free of charge¹¹. The average price of paid apps is merely around \$1¹². Therefore, a cost leadership strategy might not help much competition wise in the mobile app marketplace.

Product differentiation strives to design and make “a general class of product which differentiates from any other types of products based on any substantial characteristics and leads to a preference for one variety of the product over another” (Chamberlin 2010). The American Marketing Association (2014) defines product differentiation as one or more product attributes that make one product different from another. Several studies mainly focus on the differentiation strategies based on firms’ product innovation actions (Bhargava and Choudhary 2008; Caves and Williamson 1985; Shapiro and Varian 1998; Varian 1997). Yet, another definition states that “product differentiation is defined as a product offering which is perceived by the consumer to differ from its competition on any physical or nonphysical product characteristics including price,” and suggests that the differentiation may not be real and can be the market differentiation from the

¹¹ <https://www.appbrain.com/stats/free-and-paid-android-applications>

¹² <https://www.statista.com/statistics/267346/average-apple-app-store-price-app/>

market and contextual information (Crawford-Welch 1990). As a broad concept, product differentiation is suggestive of product uniqueness that represents its competitive advantages in its industry (Porter 1980).

Past research has categorized product differentiation strategies in different ways. For example, in management literature, Mintzberg (1988) broadly categorizes differentiation strategies into six types: quality, design, support, image, price and undifferentiated products. In the industrial organization literature, strategy researchers have explored the distinction between vertical and horizontal product differentiation (Becerra et al. 2013; Ethiraj and Zhu 2008). Vertical product differentiation tries to make the product attractive to all customers in the market, whereas horizontal differentiation only aims to attract a specific set of consumers (Makadok 2010). Existing economic literature presents two types of product differentiation strategies: product-attribute based and information based strategies (Caves and Williamson 1985). The product-attribute based differentiation strategy rests on the intrinsic complexity of the product, coupled with fixed costs of producing each variety (Caves and Williamson 1985). The nominal product offered to the market may combine some different primary attributes. Complementing the product-attribute based differentiation strategy focusing on the supplier side, an information-based differentiation strategy aims to change customers' perceived product quality and intention to purchase (Caves and Williamson 1985). Consumers may have access to various types of information such as price, advertisement, online and offline word-of-mouth, usage experiences in consumer reviews, and reviews from independent sources. This type of information may have an impact on the consumers' perceived product quality and purchase uncertainty, even if all consumers have identical preferences over the primary product attributes (Cusumano et al. 2008). A change in one type of information such as the price, with all other information kept unchanged, would induce some but

not all consumers to purchase the focal product. In the marketing literature, Miller (1986) proposed another two types of differentiation strategies: innovation-based and marketing-based strategies, which have been empirically supported (Lee and Miller 1999). An innovation-based strategy strives to create the most up-to-date and attractive products through improvements in quality, efficiency, design innovations, and style, while a marketing differentiation strategy attempts to create unique value for a product through advertising, pricing, and reputation management (Miller 1988).

Research on mobile app differentiation strategies is still at an infant stage (Dougherty 2001a; Katzmarzik 2011). Existing mobile app research predominately examines the impact of product attributes and online consumer reviews on product performance (Ghose and Han 2014; Liu et al. 2014; Wang et al. 2018). There has been some research focusing on designing better app ranking systems by mining user-generated content, as well as examining the optimal pricing strategy for app firms (Lee and Raghu 2014; Liu and Brandyberry 2014). Also, a study has examined the success model of mobile apps that result from quality updates and online consumer feedback (Lee and Raghu 2014). Hu et al. (2012) mentioned that app firms have to improve product differentiation to better meet consumer needs. However, despite its importance, the efficacy of differentiation strategies for mobile apps has not been well explored.

3.4 Research Model

Although differentiation strategies have not been extensively studied for mobile apps, research in other contexts provides useful insights about product differentiation. The economic and management literature suggests that product differentiation has an influential impact on product quality, product positioning, and ultimately product performance (Becerra et al. 2013; Ethiraj and

Zhu 2008; Miller 1988). In the mobile app industry, mobile apps achieve product innovation based differentiation through product updates. Building on prior economic and information systems (IS) studies (Chakravarti and Xie 2006a, 2006b), the assertion that app firms should pursue two differentiation strategies, a product innovation differentiation strategy, and a marketing differentiation strategy is investigated. This study aims to empirically study the impact of the two differentiation strategies on product performance in the highly competitive mobile app marketplace. Figure 3-1 illustrates the conceptual research model.

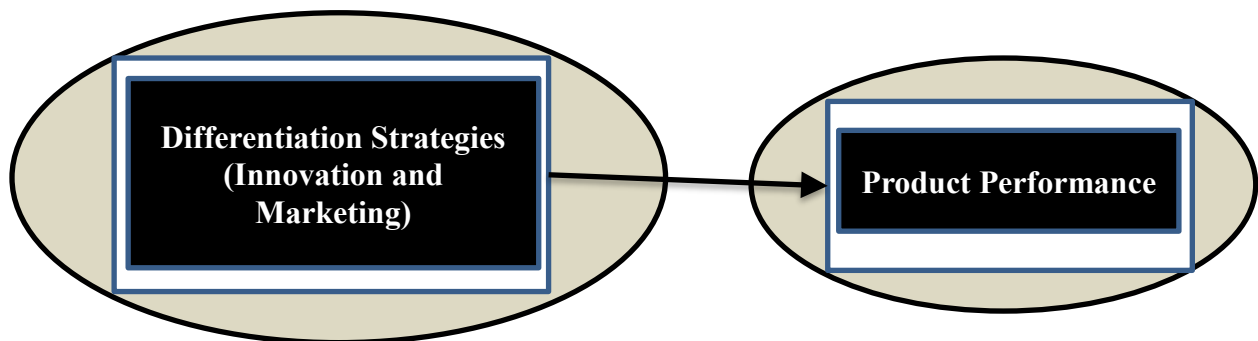


Figure 3-1 Conceptual Model

3.4.1 Product Innovation Differentiation

As reviewed earlier, product innovation differentiation refers to the competitive actions that an app firm undertakes to add new product features and improve product quality. The actions lead to incremental innovation, which includes "refinements and extensions of established designs that result in substantial price or functional benefits to users" (Kleinschmidt and Cooper 1991). It can be observed by the releases of app updates in the marketplace (Nayebi et al. 2016; Zhou et al. 2018). Incremental innovation, as a competitive action, can generally be evaluated from two aspects: quantity and complexity (Sambamurthy et al. 2003). Quantity refers to the number of

innovative actions reflected in newly released products or updates. Complexity refers to the variety and richness of innovative actions. Ferrier et al. (1999) find empirical evidence that both quantity and complexity led to increased market share. In addition, timeliness has been considered as an important success factor especially for fast-paced product development projects such as mobile apps (Cooper and Kleinschmidt 1994). In the rest of this section, research hypotheses are developed to study the impact of the three product innovation differentiation aspects, namely innovation complexity, quantity, and timeliness, on the product success of mobile apps. Figure 3-2 illustrates the complete research model.

In the IS literature, researchers suggest that firms that make complex and plentiful innovation actions create barriers for their competitors and eventually gain competitive advantage (Kim et al. 2017; Taylor Pentina 2011). Taking a variety of competitive actions could help a firm develop its dynamic capabilities on multiple facets, such as absorptive capacity and customer agility (Chen et al. 2014; Zhou et al. 2018). There are two major types of innovation actions in mobile apps: fixing bugs and addressing new feature requests (Arora et al. 2010; Chen et al. 2014; Maalej and Nabil 2015). Bug fixes focus on improving existing product quality to retain existing customers. On the other hand, adding new features can address new demands and attract new customers. Mobile app providers who are capable of performing both types of innovation actions will perform better than those who can only focus on one type.

Hypothesis 1. A mobile app's innovation complexity relative to its competitors' is positively related to app performance.

Innovation quantity refers to the total number of innovative actions taken by the firm and reflects the scale of innovation. Existing literature shows that innovation quantity is one of the important innovation measures that has a positive impact on product performance (Dougherty

2001a; Gnyawali et al. 2010). A positive relationship is expected between innovation quantity and product performance for the following reasons. First, the more competitive actions that a firm has taken, the more likely the firm is to sustain competitive advantage (Chen and MacMillan 1992; Gnyawali et al. 2006). Research shows that those firms with more innovative actions can enjoy better financial performance than those firms with fewer innovative actions (Chen et al. 1992; Smith et al. 2001). Second, innovation quantity also reflects a firm's innovation capability (Chen et al. 1992; Gnyawali et al. 2010). Firms with the capability to implement a large number of innovative actions are more likely to be competitive and attract more consumers.

Hypothesis 2a. A mobile app's innovation quantity of fixing bugs in its update relative to its competitors' has no significant impact on app performance.

Hypothesis 2b. A mobile app's innovation quantity of adding new features in its update relative to its competitors' is positively related to app performance.

Innovation timeliness refers to the number of days that have elapsed since the last product release or update. Existing literature shows that consumers are more interested in more recently released products than those released a while back (Spann et al. 2015). A negative relationship is expected between innovation timeliness and product performance for the following reasons. First, the more time that a firm takes to update its product, the less likely the firm is to sustain competitive advantage (Li et al. 2018; Lyu and Shang 2017). Compared to products in traditional industries, the development life cycle of a mobile app is much shorter. Due to the fierce competition in the mobile app marketplace, mobile app developers usually have a strong motivation to respond quickly to the competitive actions of their competitors (Chen et al. 1992; Prajogo and Ahmed 2006). A study on mobile apps shows evidence that consumers are more likely to adopt an app

product with the latest update (Comino et al. 2016; McIlroy et al. 2016). App developers will lose to their competitors if they fail to update their apps in a timely fashion (Shen et al. 2015).

Hypothesis 3. A mobile app's innovation timeliness relative to its competitors' is negatively associated with app performance.

3.4.2 Marketing Differentiation

Marketing differentiation aims to differentiate from other competitors in the same marketplace through marketing attributes (Dickson and Ginter 1987; Society 2011). It is an extension of the information-based product differentiation strategy. By releasing various types of information to consumers, such as price, advertisement, online and offline word-of-mouth, usage experiences in consumer reviews, and reviews from independent sources, firms can improve consumers' perceived product quality and reduce uncertainty in their adoption decisions (Ghose and Han 2014; Lee and Raghu 2014). In the mobile app marketplaces, much information is publicly available to consumers, app developers, and competitors, including basic app product attributes (e.g., price, file size, supported devices, maturity rating), app ranks in top charts, and ratings and reviews made by app users. The app users, constrained by information search cost and bounded rationality, normally choose a portfolio of information sources. Some but not all sources of information are controlled by product makers (Caves and Williamson 1985), who jointly adjust their prices and product qualities. A change in price may induce some but not all buyers to switch if other information sources are held constant. When the price is held constant, the buyers will rely on other information sources, such as ranking and user reviews, to infer product qualities and reputation among competing products. Three marketing attributes are considered here: pricing, perceived product quality, and product reputation.

Price differentiation strategy is a basic strategy for a firm to differentiate its product by charging a lower price than its competitors' (Mintzberg 1988). Research indicates that competition-based pricing remains dominant in pricing practice (Hinterhuber 2008). Competition-based pricing has become even more pertinent for online retailers thanks to the surge of online marketplaces (Fisher et al. 2017; Zentes et al. 2017). Johansson et al. (2012) define competition-based pricing as an approach in determining the price for a focal product based on observations of competitor prices. With the development of the Internet, competitors can easily obtain price information from different public channels. The transparency of pricing information results in more fierce competition and requires firms to monitor competitor prices constantly so as to respond to the competition. Fisher et al. (2017) show that a competition-based pricing strategy leads to significant revenue improvement in a field experiment involving a retailer in a competitive environment. Most of the mobile apps are low-cost products with an average price of ~\$3 (Distimo 2009). Firms who compete in a low-cost market usually focus on price sensitive customers in order to attract a large number of customers and increase revenue. Assuming all apps have replaceable utilities, an app with a price lower than its competing apps would receive more purchases and downloads.

Hypothesis 4. An app's price relative to its competitors' is negatively associated with app performance.

Although price is an important determinant of product performance, product quality can significantly affect product success (Balasubramanian et al. 2015; Lee and Raghu 2014). A quality differentiation strategy is an effective way to increase consumer loyalty and decrease consumer attrition (Narasimhan et al. 1993). Quality differentiation focuses on making a product better on certain features but not making it fundamentally different from competing products (Mintzberg

1988). The objective is to make the product more reliable, durable, or superior in performance than its competitors. In the mobile app marketplaces, the information channel that may reflect consumers' perceived product quality is online product ratings and reviews provided by the app users (Chen 2013; Chen and Xie 2008). A high average rating usually indicates high perceived quality for a product. Furthermore, Dey and Lahiri (2016) indicate that product quality inferred from user reviews or ratings must be understood in the context of competition. A high level of perceived quality inferred by high user ratings does not give the firm competitive advantages unless the level of quality is above that of its competing products.

Hypothesis 5. A mobile app's average user rating relative to its competitors' is positively associated with app performance.

Product ranking is an important information channel for consumers or competitors to infer product reputation (Animesh et al. 2011). Bharadwaj (2000) mention that firms who have a strong reputation tend to be ranked as a leader. Animesh et al. (2011) show that ranking creates a directional market where products listed at the top have competitive advantage over those appearing lower on the list. In the context of mobile apps, Garg and Telang (2014) show that a top-ranked app can be downloaded 150 times more than the number of downloads of the app ranked at the 200th in Apple's top paid app chart. App firms have a strong incentive to stay in the top charts in order to achieve better product performance.

Hypothesis 6. A mobile app's rank relative to its competitors' is negatively associated with app performance.

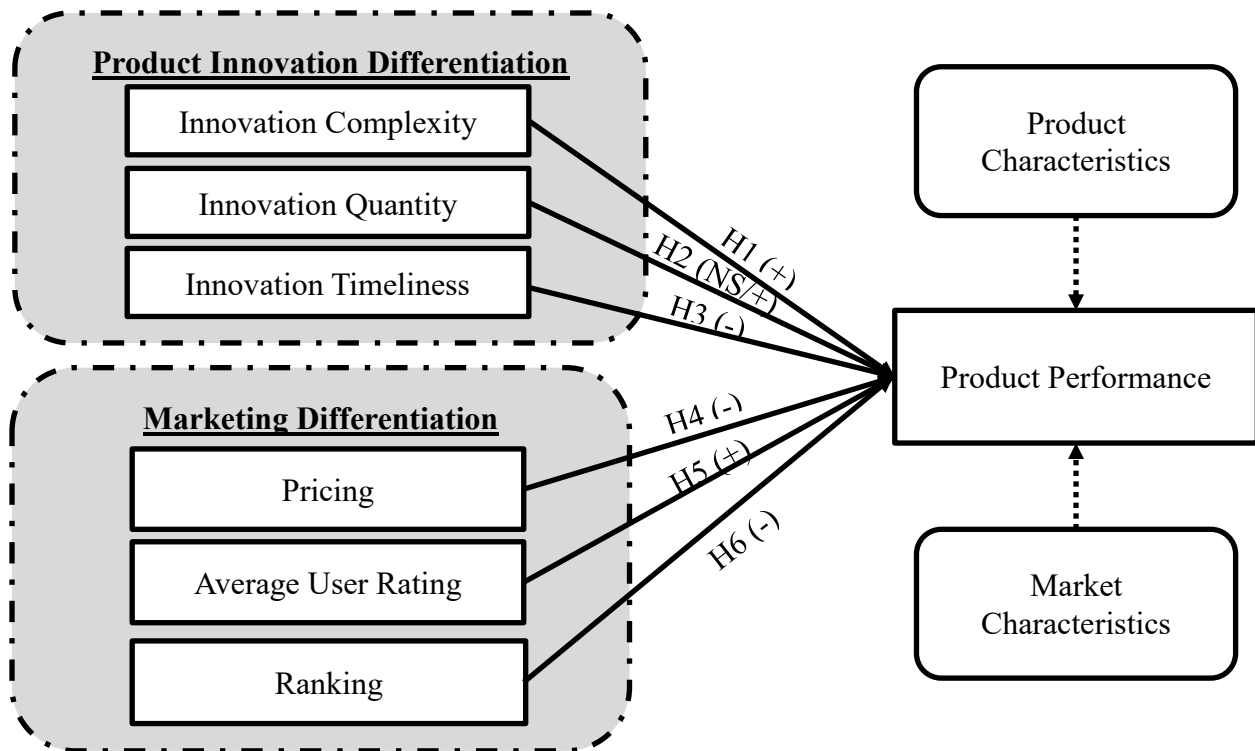


Figure 3-2 Research Model

3.5 Research Methodology

3.5.1 Data

The research model was tested empirically using data collected from the Google Play store, the largest mobile platform provider in terms of a number of apps as well as number of total downloads. Both the app products' information and user reviews are available on Google Play. Specifically, for each mobile app, customers can post reviews in thorough textual comments and an overall rating with a scale of 1 to 5 for the specific version of the product. Accordingly, app developers will update the mobile app based on the customers' requirements and post the update date and relevant update notes. In addition to app developers' update information and customer reviews, the Google Play app store also publishes various categories' top charts. The top free and top paid charts include the top 540 mobile apps in the free and paid categories based on the number of

downloads each day. Since the ranking lists are updated daily, these lists are crawled as well as their product information pages for those apps appearing on the top charts on a daily basis.

Since this study considers pricing differentiation among other differentiation strategies, the study only considered the paid apps that appeared on the top paid game chart and its subcategories. Data are collected for 620 paid game apps from the Google Play store between April 8, 2014, and September 8, 2014. The data are unbalanced panel data for a five-month period and are grouped at the app level. All apps have appeared at least once in the top paid game chart during the data collection window. Since the focus is highly competitive markets, the game subcategories that had at least 40 app products are chosen. The final dataset contains 362 paid game apps with 35,582 observations. Data collected include basic app attributes, app release/update dates, user ratings, user reviews, and developer information. Table 3-1 shows the app categories included in the dataset.

Table 3-1 Game App Categories

App Category	No. of Apps
GAME_ACTION	49
GAME_ARCADE	142
GAME_CASUAL	49
GAME_PUZZLE	80
GAME_STRATEGY	42

3.5.2 Measures

The dependent variable, product performance, used in this study is calculated as $-\text{Log}[\text{ProductRank}]$, which is the negative log of the daily rank of each app in the top paid game chart. Although a more ideal measure for the dependent variable is financial performance (e.g., revenues or profits), Google Play does not release apps' financial information or number of

downloads. App developers are not required to disclose it either. The use of product rank as a market-based performance metric can be an appropriate alternative because app download volume and sales are both factors in determining the product rank¹³. Existing studies have also used product rank as an indicator of product performance (Forman et al. 2008; Garg and Telang 2014; Ghose and Han 2014).

A first independent variable, *InnovationComplexityDiff*, refers to the relative diversity of innovation actions taken by an app firm compared to its competitors (Gnyawali et al. 2010; Taylor Pentina 2011). To calculate this variable, a nominal variable *InnovationComplexity* is defined, the value of which corresponds to three innovation complexity levels: (1) bug fixes only, (2) new features only actions, and (3) the combination of bug fixes and new features. To identify different types of innovation activities, the app update release notes are analyzed using keywords indicative to those innovation activities (Maalej and Nabil 2015). The keywords are listed in Table 3-2.

The second independent variable is relative innovation quantity. Following previous studies (Dougherty 2001b; Gnyawali et al. 2006), the relative innovation quantity is measured using two variables: *NewFeatureQuantityDiff* and *BugFixQuantityDiff*. *NewFeatureQuantityDiff* is defined as the difference between the number of new features in the last update of the focal app and the average number of new features included in the last update of its competitors. *BugFixQuantityDiff* is defined as the difference between the number of bug fixes in the last update of the focal app and the average number of bug fixes included in the last update of its competitors.

¹³ <https://appsposure.com/app-store-ranking-algorithms/>

The third independent variable is relative innovation timeliness (Spann et al. 2015), *InnovationTimelinessDiff*, which reflects the number of days that have elapsed since the last update for the focal app compared to its competitors. A small relative innovation timeliness value means that the focal app's update is relatively more recent than its competitors.

The fourth independent variable is relative price, *PriceDiff*, which reflects the focal app's price relative to its competitors. Following existing studies (Becerra et al. 2013; Fisher et al. 2017; French et al. 1997), it is measured as the difference between the focal app's price and the average price of its competitors.

The fifth independent variable is relative perceived quality, *DailyReviewRatingDiff*, which reflects the perceived quality of the focal app relative to its competitors. This measure is the difference between a focal app's average review rating and the average rating of its competitors.

The sixth independent variable is relative product reputation, *GameAppRankDiff*, which refers to a focal app's product reputation relative to its competitors'. It is calculated as the difference between a focal app's rank on the top charts and the average rank of its competitors.

In addition, a set of seven control variables are included to control for other exogenous factors that may affect app performance. Four control variables, *UpdateWordCount*, *TopChartNumber*, *AppNumByTheSeller*, and *TopDeveloper*, are used to control for app developer characteristics, which indicate the developer's available resources and innovation capability. *UpdateWordCount* is the number of words in an update release note, reflecting the effort undertaken by the developer in an update. The effort can be indicative of the resources that the developer has access to. *TopChartNumber* is the number of top charts that an app appears on. *AppNumByTheSeller* is the number of apps released by the same app developer. *TopDeveloper* is a dummy variable, indicating whether or not the developer has a top developer badge, a recognition awarded by the Google Play

store for the developers who create high quality apps. Two control variables, *MarketHorizontalDynamism* and *MarketVerticalDynamism*, are included to control for characteristics of the mobile app marketplace. *MarketHorizontalDynamism* is defined as the number of apps in the same game subcategory that enter or exit the top chart on a given day. *MarketVerticalDynamism* is calculated as the variance of all app ranks in a game subcategory on a given day. To control for the technology change's impact in this study, a variable is added, *AndroidUpdateAge*, which is the number of days that have elapsed since the last Android operating system update. Finally, the impact of seasonality is controlled by using a variable *WeekdayNumber*, which is the day of the week.

Table 3-3 summarizes all the variables included in the empirical model along with their definitions.

