ESSAYS IN NEW PRODUCT PRERELEASE MARKETING: PREDICTION AND DYNAMICS

by

PENG ZHANG

(Under the Direction of Sundar G. Bharadwaj and of Anindita Chakravarty)

ABSTRACT

My dissertation focuses on (1) firms' new product sales forecast and (2) the influence of customer and investor sentiment on firms' new product advertising adjustment both strategically and tactically during the prerelease period. In the first essay, I explore the use of dynamic online knowledge collaborative activities for prerelease new product sales forecast. Unlike customer buzz from social bonding or creativity thriving online platforms, dynamic online knowledge collaborative activities drive the quality of the co-created content about a forthcoming new product, which in turn offers necessary information for prospective customers to educate themselves before making a purchase decision. The proposed modeling process combining Functional Data Analysis (FDA) and Random Forest (RF) methods precisely captures this dynamic process. The results show that the inclusion of online knowledge collaborative activities significantly enhances the prerelease new product forecast accuracy. In the second essay, I delve into the advertising strategy and tactics of firms' new product introduction practice. I find support for both customer and investor sentiment playing a critical role affecting new products' prerelease advertising deployment. In this essay, customer sentiment reflects prospective customers' emotional attachment to the new product as the end-buyers. Investor sentiment captures the investors' confidence to further invest in the company, considering the new products' contributing role in the firm value enhancement. In response to both types of sentiment, firms dynamically adjust their prerelease advertising strategies (i.e., new product promotion and corporate branding) and utilize different advertising tactics (i.e., digital and traditional advertising) in order to effectively and efficiently communicate the benefits of the new product. The outcome model shows that dynamic prerelease advertising adjustments in response to customer sentiment are more effective in driving new product sales and firm value

INDEX WORDS: Prerelease New Product Sales Forecast, Prerelease Advertising Strategy and Tactics, Customer Sentiment, Investor Sentiment

ESSAYS IN NEW PRODUCT PRERELEASE MARKETING: PREDICTION AND DYNAMICS

by

PENG ZHANG

B.A. (Hons.), Tianjin Normal University, China, 2005

M.A., The University of Georgia, 2009

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2021

ESSAYS IN NEW PRODUCT PRERELEASE MARKETING: PREDICTION AND DYNAMICS

by

PENG ZHANG

Major Professors: Sundar G. Bharadwaj
Anindita Chakravarty

Committee: Pengyuan Wang
Elham Yazdani

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2021

DEDICATION

To my parents, Guoqiang Zhang and Shuhua Zhao, for your dedicated support and unwavering belief in me. It is your consistent encouragement that guides me through the tough time of my life.

ACKNOWLEDGEMENTS

First of all, I would like to extend my sincere appreciation to my co-advisers Sundar Bharadwaj and Anindita Chakravarty, who have provided me with endless patience and guidance over the past five years. This dissertation would not have been possible, without your support and wholehearted encouragement.

I would also like to thank my committee members – Pengyuan Wang and Elham Yazdani. I feel extremely lucky to have the smart and caring guidance from both of you toward the completion of this thesis.

Additionally, I am grateful to have my fellow Ph.D. students, past and present, to experience the ups and downs of our lives together.

Finally, I really appreciate Charlotte Mason, John Hulland, Kristy McManus and all other faculty and staff of the Marketing Department. Hope to keep in touch with everyone in the future!

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION AND OVERVIEW OF THE DISSERTATION

New product introduction plays an essential role in showcasing firms' innovativeness, maintaining their competencies, and promoting their profitability as well as market value (Eliashberg and Robertson 1988; Srinivasan and Hanssens 2009). During this process, managers often choose to make a new product preannouncement prior to the official release. Specifically, the new product preannouncement refers to "a formal, deliberate communication before a firm actually undertakes a particular marketing action" (Eliashberg and Robertson 1988 p. 282). According to the signaling theory, such a preannouncement conveys critical information to the marketplace participants (i.e., prospective customers and investors) about the superiority of the new product, resulting in new product sales and firm value enhancement (Shankar 1999; Spence 1974; Su and Rao 2010).

However, the effectiveness of new product preannouncement also involves tremendous uncertainties. First, it is unable to guarantee the favorable reactions from customers or investors. In this regard, these participants may make hasty evaluation of the new product merely based on the limited information revealed in the preannouncement. It hampers their comprehensive understanding of the superiority of the new product, thus triggering the negative consequences (e.g., sales or stock price decline). Second, it may stimulate competitors' retaliation, leading to a fiercer market competition and the new product sales decline.

Fortunately, the time between the preannouncement and the new product official release often takes a few months (even more than a year in some cases), rendering the launching firm managers plenty of opportunities to observe and learn customers' and investors' responses. Hence, managers are better able to make a more accurate prediction of their new product sales and undertake rectifying adjustments of their marketing campaigns so as to more effectively and efficiently communicate the values of the new product among those people (Chaney, Devinney, and Winer 1991; Fourt and Woodlock 1960; Guide 2000).

The first essay focuses on the prerelease new product sales forecast (prerelease NPS forecast). While most prior studies in this area explore the use of prerelease customer buzz from social bonding or creativity thriving online platforms for the prerelease NPS forecast, I aim at investigating the value of the dynamic online knowledge collaborative activities in undertaking this task.

Drawing on literature from both marketing and the information system disciplines, I differentiate online knowledge collaborative activities from customer buzz in terms of their various roles in auguring new product sales. According to Houston et al. (2018), customer buzz reflects prospective customers' emotion or enthusiasm toward a forthcoming product as well as their purchase intent, which can be found from online platforms serving the social bonding or creativity thriving functions (e.g., Facebook Youtube, TikTok, etc.). In contrast, the online knowledge collaborative activities emphasize on "the individual acts of offering knowledge to others as well as adding to, recombining, modifying, and integrating knowledge that others have contributed" (Faraj, Jarvenpaa and Majchrzak 2011 p. 1224), which occur at online platforms featuring knowledge collaboration (e.g., Wikipedia.org, Piazza.com, etc.). In the context of prerelease NPS forecast, the co-created content through online knowledge collaboration facilitates prospective customers' information search efficiency. Moreover, the dynamic process of such a knowledge collaboration drives the quality of the co-created content (measured by content informativeness, content novelty, and perspective diversity), which motivates customers to come back to educate themselves about the new product from time to time. In the end, prospective customers are able to more comprehensively evaluate the benefits of the new product and more ready to make the purchase decision.

Empirically, I employ the functional data analysis (FDA) and the machine learning based ensemble random forest (RF) techniques for the prerelease NPS forecast. The FDA technique allows me to capture the evolution of online knowledge collaborative activities and tremendously reduce the high dimensionality of the raw data, which offers great convenience for the build-up of the predictive model. The RF technique is capable of automatically determining the validity of each individual predictor and the high-level interactions (three- or four-way) amongst these predictors in terms of their contribution to the prerelease NPS forecast accuracy, without *a priori* hand-curation made by the researcher.

The proposed forecasting solution for sales six months post launch outperforms typical regression-based approaches in terms of forecast accuracy. Furthermore, the inclusion of online knowledge collaborative activities makes a critical contribution to the forecast accuracy of the new product sales, even ten weeks prior to the formal release. An accurate early forecast provides managers with the lead time to strategically adjust their marketing strategies and optimize the new products' performance. Moreover, the proposed method of current study utilizing free and publicly available data from Wikipedia is actionable and easy for managers to execute.

In the second essay, I take one step further to investigate whether managers of the launching firm may proactively undertake prerelease advertising deployment adjustment in response to both customer and investor sentiment toward the new product prerelease. Beyond that, I further explore the ultimate outcome of such adjustments in terms of new product sales and firm value.

Customer sentiment during the prerelease period is defined as "the aggregation of [emotional] expressions of anticipation by consumers for a forthcoming new product" (Houston et al. 2018 p. 349). Investor sentiment refers to the varied degrees of investors' confidence (i.e., bullish vs. bearish) about the new product in terms of its impact on firm value (Baker and Wurgler 2007; Luo, Jiang and Cai 2015). While the former directly influences the new product sales (Xiong and Bharadwaj 2014), the latter affects the launching firms' financing ability to acquire more resources from the capital market for its future operation and innovation (Chan and Hameed 2006; James 1987). Therefore, they both motivate managers to closely monitor the increased positivity (or bullishness) of customer and investor sentiment and correspondingly adjust the prerelease advertising campaigns. By doing so, these managers are able to more effectively and efficiently communicate the benefits of the new product, elicit positive emotion among prospective customers, and enhance investors' confidence about the new product's contributing role to the launching firm's future cash flow and the launching firm's marketing commitment to supporting new product success.

I examine the prerelease advertising adjustment by considering both advertising strategies and tactics. While the former attends to what to communicate in response to customer and investor sentiment, the latter handles how to communicate (i.e., through what media vehicles). Based on prior literature, I categorize advertising strategies as 1) new product promotional and 2) corporate branding campaigns. Tactically, I explore the differential roles of traditional and digital advertisings in efficiently delivering the strategical messages to the audiences (i.e., customers and investors).

Empirically, I devise a three-step modeling estimation process incorporating panel vector autoregression (PVAR), Bayesian vector autoregression (BVAR), and Bayesian Hierarchical Model. The use of PVAR allows me to holistically examine whether managers undertake prerelease advertising adjustment in response to customer and investor sentiment. The combined exploitation of BVAR and the Bayesian Hierarchical Model enables me to go down to the product level to ascertain how different prerelease advertising adjustment schemes (in response to customer and investor sentiment) may usher into any different consequences in terms of new product sales and firm value.

The second essay makes both academic and managerial contributions. First, it recognizes and effectively estimates the differential values of customer and investor sentiment as separate information sources guiding managers' prerelease advertising deployment. Second, it captures the dynamic advertising adjustment at both strategic and tactic levels in response to customer and investor sentiment during the new product prerelease period. Third, the empirically devised three-step approach quantifies the ultimate outcome of such dynamic adjustment, accounting for both endogeneity and heterogeneity concerns. Finally, it offers insightful guidelines for managers to more effectively and efficiently optimize their advertising campaigns during the critical new product release period.

REFERENCE

Baker, Malcolm and Jeffrey Wurgler (2007), "Investor Sentiment in the Stock Market," Journal of economic perspectives, 21 (2), 129-52.

Chan, Kalok and Allaudeen Hameed (2006), "Stock Price Synchronicity and Analyst Coverage in Emerging Markets," Journal of Financial Economics, 80 (1), 115-47.

Chaney, Paul K, Timothy M Devinney, and Russell S Winer (1991), "The Impact of New Product Introductions on the Market Value of Firms," Journal of business, 573-610.

Eliashberg, Jehoshua and Thomas S Robertson (1988), "New Product Preannouncing Behavior: A Market Signaling Study," Journal of Marketing Research, 25 (3), 282-92.

Faraj, Samer, Sirkka L Jarvenpaa, and Ann Majchrzak (2011), "Knowledge Collaboration in Online Communities," Organization science, 22 (5), 1224-39.

Fourt, Louis A and Joseph W Woodlock (1960), "Early Prediction of Market Success for New Grocery Products," Journal of Marketing, 25 (2), 31-38.

Guide Jr, V Daniel R (2000), "Production Planning and Control for Remanufacturing: Industry Practice and Research Needs," Journal of Operations Management, 18 (4), 467-83.

Houston, Mark B, Ann-Kristin Kupfer, Thorsten Hennig-Thurau, and Martin Spann (2018), "Pre-Release Consumer Buzz," Journal of the academy of marketing Science, 46 (2), 338-60.

James, Christopher (1987), "Some Evidence on the Uniqueness of Bank Loans," Journal of Financial Economics, 19 (2), 217-35.

Luo, Miao, Daniel X Jiang, and Jun Cai (2014), "Investor Sentiment, Product Features, and Advertising Investment Sensitivities," Asia-Pacific Journal of Financial Studies, 43 (6), 798-837.

Shankar, Venkatesh (1999), "New Product Introduction and Incumbent Response Strategies: Their Interrelationship and the Role of Multimarket Contact," Journal of Marketing Research, 36 (3), 327-44.

Spence, Michael (1974), "Competitive and Optimal Responses to Signals: An Analysis of Efficiency and Distribution," Journal of Economic theory, 7 (3), 296-332.