3.5.3 Empirical Model

Since the dependent variable is continuous, the fixed effects model with unbalanced panel data was used to conduct the analysis. The Hausman test (1978) was performed to determine whether a random effects model or a fixed effects model is more valid. The Hausman test checks for violations of the assumption that the random effect specifications do not correlate with unobserved effects and the independent variables. Results of the test reject the null hypothesis at the 0.01 significance level and are indicative of a fixed effects model. A fixed effects model allows correlation between the error term and independent variables, making the estimation more robust. Moreover, it can control the effects of the time-invariant unobserved factors, such as the app developer.

$$\begin{aligned}
-\text{Log}[\text{ProductRank}]_{it} &= \beta_0 + \beta_1 A \\
&= \beta_0 + \beta_1 \text{InnovationComplexityDiff}_{i(t-1)} \\
&+ \beta_2 \text{InnovationTimelinessDiff}_{i(t-1)} + \beta_3 \text{BugFixQuantityDiff}_{i(t-1)} \\
&+ \beta_4 \text{NewFeatureQuantityDiff}_{i(t-1)} + \beta_5 \text{PriceDiff}_{i(t-1)} \\
&+ \beta_7 \text{GameAppRankDiff}_{i(t-1)} + \beta_8 \text{DailyReviewRatingDiff}_{i(t-1)} \\
&+ \varepsilon_{it}
\end{aligned} \tag{1}$$

Table 3-2 Types of Update Based on Keyword Matching

Update Type	Keywords
Bug fixes	bug, fix, problem, issue, defect, crash, solve, optimization, improve
New features	add, need, request, new, change, want, suggest, recommend, great update

The empirical model is different from the models used in the literature that attempt to explore the impact of focal product or market characteristics on product sales. The measures of the independent variables are relative measures that compare a focal app's differentiation strategies compared to those of its competitors. Also, the potential multicollinearity problem is checked for by calculating the Variance Inflation Factor (VIF) value for every variable in all models. All VIF values are below 10, indicating no serious multicollinearity issues in the empirical results.

Table 3-3 Description of Variables

Construct	Variable	Definition
Variables regarding marketing differentiation		
	$ProductRank_{it}$	The negative natural log of the product rank of app i on day t
Reputation	$GameAppRankDiff_{i(t-1)}$	The difference between app i 's rank and the average rank of the apps in the same app subcategory on day $t-1$
Perceived Quality	$DailyReviewRatingDiff_{i(t-1)}$	The difference between app i 's average review rating and the average review rating of the apps in the same app subcategory on day $t-1$
Price	$PriceDiff_{i(t-1)}$	The difference between app i 's price and the average price of the apps in the same app subcategory on day $t-1$
Variables regarding product innovation differentiation		
Innovation Complexity	$InnovationComplexityDiff_{i(t-1)}$	The difference between app i 's $InnovationComplexity_{i(t-1)}$ and the average $InnovationComplexity_{i(t-1)}$ of the apps in the same app subcategory on day $t-1$ $InnovationComplexity_{i(t-1)}$ is a dummy variable that indicates whether app i in its last update as of day $t-1$ received bug fixes only (1), new features only (2), both bug fixes and new features (3).
Innovation Quantity	$BugFixQuantityDiff_{i(t-1)}$	The difference between the number of bug fixes that app i received in the last update and the average number of bug fixes that the apps in the same app subcategory received in their last updates as of day $t-1$
	$NewFeatureQuantityDiff_{i(t-1)}$	The difference between the number of new features that app i received in the last update and the average number of new features that the apps in the same app subcategory received in their last updates as of day $t-1$
Innovation Timeliness	$TimelinessDiff_{i(t-1)}$	The difference between app i 's update age and the average update age of the apps in the same app subcategory on day $t-1$
Control variables		
	$UpdateWordCount_{i(t-1)}$	The number of functional words (verb, noun, adjective, and adverb) in app i 's last update release notes on day $t-1$
	$TopChartNumber_{i(t-1)}$	The number of top charts where app i appears on day $t-1$
	$AppNumByTheSeller_{i(t-1)}$	The number of apps that app i 's developer has in the app store on day $t-1$
	$TopDeveloper_{i(t-1)}$	Whether app i 's developer is awarded a "Top Developer" badge by the app store on day $t-1$
	$WeekDayNumber_{i(t-1)}$	The day of the week number (0-6) for day $t-1$
	$MarketHorizontalDynamism_{i(t-1)}$	The number of apps entering or existing top charts for a specific subcategory where app i belongs to on day $t-1$
	$MarketVerticalDynamism_{i(t-1)}$	The variance of all app rank changes for a specific subcategory where app i belongs on day $t-1$
	$AndroidUpdateAge_{i(t-1)}$	The number of days that have passed since last Android system update as of day $t-1$

3.5.4 Results

Table 3-4 summarizes the descriptive statistics for all variables. Table 3-5 presents the estimation results of the fixed effects panel data regression model. The coefficient of *InnovationComplexityDiff* in Model 1 is negative and not significant at the 0.01 significance level, thus, hypothesis 1 is not supported. However, the coefficient is significant in the full model (Model 6), which means that the relationship between this variable and the dependent variable is potentially moderated by other variables. More analyses are needed to fully explore this result in the future. In Model 2, the coefficient of *NewFeatureQuantityDiff* is positive and significant at the 0.01 significance level while the coefficient of *BugFixQuantityDiff* is not significant. Hypothesis 2a and 2b are supported. In Model 3, the coefficient of *InnovationTimelinessDiff* is negative and significant, meaning that the product performance is improved when the focal app's last update is more recent than its competitors, supporting hypothesis 3. Next, the coefficient of *PriceDiff* is negative and significant in Model 4, which means the focal app's performance improves when its price is lower than the average price of its competitors. Hypothesis 4 is supported. The coefficient of *DailyReviewRatingDiff* is negative but insignificant in Model 5. The result shows that the focal app's perceived quality relative to its competitors has no real impact on its performance. Hypothesis 5 is not supported. The coefficient of *GameAppRankDiff* is negative and significant at the 0.01 significance level in Model 6, indicating that the focal app's performance improves when the perceived reputation relative to its competitors is high (i.e., a low-rank value), supporting hypothesis 6. All statistical results have been summarized in Table 3-6.

Table 3-4 Descriptive Statistics

	Mean	S.D.	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<i>1.ProductRank</i>	-5.23	1.07	-6.29	-0.69	1															
<i>2.UpdateWordCount</i>	3.9	1.94	0	6.96	0.12	1														
<i>3.TopChartNumber</i>	0.84	0.3	0	1.39	0.6	0.11	1													
<i>4.AppNumByTheSeller</i>	1.58	4	0	24	0.23	0.02	0.22	1												
<i>5.TopDeveloper</i>	0.4	0.49	0	1	0.11	0.01	0.09	0.38	1											
<i>6.WeekdayNumber</i>	2.22	1.78	0	6	0.01	-0.09	0.03	-0.16	-0.1	1										
<i>7.MarketHorizontalDynamism</i>	2.7	6.86	0	59	0	0.02	0.07	0.01	-0.02	0.01	1									
<i>8.MarketVerticalDynamism</i>	339442.56	750874.53	0	5640000	0.01	0.01	0.07	-0.01	-0.04	0.01	0.88	1								
<i>9.AndroidUpdateAge</i>	74.57	62.89	1	175	0.03	0	-0.03	0.01	-0.02	-0.01	0.26	0.2	1							
<i>10.InnovationComplexityDiff</i>	-0.01	1.25	-1.45	15.4	0.11	0.38	0.06	-0.05	-0.1	0	0	0	0	1						
<i>11.BugFixQuantityDiff</i>	0.01	1.41	-1.56	7.18	-0.01	0.46	0.01	-0.14	-0.15	0	0	0	0	0.24	1					
<i>12.NewFeatureQuantityDiff</i>	0.35	0.48	0	1	0.07	0.45	0.06	0.04	0.03	0	0.01	0	0.01	0.63	0.17	1				
<i>13.InnovationTimelinessDiff</i>	-4.08	206.33	-313.98	1097.52	-0.13	-0.14	-0.11	0.03	-0.17	0.04	0	0	0.01	-0.12	0	-0.15	1			
<i>14.PriceDiff</i>	0.11	2.59	-3.97	47.54	-0.02	0.02	0	0.13	0.07	-0.02	0	0	-0.01	0	-0.05	-0.03	-0.12	1		
<i>15.GameAppRankDiff</i>	53.46	153.64	-275.13	464	-0.72	-0.12	-0.47	-0.2	-0.08	-0.02	-0.02	-0.02	-0.06	-0.08	-0.02	-0.08	0.13	0.06	1	
<i>16.DailyReviewRatingDiff</i>	0.61	1.22	-9.07	3.26	0.11	0.03	0.11	-0.05	-0.02	0.01	0	0.01	-0.05	0.03	0.01	-0.01	-0.09	-0.01	-0.11	1

Table 3-5 Fixed Effects Estimation on Product Performance (-Log [$ProductRank_{it}$])

	(1)	(2)	(3)	(4)	(5)	(6)
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
<i>UpdateWordCount</i> _{<i>i</i>(<i>t</i>-1)}	0.02** (0.01)	0.02** (0.01)	0.01* (0.01)	0.01+ (0.01)	0.02* (0.01)	0.02** (0.01)
<i>TopChartNumber</i> _{<i>i</i>(<i>t</i>-1)}	0.49*** (0.02)	0.49*** (0.02)	0.49*** (0.02)	0.42*** (0.02)	0.48*** (0.02)	0.58*** (0.02)
<i>AppNumByTheSeller</i> _{<i>i</i>(<i>t</i>-1)}	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
<i>TopDeveloper</i> _{<i>i</i>(<i>t</i>-1)}	0.08+ (0.04)	0.08+ (0.04)	0.08+ (0.04)	0.10* (0.04)	0.11* (0.05)	0.12** (0.04)
<i>WeekdayNumber</i> _{<i>i</i>(<i>t</i>-1)}	-0.01 (0.00)	-0.01+ (0.00)	-0.00 (0.00)	-0.01+ (0.00)	-0.01+ (0.00)	-0.00 (0.00)
<i>MarketHorizontalDynamism</i> _{<i>i</i>(<i>t</i>-1)}	-0.02*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)
<i>MarketVerticalDynamism</i> _{<i>i</i>(<i>t</i>-1)}	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
<i>AndroidUpdateAge</i> _{<i>i</i>(<i>t</i>-1)}	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
<i>InnovationComplexityDiff</i> _{<i>i</i>(<i>t</i>-1)}	-0.01 (0.02)	-0.04* (0.02)	-0.05* (0.02)	-0.05* (0.02)	-0.06* (0.02)	-0.05* (0.02)
<i>NewFeatureQuantityDiff</i> _{<i>i</i>(<i>t</i>-1)}		0.03*** (0.01)	0.03*** (0.01)	0.03*** (0.01)	0.03*** (0.01)	0.03*** (0.01)
<i>BugFixQuantityDiff</i> _{<i>i</i>(<i>t</i>-1)}		0.01 (0.01)	0.00 (0.01)	0.01 (0.01)	0.01 (0.01)	0.00 (0.01)
<i>InnovationTimelinessDiff</i> _{<i>i</i>(<i>t</i>-1)}			-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00** (0.00)
<i>PriceDiff</i> _{<i>i</i>(<i>t</i>-1)}				-0.16*** (0.01)	-0.16*** (0.01)	-0.11*** (0.01)
<i>DailyReviewRatingDiff</i> _{<i>i</i>(<i>t</i>-1)}					-0.00 (0.00)	-0.01* (0.00)
<i>GameAppRankDiff</i> _{<i>i</i>(<i>t</i>-1)}						-0.00*** (0.00)
Constant	-5.85*** (0.04)	-5.84*** (0.04)	-5.83*** (0.04)	-5.76*** (0.04)	-5.78*** (0.04)	-5.78*** (0.04)
Observations	32022	32022	32022	32022	28498	28498
N_g	358.00	358.00	358.00	358.00	356.00	356.00
ll	-21836.13	-21827.70	-21812.45	-21498.92	-19928.36	-19244.71
df_m	366.00	368.00	369.00	370.00	369.00	370.00

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3-6 Explanations of Statistical Analysis

Hypothesis	Supported or Not
<i>Hypothesis 1. A mobile app's innovation complexity relative to its competitors' is positively related to app performance.</i>	Not Supported
<i>Hypothesis 2a. A mobile app's innovation quantity of fixing bugs in its update relative to its competitors' have no significant impact on app performance.</i>	Supported
<i>Hypothesis 2b. A mobile app's innovation quantity of adding new features in its update relative to its competitors' is positively related to app performance.</i>	Supported
<i>Hypothesis 3. A mobile app's innovation timeliness relative to its competitors is negatively associated with app performance.</i>	Supported
<i>Hypothesis 4. An app's price relative to its competitors is negatively associated with app performance.</i>	Supported
<i>Hypothesis 5. A mobile app's average user rating relative to its competitors is positively associated with app performance.</i>	Not Supported
<i>Hypothesis 6. A mobile app's rank relative to its competitors is negatively associated with product performance.</i>	Supported

3.6 Conclusions and Discussions

The mobile app marketplace is highly competitive due to the low technological barrier to entry and a large number of app products. It is an important research question to study the product differentiation strategies app developers may use to gain competitive advantages in this highly competitive market. This study investigates the impact of two differentiation strategies, namely product innovation differentiation and marketing differentiation, on app performance. The findings suggest that the number of new features in an app's innovation activities relative to its competitors is positively related to app performance. Moreover, an app's last update age relative to its competitors is negatively associated with app performance. Regarding the marketing differentiation strategy, the findings suggest that an app's price relative to its competitors is negatively associated with its performance. Moreover, an app's rank in the top chart relative to its

competitors is negatively associated with its performance. This empirical study makes the following theoretical and practical contributions.

3.6.1 Theoretical Contributions

This study contributes to the literature in several ways. First, this study collectively considers product innovation differentiation and marketing differentiation strategies in the context of mobile apps. Taking this approach allows me to examine how the differentiation strategies complement each other to influence app performance. Past empirical studies on differentiation strategies mostly focus on one of the two differentiation strategies. However, the mobile app marketplace is highly competitive and has multiple information channels that make publically available a wealth of information including updating information, as well as marketing information such as top charts, user reviews and ratings, and app information intended to help the users make purchasing decisions. An app developer can execute both differentiation strategies in an attempt to sustain competitive advantages. Second, the study is one of the first to examine how the two types of software innovation activities, namely bug fixes and new features, influence app performance. I find that the two innovation activities have different impacts on app performance with the number of new features having a significant impact on app performance. Such knowledge can be used to help app firms make informed differentiation decisions. Third, the study contributes to the competition-based pricing literature. The finding empirically confirms that app firms should determine a competitive price in the market. Finally, this study contributes to the traditional differentiation literature by taking advantage of new information channels available in the mobile app marketplace, such as the top chart rankings and user ratings, to infer perceived product reputation and quality.

3.6.2 Practical Implications

The study provides important practical implications. From a managerial perspective, the study underscores the business strategy value of product differentiation for app product managers. The findings indicate that different innovation activities have different impacts on product performance. To gain competitive advantage, app developers should focus on bringing more new features into their products than their competitors. Furthermore, the results also show that app product managers can benefit from updating their products more frequently than their competitors. Finally, the research suggests that it is beneficial for app product managers to keep their prices lower than their competitors. The finding on the relative price advantage is consistent with existing literature, where it has been discussed that app developers may use competitors' prices as a reference to make their pricing decisions (Chen and Iyer 2002).

These findings also have practical implications for mobile app marketplaces. The results indicate that the ranking mechanism, i.e., the top charts, as a reputation system has a significant impact on app performance. The mobile app marketplaces should design their ranking algorithms carefully in order to maintain a fair platform. Moreover, the mobile app marketplaces can learn from the findings to improve their information channels. For example, since new features and bug fixes have different impacts on app performance, the mobile app marketplaces could ask the app developers to specifically describe the two types of innovation activities in their update release notes. Also, app prices are now listed without comparison to its competitors' prices. Since a competitive price has a significant impact on app performance, the platform providers could choose to display an app price along with the average price of its competing apps. It would magnify the value of the price information to both consumers and app developers.

Some important limitations of this study also illustrate a few other directions for future research. First, the empirical results show that innovation complexity may not necessarily be associated with app performance. This is contrary to existing literature (Kim et al. 2017; Taylor Pentina 2011) and should be examined more carefully in future studies. Furthermore, it is important to note that the dataset only contains game apps. The generalizability of the findings could be improved if apps in other categories were collected and analyzed. Finally, the competitive environment of different contexts may necessitate the consideration of different innovation and marketing differentiation strategies. Future studies can replicate the model here in other domains and examine if and how propositions vary in other empirical contexts.

3.7 References

- Abrahams, A., Jiao, J., and Fan, W. 2013. “What’s Buzzing in the Blizzard of Buzz? Automotive Component Isolation in Social Media Postings,” *Decision Support Systems* (55:4), pp. 871–882.
- Abrahams, A. S., Fan, W., Wang, G. A., Zhang, Z., and Jiao, J. 2015. “An Integrated Text Analytic Framework for Product Defect Discovery,” *Production and Operations Management* (24:6), pp. 975–990.
- Alves, H., Fernandes, C., and Raposo, M. 2016. “Value Co-Creation: Concept and Contexts of Application and Study,” *Journal of Business Research* (69:5), Elsevier, pp. 1626–1633.
- Amit, R., and Zott, C. 2001. “Value Creation in E-business,” *Strategic Management Journal* (22:6-7), Wiley Online Library, pp. 493–520.
- Animesh, A., Viswanathan, S., and Agarwal, R. 2011. “Competing ‘creatively’ in Sponsored Search Markets: The Effect of Rank, Differentiation Strategy, and Competition on Performance,” *Information Systems Research* (22:1), pp. 153–169.
- Apté, C., Damerou, F., and Weiss, S. M. 1994. “Automated Learning of Decision Rules for Text Categorization,” *ACM Transactions on Information Systems* (12:3), pp. 233–251.
- Aral, S., Dellarocas, C., and Godes, D. 2013. “Introduction to the Special Issue—social Media and Business Transformation: A Framework for Research,” *Information Systems Research* (24:1), INFORMS, pp. 3–13.
- Arora, A., Krishnan, R., Telang, R., and Yang, Y. 2010. “An Empirical Analysis of Software Vendors’ Patch Release Behavior: Impact of Vulnerability Disclosure,” *Information Systems Research* (21:1), pp. 115–132.
- Bakos, Y., and Brynjolfsson, E. 1999. “Bundling Information Goods: Pricing, Profits, and Efficiency,” *Management Science* (45:12), pp. 1613–1630.
- Balasubramanian, S., Bhattacharya, S., and Krishnan, V. V. 2015. “Pricing Information Goods: A Strategic Analysis of the Selling and Pay-per-Use Mechanisms,” *Marketing Science* (34:2), INFORMS, pp. 218–234.
- Baldwin, C. Y., and Clark, K. B. 2006. “The Architecture of Participation: Does Code Architecture Mitigate Free Riding in the Open Source Development Model?,” *Management Science* (52:7), INFORMS, pp. 1116–1127.
- Basole, R. C., and Karla, J. 2011. “On the Evolution of Mobile Platform Ecosystem Structure and Strategy,” *Business & Information Systems Engineering* (3:5), Springer, p. 313.
- Basole, R. C., and Karla, J. 2012. “Value Transformation in the Mobile Service Ecosystem: A Study of App Store Emergence and Growth,” *Service Science* (4:1), INFORMS, pp. 24–41.
- Beal, R. M. 2000. “Competing Effectively: Environmental Scanning, Competitive Strategy and Organisational Performance in Small Manufacturing Firms,” *Journal of Small Business Management* (38:January), pp. 27–47.
- Becerra, M., Santaló, J., and Silva, R. 2013. “Being Better vs. Being Different: Differentiation, Competition, and Pricing Strategies in the Spanish Hotel Industry,” *Tourism Management* (34), Elsevier, pp. 71–79.
- Bengio, Y. 2009. “Learning Deep Architectures for AI,” *Foundations and Trends® in Machine Learning* (2:1), Now Publishers, Inc., pp. 1–127.
- Bengio, Y., Lamblin, P., Popovici, D., and Larochelle, H. 2007. “Greedy Layer-Wise Training of

- Deep Networks,” in *Advances in Neural Information Processing Systems*, pp. 153–160.
- Bharadwaj, A. S. 2000. “A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation,” *MIS Quarterly*, JSTOR, pp. 169–196.
- Bhargava, H. K., and Choudhary, V. 2001. “Information Goods and Vertical Differentiation,” *Journal of Management Information Systems* (18:2), Taylor & Francis, pp. 89–106.
- Bhargava, H. K., and Choudhary, V. 2008. “Research Note — When Is Versioning Optimal for Information Goods?,” *Management Science* (54:5), pp. 1029–1035.
- Boudreau, K. J. 2012. “Let a Thousand Flowers Bloom? An Early Look at Large Numbers of Software App Developers and Patterns of Innovation,” *Organization Science* (23:5), pp. 1409–1427.
- Brun, C., and Hagege, C. 2013. “Suggestion Mining: Detecting Suggestions for Improvement in Users’ Comments,” *Research in Computing Science* (70), pp. 171–181.
- Büschken, J., and Allenby, G. M. 2016. “Sentence-Based Text Analysis for Customer Reviews,” *Marketing Science* (35:6), INFORMS, pp. 953–975.
- Caves, R. E., and Williamson, P. J. 1985. “What Is Product Differentiation, Really?,” *The Journal of Industrial Economics*, JSTOR, pp. 113–132.
- Chakravarti, A., and Xie, J. 2006a. “The Impact of Standards Competition on Consumers: Effectiveness of Product Information and Advertising Formats,” (XLIII:May), pp. 224–236.
- Chakravarti, A., and Xie, J. 2006b. “The Impact of Standards Competition on Consumers: Effectiveness of Product Information and Advertising Formats,” (43:2), pp. 224–236.
- Chamberlin, E. 2010. “H.(1933) The Theory of Monopolistic Competition,” *Cambridge, Mass, Harvard University*.
- Chen, H., Chiang, R. H. L., and Storey, V. C. 2012. “Business Intelligence and Analytics: From Big Data to Big Impact,” *MIS Quarterly* (36:4), pp. 1165–1188.
- Chen, M.-J., and MacMillan, I. C. 1992. “Nonresponse and Delayed Response to Competitive Moves: The Roles of Competitor Dependence and Action Irreversibility,” *Academy of Management Journal* (35:3), Academy of Management, pp. 539–570.
- Chen, M.-J., Smith, K. G., and Grimm, C. M. 1992. “Action Characteristics as Predictors of Competitive Responses,” *Management Science* (38:3), INFORMS, pp. 439–455.
- Chen, N., Lin, J., Hoi, S., Xiao, X., and Zhang, B. 2014. “AR-Miner: Mining Informative Reviews for Developers from Mobile App Marketplace,” in *Proceedings of the 36th International Conference on Software Engineering*, ACM, pp. 767–778.
- Chen, Y. 2013. “Context-Dependent Preferences and Innovation Strategy Context-Dependent Preferences and Innovation Strategy,” (February 2014).
- Chen, Y., and Iyer, G. 2002. “Research Note Consumer Addressability and Customized Pricing,” *Marketing Science* (21:2), INFORMS, pp. 197–208.
- 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. 477–491.
- Collobert, R., and Weston, J. 2008. “A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning,” in *Proceedings of the 25th International Conference on Machine Learning*, ACM, pp. 160–167.
- Comino, S., Udine, U., Manenti, F. M., and Economiche, S. 2016. “Updates Management in Mobile Applications . iTunes vs Google Play Updates Management in Mobile Applications . iTunes vs Google Play . * .,”
- Cooper, R. G., and Kleinschmidt, E. J. 1994. “Determinants of Timeliness in Product

- Development,” *Journal of Product Innovation Management* (11:5), Elsevier, pp. 381–396.
- Crawford-Welch, S. 1990. “Market Segmentation in the Hospitality Industry,” *Hospitality Research Journal* (14:2), ICHRIE Sage CA: Los Angeles, CA, pp. 295–308.
- Cusumano, M. A., Kahl, S. j, and Suarez, F. F. 2008. “Services, Industry Evolution, and the Copetitive Strategies of Product Firms,” *Academy of Management Journal* (51:2), pp. 315–334.
- Datta, D., and Kajanana, S. 2013. “Do App Launch Times Impact Their Subsequent Commercial Success? An Analytical Approach,” in *Cloud Computing and Big Data (CloudCom-Asia), 2013 International Conference on*, IEEE, pp. 205–210.
- Dey, D., and Lahiri, A. 2016. “Versioning of Video Games: Go Vertical in a Horizontal Market?,” *Proceedings of the Annual Hawaii International Conference on System Sciences* (2016–March), pp. 5259–5268.
- Dickson, P. R., and Ginter, J. L. 1987. “Market Segmentation , Product Differentiation , and Marketing Strategy,” (51:April), pp. 1–10.
- Distimo. 2009. “Pple App Store, BlackBerry App World, Google Android MarketNokia Ovi Store & Windows Marketplace for Mobile.”
- Dougherty, D. 2001a. “Reimagining the Differentiation and Integration of Work for Sustained Product Innovation,” *Organization Science* (12:5), pp. 612–631.
- Dougherty, D. 2001b. “Reimagining the Differentiation and Integration of Work for Sustained Product Innovation,” *Organization Science* (12:5), INFORMS, pp. 612–631.
- Dougherty, D., and Dunne, D. D. 2011. “Organizing Ecologies of Complex Innovation,” *Organization Science* (22:5), pp. 1214–1223.
- Duan, W., Gu, B., and Whinston, A. B. 2008. “The Dynamics of Online Word-of-Mouth and Product sales—An Empirical Investigation of the Movie Industry,” *Journal of Retailing* (84:2), Elsevier, pp. 233–242.
- Ethiraj, S. K., and Zhu, D. H. 2008. “Performance Effects of Imitative Entry,” *Strategic Management Journal* (29:8), Wiley Online Library, pp. 797–817.
- Ferrier, W. J., Smith, K. G., and Grimm, C. M. 1999. “The Role of Competitive Action in Market Share Erosion and Industry Dethronement: A Study of Industry Leaders and Challengers,” *Academy of Management Journal* (42:4), Academy of Management, pp. 372–388.
- Fisher, M., Gallino, S., and Li, J. 2017. “Competition-Based Dynamic Pricing in Online Retailing: A Methodology Validated with Field Experiments,” *Management Science*, INFORMS.
- Forman, C., Ghose, A., and Wiesenfeld, B. 2008. “Examining the Relationship between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets,” *Information Systems Research* (19:3), INFORMS, pp. 291–313.
- French, S. A., Story, M., Jeffery, R. W., Snyder, P. A. T., Eisenberg, M., Sidebottom, A., and Murray, D. 1997. “Pricing Strategy to Promote Fruit and Vegetable Purchase in High School Cafeterias,” *Journal of the Academy of Nutrition and Dietetics* (97:9), Elsevier, pp. 1008–1010.
- Galvis Carreño, L. V, and Winbladh, K. 2013. “Analysis of User Comments: An Approach for Software Requirements Evolution,” in *Proceedings of the 2013 International Conference on Software Engineering*, pp. 582–591.
- Gan, Z., Pu, Y., Henao, R., Li, C., He, X., and Carin, L. 2017. “Learning Generic Sentence Representations Using Convolutional Neural Networks,” in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 2390–2400.

- Garg, R., and Telang, R. 2014. "Inferring App Demand from Publicly Available Data," *MIS Quarterly* (37:4), pp. 1253–1264.
- Ghose, A., and Han, S. P. 2014. "Estimating Demand for Mobile Applications in the New Economy," *Management Science* (60:6), pp. 1470–1488.
- Ghose, A., and Ipeirotis, P. G. 2011. "Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics," *IEEE Transactions on Knowledge and Data Engineering* (23:10), pp. 1498–1512.
- Gnyawali, D. R., Fan, W., and Penner, J. 2010. "Competitive Actions and Dynamics in the Digital Age: An Empirical Investigation of Social Networking Firms," *Information Systems Research* (21:3), INFORMS, pp. 594–613.
- Gnyawali, D. R., He, J., and Madhavan, R. 2006. "Impact of Co-Opetition on Firm Competitive Behavior: An Empirical Examination," *Journal of Management* (32:4), Sage Publications Sage CA: Thousand Oaks, CA, pp. 507–530.
- Goldberg, D. M., and Abrahams, A. S. 2018. "A Tabu Search Heuristic for Smoke Term Curation in Safety Defect Discovery," *Decision Support Systems* (105), Elsevier, pp. 52–65.
- Greenstein, S., and Markovich, S. 2012. "Pricing Experience Goods in Information Good Markets: The Case of eBusiness Service Providers," *International Journal of the Economics of Business* (19:1), pp. 119–139.
- Gu, B., and Ye, Q. 2014. "First Step in Social Media: Measuring the Influence of Online Management Responses on Customer Satisfaction," *Production and Operations Management* (23:4), pp. 570–582.
- Guo, X., Ling, L., Yang, C., Li, Z., and Liang, L. 2013. "Optimal Pricing Strategy Based on Market Segmentation for Service Products Using Online Reservation Systems: An Application to Hotel Rooms," *International Journal of Hospitality Management* (35), Elsevier Ltd, pp. 274–281.
- Hausman, J. A. 1978. "Specification Tests in Econometrics," *Econometrica: Journal of the Econometric Society*, JSTOR, pp. 1251–1271.
- Hevner, V. A., March, S. T., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), Springer, pp. 75–105.
- Hinterhuber, A. 2008. "Customer Value-Based Pricing Strategies: Why Companies Resist," *Journal of Business Strategy* (29:4), Emerald Group Publishing Limited, pp. 41–50.
- Hinton, G. E., and Salakhutdinov, R. R. 2006. "Reducing the Dimensionality of Data with Neural Networks," *Science* (313:5786), American Association for the Advancement of Science, pp. 504–507.
- Hippel, E. Von, and Krogh, G. Von. 2003. "Open Source Software and the ?Private-Collective? Innovation Model: Issues for Organization Science," *Organization Science* (14:2), pp. 209–223.
- Jha, N., and Mahmoud, A. 2017. "Mining User Requirements from Application Store Reviews Using Frame Semantics," in *International Working Conference on Requirements Engineering: Foundation for Software Quality*, Springer, pp. 273–287.
- Johansson, M., Hallberg, N., Hinterhuber, A., Zbaracki, M., and Liozu, S. 2012. "Pricing Strategies and Pricing Capabilities," *Journal of Revenue and Pricing Management* (11:1), Springer, pp. 4–11.
- Katzmarzik, A. 2011. "Product Differentiation for Software-as-a-Service Providers," *Business and Information Systems Engineering* (3:1), pp. 19–31.

- Kim, S. J., Kim, K. H., and Choi, J. 2017. "Design Innovation and Purchase Behavior for Augmented Products," *Journal of Business Research* (September), Elsevier, pp. 0–1.
- Kim, Y. 2014. "Convolutional Neural Networks for Sentence Classification," *arXiv Preprint arXiv:1408.5882*.
- Kleinbaum, D. G. 1998. "Survival Analysis, a Self-Learning Text," *Biometrical Journal* (40:1), Wiley Online Library, pp. 107–108.
- Kleinschmidt, E. J., and Cooper, D. R. 1991. "The Impact of Product Innovativeness on Performance," *Journal of Product Innovation Management* (8), pp. 240–251.
- Lai, H., Hsu, J. S.-C., and Wu, M.-X. 2018. "The Impact S of Requested Permission on Mobile App Adoption: The Insights Based on an Experiment in Taiwan," in *Proceedings of the 51st Hawaii International Conference on System Sciences*.
- Landis, J. R., and Koch, G. G. 1977. "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, JSTOR, pp. 159–174.
- Law, D., Gruss, R., and Abrahams, A. S. 2017. "Automated Defect Discovery for Dishwasher Appliances from Online Consumer Reviews," *Expert Systems with Applications* (67), Elsevier Ltd, pp. 84–94.
- Lee, G., and Raghu, T. S. 2014. "Determinants of Mobile Apps' Success: Evidence from the App Store Market," *Journal of Management Information Systems* (31:2), Taylor & Francis, pp. 133–170.
- Lee, J., and Miller, D. 1999. "People Matter: Commitment to Employees, Strategy and Performance in Korean Firms," *Strategic Management Journal*, JSTOR, pp. 579–593.
- Li, J., Xia, J., and Zajac, E. J. 2018. "On the Duality of Political and Economic Stakeholder Influence on Firm Innovation Performance: Theory and Evidence from Chinese Firms," *Strategic Management Journal* (39:1), Wiley Online Library, pp. 193–216.
- Liu, B. 2012. "Sentiment Analysis and Opinion Mining," *Synthesis Lectures on Human Language Technologies* (5:1), Morgan & Claypool Publishers, pp. 1–167.
- Liu, B., Hu, M., and Cheng, J. 2005. "Opinion Observer: Analyzing and Comparing Opinions on the Web," in *Proceedings of the 14th International Conference on World Wide Web*, ACM, pp. 342–351.
- Liu, C. Z., Au, Y. A., and Choi, H. S. 2014. "Effects of Freemium Strategy in the Mobile App Market: An Empirical Study of Google Play," *Journal of Management Information Systems* (31:3), Taylor & Francis, pp. 326–354.
- Liu, F., and Brandyberry, A. 2014. "Exploring the Effects of Aggregate Review Characteristics on Mobile Application Adoption," in *Thirty Fifth International Conference on Information Systems*, Auckland, New Zealand.
- Luo, X. 2009. "Quantifying the Long-Term Impact of Negative Word of Mouth on Cash Flows and Stock Prices," *Marketing Science* (28:1), pp. 148–165.
- Luo, X., Zhang, J., and Duan, W. 2013. "Social Media and Firm Equity Value," *Information Systems Research* (24:1), INFORMS, pp. 146–163.
- LYU, C., and SHANG, Y. 2017. "The Time-Space Effect Research of High-Tech Industrial Agglomeration Model and Innovation Output," *Journal of Management* (2), p. 7.
- Maalej, W., and Nabil, H. 2015. "Bug Report, Feature Request, or Simply Praise? On Automatically Classifying App Reviews," in *Requirements Engineering Conference (RE), 2015 IEEE 23rd International*, IEEE, pp. 116–125.
- Maalej, W., Nabil, H., and Stanik, C. 2017. "On the Automatic Classification of App Reviews,"

- Software Engineering 2017*, Gesellschaft Für Informatik eV.
- MacMillan, D., Burrows, P., and Ante, S. E. 2009. "Inside the App Economy," *Business Week* (22).
- Makadok, R. 2010. "The Interaction Effect of Rivalry Restraint and Competitive Advantage on Profit: Why the Whole Is Less than the Sum of the Parts," *Management Science* (56:2), INFORMS, pp. 356–372.
- McIlroy, S., Ali, N., and Hassan, A. E. 2016. "Fresh Apps: An Empirical Study of Frequently-Updated Mobile Apps in the Google Play Store," *Empirical Software Engineering* (21:3), Springer, pp. 1346–1370.
- Mikolov, T., Chen, K., Corrado, G., and Dean, J. 2013. "Efficient Estimation of Word Representations in Vector Space," *arXiv Preprint arXiv:1301.3781*.
- Miller, D. 1986. "Configurations of Strategy and Structure: Towards a Synthesis," *Strategic Management Journal* (7:3), Wiley Online Library, pp. 233–249.
- Miller, D. 1988. "Relating Porter's Business Strategies to Environment and Structure: Analysis and Performance Implications," *Academy of Management Journal* (31:2), Academy of Management, pp. 280–308.
- Mintzberg, H. 1988. "Generic Strategies: Toward a Comprehensive Framework," *Advances in Strategic Management* (5), pp. 1–67.
- Moghaddam, S. 2015. "Beyond Sentiment Analysis: Mining Defects and Improvements from Customer Feedback," in *Advances in Information Retrieval*, Springer, pp. 400–410.
- Mudambi, S. M., and Schuff, D. 2010. "What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon. Com," *MIS Quarterly* (34:1), pp. 185–200 (available at <http://ssrn.com/abstract=2175066>).
- Mummalaneni, V., Gruss, R., Goldberg, D. M., Ehsani, J. P., and Abrahams, A. S. 2018. "Social Media Analytics for Quality Surveillance and Safety Hazard Detection in Baby Cribs," *Safety Science* (104), Elsevier, pp. 260–268.
- Nagi, J., Ducatelle, F., Di Caro, G. A., Cireşan, D., Meier, U., Giusti, A., Nagi, F., Schmidhuber, J., and Gambardella, L. M. 2011. "Max-Pooling Convolutional Neural Networks for Vision-Based Hand Gesture Recognition," in *Signal and Image Processing Applications (ICSIPA), 2011 IEEE International Conference on*, IEEE, pp. 342–347.
- Nakagawa, T., Inui, K., and Kurohashi, S. 2010. "Dependency Tree-Based Sentiment Classification Using CRFs with Hidden Variables," in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Association for Computational Linguistics, pp. 786–794.
- Narasimhan, R., Ghosh, S., and Mendez, D. 1993. "A Dynamic Model of Product Quality and Pricing Decisions on Sales Response," *Decision Sciences* (24:5), Wiley Online Library, pp. 893–908.
- Nasri, L., Baghersad, M., Gruss, R., Marucchi, N. S. W., Abrahams, A. S., and Ehsani, J. P. 2018. "An Investigation into Online Videos as a Source of Safety Hazard Reports," *Journal of Safety Research* (65), Elsevier, pp. 89–99.
- Nayebi, M., Adams, B., and Ruhe, G. 2016. "Release Practices for Mobile Apps -- What Do Users and Developers Think?," *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, pp. 552–562.
- Oh, J., Kim, D., Lee, U., Lee, J.-G., and Song, J. 2013. "Facilitating Developer-User Interactions with Mobile App Review Digests," in *CHI'13 Extended Abstracts on Human Factors in*

- Computing Systems*, ACM, pp. 1809–1814.
- Pagano, D., and Maalej, W. 2013. “User Feedback in the Appstore: An Empirical Study,” in *Requirements Engineering Conference (RE), 2013 21st IEEE International*, IEEE, pp. 125–134.
- Pang, B., Lee, L., and Vaithyanathan, S. 2002. “Thumbs up?: Sentiment Classification Using Machine Learning Techniques,” in *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing-Volume 10*, Association for Computational Linguistics, pp. 79–86.
- Panichella, S., Di Sorbo, A., Guzman, E., Visaggio, C. A., Canfora, G., and Gall, H. C. 2015. “How Can I Improve My App? Classifying User Reviews for Software Maintenance and Evolution,” in *Software Maintenance and Evolution (ICSME), 2015 IEEE International Conference on*, IEEE, pp. 281–290.
- Pennington, J., Socher, R., and Manning, C. 2014. “Glove: Global Vectors for Word Representation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543.
- Porter, J. 2010. *Designing for the Social Web (Google eBook)*, Peachpit Press (available at http://books.google.com/books/about/Designing_for_the_Social_Web.html?id=CNJuvYBI_IsC&pgis=1).
- Porter, M. E. 1980. “Competitive Strategy: Techniques for Analyzing Industries and Competition,” *New York*, p. 300.
- Powers, D. M. 2011. “Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation,” Bioinfo Publications.
- Prahalad, C. K., and Ramaswamy, V. 2000. “Co-Opting Customer Competence,” *Harvard Business Review* (78:1), pp. 79–90.
- Prahalad, C. K., and Ramaswamy, V. 2004. “Co-creation Experiences: The next Practice in Value Creation,” *Journal of Interactive Marketing* (18:3), Wiley Online Library, pp. 5–14.
- Prajogo, D. I., and Ahmed, P. K. 2006. “Relationships between Innovation Stimulus, Innovation Capacity, and Innovation Performance,” *R & D Management* (36:5), pp. 499–515.
- Qiao, Z., Wang, G. A., Zhou, M., and Fan, W. 2018. “The Impact of Customer Reviews on Product Innovation: Empirical Evidence in Mobile Apps,” in *Analytics and Data Science*, Springer, pp. 95–110.
- Ragaglia, D., and Roma, P. 2014. “Understanding the Drivers of the Daily App Rank : The Role of Revenue Models,” in *Proceedings of the 26th Annual Conference of the Productions and Operations Management Society*, pp. 1–10.
- Ramanand, J., Bhavsar, K., and Pedanekar, N. 2010. “Wishful Thinking: Finding Suggestions and ‘Buy’ Wishes from Product Reviews,” in *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, Association for Computational Linguistics, pp. 54–61.
- Ramaswamy, V., and Prahalad, C. K. 2004. “Co-Creation Experiences: The Next Ractice in Value Creation,” *Journal of Interactive Marketing* (18:3), pp. 5–14.
- Ramos, J. 2003. “Using Tf-Idf to Determine Word Relevance in Document Queries,” in *Proceedings of the First Instructional Conference on Machine Learning*.
- Rehurek, R., and Sojka, P. 2010. “Software Framework for Topic Modelling with Large Corpora,” in *In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, Citeseer.

- Sak, H., Senior, A., and Beaufays, F. 2014. "Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling," in *Fifteenth Annual Conference of the International Speech Communication Association*.
- Salakhutdinov, R., and Hinton, G. 2007. "Learning a Nonlinear Embedding by Preserving Class Neighbourhood Structure," in *Artificial Intelligence and Statistics*, pp. 412–419.
- Sambamurthy, V., Bharadwaj, A., and Grover, V. 2003. "Shaping Agility through Digital Options: Reconceptualizing the Role of Information Technology in Contemporary Firms," *MIS Quarterly*, JSTOR, pp. 237–263.
- Shapiro, C., and Varian, H. R. 1998. "Versioning: The Smart Way to Sell Information," *Havard Business Review*, pp. 106–114.
- Shen, W., Hu, Y. J., and Rees, J. 2015. "Competing for Attention: An Empirical Study of Online Reviewers' Strategic Behaviors," *MIS Quarterly* (39:3), Society for Information Management and The Management Information Systems Research Center, pp. 683–696.
- Smith, K. G., Ferrier, W. J., and Ndofor, H. 2001. "Competitive Dynamics Research," *The Blackwell Handbook of Strategic Management*, Wiley Online Library, pp. 309–354.
- Society, T. E. 2011. "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," *Econometrica* (79:5), pp. 1407–1451.
- Di Sorbo, A., Panichella, S., Alexandru, C. V., Shimagaki, J., Visaggio, C. A., Canfora, G., and Gall, H. C. 2016. "What Would Users Change in My App? Summarizing App Reviews for Recommending Software Changes," in *Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering*, ACM, pp. 499–510.
- Spann, M., Fischer, M., and Tellis, G. J. 2015. "Skimming or Penetration? Strategic Dynamic Pricing for New Products," *Marketing Science* (34:2), pp. 235–249.
- Stavrianou, A., and Brun, C. 2012. "Opinion and Suggestion Analysis for Expert Recommendations," in *Proceedings of the Workshop on Semantic Analysis in Social Media*, Association for Computational Linguistics, pp. 61–69.
- Stieglitz, S., and Dang-Xuan, L. 2013a. "Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior," *Journal of Management Information Systems* (29:4), pp. 217–248.
- Stieglitz, S., and Dang-Xuan, L. 2013b. "Social Media and Political Communication: A Social Media Analytics Framework," *Social Network Analysis and Mining* (3:4), Springer, pp. 1277–1291.
- Systems, M. I. 2017. "Information Goods and Vertical Differentiation Author (S): Hemant K . Bhargava and Vidyanand Choudhary Source : Journal of Management Information Systems , Vol . 18 , No . 2 (Fall , 2001), Pp . 89-106 Stable URL : <http://www.jstor.org/stable/40398530> ," (18:2), pp. 89–106.
- Täckström, O., and McDonald, R. 2011. "Semi-Supervised Latent Variable Models for Sentence-Level Sentiment Analysis," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers-Volume 2*, Association for Computational Linguistics, pp. 569–574.
- Taylor Pentina, V. 2011. "Mobile Application Adoption by Young Adults: A Social Network Perspective ," *International Journal of Mobile Marketing* (6:2), pp. 60–70.
- Varian, H. 1997. "Versioning Information Goods," *University of California, Berkeley* (January), pp. 1–13.
- Wang, C., Jiang, F., and Yang, H. 2017. "A Hybrid Framework for Text Modeling with

- Convolutional RNN,” in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, pp. 2061–2069.
- Wang, H., Lu, Y., and Zhai, C. 2010a. “Latent Aspect Rating Analysis on Review Text Data: A Rating Regression Approach,” in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA: ACM, pp. 783–792.
- Wang, H., Lu, Y., and Zhai, C. 2010b. “Latent Aspect Rating Analysis on Review Text Data: A Rating Regression Approach,” *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, pp. 783–792.
- Wang, Q., Li, B., and Singh, P. V. 2018. “Copycats vs. Original Mobile Apps: A Machine Learning Copycat-Detection Method and Empirical Analysis,” *Information Systems Research*, INFORMS.
- Wang, X., Jiang, W., and Luo, Z. 2016. “Combination of Convolutional and Recurrent Neural Network for Sentiment Analysis of Short Texts,” in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pp. 2428–2437.
- Weston, J. F. J. F., and Besley, S. 2008. *Fundamentos de Administración Financiera*, Cengage Learning Editores.
- Winkler, M., Abrahams, A. S., Gruss, R., and Ehsani, J. P. 2016. “Toy Safety Surveillance from Online Reviews,” *Decision Support Systems* (90), Elsevier, pp. 23–32.
- Xu, B., Wang, N., Chen, T., and Li, M. 2015. “Empirical Evaluation of Rectified Activations in Convolutional Network,” *arXiv Preprint arXiv:1505.00853*.
- Xu, S. 2015. “Balancing the Two Knowledge Dimensions in Innovation Efforts: An Empirical Examination among Pharmaceutical Firms,” *Journal of Product Innovation Management* (32:4), pp. 610–621.
- Yi, Y., and Gong, T. 2013. “Customer Value Co-Creation Behavior: Scale Development and Validation,” *Journal of Business Research* (66:9), Elsevier, pp. 1279–1284.
- Zentes, J., Morschett, D., and Schramm-Klein, H. 2017. “Pricing,” in *Strategic Retail Management*, Springer, pp. 279–306.
- Zhang, Y., and Wallace, B. 2015. “A Sensitivity Analysis of (and Practitioners’ Guide to) Convolutional Neural Networks for Sentence Classification,” *arXiv Preprint arXiv:1510.03820*.
- Zhou, S., Qiao, Z., Du, Q., Wang, A. G., Fan, W., and Yan, X. 2018. “Measuring Customer Agility from Online Reviews Using Big Data Text Analytics,” *Journal of Management Information Systems*, p. Forthcoming.
- Zhu, F., and Zhang, X. 2010. “Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics,” *Journal of Marketing* (74:2), pp. 133–148.
- Zhu, J., Zou, H., Rosset, S., and Hastie, T. 2009. “Multi-Class Adaboost,” *Statistics and Its Interface* (2:3), pp. 349–360.