Srinivasan, Shuba and Dominique M Hanssens (2009), "Marketing and Firm Value: Metrics, Methods, Findings, and Future Directions," Journal of Marketing Research, 46 (3), 293-312.

Su, Meng and Vithala R Rao (2010), "New Product Preannouncement as a Signaling Strategy: An Audience-Specific Review and Analysis," Journal of Product Innovation Management, 27 (5), 658-72.

Xiong, Guiyang and Sundar Bharadwaj (2014), “Prerelease Buzz Evolution Patterns and New Product Performance,” Marketing Science, 33 (3), 401-21.

CHAPTER 2

PRERELEASE NEW PRODUCT SALES FORECAST USING ONLINE DYNAMIC KNOWLEDGE COLLABORATIVE ACTIVITIES[1]

[1] Zhang, P., S. G. Bharadwaj, and G. Xiong. To be submitted to *Journal of Marketing Research*.

ABSTRACT

This study uses the dynamics of knowledge collaborative activities on Wikipedia to build a prerelease forecasting model of new product sales. The study argues that dynamic and engaged online knowledge collaboration renders better content quality of the co-created knowledge, which assists prospective customers in educating and updating themselves about the new product benefits before making the final purchase decision. The study utilizes a unique modeling structure combining the Functional Data Analysis (FDA) and Random Forecast (RF) approaches. The FDA approach effectively captures the evolutionary paths of dynamic knowledge collaboration activities over time. The nonparametric RF method improves forecast accuracy by automatically capturing the higher order interactions among various predictors and does not require any *a priori* assumption of data distribution. The proposed forecasting solution for sales six months post launch outperforms typical regression-based approaches in terms of forecast accuracy. Furthermore, the inclusion of online knowledge collaborative activities makes a critical contribution to the forecast accuracy of the new product sales, even ten weeks prior to the formal release. An accurate early forecast provides managers with the lead time to strategically adjust their marketing strategies and optimize the new products' performance. Moreover, the proposed method of current study utilizing free and publicly available data from Wikipedia is actionable and easy for managers to execute.

INTRODUCTION

Successful new product introductions meet customers' evolving needs, maintain firms' competitiveness, and deliver organic growth for firms. However, managers of these firms are often faced with two critical questions during the new product prerelease period: how well will the new product perform? (*hereinafter* prerelease NPS forecast) and how early can they get an accurate sales forecast of their new products (Hu et al. 2017). Addressing the first question enables those managers to make necessary proactive modifications to the marketing mix, optimize the launching firms' logistics planning, and thus minimize the likelihood and the scale of new product post-release failure (Chaney, Devinney, and Winer 1991; Fourt and Woodlock 1960; Guide 2000). With respect to the second question, the earlier managers can make an accurate prediction of the new product sales, the greater the lead time for implementing strategic and tactical marketing mix changes.

An early prerelease NPS forecast has heightened relevance for entertainment products, such as video games and movies, as these products have a relatively short lifetime and most of their sales volumes occur in the first few months after official launch. Figure 2.5 of the Appendix 1 Exhibition 2.1 presents the two-year sales volume distribution for the 300 new video games launched between 2011 and 2017 in the North American market. More than 20% of the total sales volume of these games took place during the opening week and over 77% achieved within the first six months since launch. Therefore, managers of these products need to take accelerated changes to their marketing before it is too late (Neelamegham and Chintagunta 1999; Xiong and Bharadwaj 2014).

As the main objective of this study is to build a better forecast of new product sales during its prerelease period, I follow the tradition in predictive analytical research and focus more on pure prediction rather than ascertaining the causality of the predictor variables[2]. It is worth noting that rigorous predictive analytics usually serve as motivation for later researchers to build causal theory. As far as marketing discipline is concerned, there have been a large body of predictive studies of real-world phenomena that are extremely seminal and have motivated follow-up theoretical research (e.g., Chintagunta, Gopinath and Venkataraman 2010; Cui and Curry 2005; Godes and Mayzlin 2004; Neelamegham and Chintagunta 1999; Rust, Kumar, and Venkatesan 2011; Tirunillai and Tellis 2012, etc.).

Against this backdrop, this research addresses the following questions. 1) What is the relationship between the entire history of prerelease online knowledge-collaborative activities and the NPS forecast? 2) Does the addition of such dynamic information improve the accuracy of the prerelease NPS forecast, relative to the forecast exclusively based on customer social buzz? 3) Could partial history of knowledge collaborative activities, especially the early collaborative dynamics, be used to forecast sales well before product release (i.e., early forecast)?

This study makes the following contributions. First, to my best knowledge, it is among the first to leverage the dynamic knowledge collaborative activities from a public domain online community platform for prerelease NPS forecast. The proposed method is effective in enhancing predictive accuracy by leveraging online knowledge collaborative activities, over and above other data sources (i.e., customer buzz data, advertising spends, product attributes, etc.). Second, the modeling approach outperforms forecasts based on aggregated measures of the prerelease knowledge collaborative activities to predict new product sales, highlighting the importance of

[2] I adopt Shmueli and Koppius' (2011, p.555) definition of predictive analytics that "refer[s] to the building and assessment of a model aimed at making empirical predictions." In spite of this major objective, I still provide possible theoretical explanation to shed light on the mechanism between the predictors and the DV.

utilizing the dynamic history of the information. Third, the approach outperforms other modeling options, by accurately predicting the new product sales as early as three months before the formal launch (i.e., early forecasting). From a practical standpoint, not only is the prediction more accurate and can be estimated early, allowing for strategic changes to the marketing mix and new product launch timing and strategy, it also uses a readily available and free of cost source of data for the prerelease NPS forecast. In the rest of this paper, I first discuss the research context, followed by a review of the relevant literature. Then I elaborate on the theoretical foundation of online knowledge collaboration in online community platforms and its relevance to the prerelease NPS forecast. Next, I present the method, data, and results sections. I conclude with the discussion of both academic contributions and practical implications.

RESEARCH CONTEXT

Faraj, Jarvenpaa and Majchrzak (2011 p. 1224) define online community as an "open collective of dispersed individuals who are not necessarily known or identifiable but share [some] common interests [in an online setting]". Prior literature has examined the potential of using participants' activities from online communities to undertake the prerelease NPS forecast (Cheung and Thadani 2012; Duan, Gu, and Whinston 2008; Gruen, Osmonbekov, and Czaplewski 2006; Lee, Park, and Han, 2008; Xiong and Bharadwaj 2014). The logic is that online community activities foster customer-to-customer know-how exchange, which further builds awareness and enhances customer perceptions of the new product's value (cf., Rosario, Valck, and Sotgiu 2020). Given the experiential nature of new video games or movies, customers have a tendency to rely more on online community activities to acquire necessary information and evaluate the value of the new product before making purchase decisions (Chakravarty, Liu, and Mazumdar 2010; Eliashberg and Shugan 1997).

Based on user motivation (intrinsic vs. extrinsic) and the extent of audience reach (narrow- vs. broadcasting), most of online community platforms offer *social bonding* (e.g., Facebook, Twitter) and *creativity thriving* (e.g., YouTube, TikTok, online forums) functions. While intrinsic motivation (such as altruism and enjoyment) reflects a person's own desire to perform a certain task, extrinsic motivation (e.g., social capital, reputation, economic rewards or punishment, etc.) can be perceived as an external incentive encouraging some behaviors. In terms of the extent of audience reach, Barasch and Berger (2014) propose that participants while initiating any online community activity (posting, tweeting, following, etc.) may already have had an idea regarding whether they are narrowcasting (small audience reach) or broadcasting (large audience reach). When a person is narrowcasting (e.g., toward friends and families), the content of her online community activities is more other-focused, thus more beneficial for the information recipients. Otherwise, it is more self-centric with higher proportion of self-presentational content. Thus, activities on Facebook or Twitter tend to be social bonding, motivated to share similar interests, reinforce shared beliefs, and eventually foster the ideological homogeneity among 'friends' (Norris 2002). It serves as "a virtual glue for consumers to quench their desire to belong to a group" (Seraj 2012 p. 210). In contrast, online community activities on YouTube, TikTok, or online forums are more creativity thriving, which facilitates a person's free expression of her personality, creativity, and opinions in a manner that she may not usually behave due to the constraints or possible sanctions from her own social sphere (Bargh, McKenna, and Fitzsimons 2002; Derlega et al. 1993). In the context of prerelease NPS forecast, Houston et al. (2018 p. 339) have used the term "customer buzz" to describe people's activities on both social bonding and creativity thriving online community platforms and referred to it as an "aggregation of observable [emotional] expressions of anticipation by consumers for a forthcoming new product." In other words, it reflects prospective customers' enthusiasm (either positive or negative) toward the new product and their purchase intent (e.g., Xiong and Bharadwaj 2014; Gruen, Osmonbekov and Czaplewski 2006).

Sproull and Arriaga (2007) also point to a third type of online community platforms (e.g., Wikipedia.org, Piazza.com, or fandom.com) that draws on collective intelligence generation. Such platforms empower platform participants' *knowledge collaboration.* Unlike customer buzz from the former two types of platforms, participants' online knowledge collaborative activities place emphasis on "individual acts of offering knowledge to others as well as adding to, recombining, modifying, and integrating knowledge that others have contributed" (Faraj, Jarvenpaa and Majchrzak 2011 p. 1224). The following characteristics distinguish the third type of platform from the others. First, while the other types of online community platforms also play a more or less similar role in knowledge sharing, the third type differentiates itself in terms of knowledge creation through collaboration. Second, in contrast to the creativity thriving platforms which are one-way communication promoting individual ego (Martín-Perpiñá, Poch and Martin 2014; Prasarnphanich and Wagner 2009), the knowledge collaboration platform involves little individual ownership of the co-created content. Nor could participants claim any social reputation, due to the anonymity. Third, motivation-wise, online knowledge collaboration is more altruistic and intrinsic (Prasarnphanich and Wagner 2009), while social thriving platform participation is largely extrinsically motivated. Thus, the expectation is that the content is less biased. For example, Wikipedia's aspiration is to present a neutral viewpoint by asserting facts and opinions about facts (Greenstein, Gu and Zhu 2020). Finally, because of its broadcasting nature, the crowd wisdom induced collective intelligence ensures the content is more beneficial for general information recipients. Thus, they could serve as an important channel via which prospective customers gain knowledge about the new products. A recent case example is the development of Covid-19 relevant pages on Wikipedia.org with 97,088 editors actively participating and generating about 6,950 Wikipedia pages as of December 2020 (see: https://wikimediafoundation.org/covid19/data/). The accuracy of the content, which used to be the greatest concern among academia, was "ensured, thanks to the transparent and broad participation during the editing process. Any misinformation or distrustful sources in those articles were weeded out instantaneously by constant and dynamic monitoring among editors" (Andrews 2020).

LITERATURE REVIEW

Table 2.1 provides a summary of the relevant studies using platform participants' online community activities for the NPS forecast. The review indicates that most studies use customer buzz, and the new product sales forecast is limited to the post-release period (i.e., *post-release NPS* forecast). A limited number of studies have conducted *prerelease NPS* forecast based on customer buzz from social bonding or creativity thriving OC platforms (e.g., Chintagunta, Gopinath and Venkataraman 2010; Xiong and Bharadwaj 2014). However, to the best of my knowledge, little is known about whether and how prerelease knowledge collaborative online community platform activities may help improve the prerelease NPS forecast accuracy. Furthermore, knowledge collaborative activities involve the ebb and flow of the co-created content, reflecting a dynamic process where different knowledge contributors' diverse facts and opinions about facts are exchanged. In the context of the prerelease NPS forecast, such an iterative and dynamic editing and co-creating process strengthens the knowledge contributors' cognitive engagement with the new product. More importantly, the highly dynamic and engaged process entails existing knowledge to be constantly refined and updated. The perspectives of the content are less biased by accommodating different knowledge contributors to participate (Greenstein, Gu and Zhu 2020). Hence, it offers prospective customers (who do not necessarily participate in the knowledge collaborative process but actively read the co-created content) a comprehensive and unbiased picture of the new product's key attributes. It thus provides an easy access for customers to educate and empower themselves to evaluate the to-be-released new product before making purchase decisions (Wagner and Majchrzak 2006; Wasko and Faraj 2000, 2005). I present the theoretical expectations of knowledge collaborative activities and prerelease NPS sales forecast in the next section

Overall, little empirical research has been conducted using the dynamic knowledge collaborative data for prerelease NPS forecast. The closest is Liu, Singh and Srinivasan (2016), who propose an approach to predict the ratings of TV shows based on the Twitter tweet measures and the aggregate Wikipedia pageview data. They, similar to other researchers, do not incorporate the dynamic process of knowledge collaboration in informing the new product prelaunch forecast. Finally, most of the existing research in this field focuses on sales prediction on a short horizon (i.e., opening day or week). As Figure 2.5 of the Appendix 1 Exhibition 2.1, however, a longer horizon (e.g., six-month; +77% of sales volume) with an early forecast could be more meaningful for managers to modify the marketing mix strategy and improve the success of the new product after launch.

THEORETICAL FOUNDATION

Faraj, Jarvenpaa and Majchrzak (2011, p.1235) have used the term "fluidity" to describe the iterative and dynamic process of online knowledge collaborative activities in OC platforms such as Wikipedia. They view it as the differentiating attribute of knowledge collaborative OC platforms compared to the platforms focusing on social bonding and creativity thriving. In this sense, the perpetual incorporation of new information and the growing number of knowledge contributors throughout the new product prerelease period drive the convergence of diverse views from the contributors and the enhanced quality of the co-created content, in terms of content informativeness, content novelty, and perspective diversity. The three traits of the converged knowledge deliver both currency and enhanced relevance to the new product information, enabling prospective customers to undertake information search and make the purchase decision more easily (Zhang and Chang 2020). In the rest of this section, I first clarify people's motivation of online knowledge collaboration, then discuss its role in facilitating prospective customers' information search for new product purchase which in turn links to the prerelease NPS forecast, and eventually argue the necessity of using the dynamic history of the online knowledge collaborative activities for the prerelease NPS forecast.

Online Knowledge Collaboration

Knowledge refers to the "facts, information, and skills acquired through experience or education" (en.oxforddictionaries.com). In the past, researchers mainly considered "knowledge" as a private good held by an individual or organization, and its exchange and development only occurred through iterative one-to-one interaction (Davenport and Prusak 1998; Hansen et al. 1999). To explain the motivation for knowledge collaboration, I adopt the perspective that knowledge is "socially generated, maintained, and exchanged within emergent communities of practice" (Brown and Duguid 1991; Lave and Wenger 1991; Wasko and Faraj 2000 p.156). Because of this trait, knowledge is created through collective contribution and collaboration within a community. Such a collaboration is usually altruistic with no apparent or immediate compensation for the knowledge contributors *per se* (Hars and Ou 2001) Besides, community members are more motivated to share the knowledge (as a public good) with others for the welfare of the entire community instead of some personal interest, and such sharing and transferring within the community can be achieved without depleting the knowledge (Komorita and Parks 1992; Schwartz 1970; von Krogh 1998).

The advent of wikis technology facilitates the tendency to collaborate and share knowledge in an online world. A wiki is a web-based platform that enables knowledge collaborators to create, edit, and update any content posted on that platform (Wagner and Majchrzak 2006). Besides this "open editing" attribute, wiki-based websites feature an "editing preservation" function, which retains all past edits of a given content on the platform for the potential restoration of a previous version (Kane et al. 2009). These two functions of the wiki-based websites, in combination, establish a robust and transparent online knowledge collaborative environment. In the real world, there are a plethora of wiki applications for knowledge collaboration, software co-development, and peer education in both the internet and the intranet settings. For example, en.Wikipedia.org, established in 2001, is so far the best-known public online knowledge collaborative platform, created and maintained by a community of 38.44 million registered users with an average of 20 billion monthly page views worldwide (see: https://stats.wikimedia.org/#/all-projects). The IT department from the Massachusetts Institute of Technology (MIT) builds a wiki platform exclusively open to MIT students, faculty, and staff communities to collaboratively conduct academic research and software development (see https://ist.mit.edu/wikis). Piazza.com is another wiki-based public OC platform that allows students and professors worldwide to locate "the single, high-quality" questions and answers co-created by other members in the community (see https://piazza.com/).

The Relationship between Online Knowledge Collaboration and Prerelease NPS Forecast

Throughout the pre-purchase period, a customer undergoes a transition across multiple stages including need recognition, consideration, and information search based on her interactions with different touchpoints (Lemon and Verhoef 2016). Prior literature suggests that information search is an essential step for customers of either search or experiential products to reduce purchase uncertainty (Stigler 1961; Mudambi and Schuff 2010). The reduction of purchase uncertainty (at least partially) through information search incurs a cost that makes a customer choose one source over another and consider one type of information instead of another. To this end, online knowledge collaborative activities are associated with prerelease NPS forecast by providing customers with high quality content about the new product, which facilitates them to efficiently fulfill the information search stage and better evaluate the benefits of the new product.

Moreover, online platform featuring knowledge collaboration is superior to other social bonding and creativity thriving platforms in reducing customers' search cost. According to Nelson (1970), a customer's search cost has two dimensions: accessibility to the source and the capability of cognitively processing the acquired information from that source. An online knowledge collaboration platform such as Wikipedia is available 24/7 and customers can also easily query topics of their interest using keywords to search. Between customer buzz and the information generated through knowledge collaborative activities, previous studies have found that customers favor objective claims about a product (Ghose and Ipeirotis 2006; Mudambi and Schuff 2010). These studies suggest that customers are influenced less by customer buzz with extreme opinions, as these opinions may be considered as "a simple difference in taste" and disregarded. In contrast, knowledge co-created through knowledge collaborative process is viewed as less segregated, with biased opinions edited out, richer and diagnostic and thus more useful (Greenstein, Gu and Zhu 2020). Park, Shin and Xie (2021) in their recent research further find that the influence of user generated reviews on customers' purchase behavior is extremely skewed toward the first review shown on the top of the webpage. Such a striking power of the first review tremendously weakens the effect of subsequent reviews regardless of volume and valence, rendering an information-availability bias. In contrast, the Wiki collaborative process has the advantage that content created by one collaborator is subject to revision, changes, and even overall negation by others. Such a dynamic and fluid knowledge collaboration process leads to the co-created content constantly being refined, rendering prospective customers a chance to cognitively process the product information more efficiently based on product attributes, manufacturer reputation, etc. Eventually, these customers are more amenable for the purchase decision.

In short, high quality content co-created through the process of knowledge collaborative activities benefits prospective customers' information search for objective information at low search costs during the pre-purchase stage, which further assists them in making a purchase decision.

The Need to Examine the Entire Dynamic History of Online Knowledge Collaboration for Prerelease NPS Forecast

Through the lens of the "fluidity" theory proposed by Faraj, Jarvenpaa and Majchrzak (2011) and the stimulus response theory (Ferster and Skinner 1957), I recognize that the mission of a prospective customer's information search is not a one-off situation. With the old problem resolved, the new one immerges. Thus, it entails the online knowledge collaboration platform, as an information source, to constantly offer high quality content with sufficient amount of information (i.e., content informativeness), up-to-date knowledge (i.e., content novelty), and well-balanced attitudes (i.e., perspective diversity) throughout the entire new product prerelease period. In this regard, the consideration of the entire dynamic history of knowledge collaborative activities allows researchers to better evaluate the quality of the co-created content for the following reasons.

First, if the length of a new product's Wikipedia page (or article length) is constantly increasing and always greater than the average level during the prerelease period, it is more likely that this page contains greater content informativeness. Thus, it facilitates prospective customers to continually access more information and reduce their uncertainty at a lower search cost and greater convenience before making a purchase decision.

Second, what keeps customers constantly checking the co-created content on an online knowledge collaborative platform is also related to the newness and currency of the information. Knowledge contributors' rigorous content refinement (i.e., adding and removing the content) throughout the prerelease period provides content novelty, educating prospective customers about a product's strengths and weaknesses and enabling them to make an informed judgment. As Faraj, Jarvenpaa, and Majchrzak (2011) have suggested, knowledge contributors possessing new information on a given topic ardently strive to have their voices heard in the process of knowledge collaboration. In short, the availability of the new information is the sufficient condition, while the rigorousness of knowledge contributors serves as the necessary condition. Both in combination leads to the ebbs and flows in the article length and maintains the content novelty.

Turning back to the research context of the current study, if the article length of a new product on an online knowledge collaborative platform undergoes rigorous changes (ups and downs) throughout the prerelease period, it signals the co-created content remains up to date, which is beneficial for customers to keep informed with new knowledge from time to time until its official launch. If the rigorous changes of the article length occur mostly at the very beginning of the prerelease period, it implies the article contains less new knowledge when approaching to the launch time. Hence, it hinders customers' motivation to undertake follow-up information search to reduce purchase uncertainty. On the contrary, if the rigorous changes occur mostly at the end of the prerelease period, it ensures the availability of new knowledge during the launch time. However, it deprives customers' opportunity of continual information search and self-education, both of which are essential for purchase decision (Lemon and Verhoef 2016; Mudambi and Schuff 2010).

Third, knowledge collaboration accompanies periodical convergence of diverse perspectives. Instead of a sheer combination of homogenous opinions, convergence refers to a state when diverse ideas are framed, analyzed and merged into an inclusive solution (Isaksen and Treffinger 1985; Woodman, Sawyer and Griffin 1993). In this regard, Faraj, Jarvenpaa and Majckrzak (2011) argue that the ever-growing number of participants (throughout the entire prerelease period) supports convergence in two ways. First, co-created knowledge based on collaboration with a growing number of contributors' participation is more likely to contain different perspectives. Second, knowledge contributors often voluntarily take different roles in the process of knowledge collaboration. When divergent perspectives are either too contradictory or fragmented, a participant may proactively play the moderating role by integrating these perspectives in an organized and coherent manner. Therefore, an increasing and above-than-average number of distinct knowledge contributors throughout the prerelease period ensures the convergence of diverse perspectives, which in turn eases prospective customers' constant information search for new product purchase. Consequently, it augurs the new product sales.

RESEARCH CONTEXT AND EMPIRICAL STRATEGY

I selected the video game industry as the research context of this study for two reasons. First, almost 75% of U.S. households own a device that is used to play video games. More than 64% of adults and 70% of people under 18 years old in the U.S. play video games regularly. Customers spent more than $35.4 billion on video games and other relevant products in the U.S. in 2019, and this number is likely to rise in the next few years (Entertainment Software Association 2020). Given such a large market, video game companies have been releasing new games periodically, to compete for market share, revenue growth, and profit enhancement. Second, previous studies have found that most of the product sales in the entertainment industry (e.g., video games and movies) is realized in the first few weeks or months of the new product's formal release in the market (e.g., Chintagunta, Gopinath and Venkataraman 2010; Xiong and Bharadwaj 2014). Consequently, managers of these products are more motivated to undertake prerelease NPS forecast and better understand the likely performance of their to-be-release products[3]. The video games in my dataset were released by 52 distinct publishers worldwide between 2011 and 2017. Among them, 115 games are novel thus are neither sequels of the previous versions nor based on preexisting movies or TV series. The genres include "action" (11%), "adventure" (19%), "fighting" (7%), "racing" (7%), "role-playing" (16%), "shooter" (13%), "sports" (13%), "strategy" (2%), and "miscellaneous" (13%). I also account for the number of console platforms for each new video game.

Addressing the research questions of prerelease NPS forecast hinges on overcoming a few methodological challenges. First, the outcome variable, *sales volume*, is static and measured

[3] I see a similar sales distribution (see Figure 2.5 of the Exhibition 2.1 of Appendix 1) in my dataset. While the life cycles of those video games vary drastically (up to 104 weeks since launch), over 20% of the total sales volume of these new video games are realized in the first week and more than 77% in the first six months of launch.

at a single point of time. It makes the use of regular time series analysis method unviable, which requires the outcome variable to be time series (Hamilton 2020). Second, some predictors (e.g., dynamic history of knowledge collaborative activities during the prerelease period) are discrete time series data points. If I treat them as separate variables, there will be too many predictors, but too few observations in the end. This data structure limits the use of typical regression approaches due to the over-parameterization problem (Ramsay and Silverman 2008). Third, the extent of online community activities (either customer buzz or knowledge collaboration) could be conditional on product characteristics (e.g., video game genre), marketing campaigns etc. Then the regression method might be appropriate to deal with a two-way interaction. However, it is not efficient to handle multiple higher-level interactions (three- or four-way), as the number of predictors increase. Furthermore, regression requires the researcher to *ex ante* hand-curate the interactions to be included which unlike theory testing is not a requirement for predictive activities such as forecasting (Athey and Imbens 2019; Mullainathan and Spiess 2017).