4 DEEP LEARNING-BASED USER FEEDBACK CLASSIFICATION IN MOBILE APP REVIEWS

4.1 Introduction

Due to the strong competition in the mobile app industry, app quality has become an essential factor for apps to gain a competitive advantage in the mobile app market (Chen et al. 2014). The mobile app market is growing rapidly, with millions of apps and developers, billions of users, and billions of dollars in revenue. For example, the Apple App Store, one of the most competitive app markets, offered 500 apps upon its initiation in 2008 and had over 2.2 million apps by 2017 (Lai et al. 2018). Moreover, as the market competition intensifies, app developers, in order to stay competitive, usually employ an iterative process to get users involved in an app's development lifecycle by allowing users to provide feedback on the app. The Apple App Store provides a mechanism with which users can rate and review the apps they have used. Given the large volume of reviews available in the Apple App Store (Zhou et al. 2018), it is important for app developers to efficiently extract and understand user needs from user reviews (Aral et al. 2013; Chen et al. 2014; Nayebi et al. 2016).

Compared with the bug reporting and feature request mechanisms used in traditional software development, there are two outstanding challenges to extracting valuable user feedback from unstructured online reviews. First, only around one-third of app reviews contain objective statements (Abrahams et al. 2013; Law et al. 2017; Oh et al. 2013; Winkler et al. 2016). Second, manually processing a large volume of unstructured user reviews and extracting potential user needs from those reviews can be tedious. For example, the *Candy Crush Saga* game app in the Apple App Store receives on average more than 1,500 reviews per day; processing all those

reviews manually and on a daily basis would be time-consuming. Also, although the online reviews of competing apps can provide useful user feedback for app developers to make competitive decisions regarding app design, it would be difficult to manually process the user reviews of all competing apps. Thus, it is more efficient and desirable to automatically, rather than manually, extract user needs from unstructured online reviews.

To address both challenges to extracting user feedback from unstructured online reviews, existing opinion mining studies have proposed methods to automatically cluster or classify user needs in online reviews. Oh et al. (2013) use a Support Vector Machine (SVM) classifier to automatically categorize user reviews in the Google Play store into functional and non-functional requests. Chen et al. (Chen et al. 2014) propose two methods, Latent Dirichlet Allocation (LDA) and Aspect and Sentiment Unification Model (ASUM), to identify informative and non-informative topics in user reviews and find that LDA performs better than ASUM. However, the generalizability of the proposed methods to a large number of apps is questionable because the methods were only tested using reviews of four Android apps. Panichella et al. (2015) propose a system to automatically classify user reviews into two types: software maintenance and requirement evolution. However, these approaches have two shortcomings. First, although text classification methods can benefit from grammatical (e.g., Part-of-Speech Tagging), syntactical (e.g., noun phrases, verb phrases, prepositional phrases), and semantic features (e.g., word-sense) (Abrahams et al. 2015; Chen et al. 2012; Goldberg and Abrahams 2018), these methods largely consider feature independence and ignore the contextual connections between features. Second, existing approaches to classifying text often rely on dictionaries built using past reviews (Apté et al. 1994; Nasri et al. 2018; Pagano and Maalej 2013), which may not cover the new words or

perspectives in a dataset of new reviews and thus ultimately affect the performance of text classification.

To overcome these existing challenges and effectively extract user feedback from app reviews, this research proposes a deep learning-based opinion classification method to identify user needs from online reviews. The proposed method improves existing text classification methods by capturing the semantic context of words using deep text learning. In addition, in order to balance result interpretability and analysis granularity, the proposed method is implemented at the sentence level instead of the review or document level. The method helps mobile app developers automatically extract user needs from a large volume of online reviews with greater effectiveness than traditional machine learning algorithms.

This study aims to make several contributions. First, the study proposes a deep learning-based text analytics framework to conduct sentence-level opinion classification of unstructured online reviews. Deep learning helps capture the contextual information of the syntactic features and is expected to achieve better text classification performance than traditional learning algorithms. Second, this study contributes to the requirement engineering literature. As online user reviews become increasingly popular, effectively identifying user feedback in unstructured reviews can help app developers stay competitive in hypercompetitive markets, such as mobile app stores. Third, this study may also benefit the design of online review platforms. Platform providers can use the proposed method in order to improve the quality of user reviews and help app developers efficiently identify user needs. A review quality control mechanism can be implemented to help users write algorithm-friendly reviews, which will improve the proposed method's classification performance.

The rest of the chapter proceeds as follows. Section 2 reviews related work, and Section 3 states the research objective. Section 4 presents the proposed domain-oriented, deep learning method for opinion mining, and Section 5 describes the experiment and evaluation results. The last section concludes the study's findings, discusses the limitations of the study, and makes suggestions for future work.

4.2 Related Work

In this section, the literature related to customer value co-creation, text classification for opinion mining, and text analytics of mobile app reviews, is reviewed.

4.2.1 Customer Value Co-Creation

Several studies demonstrate the importance of customer involvement in product design. As described by Yi and Gong (2013), high-quality interactions between customers and firms are the key to determining novel sources of competitive advantage. Value that is co-created by firms and customers is associated with product competitive advantages, product success, and commercial success (Alves et al. 2016). Value co-creation refers to a joint creation process that involves both the company and the customers (Prahalad and Ramaswamy 2000, 2004; Ramaswamy and Prahalad 2004). Presenting a holistic perspective of value co-creation, Prahalad and Ramaswamy (2000) document the transformation of customers from “passive listeners” to “active players” over time, which is the foundation of value co-creation.

Recent studies of value co-creation in the context of mobile apps show that it is important to get the customer involved in product design in order to improve app quality and success (Chen et al. 2014; Gu and Ye 2014; Maalej and Nabil 2015). Building on the studies of von Hippel and von Krogh (2003) and Baldwin and Clark (2006), Boudreau (2012) shows that if communication costs

are relatively low compared to design costs, then app firms refrain from competing using self-design and prefer collaborative innovation with customers over independent innovation. This is in line with findings from studies that examine how users can be involved in value co-creation in the mobile app industry.

In the mobile app industry, online customer reviews are an important channel through which customers and firms can communicate regarding product quality. In fact, online customer reviews have been a significant driving force in the evolution of several apps (Pagano and Maalej 2013; Qiao et al. 2018; Zhou et al. 2018). As the number of online app reviews increases at an unprecedented speed, many app firms seek to create business opportunities by discovering business values from the reviews (Chen et al. 2014; Maalej et al. 2017; Panichella et al. 2015; Di Sorbo et al. 2016). The content of online reviews is mostly unstructured text that is often difficult to manually analyze when the volume is large. Therefore, it is necessary to develop effective and efficient ways to automatically process a large volume of text-based user reviews and extract valuable user opinions for customer value co-creation.

4.2.2 Information Types in App Reviews

Maalej and Nabil (2015) categorize app reviews into four basic types: bug reports, feature requests, user experiences, and ratings. Bug reports refer to the problems with the app that should be fixed, such as an erroneous behavior, a performance issue, or an unexpected crash. Feature requests describe new features proposed by consumers, including new functions. User experiences are the documentation of the user's interaction with the app, while ratings are sentiment text represented by different numbers of stars. This study focuses on bug reports and feature requests because they contain specific user feedback and can be used to improve product design. User experiences and

ratings are grouped together as other types because they are not directly related to the identification of user needs. Table 4-1 presents some examples of the different types of app reviews. As the examples indicate, a user review may consist of different types of information, with each sentence focusing on one specific information type. Therefore, app reviews should be analyzed at the sentence level.

Table 4-1 Examples of Different Types of App Reviews

No.	Review Content	Information Types
1.	The clock doesn't keep time like a regular clock.	Bug Reports
2.	Only problem I have with this game is that it crashes too much.	Bug Reports
3.	The connection for the game is kind of sucky.	Bug Reports
4.	I know that they already have that in <i>Need for Speed Hot Pursuit</i> , but I was hoping for something more diverse.	Feature Request
5.	Also Please add a multi player.	Feature Request
6.	Add Chevy cars and trucks too!	Feature Request
7.	I like this game a lot.	Other Types
8.	Great game!	Other Types

4.2.3 Sentence-Level User Feedback Classification

Several studies have proposed methods for sentence-level user feedback classification (Büschken and Allenby 2016; Kim 2014; Täckström and McDonald 2011). Liu (2012) explains that sentence-level classification is appropriate to classify objective information because document-level opinion mining is too coarse for applications, whereas the results of aspect-level or phrase-level opinion mining may be difficult to interpret. There are several studies in the sentence-level objective information classification literature (Liu et al. 2005; Moghaddam 2015; Stieglitz and Dang-Xuan 2013a; Wang et al. 2010). Ramanand et al. (2010) design a rule-based method to discover “wishes” sentences, in which customers make suggestions for product or service improvement. Wang et al. (2010b) propose a regression model that maps review sentences to predefined aspects of hotel reviews. However, their model relies on user ratings of those aspects and manually selected seed

words for each aspect. Nakagawa et al. (2010) propose a conditional, random field-based sentiment classification method that considers the interactions between words, which outperforms traditional methods that use only the bag-of-words feature. In a similar vein, Täckström and McDonald (2011) derive two variants of a semi-supervised method for sentence-level sentiment analysis. Their study shows that the joint use of learning sentence and document sentiment can improve prediction performance. These approaches mostly focus on sentiment analysis and categorize a given text as either positive or negative. Although distinguishing the sentiment of user reviews can help customers make purchasing decisions, it is still challenging to capture objective sentences from these reviews that can provide developers specific suggestions for product improvement.

To address this issue, recent opinion mining research focuses on discovering objective sentences that describe specific product features (Mummalaneni et al. 2018; Wang et al. 2010b) and product defects (Abrahams et al. 2015) from user-generated content. Table 4-2 provides a summary of sentence-level opinion classification studies for user-generated content. These studies can be categorized into two types: rule-based and machine learning based. Rule-based methods, such as that proposed by Brun and Hagege (2013), manually formulate linguistic rules to extract opinion sentences from customer reviews. Some machine learning-based methods, such as those proposed by Moghaddam (2015) and Galvis, Carreño, and Winbladh (2013), utilize LDA or topic modeling to extract topics from online customer reviews in an unsupervised or semi-supervised manner. LDA, or topic modeling, uses a collection of keywords to represent each topic. However, it is difficult to evaluate the quality of the topics, and the topics (i.e., the collections of words) are difficult to interpret.

Other machine learning methods apply supervised classification algorithms to opinion classification tasks. They extract linguistic features, such as bags-of-words (Pang et al. 2002) and

grammatical (e.g., Part-of-Speech Tagging), syntactical (e.g., noun phrases, verb phrases, prepositional phrases), and semantic features (e.g., word-sense) from text and apply classification algorithms, such as logistic regression, decision trees (DTs), multinomial naïve Bayes (MNB), and SVM. These methods consider that the linguistic features that can be extracted from a text are independent of each other. However, the feature extraction process ignores much of the contextual information embedded in sentence structures and word sequences. Thus, the deep learning technique has been introduced to text mining and natural language processing. One of the key components in deep text learning is word embedding, which is a language representation model that can capture the semantic and syntactic similarities between words. Deep learning captures the contextual information around words and the order between them and can help in the interpretation of textual data using a relatively holistic perspective. Moreover, deep learning has shown promising results in natural language processing applications, such as that of Stavrianou and Brun (2012). Last but not least, deep learning for user feedback classification is also understudied.

Table 4-2 Literature Reviews of Sentence-level User Feedback Classification

Study	Content Type	Domain	Method of Analysis	Classes
Brun and Hagege (2013)	Online Reviews	General	Rule-based	Suggestions for Improvement
Ramanand et al. (2010)	Product Reviews	Consumer Electronics & Banking	Rule-based	Suggestions / Wishes
Wang et al. (2010b)	Online Reviews	Hotel	Regression	Hotel Ranking / User Rating Behavior
Stavrianou and Brun (2012)	Product Reviews	Printer	SVM classifier	Recommendation
Stieglitz and Dang-Xuan (2013b)	Twitter	Politics	Naïve Bayesian & SVM	Political Opinions Summarization

Carreño and Winbladh (2013)	User Comments	Mobile Application	A Variant of Latent Dirichlet Allocation (LDA)	New / Changed Requirements
Abrahams et al. (2015)	User Postings	Automobile & Electronics	Logistic Regression	Product Defect
Moghaddam (2015)	Product Reviews	Mobile Application	Semi-supervised Latent Dirichlet Allocation (LDA)	Product Defect & Product Improvement
Panichella et al. (2015)	Product Reviews	Mobile Application	Naïve Bayes, Logistic Regression, SVM, Decision Tree	App maintenance and Evolution
Maalej et al. (2017)	Product Reviews	Mobile Application	Naïve Bayes, Decision Tree, & Maximum Entropy	Bug reports, Feature Requests, User Experiences & Ratings
Jha and Mahmoud (2017)	Product Reviews	Mobile Application	Naïve Bayesian & SVM	Question / Problem
This study	Product Reviews	Mobile Application	Word Embedding & CNN & RNN	Bug Reports, Feature Requests, & Others

4.2.4 Word Embedding

Word embedding is an efficient approach to learning word representations from a large amount of unstructured text in order to capture the syntactic and semantic similarities between words. The conventional text representation model in text mining is the vector space model, which indicates each document using a vector of word occurrence indicators or term weights. However, the vector space model suffers from data sparsity problems when the vocabulary of a sentence is highly dimensional. Moreover, the model assumes independence between words and does not capture word similarity. Word embedding projects the words in a corpus to a low-dimensional space in which semantically connected words are in close proximity. Word embedding is often trained using a large number of documents in order to fully capture the similarity relationships between words. Mikolov et al. (2013) propose *word2vec* as an efficient word-embedding training

procedure. This approach uses a vector representation for each word via a neural network architecture. Studies show that using word embedding as the input to deep text learning achieves good performance in text mining (Wang et al. 2016; Zhang and Wallace 2015).

4.2.5 Deep Learning for Text Classification

Deep learning techniques perform better than traditional machine learning algorithms because they construct features in a hierarchical way: in other words, the higher-level features contain semantic connections extracted from the lower-level features such as words. The theoretical deep learning literature suggests that, in order to learn the various complex functions that can characterize high-level abstractions (e.g., image, language, and audio), researchers may need deep architectures (Bengio et al. 2007). Deep learning can use a broad collection of deep architectures (Bengio 2009), including graphical models with many levels of hidden variables (Hinton and Salakhutdinov 2006), neural networks with several hidden layers (Collobert and Weston 2008), and others (Zhu et al. 2009). The recent surge in deep learning and artificial intelligence research has demonstrated the superiority of deep learning over traditional machine learning techniques in computer vision (Ranzato et al., 2008; Lee et al., 2009; Mobahi et al., 2009), natural language processing (Collobert and Weston 2008; Weston and Besley 2008), and information retrieval (Salakhutdinov and Hinton 2007).

Two variants of deep learning techniques are widely used: a convolutional neural network (CNN) and a recurrent neural network (RNN). A CNN can efficiently capture the contextual information around words. Specifically, a CNN is used to denote big context sizes, such as unigram, bigram (a two-word sequence), and trigram (a three-word sequence), and to extract salient features within larger contexts through convolution and max-pooling operations. However,

a CNN does not consider the order of words in a sentence, which is important for understanding the semantics among the features. To solve this problem, an RNN views the input as a sequential structure and requires a series of linear operations (Wang et al. 2017). An RNN is well designed for sequence modeling. Long Short-Term Memory (LSTM), a variant of an RNN, provides an effective way of sequentially composing the semantic understanding in texts. The key units in LSTM are gates, which are implemented by a sigmoid function. Using sequence data, $\{w_1, \dots, w_n\}$, the gates can help control how much new information, w_t , from the current step, t , is added, how many long memories from the previous step are needed to establish new memories, and how much information is needed as features to generate the output at the current step. In this way, LSTM can decide the amount of information that can pass through gates automatically and dynamically based on different inputs at different steps. The training process contains sequences that activate the next hidden layer using a previous time step as the input to the current layer to influence predictions at the current time step (Sak et al. 2014). Studies have shown that applying an RNN or a CNN to generic sentence classification demonstrates outstanding performance in terms of classification performance metrics (e.g., F-measure, precision and recall) (Gan et al. 2017; Wang et al. 2017).

4.3 Research Objectives

In this research, a deep learning based user feedback classification framework is proposed for identifying information types from user reviews about mobile apps. Two types of information are emphasized: bug reports and new feature requests, which are helpful for customer value co-creation. Text classification will be conducted at the sentence level because it provides a good balance between analysis granularity and interpretability. It is expected that the deep learning

based approach will outperform existing text classification methods because deep learning can capture more semantic and contextual relationships between words. A comprehensive evaluation will be conducted to evaluate the proposed framework. The framework provides a useful and efficient way for mobile app developers to analyze user feedback from the large volume of online user reviews and maintain their competitive advantages.

4.4 Research Design

4.4.1 A Deep Learning-Based, Sentence-Level User Feedback Classification Framework

This section proposes an user feedback classification framework based on deep learning. Text documents, such as user reviews, must be preprocessed before the text analysis. Each review is segmented into individual sentences using a sentence tokenizer. A word tokenizer breaks each sentence into a sequenced collection of words. Punctuations and stop-words¹⁴ are removed. All letters are converted into lower-case letters. The proposed userfeedback classification framework consists of three main processing layers: a word-embedding layer, a CNN layer, and an RNN layer. Notations used in the framework are listed in Table 4-3 and a summarization of the computing process is mentioned in Algorithm 1. The three processing layers are described as follows.

¹⁴ <https://nlp.stanford.edu/software/>

1. **The word-embedding layer:** The word-embedding layer is used to learn word representation from a large corpus. The word-embedding model captures the semantic similarity and connection among words by projecting words to a low-dimensional space where semantically connected words are in close proximity. Figure 4-1 illustrates how word-embedding is used to predict contextual information. The *word2vec* *gensim* package (Rehurek and Sojka 2010) provides a domain-independent word-embedding model. It can be trained on the entire collection of Wikipedia documents (Pennington et al. 2014). Alternatively, a domain-dependent, word-embedding model can be trained using a domain-related corpus, such as mobile app reviews. An input sentence, S , is comprised of a sequence of words, $[w_i, \dots, w_{|S|}]$, that belong to a vocabulary, V , with size $|V|$. Before feeding into the next layer, each word, w_i ,

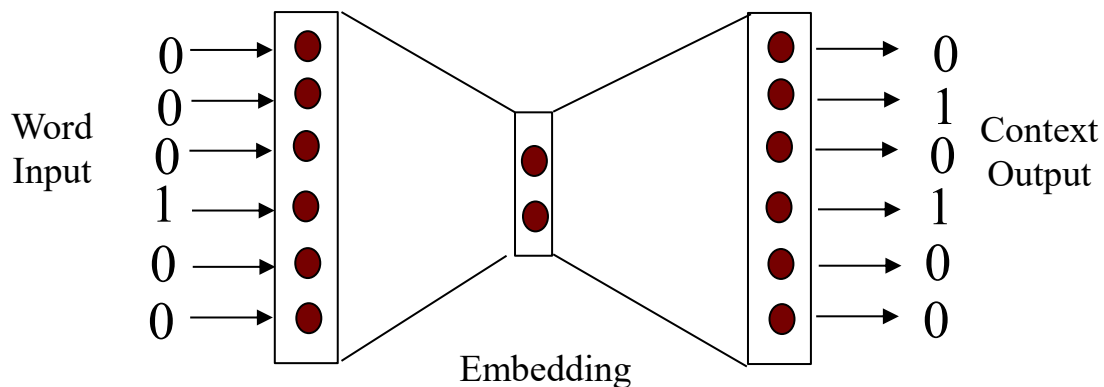


Figure 4-1 Word Embedding for Predicting Contextual Information

is transformed into a low-dimensional dense vector, v_i , by determining a word-embedding matrix, E . As a result, the input sentence, S , is represented as a matrix in which each column corresponds to a low-dimensional, word-embedding vector.

2. **The CNN layer:** The CNN layer is used to create a latent layer of features that capture the contextual information between a unigram, bigram, or trigram of words. It consists of a convolution layer and a pooling layer. Figure 4-2 shows illustrates how the CNN works.

- a. **The convolution layer:** The convolution layer is the first step to extract salient features from a sentence matrix represented by word embedding. The sentence matrix generated from the word-embedding layer is the input of the convolution layer. The convolution layer uses a sliding window to process each of the word-embedding vectors through a linear convolution operation. The objective of the layer is to filter the important features that contain the contextual information of each focal word within the sliding window. A non-linear activation function is employed to identify important and non-important features. In this paper, a rectified linear function is used as a non-linear activation function in order to determine more features (Xu et al. 2015).
- b. **The pooling layer:** Pooling is used to reduce the dimensionality of the feature maps generated by the convolution layer and extract important features. Pooling has several variants: max-pooling, min-pooling, and average-pooling. As Nagi et al. (2011) point out, max-pooling is often preferred over the other two variants because it usually achieves the best performance. Following max-pooling, the max value is extracted from each row of C_s , which generates the final representation vector, X_s , for the input sentence S .

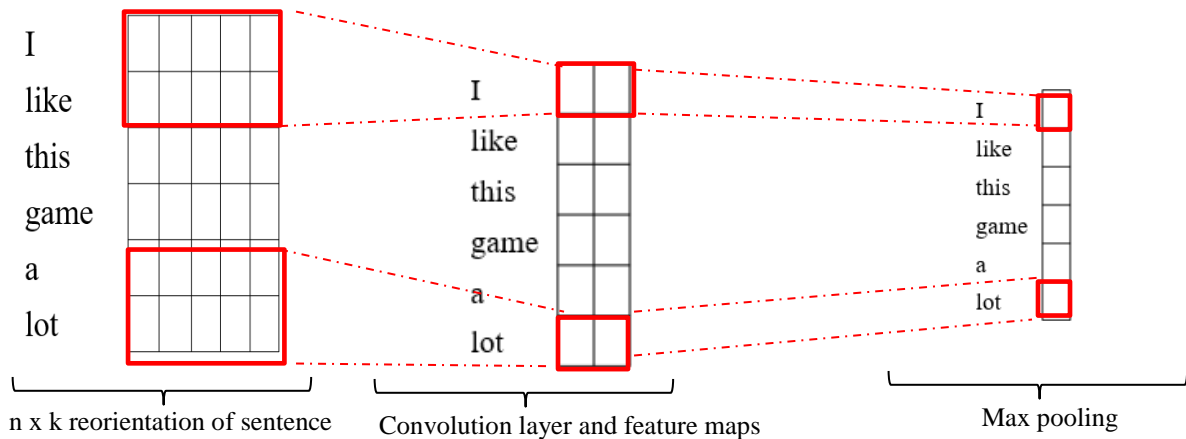


Figure 4-2 CNN for Capturing N-Grams Contextual Information

- The RNN layer:** Using the output of the CNN layer and X_s as the layer's input, LSTM is used to produce a sentence representation, R_s , from word representations. In LSTM, the core units are gates, which are implemented by the activation function. By means of a data sequence, $\{w_1, \dots, w_n\}$, the gates can decide how much new information, w_t , from the current step, t , is added; how many long memories from the previous step are required to launch new memories; and how much information is needed as features to produce the output at the current step. In this way, LSTM can decide the amount of information that can pass through gates automatically and dynamically based on different inputs at at different steps. The output of the layer is to generate a more condensed feature map after considering the word order. Figure 4-3 shows the process of the RNN's work flow.

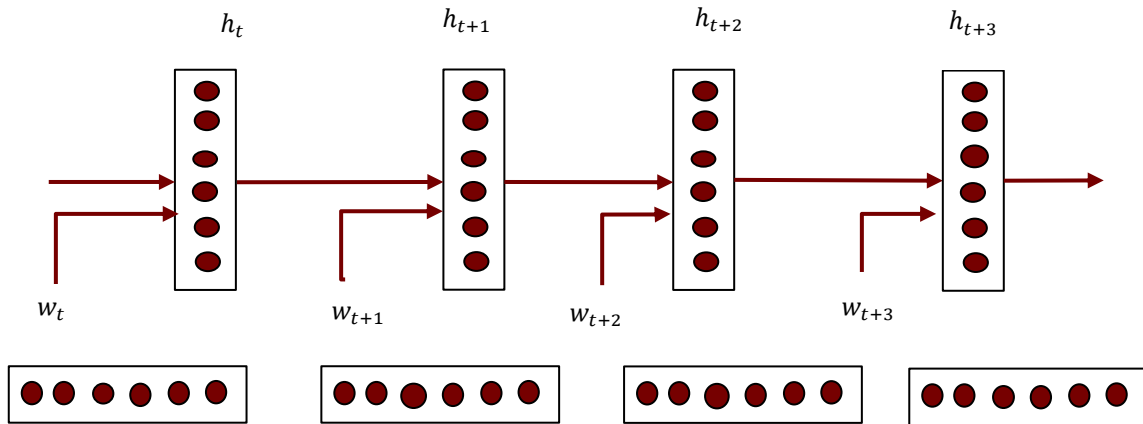


Figure 4-3 RNN for Capturing Word Sequence-based Contextual Information

According to the output of LSTM, an argmax function is then used to adaptively encode the class labels of sentences and generate the label vector, L_s , for the input sentence. These representations are considered features for user feedback classification. Figure 4-4 illustrates the major processes of the proposed framework.

Table 4-3 Notations in the Proposed Framework

Notation	Description
V	The vocabulary
$ V $	The vocabulary size
v_i	The word embedding word w_i
E	The word embedding matrix
w_i	The i th word in the input sentence
$ s $	The length of the input sentence
S	The input sentence
dim_e	The dimension of word embedding
C_s	The output of convolution layer for the input sentence S
X_s	The output of max pooling for the input sentence S
R_s	The final semantic representation of input sentence S
L_s	The final label of input sentence S

Algorithm 1: Deep learning based user feedback classification

Input: The input sentence contains a series of words: $[w_i, \dots, w_{|s|}]$, where w_i is chosen from a vocabulary V

1. Represent w_i using its word embedding v_i by looking up word embedding matrix E . Define $S = [v_i, \dots, v_{|s|}]$ as the input sentence embedding matrix R with dimension $dim_e \times |s|$.
2. Apply CNN to process S to get outputs of convolution C_s .
3. Apply max pooling to process C_s and get X_s .
4. Apply LSTM to process X_s and get R_s
5. Apply argmax function to R_s and get L_s where $L_s[i] = \text{argmax}(R[i, :])$

Output: Return L_s

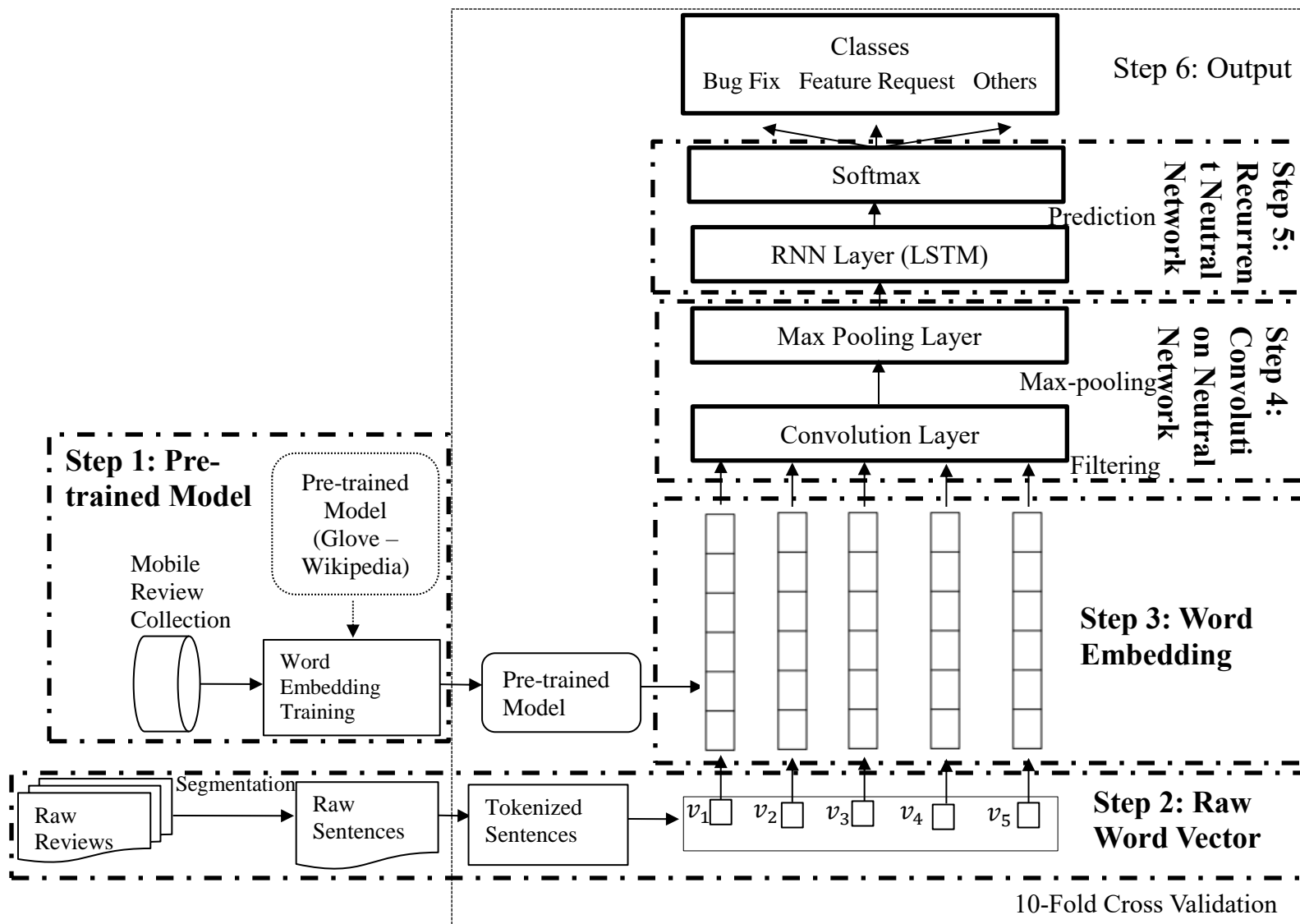


Figure 4-4 Deep Learning Based User Feedback Classification Framework

4.5 Experiments

In this section, the experiment used to evaluate the performance of the proposed user feedback classification framework to identify user needs in app user reviews is described. Baseline methods include several traditional text classification methods, such as MNB, SVM, k-nearest neighbors (KNN), DTs, logistic regression (LG), random forest (RF), and ridge regression (RR). For these baseline methods, the TF-IDF (term-frequency-inversed-document-frequency) vector space model is used as the text representation model, which is commonly used in text classification (Ramos 2003).

4.5.1 Data Description

The Apple App Store offers more than 2.8 million apps and is the largest app store in terms of its total generated revenue (Lai et al. 2018). 18,261,515 app reviews were collected for 4,602 game apps from their date of release to November 29, 2015. Each user review consists of review text content, review time, review rating, review title, app version, reviewer identity, and reviewer country. To overcome potential selection bias, five apps out of the 4,602 game apps and 3,000 user reviews made for the five apps were randomly selected. To obtain the ground truth of user feedback classification, 26 undergraduate and graduate students were recruited to label these reviews with 12,864 sentences. Each sentence could be labeled as three information types: bug fixes, feature requests, and others. The final dataset contained 6,915 sentences that were tagged by at least two taggers. The agreement rate was 79.9%, while the inter-rater reliability score was 74.5% which is fair good (Landis and Koch 1977). The feature request information type was 444 agreed sentence labels, which was the least across the three information types. Therefore, 444 sentences labeled as bug reports and 444 sentences labeled as other types were drawn in order to build a balanced evaluation data set.

4.5.2 Performance Metrics

The performance of the proposed framework was measured using the following four measures: precision, recall, and the F-measure. These measures are broadly used in information retrieval and text mining evaluations (Powers 2011). Precision is defined as the percentage of correctly identified instances in all the instances identified by the framework. Recall is the percentage of the instances that the framework has correctly identified over the total number of relevant instances, and the F-measure is the weighted average of precision and recall.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (1)$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2)$$

$$F - measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (3)$$

For the performance evaluation of multi-class classification, the macro-average method was used to calculate the average precision, recall, and F-measure.

$$Recall_{avg} = \sum_n(Recall) / n \quad (4)$$

$$Precision_{avg} = \sum_n(Precision) / n \quad (5)$$

$$F - measure_{avg} = \sum_n(F - measure) / n \quad (6)$$

4.5.3 Experiment

4.5.3.1 Training Word-Embedding Models

Studies have confirmed that initializing word vectors with pre-trained word embedding can improve the performance of text classification in the absence of a large, supervised training set (Collobert et al., 2011; Socher et al., 2011; Iyyer et al., 2014). Word embedding can be trained using a domain-independent or a domain-dependent corpus, which were both tested in this experiment. First, the publicly available *word2vec* vectors that had been obtained from training using 100 billion tokens from Google News were used. Alternatively, training using word embedding with a domain-specific corpus, 1.8 billion tokens from the app reviews collected from the Apple App Store, was employed. To achieve the optimal training performance, I tested several parameter values used in the proposed framework. For example, window size parameter, which indicates the maximum number of words between the current and predicted word in a sentence, was tested. Its values include 3, 4, and 5 words. The results indicated that the window size was 4 when the training performance (measured by accuracy) was the best. Similarly, the vector size of each word, which means the dimensionality of the feature vectors, was tested. Its values include 100, 200, 300, 400, and 500. The classification performance shows that vector size 300 made the training classification performance the best. The word frequency is a minimum threshold to determine the word as features. The experiment shows when the word frequency was 5, the classification performance was the best. The vectors were trained via the continuous bag-of-words model (Mikolov et al. 2013). Words not occurring in the set of pre-trained words were initialized randomly. The *word2vec* *gensim* library was utilized to train the model, tuning one parameter at a time with the other parameters held constant. Figures 4-2~4 show the training accuracy based on different parameters. In summary, the best accuracy was achieved by having 4 as the window size, 300 as the vector size, and 5 as the word frequency.

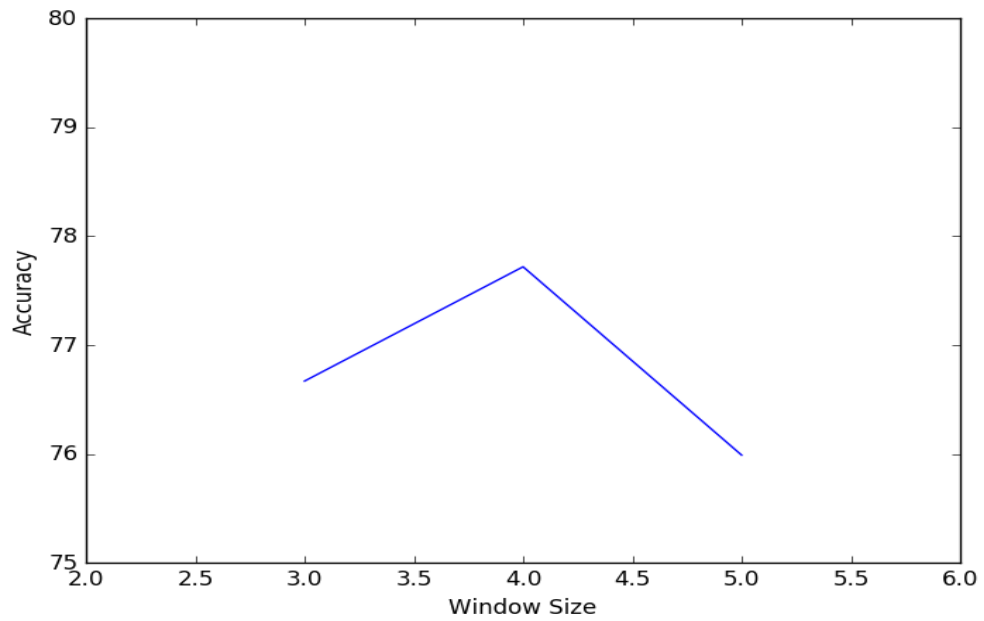


Figure 4-5 Training Accuracy with Different Window Sizes (word count=5 and dimension=300)

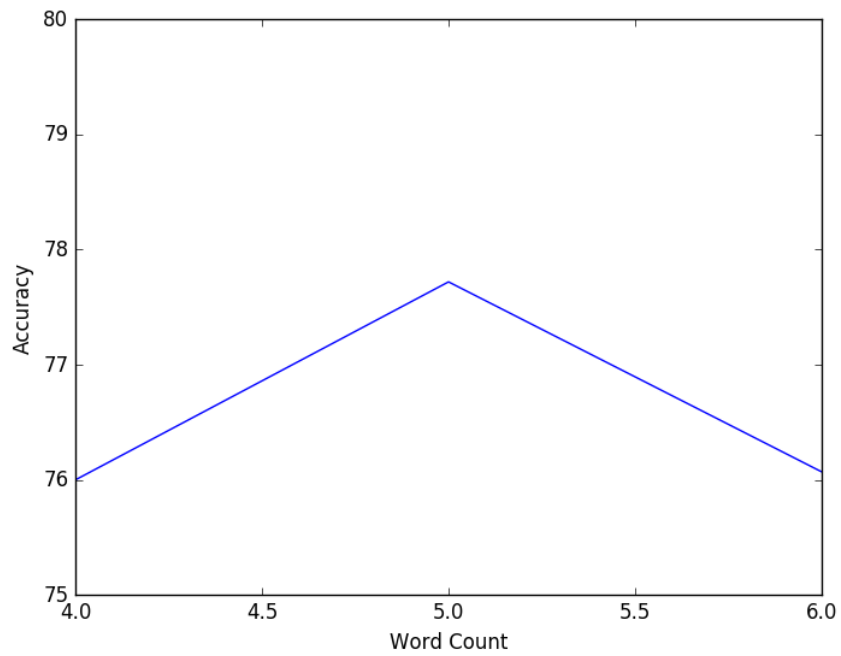


Figure 4-6 Training Accuracy with Different Word Counts (window size=4 and dimension=300)

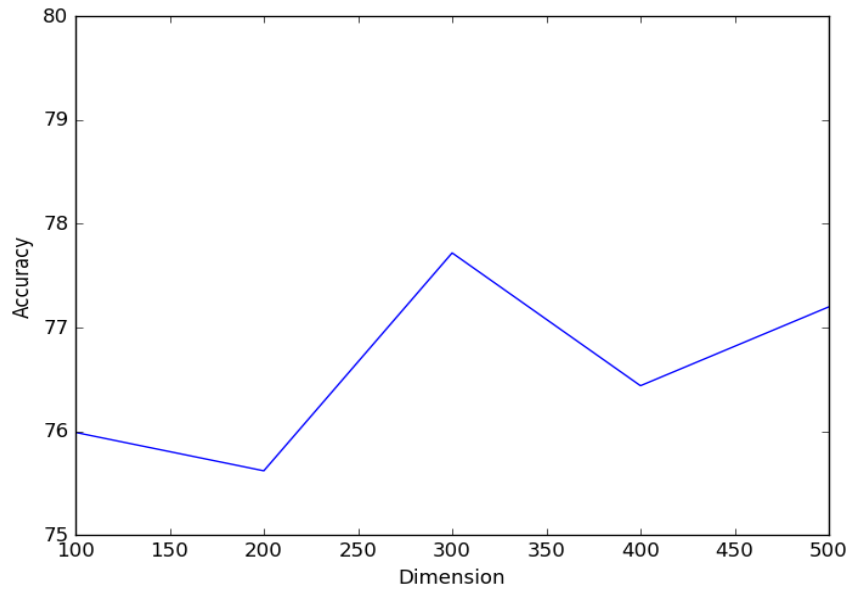


Figure 4-7 Training Accuracy with Different Dimensions (window size=4 and word count= 5)

4.5.3.2 Experimental Procedure

To evaluate the classification performance of the proposed framework, a 10-fold cross-validation procedure was used. The dataset, which contained 1,332 labeled review sentences, was randomly divided into 10 subsets. Each time, nine subsets were used for training, while the other subset was used for testing. The training process was used to tune the parameters in the convolution layer and the RNN layer. The testing process was used to predict the class label of each sentence in the testing subset. The classification performance was averaged over the ten folds.

4.5.3.3 CNN and RNN Parameter Training

Hyper-parameters were used to identify the appropriate configurations to achieve the best performance based on the different datasets. The procedure in Wang et al.'s (2017) study was followed to configure the hyper-parameters. In addition, because this study introduces a multi-class classification rather than a binary classification, the hyper-parameter – the loss function of

the RNN was set to “categorical entropy.” Through tuning the hyper-parameter of the optimizer, the “adam” optimizer was chosen.

4.5.4 Experimental Results

4.5.4.1 Word Embedding Evaluations

The study follows the empirical evaluations proposed by Mikolov et al. (2013). The results in Table 4-4 and Table 4-5 show some intuitive examples to indicate that the trained word embedding models based on mobile app reviews can capture both semantic and syntactic relationships of words. Table 4-4 presents the semantic relationships and Table 4-5 shows the syntactic relationships captured by word-embedding.

Table 4-4 Word Embedding for Capturing Semantic Relationships between Words

Information Types	Type of Relationship	Keyword	Top Similar Contextual Words
Bug Reports	Synonyms	bug	1.glitch, 2. bug/glitch, 3. glich, 4. problem, 5. issue
		fix	1.resolve, 2. address, 3. rectify, 4.remedy, 5. correct
	Misspelling	glitch	1.glich, 2. glitch: 3.glitchi, 4.glitche, 5.glicth
		fix	1.fixe, 2.fixs, 3.fixx, 4.fux, 5.pleasefix
Feature Request	Synonyms	change	1.adjust, 2.switch, 3.alter, 4. modify, 5.add
		add	1.include, 2.make, 3.create, 4.implement, 5.incorporate
	Misspelling	request	1.request:, 2.request:, 3.req, 4.gifts/requests, 5. request-
		new	1.mew, 2.gamenew, 3.funnew, 4.*new, 5.more/new

Table 4-5 Word Embedding for Capturing Syntactic Relationships between Words

Type of Relationship	Word Pair 1		Word Pair 2	
	Plural Verbs	add	adds	fix
Plural Nouns	bug	bugs	feature	features
Comparative	new	newer	old	younger
Superlative	new	newest	old	oldest
Past Tense	fix	fixed	add	added
opposite	good	bad	old	young

4.5.4.2 Classification Results

Table 4-4 summarizes the performance of the benchmark methods and the deep learning-based user feedback classification method. The proposed deep learning-based method outperformed all the traditional text classification methods for identifying bug reports and new feature requests from user reviews. Among these classes, bug reports achieved the best performance, and the F-measure reached up to 0.83. By contrast, the F-measure of the feature request classification performance was only 0.74. It is possible that the features of the bug fixes were more focused than those of the feature requests. Regarding word-embedding training, the word-embedding models trained by the Wikipedia corpus and app review corpus both achieved better performance than traditional text classification methods. However, the word-embedding model trained using apps reviews achieved better performance than that trained using the domain-independent Wikipedia corpus. This showed that the domain-specific corpus helped build a better word-embedding model than the domain-independent corpus.