To address these challenges, I propose a novel modeling process comprised of the Functional Data Analysis (FDA) approach and the ensemble Random Forest method. I employ FDA to capture the dynamic history of knowledge collaborative activities and crystalize them into a few principle component scores to achieve dimensionality reduction (Ramsay and Silver 2007; Xiong and Bharadwaj 2014). I then employ a non-parametric ensemble Random Forest method to look beyond the main effects and investigate all possible interaction effects of emotional customer buzz and cognitively engaging knowledge collaborative activities, with respect to the prerelease NPS forecast accuracy, along with other covariates (or controls).

Data and Measures

I collected sales volume data from VGChartz.com for 300 new video games historically launched in the North American market between 2011 and 2017. A closer look at the sales volume distribution of these new video games shows that over 20% of the sales volume was generated in the opening week and over 77% was achieved within the first six months since official release. For the sake of managerial relevancy and a better examination of the proposed method's predictive power, I used logarithm transformed six-month sales volume of these video games as the dependent variable.

Predictor-wise, I collected data of the dynamic knowledge collaborative activities of these 300 new video games on en.Wikipedia.org during the prerelease period. In addition, I also considered other predictors including: the dynamic evolution of Wikipedia pageview data, the dynamic customer buzz data about the new games, the dynamic evolution of the corporate updates of those video games, the intrinsic product attributes (i.e., video game genres, the number of available console platforms, sequel games or not), and the launching firms' advertising expenditures and PR updates during the prerelease period.

I select Wikipedia's English website (*en.wikipedia.org*) as the OC platform of knowledge collaboration for the following reasons. First, Wikipedia is so far one of the largest knowledge collaborative platforms worldwide. Statistics from en.Wikipedia.org show that there have been more than 6 million articles and almost 1 billion edits made by 38.7 million registered users as of April 2020. On average, more than 18 billion page-views are generated every month (according to Alexa.com). Second, it has an established linkage with the video game industry. A query of 'video game' on *en.wikipedia.org* generated a list of 121,192 related articles. On the one hand, engaged knowledge contributors with expertise and enthusiasm in the video game industry spontaneously work as groups "dedicated to improving Wikipedia's coverage of topics related to video games"[4]. They not only collaboratively create articles for those existing video games, but also cover the games that are yet to be released. On the other hand, prospective customers also proactively go to Wikipedia to learn about those to-be-release video games throughout the entire prerelease period (see Figure 2.6 of the Appendix 1 Exhibition 2.2). I acknowledge that new video games' Wikipedia pages may not be 100% original. Instead, they are built up by integrating facts and facts about opinions from different sources of information. Moreover, it serves as a one-stop platform where customers can efficiently acquire all information required for a purchase decision, which otherwise may not be freely accessible. Furthermore, such an integration is not some simple copy and paste. It is based on knowledge contributor's engaged learning and internalization outcome where one intellectually accommodates new information with her existing knowledge (Hodge, Magolda and Haynes 2009).[5]

I collected the editing history data for each game from the "view history" tab of the corresponding Wikipedia pages. Automatic robotic editing activities have been removed before the modeling analysis.[6] In summary, 52,522 distinct editors made a total of 178,709 edits before the 300 games' formal launch. On average, most of these video games began having their own Wikipedia pages 58 weeks before the launch.[7] Figure 2.1 illustrates the knowledge collaborative

[4] https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Video_games

[5] The researcher painstakingly examined information sources of all 300 new video games' Wikipedia pages during their prerelease period. In general, knowledge contributors collect information from four main areas: launching firm's official website or PR news release, launching firms' social media public pages (e.g., Facebook, Twitter, etc.), professional reviews, as well as other media coverages. Information collected from all these sources are not simply copied and pasted to Wikipedia. Instead, the entire knowledge collaboration community closely monitor how information from these sources are added to Wikipedia. To take video game Tokyo Xanadu (launched on June 30th, 2017) for example, one knowledge contributor entirely reversed another's 'contribution' (1,224 characters) on June 22nd, 2017 and commented "you can't just copy/paste entire Gematsu articles/ press releases." For more detailed information on Wikipedia page sources, please refer to Figure 2.7 of Exhibition 2.3 of Appendix 1.

[6] On Wikipedia.org, if an edit is made by automatic robots, it will be labeled as "Robot." Based on this information, I remove all editing activities with a "Robot" label.

7 There is one video game, *Metal Gear Rising: Revengeance*, whose Wikipedia page was created on 6/1/2009, following an official preannouncement at the Electronic Entertainment Expo 2009. However, the video game was

activities for those new video games in my dataset throughout their prerelease period. The top graph of Figure 2.1 shows the weekly fluctuation of the article length of each new game's Wikipedia page (measured by the logarithm-transformed number of characters).[8] The middle graph reflects the number of characters added or removed for each video game's Wikipedia page (i.e., editing activeness) on the weekly basis, and the bottom graph is the trajectories of the number of distinct knowledge contributors (logarithm-transformed) throughout those new games' prerelease period.

Predictor Measurement

In line with prior literature (Kimmons 2015; Lih 2004) and my discussion in previous section, I measure the quality of the content co-created by means of knowledge collaboration through three distinct dimensions: content informativeness (*Article Length*); content novelty (*Editing Rigor*), and perspective diversity (*Diversity*).

Article Length is a time-varying variable measured by the total number of characters of the Wikipedia page for a new video game *i* in week *t*. *Editing Rigor* is a time-varying variable measured by the number of characters added to or removed from a new video game *i*'s Wikipedia article in week *t*. *Diversity* is a time-varying variable measuring the number of distinct editors contributing to the Wikipedia page build-up for a video game *i* in week *t* (Lih 2004).

Figure 2.2 presents six new video games (two on each panel) with quite different trends and bursts in their knowledge collaborative histories on Wikipedai.org (i.e., in the length of the article, the changes of the articles length, and the number of distinct knowledge contributors for the article) throughout the prerelease period. If I merely use aggregated measures, games on each

---

not officially released until 194 weeks later (i.e., 2/19/2013). In the final modeling process, I truncated the length of the editing histories for all 300 video games to 58 weeks prior to launch (the mean value) to both keep the essential information of the dynamic knowledge collaborative activities and constrain any possible interferences due to outliers.

[8] I log-transformed two Wikipedia measures to deal with the non-negative distribution of the raw data and make them a better fit for the FDA process.

panel have very similar values. However, an aggregated measure fails to capture the heterogeneity over time and thus results in information loss (Chintagunta and Lee 2012; Xiong and Bharadwaj 2014).

To enhance the predictability of those knowledge collaborative measures, I also include a list of control variables. First, I considered the full history of pageview data (FDA-processed) for each of the 300 new videos games' Wikipedia pages. Second, I created "specialist %" to account for the expertise of those knowledge contributors when contributing knowledge to a new video game genres. Third, I included the dynamic trend of positive customer buzz of those new video games during their prerelease period. To do so, I selected NeoGaf, an online gaming forum, to collect customer buzz data about those sampled new video games before their release.[9] I collected 193,930 unique posts. I then trained a Recurrent-Neural-Network-based Gated Recurrent Unit (GRU) model based on a prelabeled IMDb dataset[10] [11] [12], and applied the trained model to estimate the valence of each post by labeling them as positive or non-positive.[13] Eventually, I followed Xiong and Bharadwaj (2014) by employing the FDA and fPCA methods

---

[9] Founded in 1999, NeoGaf is considered as one of the largest online gaming forums in the world (according to Alexa.com). As of April 2020, there have been more than 143,000 registered users on the forum, who generated more than 878,096 threads and 130 million video game-related posts.

[10] Please refer to the Exhibition 2.4 of Appendix 1 for more detailed discussion of my selection of the RNN-GRU model based on its comparative advantages against other natural language processing methods.

[11] This prelabeled dataset contains 25,000 highly polarized (i.e., either positive or negative) movie reviews as the training set, and another 25,000 highly polarized reviews as test set. Researchers Maas and colleagues (2011) first used it and offered it on their website (http://ai.stanford.edu/~amaas/data/sentiment/). Given the large number of my posts, a 10% sample is 19,393 posts. If I use human raters for coding, it will greatly slow down the efficiency. However, I did conduct a robustness check for the accuracy of my sentiment analysis based on the pre-labeled IMDb dataset. The detailed process and result can be found in Appendix 1.

[12] I also evaluated the Sentiment 140 dataset containing 1.6 million prelabeled Twitter tweets (positive versus negative) provided by Go, Bhayani and Huang (2009) on http://help.sentiment140.com/home. I initially used that dataset for the GRU model training. However, the result is not satisfactory compared with the one based on the IMDb dataset. I assume this is due to the length of a Twitter tweet, which is constrained to 140 characters. Relative to it, the average number of words of my collected posts is 236.83. Thus, it may contain a more complex idea. At this point, the movie reviews from the IMDb dataset are on average longer and contain more complex information for the model training.

[13] I followed Chintagunta, Gopinath and Venkataraman (2010) by examining the valence (i.e., positive) of customer buzz and its predicting power in new product sales. Details of the GRU modeling procedure are provided in Appendix 1.

to recover and decompose the dynamic trend of customers' positive prerelease buzz and used the crystalized fPCA scores as the control.

Finally, I also included a number of other control variables to robustly examine the prediction power of the dynamic knowledge collaborative measures in prerelease NPS forecast (Chintagunta, Gopinath, and Venkataraman 2010; Zhu and Zhang 2010; Xiong and Bharadwaj 2014). First, I measured firms' marketing activities with their advertising expenditures and the launching firms' PR updates (FDA-processed) on each new video game before the formal release. Second, I set game genre dummies to control for their potential influence on the prerelease NPS forecast. All information about the game genre has been cross verified on the major video game sites, including VGChartz, NeoGaf, IGN, and GameSpot. Third, I created a dummy variable to accommodate the potential effect of the sequel games. Fourth, I controlled the total number of available console platforms for the new video games, because the more platforms that are available for the new game, the broader its potential customer base will be. Finally, I also set up year fixed effects to absorb the influence from the macro-market.[14] Table 2.2 summarizes all these variables and the description of their measures.

METHOD

Functional Data Analysis (FDA)

A significant methodological challenge in examining the effect of the knowledge collaboration dynamics on the prerelease NPS forecast is that the predictors are discrete time-series data capturing all the iterative and dynamic knowledge collaborative activities, while the dependent variable is cross-sectional (i.e., a point estimate) capturing the sales volume. In this

[14] The time of the year at which a new video game is launched may potentially influence its opening sales, too. For example, the holiday season (mid-November, December, and early January) is often considered as peak shopping season. However, the holiday season dummy does not significantly improve the predictive power and is thus dropped in the final model.

study, I observed knowledge collaborative activities on en.Wikipedia.org for the 300 new video games occurred 58 weeks prior to the formal launch but the dependent variable is a single-vector of 300 video games' sales volume six-months post launch (accounting for more than 77% of the total sales volume). This results in the three predictors forming a $300 \times 58 \times 3$ matrix plus other controls to be considered while the dependent variable is merely a 300×1 vector, which leads to over-parameterization.

I solve the over-parameterization challenge through dimensionality reduction. To achieve that goal, I employ the functional data analysis method (FDA). FDA is able to capture the trajectory (both trend and bursts) in discrete time-series data (regardless of the density or sparsity of the data distribution) by extracting the underlying functional object in a nonparametric way and reduce dimensionality in this process (Ramsay and Silverman 2005). This enables the use of a single function as the base unit for the follow-up analysis rather than wrestling with the volume of temporally discrete data points.

Compared with traditional regression and time series analysis models, FDA has three advantages. First, as a non-parametric approach, it does not require any presumed parameter distribution specifications (i.e., mean and variance) (Ramsay and Silverman 2005; Xiong and Bharadwaj 2014). Figure 2.1 illustrates the distribution of knowledge collaborative activities measured by *Article Length*, *Editing Rigor*, and *Diversity* over time in my data. None of them follows any predefined distribution. Second, FDA can directly fit either stationary or non-stationary time series data (Ramsay and Dalzell 1991). For example, Wang, Jank and Shmueli (2008) observed that customers' bidding activities on eBay.com tend to grow over time and peak at the end of the auction period. Such a growth trend is not stationary and violates the requirement of the regular time series analysis. As a result, they used FDA to fit this non-stationary process and applied the results to predicting future bidding price. I observe similar trajectories in the knowledge collaborative process on Wikipedia. As shown by the solid line (i.e., the mean trend) shown in Figure 2.3, the knowledge collaborative activities in general grew more drastically when it approaches the launch time. Third, the non-parametric FDA approach does not assume the uniform distribution of the occurrence of any events in a time series. This is critically beneficial in dealing with knowledge collaborative (or customer buzz) data. Imagine, a few hundred people may collaboratively contribute knowledge of a new product on one day, but abruptly pause any editing activities for a few days or weeks due to a temporary opinion convergence. This leads to the time series of knowledge collaborative activities to be extremely sparse during some periods, but very dense in other periods. This data generation process rules out traditional time series methods that assume that the probability of all time-events (i.e., *Article Length*, *Editing Rigor*, and *Diversity* measures in this study) is uniformly distributed (Hamilton 2020; Jank and Shmueli 2006).

Recovering the Underlying Curves of Customers' Dynamic Knowledge Collaboration

The first step of FDA is to recover the underlying function of a curve, an image, or a complex object moving through both space and time. According to Ramsay and Silverman (2007), if the discrete raw data are errorless, it can be achieved by interpolation. Otherwise, researchers in the past have suggested the smoothing method to accommodate those random and measurement errors with the raw data (Fan et al. 2014; Ramsay and Silverman 2007; Shmueli, Russo and Jank 2004; Sood, James and Tellis, 2009).

For this study, I implemented the penalized smoothing splines to recover the underlying curves of customers' dynamic knowledge collaboration: *Article Length* ($A_i(T_t)$), *Editing Rigor* ($E_i(T_t)$), and *Diversity* ($D_i(T_t)$) for each new video game i ($i = 1, 2, \ldots, I; I = 300$), where t

denotes the period from 58 weeks until one week before the formal launch. To take the *Article Length* ($A_i(T_t)$) variable for illustration, my goal is to identify a continuous and smooth function $f_i(\cdot)$ to fit the interval between two discrete time-series data points for each video game. However, a highly smooth curve may lead to the over-fitting of the raw data. Following Ramsay and Silverman (2005), I included a penalty component to balance between the roughness and the smoothness of the fitted curve:

$$PENSS_i = \sum_{t=-58}^{57} \left(A_i(T_t) - f_i(T_t)\right)^2 + \lambda \int [D^2 f_i(T_t)]^2 dt$$ Equation [1]

, where T_t refers to the t^{th} week in the prerelease period ($t = -58, -57, \ldots, -1$). $A_i(T_t)$ denotes the *Article Length* of game "i" in week "*t*". $f_i(T_t)$ is the non-parametric smoothing function. $\left(A_i(T_t) - f_i(T_t)\right)^2$ measures the fit of the function. $[D^2 f_i(T_t)]^2$ refers to the squared term of the second-order derivative of the function $f_i(T_t)$ at time "*t*", which is also called "curvature." The second-order derivative of a straight line is zero, signifying non-curvature. Therefore, $\int [D^2 f_i(T_t)]^2 dt$, the integral of the curvature, is a "natural measure" of a functional curve's roughness (Ramsay and Silverman 2005). Finally, a smoothing parameter λ ranging from 0 to ∞ is evaluated to tradeoff between the fit and the roughness. Specifically, when λ approaches to 0, the functional curve achieves the highest smoothness by fitting all the raw data points. When λ is larger, the functional curve will tend to be rougher. I used cross-validation to identify the value of λ by following prior literature (Ramsay and Silverman 2007). Figure 2.1 plots the FDA-recovered curves of *Article Length*, *Editing Rigor* and *Diversity* for each game in my data.

Dimensionality Reduction through fPCA

Based on the recovered functional curves in the previous step, I conducted the functional principal component analysis (fPCA) to further capture the essential features (patterns) of those curves for *Article Length*, *Editing Rigor*, and *Diversity* measures. Similar to the regular principal component analysis, the objective of the fPCA is to generate principal component scores. Usually, the top few scores are good enough to capture most features of the curves (See: Hyde, Moore and Hodge 2004; Ramsay and Silverman 2005, 2007). Therefore, I further reduce the dimensionalities of the two measures before the final predictive analysis.

To take the *Article Length* measure for illustration, the recovered functional curve $A_i(T_t)$ can be further decomposed by fPCA as:

$$A_i(T_t) = \mu(T_t) + \sum c_{ik}\theta_k(T_t) \quad \text{Equation [2]}$$

, where $\mu(T_t)$ refers to the mean curve of the measure *Article Length* of the overall 300 new video games. $\theta_k(T_t)$ is the k^{th} functional principal component (FPC), revealing the variability in the functional curves from the mean. In this illustration, each $\theta_k(T_t)$ presents the distinguishing pattern of the variations among the *Article Length* curves. The estimation of $\theta_k(T_t)$ is subject to the orthogonality constraints where $\int \theta_k(T_t)^2 dT = 1$ and $\int \theta_k(T_t)\theta_n(T_t)\, dT = 0 \; \forall \; k \neq n$ and $k = 1, 2, 3, ...$ By definition, the first $\theta_k(T_t)$ captures the largest variability, followed by the second. Usually, the top three or four are able to identify the key patterns of the curve (Hyde, Moore and Hodge 2004; Ramsay and Silverman 2005, 2007). c_{ik} measures the degree that the key pattern of each new video game's *Article Length* curve $A_i(T_t)$ may deviate from the one described by the k^{th} component $\theta_k(T_t)$. If c_{ik} is positive (negative), that shape of that game's *Article Length* curve $A_i(T_t)$ will change in the same (opposite) direction denoted by $\theta_k(T_t)$. In sum, $\mu(T_t) + \sum c_{ik}\theta_k(T_t)$ reflects both the mean trend across all the games and the heterogeneity of each game.

Panel A, B, and C of Figure 2.3 present the first four FPC scores (*hereafter* FPCS) of the dynamic knowledge collaborative activities measured by *Article Length*, *Editing Rigor*, and *Diversity* respectively. The solid curve in each graph denotes the mean curve $\mu(T_t)$ for each measure. The + + + $(- - -)$ captures the variances of each video game in terms of its deviation from the mean curve measured by the three variables.

For *Article Length*, the top four FPC scores collectively account for 94.5% of the variance for all the 300 sampled video games. If a video game's first FPC score is positive, according to the graph, the trend of the video game's Wikipedia page length was ever growing and the sheer volume was always longer than the average level throughout the prerelease period. If a video game had a positive second FPC score, the length of its Wikipedia page was usually lower (higher) than the average at the beginning (-58 to -40 weeks prior to launch) but rose tremendously in the middle of the prerelease period (approximately starting from the 40^{th} week before launch) and eventually surpassed the average level. I refrain from offering any interpretation of the third and the fourth FPC scores for *Article Length*, considering their trivial contribution to the overall trend (lower than 7% in combination).

In terms of *Editing Rigor*, its top four FPC scores collectively capture 52.6% of the overall variances of each video game. If the first FPC score is positive, it describes a situation when there were rigorous changes with the co-created content starting from the mid of the prerelease period. In the similar vein, the fourth FPC score corresponds to a tendency where there was a greater degree of activeness in the editing of Wikipedia content (either addition or deletion) throughout the video game's prerelease period. The second FPC score shows that there were more changes with a new video game's Wikipedia page length at the beginning of its prelease period. Correspondingly, the third FPC score shows that there were more changes in the middle and the end of a video game's prerelease period, while the rest of this period is fairly "quiet" with little new content at all.

For *Diversity*, the top four FPC scores for this measure explain 74.5% of the variance. The expectation is that the higher the number of distinct contributors participating in the co-creation of a video game's Wikipedia page, the more diverse perspectives that Wikipedia page may embody. If a video game's first FPC score is positive (relative to negative), according to the graph, that video game had attracted a higher-than-average number of distinct content contributors for its Wiki article throughout the prerelease period. If the video game's second FPC score is positive (relative to negative), then it might not have attracted a lot of content contributors at the beginning (58 to 30 weeks before launch), but this number began to grow and eventually outperformed the average level when the video game was close to launch. Similar to *Article Length*, I consider the third and the fourth FPC scores too trivial and avoid presenting any possibly misleading interpretation (lower than 12% of contribution in combination).

Predictive Analysis Based on Ensemble Random Forest Method

Following the FDA analysis, I adopted a machine-learning-based ensemble Random Forecast (RF) approach to examine how well the addition of knowledge collaborative activities (along with other covariates) may contribute to the forecasting accuracy enhancement. Unlike regular regression method, one comparative advantage of the RF approach lies in its non-parametric nature, which has no *a priori* assumption of the data distribution and renders substantial flexibility for analysts to dig out information purely from the data (Elgammal, Harwood and Davis 2000). Besides, the ensemble approach (in particular "bagging" or “boosting”) improves forecast accuracy by leveraging the collective predictive powers of modeling results based on different sub-datasets (Shmueli and Koppius 2011). Another advantage of the RF approach is that it does not assume certain functional relationship between the dependent variable (DV) and the predictors (see Wooldridge 2010, p. 35 - 43). Instead, it automatically detects and models the best DV-predictor relationship, with the consideration of both additive and multiplicative combinations of the predictors, such as higher-order interactions (Breiman 2001; Lunetta et al. 2004).

Tracing its origins from the regression tree method developed by Breiman et al. (1984), the ensemble Random Forecast approach outperforms the single tree method by overcoming the overfitting issue (See: Strobl, Malley and Tutz 2009 and Appendix 1 Exhibition 2.5 for regression tree method description). Moreover, a minor change in the predicting variables will lead to extremely different estimating results and predicting errors (See: Lewis 2000; Segal 2004). One solution is to build multiple ‘trees’ to offset the biasness from one single tree. Such an ensemble approach is labelled as Random Forest (RF) (See: Breiman 2001, 2003). Instead of building one tree, Breiman (2001) has suggested to grow multiple trees to make dependent variable estimations and take the mean value of all these estimation results to report. Such an ensemble approach is more beneficial to reduce the variance of the results and overcome the potential overfitting issue from a single tree (See: Breiman 2001, 2003; Gislason, Benediktsson and Sveinsson 2006)

I considered two types of ensemble RF models. First, a “bagging” approach, which is a bootstrapping approach first proposed by Breiman (1994). The bagging approach constructs multiple trees with bootstrapped sub-sample. Specifically, one can randomly draw a subsample from the entire dataset and use it to build a regression tree. Repeating this procedure, one gets multiple trees, each based on a randomly sampled sub-dataset. When given a new set of independent variables, one takes the average of the estimators from each tree as the final estimator.

An alternative approach for ensemble RF is the "boosting" approach. Different from the "bagging" method where each trained tree is independent of one another, the boosting method entails each tree to be built based on the learnings from the previous tree (See: Freund and Schapire 1996; Schapire 1999). More specifically, during each iteration, the following tree for further examination chooses the incorrectly predicted incidences by the previous tree more often. It attempts to identify new trees that perform better than the previous ones in dealing with those "hard" incidences (Freund and Schapire 1996; Schapire et al. 1998).

The main difference between the two approaches lies in the fact that the distribution of each bootstrapped subsample in the "bagging" procedure is independent and identically distributed (*i.i.d.*), while the "boosting" method adaptively changes the distribution of the dataset based on the result of the previous tree. Past studies have found that both methods are effective in dealing with prediction bias and lowering the variance. However, "boosting" performs better for noise-free data, while "bagging" is better in handling "noisy" situations (Bauer and Kohavi 1999; Breiman 2001; Rodriguez-Galiano, Ghimire and Rogan 2012). In this study, I use both for the prerelease NPS forecast.