Table 4-6 Performance of the Proposed Framework vs. Baseline Methods

Method Name	Information Type	Performance Metrics		
		F-Measure	Precision	Recall
MNB	Bug Report	0.764	0.709	0.837
	Feature Request	0.626	0.725	0.567
	Others	0.695	0.684	0.717
	Combined	0.696	0.711	0.702
SVM	Bug Report	0.762	0.726	0.802
	Feature Request	0.638	0.679	0.609
	Others	0.718	0.716	0.722
	Combined	0.706	0.711	0.708
KNN	Bug Report	0.645	0.632	0.664
	Feature Request	0.572	0.519	0.644
	Others	0.589	0.711	0.507
	Combined	0.604	0.626	0.603
DT	Bug Report	0.561	0.605	0.525
	Feature Request	0.479	0.463	0.508
	Others	0.544	0.542	0.56
	Combined	0.530	0.541	0.531
LG	Bug Report	0.694	0.598	0.846
	Feature Request	0.517	0.763	0.394
	Others	0.690	0.673	0.717
	Combined	0.633	0.682	0.647
RF	Bug Report	0.694	0.598	0.846
	Feature Request	0.517	0.763	0.394
	Others	0.690	0.673	0.717
	Combined	0.633	0.682	0.647
RR	Bug Report	0.722	0.629	0.86
	Feature Request	0.509	0.761	0.385
	Others	0.713	0.672	0.762
	Combined	0.647	0.691	0.664
DL (Wikipedia)	Bug Report	0.805	0.824	0.789
	Feature Request	0.710	0.731	0.697
	Others	0.688	0.664	0.718
	Combined	0.738	0.742	0.737
DL (Mobile app reviews)	Bug Report	0.828	0.815	0.846
	Feature Request	0.743	0.773	0.719
	Others	0.725	0.719	0.736
	Combined	0.766	0.769	0.768

4.6 Conclusion and Discussion

In this paper, a new user feedback classification framework was proposed that automatically identified bug reports and feature requests from massive volumes of online app reviews. It was demonstrated that the proposed deep learning-based framework outperformed existing text classification baseline methods.

This study makes several methodological and theoretical contributions. First, the study proposes a novel deep learning framework that incorporates a CNN and an RNN to identify user needs and relevant details from unstructured textual data. The proposed model provides an effective framework to incorporate contextual information (syntactic features and domain background) to uncover sentences related to user needs. Second, the study confirms previous findings that word embedding can achieve better performance in terms of classification accuracy through different configurations of hyper-configurations. Experimental results on word embedding show that the proposed model outperforms the competing traditional classification methods and discovers more meaningful and accurate user needs. Third, this study explores the domain adaption problem, and the results show that the pre-trained word embedding model that is trained on the online app dataset outperforms all the benchmarks.

This research contributes to the rich body of research on customer value co-creation by providing an automated tool to classify user feedback in user reviews. Researchers can use the proposed methodology to identify and understand the user feedback embedded in the large volume of online, user-generated content. Firms facing hyper-competition can use the proposed method to automatically identify product issues from customer feedback and improve their product design by addressing those issues. Hence, managers are urged to see the benefits of quickly understanding customer feedback and transforming collaborative inputs from users into new business

opportunities. The proposed method provides a feasible way for users to be involved in value co-creation (Ramaswamy and Prahalad 2004), which can create business value.

This research has significant managerial implications for app developers, users, and platform providers. For example, app developers can use information regarding user needs to improve product quality. By receiving attention from app developers, app customers will be more inclined to contribute valuable feedback regarding the developers' products. App platform providers can design new features to categorize user information and incorporate more innovative information based on app developers' needs and customers' feedback.

Despite its findings and implications, this study has several limitations. First, this research uses only one public data source and only one method of analysis (text analysis). Thus, an empirical study that incorporates other sources of data from manufacturers may yield more valuable and practical insights regarding quality improvement and product innovation opportunities. Second, this study only examines objective information classification problems in user reviews. Future studies should incorporate other perspectives, such as those regarding product advantages, which can also help managers understand customers' preferences and demands and thus allow managers to better position their products in the right customer segments. Third, although this study evaluates the results based on the classification performance of information types, it is still necessary to show qualitative measurements (e.g., key word lists) in the future for practical purposes.

4.7 References

- Abrahams, A., Jiao, J., and Fan, W. 2013. “What’s Buzzing in the Blizzard of Buzz? Automotive Component Isolation in Social Media Postings,” *Decision Support Systems* (55:4), pp. 871–882.
- Abrahams, A. S., Fan, W., Wang, G. A., Zhang, Z., and Jiao, J. 2015. “An Integrated Text Analytic Framework for Product Defect Discovery,” *Production and Operations Management* (24:6), pp. 975–990.
- Alves, H., Fernandes, C., and Raposo, M. 2016. “Value Co-Creation: Concept and Contexts of Application and Study,” *Journal of Business Research* (69:5), Elsevier, pp. 1626–1633.
- Amit, R., and Zott, C. 2001. “Value Creation in E-business,” *Strategic Management Journal* (22:6-7), Wiley Online Library, pp. 493–520.
- Animesh, A., Viswanathan, S., and Agarwal, R. 2011. “Competing ‘creatively’ in Sponsored Search Markets: The Effect of Rank, Differentiation Strategy, and Competition on Performance,” *Information Systems Research* (22:1), pp. 153–169.
- Apté, C., Damerau, F., and Weiss, S. M. 1994. “Automated Learning of Decision Rules for Text Categorization,” *ACM Transactions on Information Systems* (12:3), pp. 233–251.
- Aral, S., Dellarocas, C., and Godes, D. 2013. “Introduction to the Special Issue—social Media and Business Transformation: A Framework for Research,” *Information Systems Research* (24:1), INFORMS, pp. 3–13.
- Arora, A., Krishnan, R., Telang, R., and Yang, Y. 2010. “An Empirical Analysis of Software Vendors’ Patch Release Behavior: Impact of Vulnerability Disclosure,” *Information Systems Research* (21:1), pp. 115–132.
- Bakos, Y., and Brynjolfsson, E. 1999. “Bundling Information Goods: Pricing, Profits, and Efficiency,” *Management Science* (45:12), pp. 1613–1630.
- Balasubramanian, S., Bhattacharya, S., and Krishnan, V. V. 2015. “Pricing Information Goods: A Strategic Analysis of the Selling and Pay-per-Use Mechanisms,” *Marketing Science* (34:2), INFORMS, pp. 218–234.
- Baldwin, C. Y., and Clark, K. B. 2006. “The Architecture of Participation: Does Code Architecture Mitigate Free Riding in the Open Source Development Model?,” *Management Science* (52:7), INFORMS, pp. 1116–1127.
- Basole, R. C., and Karla, J. 2011. “On the Evolution of Mobile Platform Ecosystem Structure and Strategy,” *Business & Information Systems Engineering* (3:5), Springer, p. 313.
- Basole, R. C., and Karla, J. 2012. “Value Transformation in the Mobile Service Ecosystem: A Study of App Store Emergence and Growth,” *Service Science* (4:1), INFORMS, pp. 24–41.
- Beal, R. M. 2000. “Competing Effectively: Environmental Scanning, Competitive Strategy and Organisational Performance in Small Manufacturing Firms,” *Journal of Small Business Management* (38:January), pp. 27–47.
- Becerra, M., Santaló, J., and Silva, R. 2013. “Being Better vs. Being Different: Differentiation, Competition, and Pricing Strategies in the Spanish Hotel Industry,” *Tourism Management* (34), Elsevier, pp. 71–79.
- Bengio, Y. 2009. “Learning Deep Architectures for AI,” *Foundations and Trends® in Machine Learning* (2:1), Now Publishers, Inc., pp. 1–127.
- Bengio, Y., Lamblin, P., Popovici, D., and Larochelle, H. 2007. “Greedy Layer-Wise Training of Deep Networks,” in *Advances in Neural Information Processing Systems*, pp. 153–160.
- Bharadwaj, A. S. 2000. “A Resource-Based Perspective on Information Technology Capability

- and Firm Performance: An Empirical Investigation,” *MIS Quarterly*, JSTOR, pp. 169–196.
- Bhargava, H. K., and Choudhary, V. 2001. “Information Goods and Vertical Differentiation,” *Journal of Management Information Systems* (18:2), Taylor & Francis, pp. 89–106.
- Bhargava, H. K., and Choudhary, V. 2008. “Research Note — When Is Versioning Optimal for Information Goods?,” *Management Science* (54:5), pp. 1029–1035.
- Boudreau, K. J. 2012. “Let a Thousand Flowers Bloom? An Early Look at Large Numbers of Software App Developers and Patterns of Innovation,” *Organization Science* (23:5), pp. 1409–1427.
- Brun, C., and Hagege, C. 2013. “Suggestion Mining: Detecting Suggestions for Improvement in Users’ Comments,” *Research in Computing Science* (70), pp. 171–181.
- Büschken, J., and Allenby, G. M. 2016. “Sentence-Based Text Analysis for Customer Reviews,” *Marketing Science* (35:6), INFORMS, pp. 953–975.
- Caves, R. E., and Williamson, P. J. 1985. “What Is Product Differentiation, Really?,” *The Journal of Industrial Economics*, JSTOR, pp. 113–132.
- Chakravarti, A., and Xie, J. 2006a. “The Impact of Standards Competition on Consumers: Effectiveness of Product Information and Advertising Formats,” (XLIII:May), pp. 224–236.
- Chakravarti, A., and Xie, J. 2006b. “The Impact of Standards Competition on Consumers: Effectiveness of Product Information and Advertising Formats,” (43:2), pp. 224–236.
- Chamberlin, E. 2010. “H.(1933) The Theory of Monopolistic Competition,” *Cambridge, Mass, Harvard University*.
- Chen, H., Chiang, R. H. L., and Storey, V. C. 2012. “Business Intelligence and Analytics: From Big Data to Big Impact.,” *MIS Quarterly* (36:4), pp. 1165–1188.
- Chen, M.-J., and MacMillan, I. C. 1992. “Nonresponse and Delayed Response to Competitive Moves: The Roles of Competitor Dependence and Action Irreversibility,” *Academy of Management Journal* (35:3), Academy of Management, pp. 539–570.
- Chen, M.-J., Smith, K. G., and Grimm, C. M. 1992. “Action Characteristics as Predictors of Competitive Responses,” *Management Science* (38:3), INFORMS, pp. 439–455.
- Chen, N., Lin, J., Hoi, S., Xiao, X., and Zhang, B. 2014. “AR-Miner: Mining Informative Reviews for Developers from Mobile App Marketplace,” in *Proceedings of the 36th International Conference on Software Engineering*, ACM, pp. 767–778.
- Chen, Y. 2013. “Context-Dependent Preferences and Innovation Strategy Context-Dependent Preferences and Innovation Strategy,” (February 2014).
- Chen, Y., and Iyer, G. 2002. “Research Note Consumer Addressability and Customized Pricing,” *Marketing Science* (21:2), INFORMS, pp. 197–208.
- 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. 477–491.
- Collobert, R., and Weston, J. 2008. “A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning,” in *Proceedings of the 25th International Conference on Machine Learning*, ACM, pp. 160–167.
- Comino, S., Udine, U., Manenti, F. M., and Economiche, S. 2016. “Updates Management in Mobile Applications . iTunes vs Google Play Updates Management in Mobile Applications . iTunes vs Google Play . *,”
- Cooper, R. G., and Kleinschmidt, E. J. 1994. “Determinants of Timeliness in Product Development,” *Journal of Product Innovation Management* (11:5), Elsevier, pp. 381–396.
- Crawford-Welch, S. 1990. “Market Segmentation in the Hospitality Industry,” *Hospitality Research Journal* (14:2), ICHRIE Sage CA: Los Angeles, CA, pp. 295–308.

- Cusumano, M. A., Kahl, S. j, and Suarez, F. F. 2008. "Services, Industry Evolution, and the Copetitive Strategies of Product Firms," *Academy of Management Journal* (51:2), pp. 315–334.
- Datta, D., and Kajan, S. 2013. "Do App Launch Times Impact Their Subsequent Commercial Success? An Analytical Approach," in *Cloud Computing and Big Data (CloudCom-Asia), 2013 International Conference on*, IEEE, pp. 205–210.
- Dey, D., and Lahiri, A. 2016. "Versioning of Video Games: Go Vertical in a Horizontal Market?," *Proceedings of the Annual Hawaii International Conference on System Sciences* (2016–March), pp. 5259–5268.
- Dickson, P. R., and Ginter, J. L. 1987. "Market Segmentation , Product Differentiation , and Marketing Strategy," (51:April), pp. 1–10.
- Distimo. 2009. "Pple App Store, BlackBerry App World, Google Android MarketNokia Ovi Store & Windows Marketplace for Mobile.,"
- Dougherty, D. 2001a. "Reimagining the Differentiation and Integration of Work for Sustained Product Innovation," *Organization Science* (12:5), pp. 612–631.
- Dougherty, D. 2001b. "Reimagining the Differentiation and Integration of Work for Sustained Product Innovation," *Organization Science* (12:5), INFORMS, pp. 612–631.
- Dougherty, D., and Dunne, D. D. 2011. "Organizing Ecologies of Complex Innovation," *Organization Science* (22:5), pp. 1214–1223.
- Duan, W., Gu, B., and Whinston, A. B. 2008. "The Dynamics of Online Word-of-Mouth and Product sales—An Empirical Investigation of the Movie Industry," *Journal of Retailing* (84:2), Elsevier, pp. 233–242.
- Ethiraj, S. K., and Zhu, D. H. 2008. "Performance Effects of Imitative Entry," *Strategic Management Journal* (29:8), Wiley Online Library, pp. 797–817.
- Ferrier, W. J., Smith, K. G., and Grimm, C. M. 1999. "The Role of Competitive Action in Market Share Erosion and Industry Dethronement: A Study of Industry Leaders and Challengers," *Academy of Management Journal* (42:4), Academy of Management, pp. 372–388.
- Fisher, M., Gallino, S., and Li, J. 2017. "Competition-Based Dynamic Pricing in Online Retailing: A Methodology Validated with Field Experiments," *Management Science*, INFORMS.
- Forman, C., Ghose, A., and Wiesenfeld, B. 2008. "Examining the Relationship between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets," *Information Systems Research* (19:3), INFORMS, pp. 291–313.
- French, S. A., Story, M., Jeffery, R. W., Snyder, P. A. T., Eisenberg, M., Sidebottom, A., and Murray, D. 1997. "Pricing Strategy to Promote Fruit and Vegetable Purchase in High School Cafeterias," *Journal of the Academy of Nutrition and Dietetics* (97:9), Elsevier, pp. 1008–1010.
- Galvis Carreño, L. V, and Winbladh, K. 2013. "Analysis of User Comments: An Approach for Software Requirements Evolution," in *Proceedings of the 2013 International Conference on Software Engineering*, pp. 582–591.
- Gan, Z., Pu, Y., Henao, R., Li, C., He, X., and Carin, L. 2017. "Learning Generic Sentence Representations Using Convolutional Neural Networks," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 2390–2400.
- Garg, R., and Telang, R. 2014. "Inferring App Demand from Publicly Available Data," *MIS Quarterly* (37:4), pp. 1253–1264.
- Ghose, A., and Han, S. P. 2014. "Estimating Demand for Mobile Applications in the New Economy," *Management Science* (60:6), pp. 1470–1488.

- Ghose, A., and Ipeirotis, P. G. 2011. "Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics," *IEEE Transactions on Knowledge and Data Engineering* (23:10), pp. 1498–1512.
- Gnyawali, D. R., Fan, W., and Penner, J. 2010. "Competitive Actions and Dynamics in the Digital Age: An Empirical Investigation of Social Networking Firms," *Information Systems Research* (21:3), INFORMS, pp. 594–613.
- Gnyawali, D. R., He, J., and Madhavan, R. 2006. "Impact of Co-Opetition on Firm Competitive Behavior: An Empirical Examination," *Journal of Management* (32:4), Sage Publications Sage CA: Thousand Oaks, CA, pp. 507–530.
- Goldberg, D. M., and Abrahams, A. S. 2018. "A Tabu Search Heuristic for Smoke Term Curation in Safety Defect Discovery," *Decision Support Systems* (105), Elsevier, pp. 52–65.
- Greenstein, S., and Markovich, S. 2012. "Pricing Experience Goods in Information Good Markets: The Case of eBusiness Service Providers," *International Journal of the Economics of Business* (19:1), pp. 119–139.
- Gu, B., and Ye, Q. 2014. "First Step in Social Media: Measuring the Influence of Online Management Responses on Customer Satisfaction," *Production and Operations Management* (23:4), pp. 570–582.
- Guo, X., Ling, L., Yang, C., Li, Z., and Liang, L. 2013. "Optimal Pricing Strategy Based on Market Segmentation for Service Products Using Online Reservation Systems: An Application to Hotel Rooms," *International Journal of Hospitality Management* (35), Elsevier Ltd, pp. 274–281.
- Hausman, J. A. 1978. "Specification Tests in Econometrics," *Econometrica: Journal of the Econometric Society*, JSTOR, pp. 1251–1271.
- Hevner, V. A., March, S. T., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), Springer, pp. 75–105.
- Hinterhuber, A. 2008. "Customer Value-Based Pricing Strategies: Why Companies Resist," *Journal of Business Strategy* (29:4), Emerald Group Publishing Limited, pp. 41–50.
- Hinton, G. E., and Salakhutdinov, R. R. 2006. "Reducing the Dimensionality of Data with Neural Networks," *Science* (313:5786), American Association for the Advancement of Science, pp. 504–507.
- Hippel, E. Von, and Krogh, G. Von. 2003. "Open Source Software and the ?Private-Collective? Innovation Model: Issues for Organization Science," *Organization Science* (14:2), pp. 209–223.
- Jha, N., and Mahmoud, A. 2017. "Mining User Requirements from Application Store Reviews Using Frame Semantics," in *International Working Conference on Requirements Engineering: Foundation for Software Quality*, Springer, pp. 273–287.
- Johansson, M., Hallberg, N., Hinterhuber, A., Zbaracki, M., and Liozu, S. 2012. "Pricing Strategies and Pricing Capabilities," *Journal of Revenue and Pricing Management* (11:1), Springer, pp. 4–11.
- Katzmarzik, A. 2011. "Product Differentiation for Software-as-a-Service Providers," *Business and Information Systems Engineering* (3:1), pp. 19–31.
- Kim, S. J., Kim, K. H., and Choi, J. 2017. "Design Innovation and Purchase Behavior for Augmented Products," *Journal of Business Research* (September), Elsevier, pp. 0–1.
- Kim, Y. 2014. "Convolutional Neural Networks for Sentence Classification," *arXiv Preprint arXiv:1408.5882*.
- Kleinbaum, D. G. 1998. "Survival Analysis, a Self-Learning Text," *Biometrical Journal* (40:1),

- Wiley Online Library, pp. 107–108.
- Kleinschmidt, E. J., and Cooper, D. R. 1991. “The Impact of Product Innovativeness on Performance,” *Journal of Product Innovation Management* (8), pp. 240–251.
- Lai, H., Hsu, J. S.-C., and Wu, M.-X. 2018. “The Impact S of Requested Permission on Mobile App Adoption: The Insights Based on an Experiment in Taiwan,” in *Proceedings of the 51st Hawaii International Conference on System Sciences*.
- Landis, J. R., and Koch, G. G. 1977. “The Measurement of Observer Agreement for Categorical Data,” *Biometrics*, JSTOR, pp. 159–174.
- Law, D., Gruss, R., and Abrahams, A. S. 2017. “Automated Defect Discovery for Dishwasher Appliances from Online Consumer Reviews,” *Expert Systems with Applications* (67), Elsevier Ltd, pp. 84–94.
- Lee, G., and Raghu, T. S. 2014. “Determinants of Mobile Apps’ Success: Evidence from the App Store Market,” *Journal of Management Information Systems* (31:2), Taylor & Francis, pp. 133–170.
- Lee, J., and Miller, D. 1999. “People Matter: Commitment to Employees, Strategy and Performance in Korean Firms,” *Strategic Management Journal*, JSTOR, pp. 579–593.
- Li, J., Xia, J., and Zajac, E. J. 2018. “On the Duality of Political and Economic Stakeholder Influence on Firm Innovation Performance: Theory and Evidence from Chinese Firms,” *Strategic Management Journal* (39:1), Wiley Online Library, pp. 193–216.
- Liu, B. 2012. “Sentiment Analysis and Opinion Mining,” *Synthesis Lectures on Human Language Technologies* (5:1), Morgan & Claypool Publishers, pp. 1–167.
- Liu, B., Hu, M., and Cheng, J. 2005. “Opinion Observer: Analyzing and Comparing Opinions on the Web,” in *Proceedings of the 14th International Conference on World Wide Web*, ACM, pp. 342–351.
- Liu, C. Z., Au, Y. A., and Choi, H. S. 2014. “Effects of Freemium Strategy in the Mobile App Market: An Empirical Study of Google Play,” *Journal of Management Information Systems* (31:3), Taylor & Francis, pp. 326–354.
- Liu, F., and Brandyberry, A. 2014. “Exploring the Effects of Aggregate Review Characteristics on Mobile Application Adoption,” in *Thirty Fifth International Conference on Information Systems*, Auckland, New Zealand.
- Luo, X. 2009. “Quantifying the Long-Term Impact of Negative Word of Mouth on Cash Flows and Stock Prices,” *Marketing Science* (28:1), pp. 148–165.
- Luo, X., Zhang, J., and Duan, W. 2013. “Social Media and Firm Equity Value,” *Information Systems Research* (24:1), INFORMS, pp. 146–163.
- LYU, C., and SHANG, Y. 2017. “The Time-Space Effect Research of High-Tech Industrial Agglomeration Model and Innovation Output,” *Journal of Management* (2), p. 7.
- Maalej, W., and Nabil, H. 2015. “Bug Report, Feature Request, or Simply Praise? On Automatically Classifying App Reviews,” in *Requirements Engineering Conference (RE), 2015 IEEE 23rd International*, IEEE, pp. 116–125.
- Maalej, W., Nabil, H., and Stanik, C. 2017. “On the Automatic Classification of App Reviews,” *Software Engineering 2017*, Gesellschaft Für Informatik eV.
- MacMillan, D., Burrows, P., and Ante, S. E. 2009. “Inside the App Economy,” *Business Week* (22).
- Makadok, R. 2010. “The Interaction Effect of Rivalry Restraint and Competitive Advantage on Profit: Why the Whole Is Less than the Sum of the Parts,” *Management Science* (56:2), INFORMS, pp. 356–372.