Regression-based Model for Further Interpretation of the Result

Despite the comparative advantage in its predictive power, results based on the non-parametric ensemble RF model is lack of interpretability to uncover the relationship between customers' knowledge collaborative activities and the pre-release NPS forecast. To overcome this barrier, I further specify the following regression model below:

$$y_i = \beta_0 + \beta(A_i(T_t), E_i(T_t), D_i(T_t)) + \boldsymbol{\varphi Other\ Controls_{i.t}} + \epsilon_i \qquad \text{Equation [3]}$$

, where $A_i(T_t)$, $E_i(T_t)$, and $D_i(T_t)$ respectively denote the top four FPC scores of each of the three knowledge collaborative measures: *Article Length*, *Editing Rigor*, and *Diversity* for a specific new video game 'i'. $\boldsymbol{Other\ Controls_{i.t}}$ is a vector all other time-varying and time-invariant control variables discussed in the earlier section. y_i denotes the six-month sales volume of those 300 games, and ϵ_i is the error term.

RESULTS

Recall that I am interested in the role of the dynamics (i.e., trend and bursts) of knowledge collaborative activities during a new product's prerelease period in predicting the sales volume of that new video game after launch. Based on the regression results presented in Table 2.3, all three knowledge collaboration measures are significant, with respect to new product sales forecast (i.e., six-month post-release). *Article Length* denotes the informativeness of each video game's Wikipedia page. Based on the full regression model results, video games with longer than average Wikipedia articles throughout the prerelease period will have lower sales volume ($\beta = -0.010\ p < .001$). This result is in contrary with my expectation presented earlier. I will discuss and offer my explanation in the discussion section.

Editing Rigor measures the co-created content novelty. The expectation is that the constant maintenance of content novelty during the prerelease period meets customers' need to conduct constant information search for new knowledge about the forthcoming product, thus influencing sales. The model shows that both its first ($\beta = 0.069\ p < .05$) and fourth ($\beta = 0.160\ p < .01$) FPC scores are significant and positive for new video game sales volume forecast. Linking the regression coefficient results and the patterns of its FPC scores in Panel B of Figure 2.3, I find that both the editing activeness and the timing of content changes (i.e., recency) are critical. As the launch time approaches, if there are more content changes (relative to the average level), I could expect to see a higher sales volume for the new video game. On the contrary, early content change (e.g., FPCS 2) or late and abrupt content change (e.g., FPCS 3) do not significantly possess any prediction power. *Diversity* reflects different opinions the co-created content embodies during the prerelease period. Hereby, the first FPC score of this measure ($\beta = 0.037\ p < .001$) is significant and positive in predicting a new video game's sales volume. Based on the pattern of its first FPC score in Panel C of Figure 2.3, I consider that an above-than-average level of distinct knowledge contributors participating in the knowledge collaborative process throughout the prerelease period is positively associated with the sales volume increase of a new video game. In addition, the second FPC score of the *Diversity* measure is also significant and positive ($\beta = 0.019\ p < .1$). The results imply that there could be again a recency effect. Examining the illustration of Panel C of Figure 2.3, I find that if a video game attracted more knowledge contributors (relative to the mean level) suggesting better quality of information available to prospective consumers, starting from approximately the 28^{th} week prior to launch, its sales volume also tends to be higher.

Consistent with my expectation, I also find that the trend of Wikipedia pageview, the trend of positive customer buzz, launching firm's marketing activities (i.e., advertising expenditure and corporate updates of the new product), pre-order volume, and video game attributes (console platform available, sequel or not, genre) all serve as significant predictors of a new video game's sales volume.

Forecast Accuracy Improvement with the Inclusion of Knowledge Collaborative Measures

To illustrate the benefit of considering knowledge collaborative activities in prerelease NPS forecast, I compare alternative model methods and present the results in Table 2.4 and 2.5.

For the regression method, Model 1 is based on game characteristics predictors only. Model 2 consists of game characteristics, launching firm's marketing activities, and customer buzz predictors. For Model 3, I have two versions. Model 3 – a is built upon predictors of game characteristics and the aggregated knowledge collaboration measures. Model 3 – b is based on predictors of characteristics and the dynamic knowledge collaboration measures[15]. Finally, the full model considers all predictors. As expected, the inclusion of dynamic knowledge collaborative measures significantly improves the R^2 by 20.1% (Model 3 – b vs. Model 2) and 20.7% (full model vs. Model 2). Table 2.4 presents the *F*-test results with respect to the adjusted R^2 comparisons between each regression model (i.e., Model 1, Model 2, and Model 3 – a / b) and the full model. For the RF models, I did not *a priori* set up different versions (i.e., predictor sets) for comparison. Instead, I let the RF algorithm determine which predictors and/or their combinations may usher into the best result. As the Figure 2.4 shows, almost all knowledge collaborative measures are among the top contributors to the improvement of the prediction accuracy.

To examine how much the inclusion of knowledge collaborative measures improves the prerelease NPS forecast accuracy, I calculated the Mean Absolute Percentage Error (MAPE) and the Mean Squared Error (MSE) of models based on the regression method and that based on the ensemble RF method. Specifically, I followed Foutz and Jank (2010) by holding out one video game *i* at a time and train the regression and the RF models based on observations of the remaining 299 video games. Then, I applied the model results to predicting the six-month sales volume of the video game *i* and calculate both the MAPE and MSE based on that prediction. Next, I repeated the process for the rest of the 299 video games one at a time. Finally, I

[15] Both Model 3 versions also include dynamic pageview measures and the percentage of specialists among knowledge contributors of those co-created Wikipedia articles throughout the prerelease period as control.

calculated the mean MAPE and MSE across all regression and RF models to evaluate their forecasting accuracy. Table 2.5 summarizes the prediction accuracy of these models based on MAPE and MSE. In sum, the "bagging" ensemble RF model outperforms all other model options ($MAPE = 5.99\%, MSE = 0.88$).

Prerelease NPS Early Forecast based on Partial History of Knowledge Collaboration

From the managerial perspective, an accurate new product sales forecast is critical. What is more valuable, however, is how early the accurate sales forecast is available. It gives managers more flexibility to adjust their marketing strategy to cope with the demand potential. In this regard, I examined the forecast accuracy of the proposed ensemble RF model with partial history of knowledge collaborative activities (10 weeks and 5 weeks prior to formal launch) and compared it with regression model without knowledge collaborative measures (i.e., Model 2). To take the 10-week prior to formal launch early forecast for example, I only included knowledge collaborative activities as well as other time-varying activities (i.e., Wikipedia pageview, advertising spends, firms' corporate updates, and positive customer buzz) accrued from the 58^{th} week until the 11^{th} week before launch, with all time-invariant factors unchanged (e.g., game genre, sequel, etc.). Next, I undertook the FDA and fPCA analyses to extract their FPC scores for the relevant predictors. Then, I followed the same steps summarized in the earlier subsections (i.e., leave one out approach) and evaluated the mean MAPE and MSE metrics based on both the regression and the ensemble RF methods.

As presented in Table 2.6, it is not surprising to see that early forecasts based on the ensemble RF model with closer-to-launch data (i.e., 5 weeks prior to launch) perform better than those based on data 10 weeks prior. However, the ensemble bagging RF model based on the partial history of knowledge collaborative activities (10 weeks before launch) still outperforms the regression model that does not include knowledge collaborative measures (i.e., Model 2), but with the full-time-range data. In short, my proposed modeling approach is superior for both full-time-range and early forecast, which allows managers to execute their marketing plan and potentially optimize the new product sales performance.

ROBUSTNESS CHECK

To further examine the utility of considering the dynamic knowledge collaborative activities in the prerelease NPS forecast and the necessity of applying the FDA method to pre-process such dynamicity, I conducted several robustness checks, by exploring alternative methods to restore the trajectories of the knowledge collaborative history. Briefly, I employed three approaches. First, I regressed the weekly volume of the three main Wikipedia measures respectively over a sequential number from 1 through 58 (constraining the intercept as zero), in order to approximate the linear trend on how the knowledge collaboration may evolve. I then replaced all the FDA-preprocessed variables with their aggregated volumes along with their linear trends measured by the coefficients of the sequential number. Next, I followed the previous steps to calculate the mean MAPE and MSE metrics. The second approach originates from the restoration of the recency effects of those knowledge collaborative measures. I tried using the aggregated volumes of those variables only four weeks before new product launch or using the difference between the last week (before launch) volume and the first week volume of those aforementioned measures. Then, I calculated their mean MAPE and MSE metrics. Table 2.7 summarizes the results in terms of all these robustness checks. In sum, the use of FDA approach is necessary to restore the dynamics of the knowledge collaboration and enhance forecast accuracy for all model methods (both the regression and the ensemble RF).

DISCUSSION AND CONCLUSION

The focus of this study is on the effective use of dynamic online knowledge collaborative activities in the prerelease NPS forecast. The co-created content through such a dynamic collaboration is novel and useful. It benefits prospective customers' access to the content from time to time to undertake information search. Eventually it facilitates those prospective customers to make purchase decision, supporting the use of the dynamic knowledge collaborative activities in the prerelease NPS forecast. The proposed modeling process based on the FDA and the ensemble Random Forest techniques efficiently captures such dynamics and improves the forecast accuracy by leveraging it and other relevant factors (e.g., customer buzz, firm's marketing activities, new product attributes, etc.). The finding shows that the inclusion of the dynamic knowledge collaboration variables significantly improves the accuracy of the prerelease NPS forecast. Moreover, the superiority of the proposed method in terms of forecasting accuracy still holds when the forecast is conducted 10 weeks prior to the final release. Thus, it offers managers plenty of time and flexibility to undertake changes to the launch marketing strategy.

In the study, the quality of the co-created content through this dynamic and active knowledge collaboration is measured through three distinctive dimensions: content informativeness (*Article Length*), content novelty (*Editing Rigor*), and content diversity (*Diversity*). Based on the final results and the patterns of each variable's dynamic trajectory, I learn that video games with longer Wikipedia page during the prerelease period (relative to the mean value) tend to have lower post-release sales volume. This is against what my prior belief. However, a careful survey of previous studies suggests that too much information presented in a 'lengthy' article may increase customers' cognitive load, deteriorate their comprehension, and hinder the psychological excitement which is supposed to be realized through video game consumption (Johnson and Rosenbaum 2015; Mousavi, Low and Sweller 1995; Sweller 2011; Zillmann, Hay and Bryant 1975).

When knowledge contributors are more active in editing the co-created content (either adding new information, remove existing information, or both), the forecasting accuracy improves. Such a link is more salient when the editing activeness occurs starting from the middle or throughout the entire prerelease period. Finally, when there are more knowledge contributors participating in the co-creating process throughout the entire prerelease period or starting from the middle of the period, the NPS forecast accuracy gets better enhanced.

This study provides novel implications for marketing practitioners. First, it demonstrates the value of using the dynamic knowledge collaborative activities in forecasting the prerelease NPS accurately, a valuable and free data source that has not been capitalized on in the past. Second, the modeling process that I propose efficiently handles the high dimensionality of the data and comprehensively accounts for both main and higher order interaction effects during the predictive modeling stage. Third, the marketing department has the predominant responsibility for new product forecasting (Kahn 2002). The superior forecast accuracy of the proposed approach, especially the early forecast, offers managers both time and flexibility to make appropriate modifications to the new product launch marketing strategy, resulting in the improvement of the new product performance.

Nevertheless, I acknowledge that this research has a few limitations that future studies may focus on. First, while (1) the dynamic knowledge collaborative activities precede new product sales temporally and (2) I account for other potentially alternative explanations by using control variables, the study does not make any causal claim. Second, my research context is the video game industry. As an experiential product, customers of video games often face high purchase uncertainty and require additional information to evaluate the benefits of the new product before purchase. However, I do not know whether the online dynamic knowledge collaborative activities help improve the prerelease NPS forecast accuracy for other product categories with more utilitarian characteristics. Third, the data I extracted for this study is from the en.Wikipedia.org. However, there are other knowledge collaborative OC platforms available on the internet dedicated to video game (e.g., fandom.com, xinglugu.huijiwiki.com/wiki, etc.). Similar to Wikipedia.org, these platforms are also maintained by video game enthusiasts. Future research could examine whether the dynamic knowledge collaborative activities on these platforms can enhance NPS forecasting accuracy.

TABLES

Table 2.1
Literature Review of Representative Studies

Studies	OC Platform Types as Data Source	Knowledge Collaboration Data Included	Dynamic Knowledge Collaboration Process Accounted	Dynamic Customer Buzz	Pre/Post-NPS Forecast	Research Context
Asur and Huberman (2010)	1 & 2	No	No	No	Post	The impact of Twitter tweets on movies' box office performance
Chevalier and Mayzlin (2006)	1 & 2	No	No	Yes	Pre	Impact of customer reviews (both volume and stars) on book sales rank
Chintagunta, Gopinath and Venkataraman (2010)	1 & 2	No	No	No	Pre	The effect of prerelease movie review volume, rating scores, and rating score variances on the movies' opening week box office performance.
Craig, Greene and Versaci (2015)	1 & 2	No	No	No	Post	Viewership of online comments and their intent to watch a movie vs. the launch week box office performance of the movie
Cui, Lui and Guo (2012)	1 & 2	No	No	Yes	Post	The link between customer review volume and viewership and new products' post-release sales.
Dellarocas, Zhang and Awad (2007)	1 & 2	No	No	No	Post	Customer prerelease eWOM's effect on movies' box office performance.
Dewan and Ramaprasad (2014)	1 & 2	No	No	No	Post	Customer buzz and music album (or song) sales
Duan, Gu and Whinston (2008)	1 & 2	No	No	Yes	Post	Impact of movie review volume and valence on movies' box office performance

Studies	OC Platform Types as Data Source	Knowledge Collaboration Data Included	Dynamic Knowledge Collaboration Process Accounted	Dynamic Customer Buzz	Pre/Post-NPS Forecast	Research Context
Godes and Mayzlin (2004)	1 & 2	No	No	No	Post	Volume and variance of customer buzz and their impact on TV show ratings
Liu, Singh, and Srinivasan (2016)	1 & 3	No[a]	No	Yes	Post	The impact of both customer buzz and Wikipedia pageview on TV show ratings.
Mestyán, Yasseri and Kertész (2013)	1 & 3	Yes	No	No	Post	The link between cumulative Wikipedia editing activities and pageview and the launch week movie box office performance.
Xiong and Bharadwaj (2014)	1 & 2	No	No	Yes	Pre	The evolution of customer buzz trajectories and new video games' sales volume prediction.
This Study	***1 & 2 & 3***	***Yes***	***Yes***	***Yes***	***Pre***	***The evolution of online knowledge collaborative activities (along with other predictors) and new video games' sales volume prediction.***

Note: 1 – social bonding; 2 – creativity thriving; 3 – knowledge collaboration; a. Liu, Singh and Srinivasan (2016) only included Wikipedia pageview data in their final model.

Table 2.2
Key Variables Measurement

Variable Type	Variable Name	Data Source	Measurement
Dependent Variables	Sales	VGChartz.com	Total number of orders (*log-transformed*) a new video game '*i*' received within the first six months since its formal launch
Knowledge Collaboration Predictors	Content Informativeness (i.e., Article Length)	Wikipedia.org	Time-varying variable measuring the length of the Wikipedia page (*in characters and log-transformed*) for the game '*i*', since the page was initially created till day '*t*' when the video game was launched.
	Content Novelty (i.e., Editing Rigor)		Time-varying variable measuring the length of the Wikipedia page (*in character*) for the game '*i*', since the page was initially created till day '*t*' when the video game was launched.
	Perspective Balance (i.e., Diversity)		Time-varying variable measuring the number of distinct editors (*log-transformed*) who made edits on the Wikipedia page of a new video game '*i*', since the page was initially created till day '*t*' when the video game was launched.
	Specialist %		The percentage of knowledge contributors who had been dedicated to merely one game genre as of the time she participated in the knowledge co-creation for the current video game '*i*'.
Other Predictors	Pageview	Wikipedia.org	Time-varying variable measuring the number of pageviews (*log-transformed*) for each new video game '*i*' throughout the prerelease period.
	Positive Customer Buzz	NeoGaf.com	Time-varying variable measuring the number of sentimentally positive posts (*log-transformed*) for a video game '*i*' throughout its prerelease period.
	Advertising Expenditures	Kantar Ad$pender	Total advertising spends (*log-transformed*) for a video game '*i*' made by the launch firm throughout the prerelease period
	Corporate Updates	Factiva	Time-vary variable measuring the number of corporate updated announcements about a new video game '*i*' during the prerelease period.
	Number of Platforms Available	VGChartz.com	The number of console platforms where the new video game is available
	Sequel	Multiple Sources	Dummy variable. 1 denotes the new video game is a sequel; 0 otherwise.
	Genre	Multiple Sources	Eight dummy variables (1 vs. 0) to reflect each new video game's genre. The base genre to be compared against is 'miscellaneous'.

Table 2.3
Regression Model Results

	Model 1		Model 2		Model 3		Full Model	
Knowledge Collaborative Activities								
Article Length FPCS 1					-0.008	**	-0.010	***
Article Length FPCS 2					-0.011	^	-0.007	
Article Length FPCS 3					-0.002		-0.003	
Article Length FPCS 4					0.016		0.013	
Editing Rigor FPCS 1					0.100	**	0.078	**
Editing Rigor FPCS 2					0.181		0.160	
Editing Rigor FPCS 3					0.145	**	0.063	
Editing Rigor FPCS 4					0.185	**	0.170	**
Diversity FPCS 1					0.045	***	0.038	***
Diversity FPCS 2					0.039	**	0.021	^
Diversity FPCS 3					0.007		0.001	
Diversity FPCS 4					0.003		0.001	
Specialist %					-0.223		-0.034	
Prospective Customer Involvement								
Pageview FPCS 1					-0.004		0.241	
Pageview FPCS 2					0.050	***	-0.007	^
Pageview FPCS 3					0.010		0.047	***
Pageview FPCS 4					-0.041	*	0.006	
Customer Buzz								
Pos. Post Vol. FPCS1			0.019	*			0.015	*
Pos. Post Vol. FPCS2			-0.008				-0.017	
Pos. Post Vol. FPCS3			0.031				0.008	
Pos. Post Vol. FPCS4			0.027				0.019	
Corporate Marketing Activities								
Corp. Updates. FPCS1			0.071	***			0.049	***
Corp. Updates. FPCS2			0.000				0.009	
Corp. Updates. FPCS3			-0.016				-0.033	
Corp. Updates. FPCS4			-0.043				-0.001	
Advertising Expenditures (*log*)			0.123	***			0.091	***
Game Attributes and Genres								
No. of Platforms Available	0.418	***	0.298	***	0.277	***	0.237	***
Sequel	0.560	**	0.516	**	0.233		0.320	*
Action	0.318		0.126		0.062		-0.036	
Adventure	-0.061		-0.149		-0.073		-0.176	
Fighting	-0.320		0.009		-0.479		-0.226	
Racing	-0.515		-0.371		-0.357		-0.368	
Role-play	-0.073		0.115		-0.119		0.042	
Shooter	0.245		0.008		0.009		-0.146	
Sports	0.282		0.093		0.294		0.051	
Strategy	-1.432	*	-1.612	**	-1.544	**	-1.658	***
Year Dummies Included	Yes		Yes		Yes		Yes	
Intercept	10.832	***	11.173	***	11.609	***	11.424	***
Adj. R^2	0.25		0.45		0.48		0.56	

Note: *** $p < .001$; ** $p < .01$; * $p < .05$; ^ $p < .1$; In the exploratory analysis, I examined regression models with Customer Buzz variables only and the model with dynamic Wikipedia knowledge collaboration variables only. The latter is more than three times better in terms of the adjusted R^2. I also replaced those dynamic (i.e., FDA-processed) Wikipedia measures in Model 3 with their aggregated values. The R^2 is 0.41 (c.f. 0.48 for Model 3). For the sake of brevity, I do not present the results here, but detailed information is available upon request.

Table 2.4
***Adj.* R^2 Comparison**

	Variables Considered[a]	$Adj.\ R^2$	Difference c.f. Full Model	F-Test[b]	
Model 1	1	0.29	-48.21%	5.82	***
Model 2	1, 2, 3	0.45	-19.64%	3.56	***
Model 3 – a	1, 4	0.41	-26.79%	3.96	***
Model 3 – b	1, 5	0.48	-14.29%	4.65	***
Full Model	1, 2, 3, 4	0.56	N/A	N/A	

Note: *** $p < .001$; ** $p < .01$; * $p < .05$; a: 1. Game Characteristics 2. Launching Firm Marketing Activities; 3. Customer Buzz; 4. Knowledge Collaboration and Wikipedia pageview Measures (aggregated); 5. Knowledge Collaboration and Wikipedia pageview Measures (dynamic); b: The F-Test is undertaken based on Geladi and Kowalski (1986), which has been widely adopted by researchers in various disciplines for the comparison of R^2 of the partial regression models and their full model version.

Table 2.5
Prediction Accuracy

	Variables Considered[a]	MAPE	MSE
Model 1	1	9.10%	1.80
Model 2	1, 2, 3	7.76%	1.35
Model 3 – a	1, 4	7.99%	1.50
Model 3 – b	1, 5	6.78%	1.10
Full Model	1, 2, 3, 4	6.75%	1.09
Ensemble RF (Boosting)	1, 2, 3, 4	6.34%	0.93
Ensemble RF (Bagging)	1, 2, 3, 4	5.99%	0.88

Note: a: 1. Game Characteristics 2. Launching Firm Marketing Activities; 3. Customer Buzz; 4. Knowledge Collaboration Measures (aggregated); 5. Knowledge Collaboration Measures (dynamic)

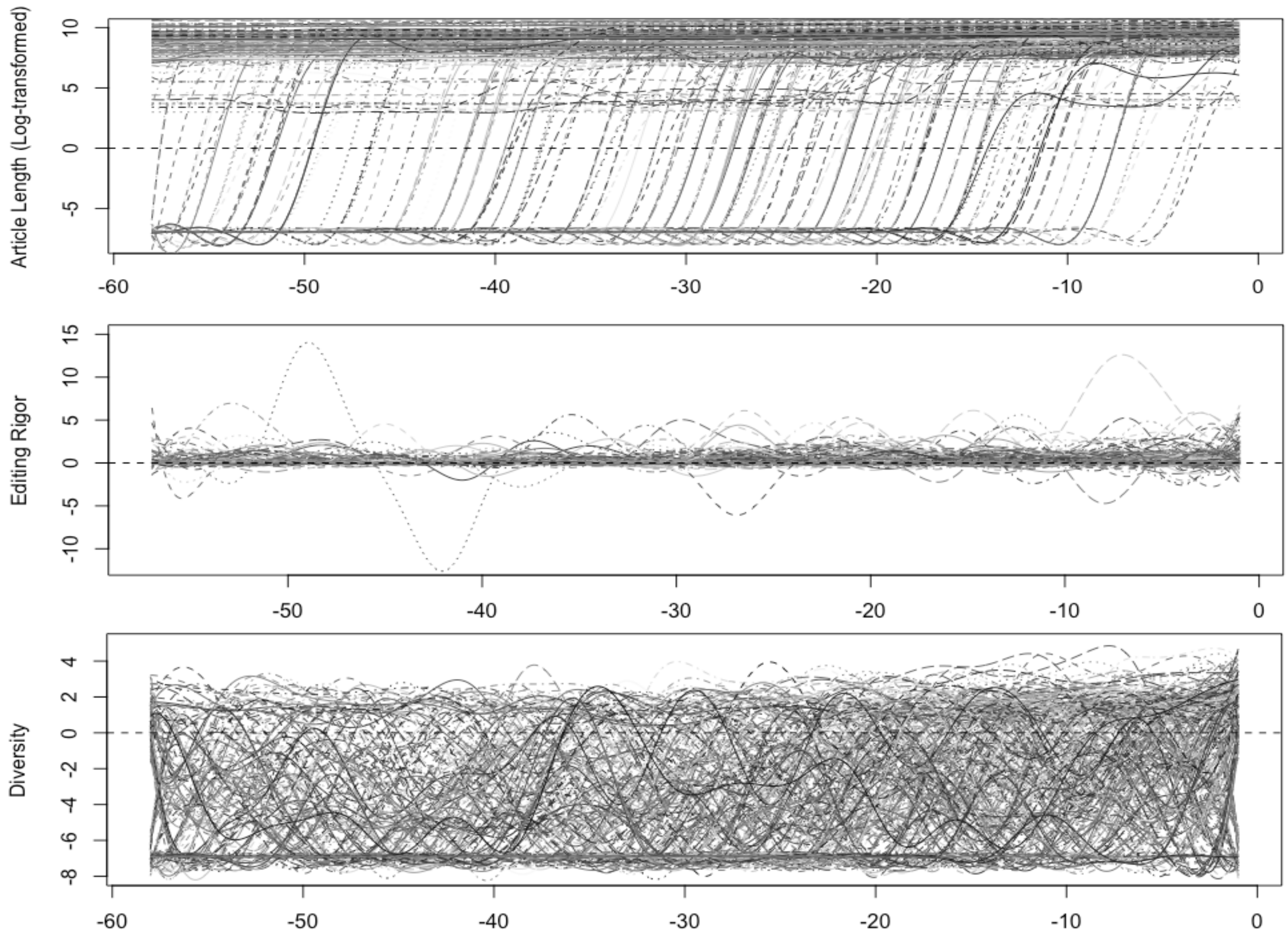
Table 2.6
Early Forecast Accuracy

Early Forecast	RF(Bagging)		RF(Boosting)	
	MAPE	MSE	MAPE	MSE
10 weeks prior to launch	7.02%	1.15	7.20%	1.18
5 weeks prior to launch	6.40%	0.97	6.97%	1.09
Full time-range data	5.99%	0.88	6.34%	0.93