- McIlroy, S., Ali, N., and Hassan, A. E. 2016. “Fresh Apps: An Empirical Study of Frequently-Updated Mobile Apps in the Google Play Store,” *Empirical Software Engineering* (21:3), Springer, pp. 1346–1370.
- Mikolov, T., Chen, K., Corrado, G., and Dean, J. 2013. “Efficient Estimation of Word Representations in Vector Space,” *arXiv Preprint arXiv:1301.3781*.
- Miller, D. 1986. “Configurations of Strategy and Structure: Towards a Synthesis,” *Strategic Management Journal* (7:3), Wiley Online Library, pp. 233–249.
- Miller, D. 1988. “Relating Porter’s Business Strategies to Environment and Structure: Analysis and Performance Implications,” *Academy of Management Journal* (31:2), Academy of Management, pp. 280–308.
- Mintzberg, H. 1988. “Generic Strategies: Toward a Comprehensive Framework,” *Advances in Strategic Management* (5), pp. 1–67.
- Moghaddam, S. 2015. “Beyond Sentiment Analysis: Mining Defects and Improvements from Customer Feedback,” in *Advances in Information Retrieval*, Springer, pp. 400–410.
- Mudambi, S. M., and Schuff, D. 2010. “What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon. Com,” *MIS Quarterly* (34:1), pp. 185–200 (available at <http://ssrn.com/abstract=2175066>).
- Mummalaneni, V., Gruss, R., Goldberg, D. M., Ehsani, J. P., and Abrahams, A. S. 2018. “Social Media Analytics for Quality Surveillance and Safety Hazard Detection in Baby Cribs,” *Safety Science* (104), Elsevier, pp. 260–268.
- Nagi, J., Ducatelle, F., Di Caro, G. A., Cireşan, D., Meier, U., Giusti, A., Nagi, F., Schmidhuber, J., and Gambardella, L. M. 2011. “Max-Pooling Convolutional Neural Networks for Vision-Based Hand Gesture Recognition,” in *Signal and Image Processing Applications (ICSIPA), 2011 IEEE International Conference on*, IEEE, pp. 342–347.
- Nakagawa, T., Inui, K., and Kurohashi, S. 2010. “Dependency Tree-Based Sentiment Classification Using CRFs with Hidden Variables,” in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Association for Computational Linguistics, pp. 786–794.
- Narasimhan, R., Ghosh, S., and Mendez, D. 1993. “A Dynamic Model of Product Quality and Pricing Decisions on Sales Response,” *Decision Sciences* (24:5), Wiley Online Library, pp. 893–908.
- Nasri, L., Baghersad, M., Gruss, R., Marucchi, N. S. W., Abrahams, A. S., and Ehsani, J. P. 2018. “An Investigation into Online Videos as a Source of Safety Hazard Reports,” *Journal of Safety Research* (65), Elsevier, pp. 89–99.
- Nayebi, M., Adams, B., and Ruhe, G. 2016. “Release Practices for Mobile Apps -- What Do Users and Developers Think?,” *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, pp. 552–562.
- Oh, J., Kim, D., Lee, U., Lee, J.-G., and Song, J. 2013. “Facilitating Developer-User Interactions with Mobile App Review Digests,” in *CHI’13 Extended Abstracts on Human Factors in Computing Systems*, ACM, pp. 1809–1814.
- Pagano, D., and Maalej, W. 2013. “User Feedback in the Appstore: An Empirical Study,” in *Requirements Engineering Conference (RE), 2013 21st IEEE International*, IEEE, pp. 125–134.
- Pang, B., Lee, L., and Vaithyanathan, S. 2002. “Thumbs up?: Sentiment Classification Using Machine Learning Techniques,” in *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing-Volume 10*, Association for Computational

- Linguistics, pp. 79–86.
- Panichella, S., Di Sorbo, A., Guzman, E., Visaggio, C. A., Canfora, G., and Gall, H. C. 2015. “How Can I Improve My App? Classifying User Reviews for Software Maintenance and Evolution,” in *Software Maintenance and Evolution (ICSME), 2015 IEEE International Conference on*, IEEE, pp. 281–290.
- Pennington, J., Socher, R., and Manning, C. 2014. “Glove: Global Vectors for Word Representation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543.
- Porter, J. 2010. *Designing for the Social Web (Google eBook)*, Peachpit Press (available at http://books.google.com/books/about/Designing_for_the_Social_Web.html?id=CNJuvYBI_IsC&pgis=1).
- Porter, M. E. 1980. “Competitive Strategy: Techniques for Analyzing Industries and Competition,” *New York*, p. 300.
- Powers, D. M. 2011. “Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation,” Bioinfo Publications.
- Prahalad, C. K., and Ramaswamy, V. 2000. “Co-Opting Customer Competence,” *Harvard Business Review* (78:1), pp. 79–90.
- Prahalad, C. K., and Ramaswamy, V. 2004. “Co-creation Experiences: The next Practice in Value Creation,” *Journal of Interactive Marketing* (18:3), Wiley Online Library, pp. 5–14.
- Prajogo, D. I., and Ahmed, P. K. 2006. “Relationships between Innovation Stimulus, Innovation Capacity, and Innovation Performance,” *R & D Management* (36:5), pp. 499–515.
- Qiao, Z., Wang, G. A., Zhou, M., and Fan, W. 2018. “The Impact of Customer Reviews on Product Innovation: Empirical Evidence in Mobile Apps,” in *Analytics and Data Science*, Springer, pp. 95–110.
- Ragaglia, D., and Roma, P. 2014. “Understanding the Drivers of the Daily App Rank : The Role of Revenue Models,” in *Proceedings of the 26th Annual Conference of the Productions and Operations Management Society*, pp. 1–10.
- Ramanand, J., Bhavsar, K., and Pedanekar, N. 2010. “Wishful Thinking: Finding Suggestions and ‘Buy’ Wishes from Product Reviews,” in *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, Association for Computational Linguistics, pp. 54–61.
- Ramaswamy, V., and Prahalad, C. K. 2004. “Co-Creation Experiences: The Next Ractice in Value Creation,” *Journal of Interactive Marketing* (18:3), pp. 5–14.
- Ramos, J. 2003. “Using Tf-Idf to Determine Word Relevance in Document Queries,” in *Proceedings of the First Instructional Conference on Machine Learning*.
- Rehurek, R., and Sojka, P. 2010. “Software Framework for Topic Modelling with Large Corpora,” in *In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, Citeseer.
- Sak, H., Senior, A., and Beaufays, F. 2014. “Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling,” in *Fifteenth Annual Conference of the International Speech Communication Association*.
- Salakhutdinov, R., and Hinton, G. 2007. “Learning a Nonlinear Embedding by Preserving Class Neighbourhood Structure,” in *Artificial Intelligence and Statistics*, pp. 412–419.
- Sambamurthy, V., Bharadwaj, A., and Grover, V. 2003. “Shaping Agility through Digital Options: Reconceptualizing the Role of Information Technology in Contemporary Firms,” *MIS Quarterly*, JSTOR, pp. 237–263.

- Shapiro, C., and Varian, H. R. 1998. "Versioning: The Smart Way to Sell Information," *Harvard Business Review*, pp. 106–114.
- Shen, W., Hu, Y. J., and Rees, J. 2015. "Competing for Attention: An Empirical Study of Online Reviewers' Strategic Behaviors," *MIS Quarterly* (39:3), Society for Information Management and The Management Information Systems Research Center, pp. 683–696.
- Smith, K. G., Ferrier, W. J., and Ndofor, H. 2001. "Competitive Dynamics Research," *The Blackwell Handbook of Strategic Management*, Wiley Online Library, pp. 309–354.
- Society, T. E. 2011. "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," *Econometrica* (79:5), pp. 1407–1451.
- Di Sorbo, A., Panichella, S., Alexandru, C. V., Shimagaki, J., Visaggio, C. A., Canfora, G., and Gall, H. C. 2016. "What Would Users Change in My App? Summarizing App Reviews for Recommending Software Changes," in *Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering*, ACM, pp. 499–510.
- Spann, M., Fischer, M., and Tellis, G. J. 2015. "Skimming or Penetration? Strategic Dynamic Pricing for New Products," *Marketing Science* (34:2), pp. 235–249.
- Stavrianou, A., and Brun, C. 2012. "Opinion and Suggestion Analysis for Expert Recommendations," in *Proceedings of the Workshop on Semantic Analysis in Social Media*, Association for Computational Linguistics, pp. 61–69.
- Stieglitz, S., and Dang-Xuan, L. 2013a. "Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior," *Journal of Management Information Systems* (29:4), pp. 217–248.
- Stieglitz, S., and Dang-Xuan, L. 2013b. "Social Media and Political Communication: A Social Media Analytics Framework," *Social Network Analysis and Mining* (3:4), Springer, pp. 1277–1291.
- Systems, M. I. 2017. "Information Goods and Vertical Differentiation Author (S): Hemant K . Bhargava and Vidyanand Choudhary Source : Journal of Management Information Systems , Vol . 18 , No . 2 (Fall , 2001), Pp . 89-106 Stable URL : <http://www.jstor.org/stable/40398530> ," (18:2), pp. 89–106.
- Täckström, O., and McDonald, R. 2011. "Semi-Supervised Latent Variable Models for Sentence-Level Sentiment Analysis," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers-Volume 2*, Association for Computational Linguistics, pp. 569–574.
- Taylor Pentina, V. 2011. "Mobile Application Adoption by Young Adults: A Social Network Perspective ," *International Journal of Mobile Marketing* (6:2), pp. 60–70.
- Varian, H. 1997. "Versioning Information Goods," *University of California, Berkeley* (January), pp. 1–13.
- Wang, C., Jiang, F., and Yang, H. 2017. "A Hybrid Framework for Text Modeling with Convolutional RNN," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, pp. 2061–2069.
- Wang, H., Lu, Y., and Zhai, C. 2010a. "Latent Aspect Rating Analysis on Review Text Data: A Rating Regression Approach," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA: ACM, pp. 783–792.
- Wang, H., Lu, Y., and Zhai, C. 2010b. "Latent Aspect Rating Analysis on Review Text Data: A Rating Regression Approach," *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, pp. 783–792.

- Wang, Q., Li, B., and Singh, P. V. 2018. "Copycats vs. Original Mobile Apps: A Machine Learning Copycat-Detection Method and Empirical Analysis," *Information Systems Research*, INFORMS.
- Wang, X., Jiang, W., and Luo, Z. 2016. "Combination of Convolutional and Recurrent Neural Network for Sentiment Analysis of Short Texts," in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pp. 2428–2437.
- Weston, J. F. J. F., and Besley, S. 2008. *Fundamentos de Administración Financiera*, Cengage Learning Editores.
- Winkler, M., Abrahams, A. S., Gruss, R., and Ehsani, J. P. 2016. "Toy Safety Surveillance from Online Reviews," *Decision Support Systems* (90), Elsevier, pp. 23–32.
- Xu, B., Wang, N., Chen, T., and Li, M. 2015. "Empirical Evaluation of Rectified Activations in Convolutional Network," *arXiv Preprint arXiv:1505.00853*.
- Xu, S. 2015. "Balancing the Two Knowledge Dimensions in Innovation Efforts: An Empirical Examination among Pharmaceutical Firms," *Journal of Product Innovation Management* (32:4), pp. 610–621.
- Yi, Y., and Gong, T. 2013. "Customer Value Co-Creation Behavior: Scale Development and Validation," *Journal of Business Research* (66:9), Elsevier, pp. 1279–1284.
- Zentes, J., Morschett, D., and Schramm-Klein, H. 2017. "Pricing," in *Strategic Retail Management*, Springer, pp. 279–306.
- Zhang, Y., and Wallace, B. 2015. "A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification," *arXiv Preprint arXiv:1510.03820*.
- Zhou, S., Qiao, Z., Du, Q., Wang, A. G., Fan, W., and Yan, X. 2018. "Measuring Customer Agility from Online Reviews Using Big Data Text Analytics," *Journal of Management Information Systems*, p. Forthcoming.
- Zhu, F., and Zhang, X. 2010. "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing* (74:2), pp. 133–148.
- Zhu, J., Zou, H., Rosset, S., and Hastie, T. 2009. "Multi-Class Adaboost," *Statistics and Its Interface* (2:3), pp. 349–360.

5 CONCLUSIONS

This dissertation emphasizes the strategic value of online customer reviews, shifting the focus from their impacts on customer purchase behavior to their impacts on app firms' operational behaviors. Guided by signaling theory, product differentiation, and co-creation customer value, this dissertation proposes a framework for understanding: how online customer reviews as quality signals influence app firms' product innovation decisions; how product innovation differentiation strategies with marketing differentiation strategies influence app product performance; and what type of actionable information customers have talked about in online reviews.

The first research question, addressed in Chapter 2, examines the impact of quality signals in online reviews and product characteristics on firms' product innovation decisions. Product innovation (i.e., updating) plays a significant role in maintaining mobile apps' competitiveness. The focus of this dissertation was on the external online information environment, which may have significant impacts on app developers' product innovation decisions. The information environment is characterized by the recurrent nature of app updates and the real-time online customer feedback loop. The findings show that the app updating rate is faster for apps whose app makers signal quality in the app information page or when the novelty of the app decreases, and when customer feedback is negative or diverse (indicating a variety of consumer desires or complaints). It is found that the app updating rate gets slower as the app and app maker mature and succeed, as the app supports more devices, and when consumer feedback indicates the app is popular and of sufficiently high quality. The chance of an update is also diminished when the consumer feedback is addressed in the last update. The findings indicate marketplace signals can indeed help predict updating of apps, and thus provide indications of strategy to market players.

The second question, addressed in Chapter 3, emphasizes product innovation as a differentiation strategy and examines its impact on app performance. Product differentiation strategies are key to firms' survival and success, especially in a highly competitive industry. Compared to traditional software products and other digital goods, mobile apps have some unique characteristics that increase the importance and necessity of product differentiation strategies. However, existing studies lack insight regarding how and why online consumer reviews help firms achieve better performance or how firms' innovation strategies may cause online consumers to contribute more reviews. Adopting a fixed effects model, Chapter 3 conducted a longitudinal study of two-sided (demand-side and supply-side) coordinating efforts in a mobile app marketplace. This study illustrates the value of applying product innovation differentiation and marketing differentiation for a focal firm relative to its competitors to gain competitive advantages.

The third question, answered in Chapter 4, seeks to determine the types of actionable information from online customer reviews. To achieve that goal, the dissertation introduced a novel text mining framework that includes a domain-adapted deep learning approach to categorize two types of latent innovation opportunities, namely bug reports (e.g., exploitative information) and new feature requests (e.g., explorative information), embedded in online product reviews. The benefits of the proposed method compared to traditional supervised learning methods include better precision, recall, and F-measure score.

The work studied in this dissertation demonstrated significant impacts of online customer reviews on app firms' strategic decisions and their impacts on app performance. This dissertation has presented the possibility and value of consumer-centric strategy in the mobile app industry, using various quantitative and analytic studies. Now that the impacts of online customer reviews' on app firms' operational behaviors have been proven, the next step of understanding the dynamics

of online customer reviews and their impacts on other domains can be pursued. In particular, plans are to continue the research stream on how consumer feedback influences operational decision-making in the mobile app marketplace, and to continue to develop advanced deep learning techniques that can discover actionable information (e.g., consumer needs) more effectively. Rapidly understanding and responding to consumer needs is expected to improve customer satisfaction and strengthen firms' competitive advantages. Moreover, the further examination of the application of deep learning techniques in other domains such as drug event, service quality, and disruption event analysis is expected to be particularly promising. Finally, inspecting the impact of how consumers' social networks influence organizations' innovation decisions in hypercompetitive environments should be very fruitful. Future research will also include studies on marketing strategies related to demand estimation and competition.

6 BIBLIOGRAPHY

- Abrahams, A., Jiao, J., and Fan, W. 2013. "What's Buzzing in the Blizzard of Buzz? Automotive Component Isolation in Social Media Postings," *Decision Support Systems* (55:4), pp. 871–882.
- Abrahams, A. S., Fan, W., Wang, G. A., Zhang, Z., and Jiao, J. 2015. "An Integrated Text Analytic Framework for Product Defect Discovery," *Production and Operations Management* (24:6), pp. 975–990.
- Alves, H., Fernandes, C., and Raposo, M. 2016. "Value Co-Creation: Concept and Contexts of Application and Study," *Journal of Business Research* (69:5), Elsevier, pp. 1626–1633.
- Amit, R., and Zott, C. 2001. "Value Creation in E-business," *Strategic Management Journal* (22:6-7), Wiley Online Library, pp. 493–520.
- Animesh, A., Viswanathan, S., and Agarwal, R. 2011. "Competing 'creatively' in Sponsored Search Markets: The Effect of Rank, Differentiation Strategy, and Competition on Performance," *Information Systems Research* (22:1), pp. 153–169.
- Apté, C., Damerau, F., and Weiss, S. M. 1994. "Automated Learning of Decision Rules for Text Categorization," *ACM Transactions on Information Systems* (12:3), pp. 233–251.
- Aral, S., Dellarocas, C., and Godes, D. 2013. "Introduction to the Special Issue—social Media and Business Transformation: A Framework for Research," *Information Systems Research* (24:1), INFORMS, pp. 3–13.
- Arora, A., Krishnan, R., Telang, R., and Yang, Y. 2010. "An Empirical Analysis of Software Vendors' Patch Release Behavior: Impact of Vulnerability Disclosure," *Information Systems Research* (21:1), pp. 115–132.
- Bakos, Y., and Brynjolfsson, E. 1999. "Bundling Information Goods: Pricing, Profits, and Efficiency," *Management Science* (45:12), pp. 1613–1630.
- Balasubramanian, S., Bhattacharya, S., and Krishnan, V. V. 2015. "Pricing Information Goods: A Strategic Analysis of the Selling and Pay-per-Use Mechanisms," *Marketing Science* (34:2), INFORMS, pp. 218–234.
- Baldwin, C. Y., and Clark, K. B. 2006. "The Architecture of Participation: Does Code Architecture Mitigate Free Riding in the Open Source Development Model?," *Management Science* (52:7), INFORMS, pp. 1116–1127.
- Basole, R. C., and Karla, J. 2011. "On the Evolution of Mobile Platform Ecosystem Structure and Strategy," *Business & Information Systems Engineering* (3:5), Springer, p. 313.
- Basole, R. C., and Karla, J. 2012. "Value Transformation in the Mobile Service Ecosystem: A Study of App Store Emergence and Growth," *Service Science* (4:1), INFORMS, pp. 24–41.
- Beal, R. M. 2000. "Competing Effectively: Environmental Scanning, Competitive Strategy and Organisational Performance in Small Manufacturing Firms," *Journal of Small Business Management* (38:January), pp. 27–47.
- Becerra, M., Santaló, J., and Silva, R. 2013. "Being Better vs. Being Different: Differentiation, Competition, and Pricing Strategies in the Spanish Hotel Industry," *Tourism Management* (34), Elsevier, pp. 71–79.
- Bengio, Y. 2009. "Learning Deep Architectures for AI," *Foundations and Trends® in Machine Learning* (2:1), Now Publishers, Inc., pp. 1–127.
- Bengio, Y., Lamblin, P., Popovici, D., and Larochelle, H. 2007. "Greedy Layer-Wise Training of Deep Networks," in *Advances in Neural Information Processing Systems*, pp. 153–160.
- Bharadwaj, A. S. 2000. "A Resource-Based Perspective on Information Technology Capability

- and Firm Performance: An Empirical Investigation,” *MIS Quarterly*, JSTOR, pp. 169–196.
- Bhargava, H. K., and Choudhary, V. 2001. “Information Goods and Vertical Differentiation,” *Journal of Management Information Systems* (18:2), Taylor & Francis, pp. 89–106.
- Bhargava, H. K., and Choudhary, V. 2008. “Research Note — When Is Versioning Optimal for Information Goods?,” *Management Science* (54:5), pp. 1029–1035.
- Boudreau, K. J. 2012. “Let a Thousand Flowers Bloom? An Early Look at Large Numbers of Software App Developers and Patterns of Innovation,” *Organization Science* (23:5), pp. 1409–1427.
- Brun, C., and Hagege, C. 2013. “Suggestion Mining: Detecting Suggestions for Improvement in Users’ Comments,” *Research in Computing Science* (70), pp. 171–181.
- Büschken, J., and Allenby, G. M. 2016. “Sentence-Based Text Analysis for Customer Reviews,” *Marketing Science* (35:6), INFORMS, pp. 953–975.
- Caves, R. E., and Williamson, P. J. 1985. “What Is Product Differentiation, Really?,” *The Journal of Industrial Economics*, JSTOR, pp. 113–132.
- Chakravarti, A., and Xie, J. 2006a. “The Impact of Standards Competition on Consumers: Effectiveness of Product Information and Advertising Formats,” (XLIII:May), pp. 224–236.
- Chakravarti, A., and Xie, J. 2006b. “The Impact of Standards Competition on Consumers: Effectiveness of Product Information and Advertising Formats,” (43:2), pp. 224–236.
- Chamberlin, E. 2010. “H.(1933) The Theory of Monopolistic Competition,” *Cambridge, Mass, Harvard University*.
- Chen, H., Chiang, R. H. L., and Storey, V. C. 2012. “Business Intelligence and Analytics: From Big Data to Big Impact.,” *MIS Quarterly* (36:4), pp. 1165–1188.
- Chen, M.-J., and MacMillan, I. C. 1992. “Nonresponse and Delayed Response to Competitive Moves: The Roles of Competitor Dependence and Action Irreversibility,” *Academy of Management Journal* (35:3), Academy of Management, pp. 539–570.
- Chen, M.-J., Smith, K. G., and Grimm, C. M. 1992. “Action Characteristics as Predictors of Competitive Responses,” *Management Science* (38:3), INFORMS, pp. 439–455.
- Chen, N., Lin, J., Hoi, S., Xiao, X., and Zhang, B. 2014. “AR-Miner: Mining Informative Reviews for Developers from Mobile App Marketplace,” in *Proceedings of the 36th International Conference on Software Engineering*, ACM, pp. 767–778.
- Chen, Y. 2013. “Context-Dependent Preferences and Innovation Strategy Context-Dependent Preferences and Innovation Strategy,” (February 2014).
- Chen, Y., and Iyer, G. 2002. “Research Note Consumer Addressability and Customized Pricing,” *Marketing Science* (21:2), INFORMS, pp. 197–208.
- 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. 477–491.
- Collobert, R., and Weston, J. 2008. “A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning,” in *Proceedings of the 25th International Conference on Machine Learning*, ACM, pp. 160–167.
- Comino, S., Udine, U., Manenti, F. M., and Economiche, S. 2016. “Updates Management in Mobile Applications . iTunes vs Google Play Updates Management in Mobile Applications . iTunes vs Google Play . *,”
- Cooper, R. G., and Kleinschmidt, E. J. 1994. “Determinants of Timeliness in Product Development,” *Journal of Product Innovation Management* (11:5), Elsevier, pp. 381–396.
- Crawford-Welch, S. 1990. “Market Segmentation in the Hospitality Industry,” *Hospitality Research Journal* (14:2), ICHRIE Sage CA: Los Angeles, CA, pp. 295–308.

- Cusumano, M. A., Kahl, S. j, and Suarez, F. F. 2008. "Services, Industry Evolution, and the Copetitive Strategies of Product Firms," *Academy of Management Journal* (51:2), pp. 315–334.
- Datta, D., and Kajanana, S. 2013. "Do App Launch Times Impact Their Subsequent Commercial Success? An Analytical Approach," in *Cloud Computing and Big Data (CloudCom-Asia), 2013 International Conference on*, IEEE, pp. 205–210.
- Dey, D., and Lahiri, A. 2016. "Versioning of Video Games: Go Vertical in a Horizontal Market?," *Proceedings of the Annual Hawaii International Conference on System Sciences* (2016–March), pp. 5259–5268.
- Dickson, P. R., and Ginter, J. L. 1987. "Market Segmentation , Product Differentiation , and Marketing Strategy," (51:April), pp. 1–10.
- Distimo. 2009. "Pple App Store, BlackBerry App World, Google Android MarketNokia Ovi Store & Windows Marketplace for Mobile.,"
- Dougherty, D. 2001a. "Reimagining the Differentiation and Integration of Work for Sustained Product Innovation," *Organization Science* (12:5), pp. 612–631.
- Dougherty, D. 2001b. "Reimagining the Differentiation and Integration of Work for Sustained Product Innovation," *Organization Science* (12:5), INFORMS, pp. 612–631.
- Dougherty, D., and Dunne, D. D. 2011. "Organizing Ecologies of Complex Innovation," *Organization Science* (22:5), pp. 1214–1223.
- Duan, W., Gu, B., and Whinston, A. B. 2008. "The Dynamics of Online Word-of-Mouth and Product sales—An Empirical Investigation of the Movie Industry," *Journal of Retailing* (84:2), Elsevier, pp. 233–242.
- Ethiraj, S. K., and Zhu, D. H. 2008. "Performance Effects of Imitative Entry," *Strategic Management Journal* (29:8), Wiley Online Library, pp. 797–817.
- Ferrier, W. J., Smith, K. G., and Grimm, C. M. 1999. "The Role of Competitive Action in Market Share Erosion and Industry Dethronement: A Study of Industry Leaders and Challengers," *Academy of Management Journal* (42:4), Academy of Management, pp. 372–388.
- Fisher, M., Gallino, S., and Li, J. 2017. "Competition-Based Dynamic Pricing in Online Retailing: A Methodology Validated with Field Experiments," *Management Science*, INFORMS.
- Forman, C., Ghose, A., and Wiesenfeld, B. 2008. "Examining the Relationship between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets," *Information Systems Research* (19:3), INFORMS, pp. 291–313.
- French, S. A., Story, M., Jeffery, R. W., Snyder, P. A. T., Eisenberg, M., Sidebottom, A., and Murray, D. 1997. "Pricing Strategy to Promote Fruit and Vegetable Purchase in High School Cafeterias," *Journal of the Academy of Nutrition and Dietetics* (97:9), Elsevier, pp. 1008–1010.
- Galvis Carreño, L. V, and Winbladh, K. 2013. "Analysis of User Comments: An Approach for Software Requirements Evolution," in *Proceedings of the 2013 International Conference on Software Engineering*, pp. 582–591.
- Gan, Z., Pu, Y., Henao, R., Li, C., He, X., and Carin, L. 2017. "Learning Generic Sentence Representations Using Convolutional Neural Networks," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 2390–2400.
- Garg, R., and Telang, R. 2014. "Inferring App Demand from Publicly Available Data," *MIS Quarterly* (37:4), pp. 1253–1264.
- Ghose, A., and Han, S. P. 2014. "Estimating Demand for Mobile Applications in the New Economy," *Management Science* (60:6), pp. 1470–1488.

- Ghose, A., and Ipeirotis, P. G. 2011. "Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics," *IEEE Transactions on Knowledge and Data Engineering* (23:10), pp. 1498–1512.
- Gnyawali, D. R., Fan, W., and Penner, J. 2010. "Competitive Actions and Dynamics in the Digital Age: An Empirical Investigation of Social Networking Firms," *Information Systems Research* (21:3), INFORMS, pp. 594–613.
- Gnyawali, D. R., He, J., and Madhavan, R. 2006. "Impact of Co-Opetition on Firm Competitive Behavior: An Empirical Examination," *Journal of Management* (32:4), Sage Publications Sage CA: Thousand Oaks, CA, pp. 507–530.
- Goldberg, D. M., and Abrahams, A. S. 2018. "A Tabu Search Heuristic for Smoke Term Curation in Safety Defect Discovery," *Decision Support Systems* (105), Elsevier, pp. 52–65.
- Greenstein, S., and Markovich, S. 2012. "Pricing Experience Goods in Information Good Markets: The Case of eBusiness Service Providers," *International Journal of the Economics of Business* (19:1), pp. 119–139.
- Gu, B., and Ye, Q. 2014. "First Step in Social Media: Measuring the Influence of Online Management Responses on Customer Satisfaction," *Production and Operations Management* (23:4), pp. 570–582.
- Guo, X., Ling, L., Yang, C., Li, Z., and Liang, L. 2013. "Optimal Pricing Strategy Based on Market Segmentation for Service Products Using Online Reservation Systems: An Application to Hotel Rooms," *International Journal of Hospitality Management* (35), Elsevier Ltd, pp. 274–281.
- Hausman, J. A. 1978. "Specification Tests in Econometrics," *Econometrica: Journal of the Econometric Society*, JSTOR, pp. 1251–1271.
- Hevner, V. A., March, S. T., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *MIS Quarterly* (28:1), Springer, pp. 75–105.
- Hinterhuber, A. 2008. "Customer Value-Based Pricing Strategies: Why Companies Resist," *Journal of Business Strategy* (29:4), Emerald Group Publishing Limited, pp. 41–50.
- Hinton, G. E., and Salakhutdinov, R. R. 2006. "Reducing the Dimensionality of Data with Neural Networks," *Science* (313:5786), American Association for the Advancement of Science, pp. 504–507.
- Hippel, E. Von, and Krogh, G. Von. 2003. "Open Source Software and the ?Private-Collective? Innovation Model: Issues for Organization Science," *Organization Science* (14:2), pp. 209–223.
- Jha, N., and Mahmoud, A. 2017. "Mining User Requirements from Application Store Reviews Using Frame Semantics," in *International Working Conference on Requirements Engineering: Foundation for Software Quality*, Springer, pp. 273–287.
- Johansson, M., Hallberg, N., Hinterhuber, A., Zbaracki, M., and Liozu, S. 2012. "Pricing Strategies and Pricing Capabilities," *Journal of Revenue and Pricing Management* (11:1), Springer, pp. 4–11.
- Katzmarzik, A. 2011. "Product Differentiation for Software-as-a-Service Providers," *Business and Information Systems Engineering* (3:1), pp. 19–31.
- Kim, S. J., Kim, K. H., and Choi, J. 2017. "Design Innovation and Purchase Behavior for Augmented Products," *Journal of Business Research* (September), Elsevier, pp. 0–1.
- Kim, Y. 2014. "Convolutional Neural Networks for Sentence Classification," *arXiv Preprint arXiv:1408.5882*.
- Kleinbaum, D. G. 1998. "Survival Analysis, a Self-Learning Text," *Biometrical Journal* (40:1),

- Wiley Online Library, pp. 107–108.
- Kleinschmidt, E. J., and Cooper, D. R. 1991. “The Impact of Product Innovativeness on Performance,” *Journal of Product Innovation Management* (8), pp. 240–251.
- Lai, H., Hsu, J. S.-C., and Wu, M.-X. 2018. “The Impact S of Requested Permission on Mobile App Adoption: The Insights Based on an Experiment in Taiwan,” in *Proceedings of the 51st Hawaii International Conference on System Sciences*.
- Landis, J. R., and Koch, G. G. 1977. “The Measurement of Observer Agreement for Categorical Data,” *Biometrics*, JSTOR, pp. 159–174.
- Law, D., Gruss, R., and Abrahams, A. S. 2017. “Automated Defect Discovery for Dishwasher Appliances from Online Consumer Reviews,” *Expert Systems with Applications* (67), Elsevier Ltd, pp. 84–94.
- Lee, G., and Raghu, T. S. 2014. “Determinants of Mobile Apps’ Success: Evidence from the App Store Market,” *Journal of Management Information Systems* (31:2), Taylor & Francis, pp. 133–170.
- Lee, J., and Miller, D. 1999. “People Matter: Commitment to Employees, Strategy and Performance in Korean Firms,” *Strategic Management Journal*, JSTOR, pp. 579–593.
- Li, J., Xia, J., and Zajac, E. J. 2018. “On the Duality of Political and Economic Stakeholder Influence on Firm Innovation Performance: Theory and Evidence from Chinese Firms,” *Strategic Management Journal* (39:1), Wiley Online Library, pp. 193–216.
- Liu, B. 2012. “Sentiment Analysis and Opinion Mining,” *Synthesis Lectures on Human Language Technologies* (5:1), Morgan & Claypool Publishers, pp. 1–167.
- Liu, B., Hu, M., and Cheng, J. 2005. “Opinion Observer: Analyzing and Comparing Opinions on the Web,” in *Proceedings of the 14th International Conference on World Wide Web*, ACM, pp. 342–351.
- Liu, C. Z., Au, Y. A., and Choi, H. S. 2014. “Effects of Freemium Strategy in the Mobile App Market: An Empirical Study of Google Play,” *Journal of Management Information Systems* (31:3), Taylor & Francis, pp. 326–354.
- Liu, F., and Brandyberry, A. 2014. “Exploring the Effects of Aggregate Review Characteristics on Mobile Application Adoption,” in *Thirty Fifth International Conference on Information Systems*, Auckland, New Zealand.
- Luo, X. 2009. “Quantifying the Long-Term Impact of Negative Word of Mouth on Cash Flows and Stock Prices,” *Marketing Science* (28:1), pp. 148–165.
- Luo, X., Zhang, J., and Duan, W. 2013. “Social Media and Firm Equity Value,” *Information Systems Research* (24:1), INFORMS, pp. 146–163.
- LYU, C., and SHANG, Y. 2017. “The Time-Space Effect Research of High-Tech Industrial Agglomeration Model and Innovation Output,” *Journal of Management* (2), p. 7.
- Maalej, W., and Nabil, H. 2015. “Bug Report, Feature Request, or Simply Praise? On Automatically Classifying App Reviews,” in *Requirements Engineering Conference (RE), 2015 IEEE 23rd International*, IEEE, pp. 116–125.
- Maalej, W., Nabil, H., and Stanik, C. 2017. “On the Automatic Classification of App Reviews,” *Software Engineering 2017*, Gesellschaft Für Informatik eV.
- MacMillan, D., Burrows, P., and Ante, S. E. 2009. “Inside the App Economy,” *Business Week* (22).
- Makadok, R. 2010. “The Interaction Effect of Rivalry Restraint and Competitive Advantage on Profit: Why the Whole Is Less than the Sum of the Parts,” *Management Science* (56:2), INFORMS, pp. 356–372.

- McIlroy, S., Ali, N., and Hassan, A. E. 2016. “Fresh Apps: An Empirical Study of Frequently-Updated Mobile Apps in the Google Play Store,” *Empirical Software Engineering* (21:3), Springer, pp. 1346–1370.
- Mikolov, T., Chen, K., Corrado, G., and Dean, J. 2013. “Efficient Estimation of Word Representations in Vector Space,” *arXiv Preprint arXiv:1301.3781*.
- Miller, D. 1986. “Configurations of Strategy and Structure: Towards a Synthesis,” *Strategic Management Journal* (7:3), Wiley Online Library, pp. 233–249.
- Miller, D. 1988. “Relating Porter’s Business Strategies to Environment and Structure: Analysis and Performance Implications,” *Academy of Management Journal* (31:2), Academy of Management, pp. 280–308.
- Mintzberg, H. 1988. “Generic Strategies: Toward a Comprehensive Framework,” *Advances in Strategic Management* (5), pp. 1–67.
- Moghaddam, S. 2015. “Beyond Sentiment Analysis: Mining Defects and Improvements from Customer Feedback,” in *Advances in Information Retrieval*, Springer, pp. 400–410.
- Mudambi, S. M., and Schuff, D. 2010. “What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon. Com,” *MIS Quarterly* (34:1), pp. 185–200 (available at <http://ssrn.com/abstract=2175066>).
- Mummalaneni, V., Gruss, R., Goldberg, D. M., Ehsani, J. P., and Abrahams, A. S. 2018. “Social Media Analytics for Quality Surveillance and Safety Hazard Detection in Baby Cribs,” *Safety Science* (104), Elsevier, pp. 260–268.
- Nagi, J., Ducatelle, F., Di Caro, G. A., Cireşan, D., Meier, U., Giusti, A., Nagi, F., Schmidhuber, J., and Gambardella, L. M. 2011. “Max-Pooling Convolutional Neural Networks for Vision-Based Hand Gesture Recognition,” in *Signal and Image Processing Applications (ICSIPA), 2011 IEEE International Conference on*, IEEE, pp. 342–347.
- Nakagawa, T., Inui, K., and Kurohashi, S. 2010. “Dependency Tree-Based Sentiment Classification Using CRFs with Hidden Variables,” in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Association for Computational Linguistics, pp. 786–794.
- Narasimhan, R., Ghosh, S., and Mendez, D. 1993. “A Dynamic Model of Product Quality and Pricing Decisions on Sales Response,” *Decision Sciences* (24:5), Wiley Online Library, pp. 893–908.
- Nasri, L., Baghersad, M., Gruss, R., Marucchi, N. S. W., Abrahams, A. S., and Ehsani, J. P. 2018. “An Investigation into Online Videos as a Source of Safety Hazard Reports,” *Journal of Safety Research* (65), Elsevier, pp. 89–99.
- Nayebi, M., Adams, B., and Ruhe, G. 2016. “Release Practices for Mobile Apps -- What Do Users and Developers Think?,” *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, pp. 552–562.
- Oh, J., Kim, D., Lee, U., Lee, J.-G., and Song, J. 2013. “Facilitating Developer-User Interactions with Mobile App Review Digests,” in *CHI’13 Extended Abstracts on Human Factors in Computing Systems*, ACM, pp. 1809–1814.
- Pagano, D., and Maalej, W. 2013. “User Feedback in the Appstore: An Empirical Study,” in *Requirements Engineering Conference (RE), 2013 21st IEEE International*, IEEE, pp. 125–134.
- Pang, B., Lee, L., and Vaithyanathan, S. 2002. “Thumbs up?: Sentiment Classification Using Machine Learning Techniques,” in *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing-Volume 10*, Association for Computational

- Linguistics, pp. 79–86.
- Panichella, S., Di Sorbo, A., Guzman, E., Visaggio, C. A., Canfora, G., and Gall, H. C. 2015. “How Can I Improve My App? Classifying User Reviews for Software Maintenance and Evolution,” in *Software Maintenance and Evolution (ICSME), 2015 IEEE International Conference on*, IEEE, pp. 281–290.
- Pennington, J., Socher, R., and Manning, C. 2014. “Glove: Global Vectors for Word Representation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543.
- Porter, J. 2010. *Designing for the Social Web (Google eBook)*, Peachpit Press (available at http://books.google.com/books/about/Designing_for_the_Social_Web.html?id=CNJuvYBI_IsC&pgis=1).
- Porter, M. E. 1980. “Competitive Strategy: Techniques for Analyzing Industries and Competition,” *New York*, p. 300.
- Powers, D. M. 2011. “Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation,” Bioinfo Publications.
- Prahalad, C. K., and Ramaswamy, V. 2000. “Co-Opting Customer Competence,” *Harvard Business Review* (78:1), pp. 79–90.
- Prahalad, C. K., and Ramaswamy, V. 2004. “Co-creation Experiences: The next Practice in Value Creation,” *Journal of Interactive Marketing* (18:3), Wiley Online Library, pp. 5–14.
- Prajogo, D. I., and Ahmed, P. K. 2006. “Relationships between Innovation Stimulus, Innovation Capacity, and Innovation Performance,” *R & D Management* (36:5), pp. 499–515.
- Qiao, Z., Wang, G. A., Zhou, M., and Fan, W. 2018. “The Impact of Customer Reviews on Product Innovation: Empirical Evidence in Mobile Apps,” in *Analytics and Data Science*, Springer, pp. 95–110.
- Ragaglia, D., and Roma, P. 2014. “Understanding the Drivers of the Daily App Rank : The Role of Revenue Models,” in *Proceedings of the 26th Annual Conference of the Productions and Operations Management Society*, pp. 1–10.
- Ramanand, J., Bhavsar, K., and Pedanekar, N. 2010. “Wishful Thinking: Finding Suggestions and ‘Buy’ Wishes from Product Reviews,” in *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, Association for Computational Linguistics, pp. 54–61.
- Ramaswamy, V., and Prahalad, C. K. 2004. “Co-Creation Experiences: The Next Ractice in Value Creation,” *Journal of Interactive Marketing* (18:3), pp. 5–14.
- Ramos, J. 2003. “Using Tf-Idf to Determine Word Relevance in Document Queries,” in *Proceedings of the First Instructional Conference on Machine Learning*.
- Rehurek, R., and Sojka, P. 2010. “Software Framework for Topic Modelling with Large Corpora,” in *In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, Citeseer.
- Sak, H., Senior, A., and Beaufays, F. 2014. “Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling,” in *Fifteenth Annual Conference of the International Speech Communication Association*.
- Salakhutdinov, R., and Hinton, G. 2007. “Learning a Nonlinear Embedding by Preserving Class Neighbourhood Structure,” in *Artificial Intelligence and Statistics*, pp. 412–419.
- Sambamurthy, V., Bharadwaj, A., and Grover, V. 2003. “Shaping Agility through Digital Options: Reconceptualizing the Role of Information Technology in Contemporary Firms,” *MIS Quarterly*, JSTOR, pp. 237–263.

- Shapiro, C., and Varian, H. R. 1998. "Versioning: The Smart Way to Sell Information," *Harvard Business Review*, pp. 106–114.
- Shen, W., Hu, Y. J., and Rees, J. 2015. "Competing for Attention: An Empirical Study of Online Reviewers' Strategic Behaviors," *MIS Quarterly* (39:3), Society for Information Management and The Management Information Systems Research Center, pp. 683–696.
- Smith, K. G., Ferrier, W. J., and Ndofor, H. 2001. "Competitive Dynamics Research," *The Blackwell Handbook of Strategic Management*, Wiley Online Library, pp. 309–354.
- Society, T. E. 2011. "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," *Econometrica* (79:5), pp. 1407–1451.
- Di Sorbo, A., Panichella, S., Alexandru, C. V., Shimagaki, J., Visaggio, C. A., Canfora, G., and Gall, H. C. 2016. "What Would Users Change in My App? Summarizing App Reviews for Recommending Software Changes," in *Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering*, ACM, pp. 499–510.
- Spann, M., Fischer, M., and Tellis, G. J. 2015. "Skimming or Penetration? Strategic Dynamic Pricing for New Products," *Marketing Science* (34:2), pp. 235–249.
- Stavrianou, A., and Brun, C. 2012. "Opinion and Suggestion Analysis for Expert Recommendations," in *Proceedings of the Workshop on Semantic Analysis in Social Media*, Association for Computational Linguistics, pp. 61–69.
- Stieglitz, S., and Dang-Xuan, L. 2013a. "Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior," *Journal of Management Information Systems* (29:4), pp. 217–248.
- Stieglitz, S., and Dang-Xuan, L. 2013b. "Social Media and Political Communication: A Social Media Analytics Framework," *Social Network Analysis and Mining* (3:4), Springer, pp. 1277–1291.
- Systems, M. I. 2017. "Information Goods and Vertical Differentiation Author (S): Hemant K . Bhargava and Vidyanand Choudhary Source : Journal of Management Information Systems , Vol . 18 , No . 2 (Fall , 2001), Pp . 89-106 Stable URL : <http://www.jstor.org/stable/40398530> ," (18:2), pp. 89–106.
- Täckström, O., and McDonald, R. 2011. "Semi-Supervised Latent Variable Models for Sentence-Level Sentiment Analysis," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers-Volume 2*, Association for Computational Linguistics, pp. 569–574.
- Taylor Pentina, V. 2011. "Mobile Application Adoption by Young Adults: A Social Network Perspective ," *International Journal of Mobile Marketing* (6:2), pp. 60–70.
- Varian, H. 1997. "Versioning Information Goods," *University of California, Berkeley* (January), pp. 1–13.
- Wang, C., Jiang, F., and Yang, H. 2017. "A Hybrid Framework for Text Modeling with Convolutional RNN," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, pp. 2061–2069.
- Wang, H., Lu, Y., and Zhai, C. 2010a. "Latent Aspect Rating Analysis on Review Text Data: A Rating Regression Approach," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA: ACM, pp. 783–792.
- Wang, H., Lu, Y., and Zhai, C. 2010b. "Latent Aspect Rating Analysis on Review Text Data: A Rating Regression Approach," *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, pp. 783–792.

- Wang, Q., Li, B., and Singh, P. V. 2018. "Copycats vs. Original Mobile Apps: A Machine Learning Copycat-Detection Method and Empirical Analysis," *Information Systems Research*, INFORMS.
- Wang, X., Jiang, W., and Luo, Z. 2016. "Combination of Convolutional and Recurrent Neural Network for Sentiment Analysis of Short Texts," in *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pp. 2428–2437.
- Weston, J. F. J. F., and Besley, S. 2008. *Fundamentos de Administración Financiera*, Cengage Learning Editores.
- Winkler, M., Abrahams, A. S., Gruss, R., and Ehsani, J. P. 2016. "Toy Safety Surveillance from Online Reviews," *Decision Support Systems* (90), Elsevier, pp. 23–32.
- Xu, B., Wang, N., Chen, T., and Li, M. 2015. "Empirical Evaluation of Rectified Activations in Convolutional Network," *arXiv Preprint arXiv:1505.00853*.
- Xu, S. 2015. "Balancing the Two Knowledge Dimensions in Innovation Efforts: An Empirical Examination among Pharmaceutical Firms," *Journal of Product Innovation Management* (32:4), pp. 610–621.
- Yi, Y., and Gong, T. 2013. "Customer Value Co-Creation Behavior: Scale Development and Validation," *Journal of Business Research* (66:9), Elsevier, pp. 1279–1284.
- Zentes, J., Morschett, D., and Schramm-Klein, H. 2017. "Pricing," in *Strategic Retail Management*, Springer, pp. 279–306.
- Zhang, Y., and Wallace, B. 2015. "A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification," *arXiv Preprint arXiv:1510.03820*.
- Zhou, S., Qiao, Z., Du, Q., Wang, A. G., Fan, W., and Yan, X. 2018. "Measuring Customer Agility from Online Reviews Using Big Data Text Analytics," *Journal of Management Information Systems*, p. Forthcoming.
- Zhu, F., and Zhang, X. 2010. "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing* (74:2), pp. 133–148.
- Zhu, J., Zou, H., Rosset, S., and Hastie, T. 2009. "Multi-Class Adaboost," *Statistics and Its Interface* (2:3), pp. 349–360.