Table 2.7
Robustness Check Results

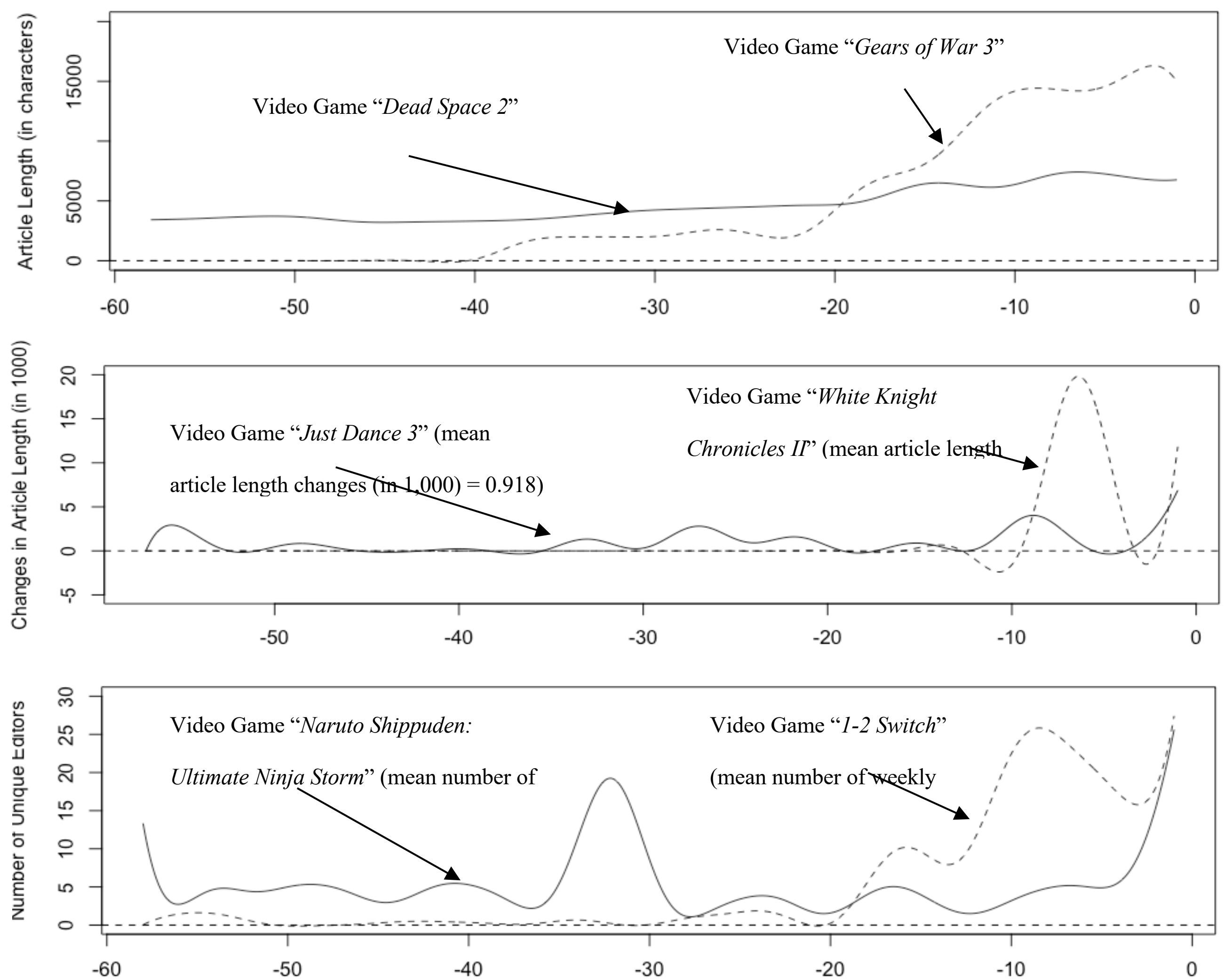
Scenarios	Methods	MAPE	MSE
Volume of the Week before Launch	Regression	11.42%	1.80
	Ensemble RF (Boosting)	9.87%	2.07
	Ensemble RF (Bagging)	7.73%	1.31
Late volume *less* Early Volume	Regression	10.64%	1.58
	Ensemble RF (Boosting)	9.49%	1.93
	Ensemble RF (Bagging)	7.41%	1.18
Linear Trend *plus* Total Volume	Regression	10.80%	1.66
	Ensemble RF (Boosting)	9.26%	1.78
	Ensemble RF (Bagging)	7.48%	1.30
FDA Processed Measures	Regression	6.75%	1.09
	Ensemble RF (Boosting)	6.34%	0.93
	Ensemble RF (Bagging)	5.99%	0.88

FIGURES



Note: From the top to the bottom, the above three graphs illustrate the spline smoothed curves of the three knowledge collaborative measures of this study's interest (i.e., Article Length, Editing Rigor, and Diversity). The horizontal axes of the three graphs denotes the prerelease period of the video games from -58th to -1st week before launch.

Figure 2.1
FDA-recovered Trajectories of Knowledge Collaborative Activities for the 300 Video Games during the Prerelease Period



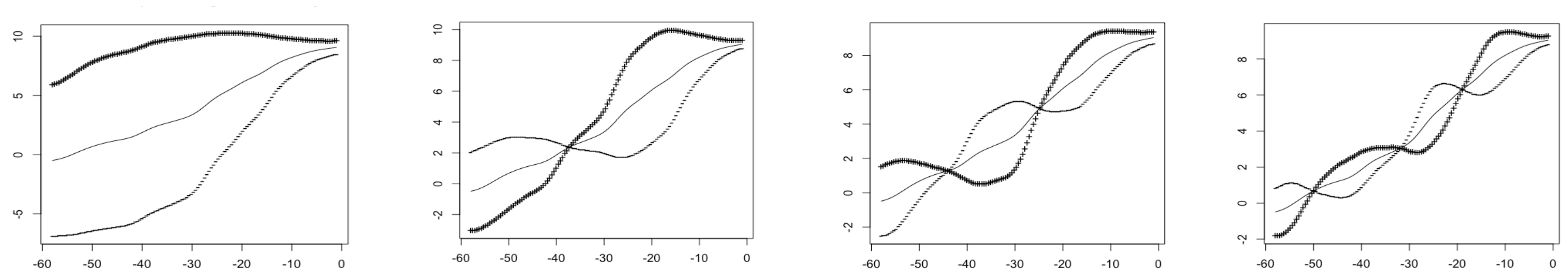
Note: From the top to the bottom, the above three graphs illustrate the comparison between two video games (two games per graph) that cumulatively have the same values of the people's knowledge collaborative activities (i.e., article length, editing rigor, and diversity), but dynamically present quite different patterns. The horizontal axes of the three graphs denotes the prerelease period of the video games from -58th to -1st week before launch.

Figure 2.2

Comparison between the Cumulative and Dynamic Trends of the Knowledge Collaborative Histories

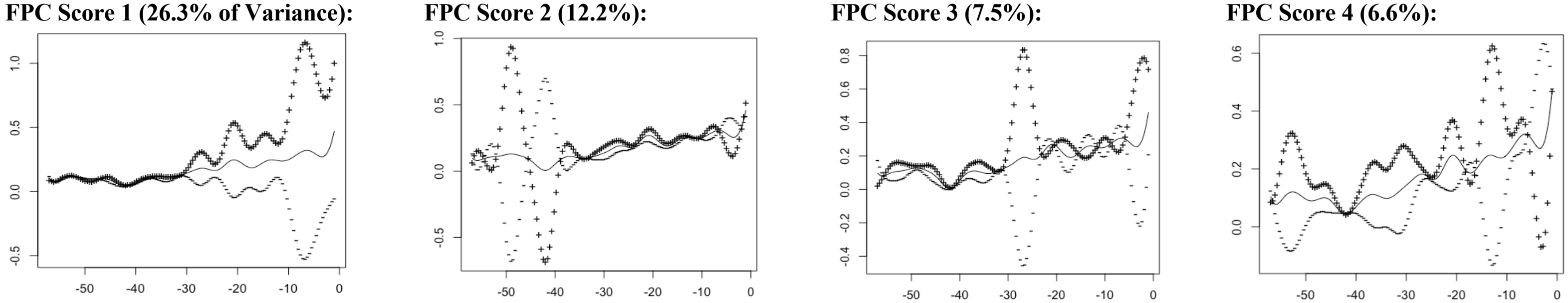
Panel A: *Article Length*

FPC Score 1 (76.8% of Variance): **FPC Score 2 (10.9%):** **FPC Score 3 (4.7%):** **FPC Score 4 (2.1%):**



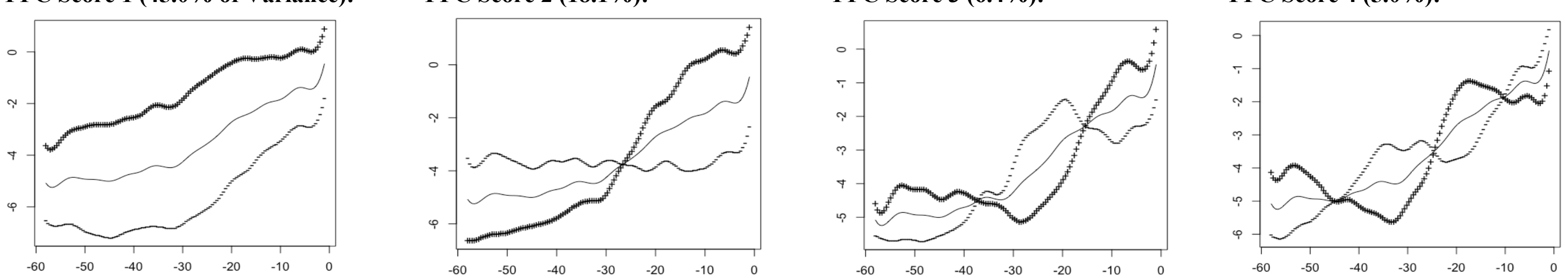
Panel B: *Editing Rigor*

FPC Score 1 (26.3% of Variance): **FPC Score 2 (12.2%):** **FPC Score 3 (7.5%):** **FPC Score 4 (6.6%):**



Panel C: *Diversity*

FPC Score 1 (45.0% of Variance): **FPC Score 2 (18.1%):** **FPC Score 3 (6.4%):** **FPC Score 4 (5.0%):**



Note: The horizontal axis denotes the length of the prerelease period (-58 to -1 week prior to launch); The mean curves (solid curve) of the above graphs represent the mean trends of the three knowledge collaborative measures respectively during the prerelease period (i.e., -58 to -1 weeks before launch). The '+ + +' or '— — —'curves suggest the variances of those 300 video games from the mean levels of by adding or subtracting a small amount from the mean values.

Figure 2.3
Functional Principle Component Analysis Results

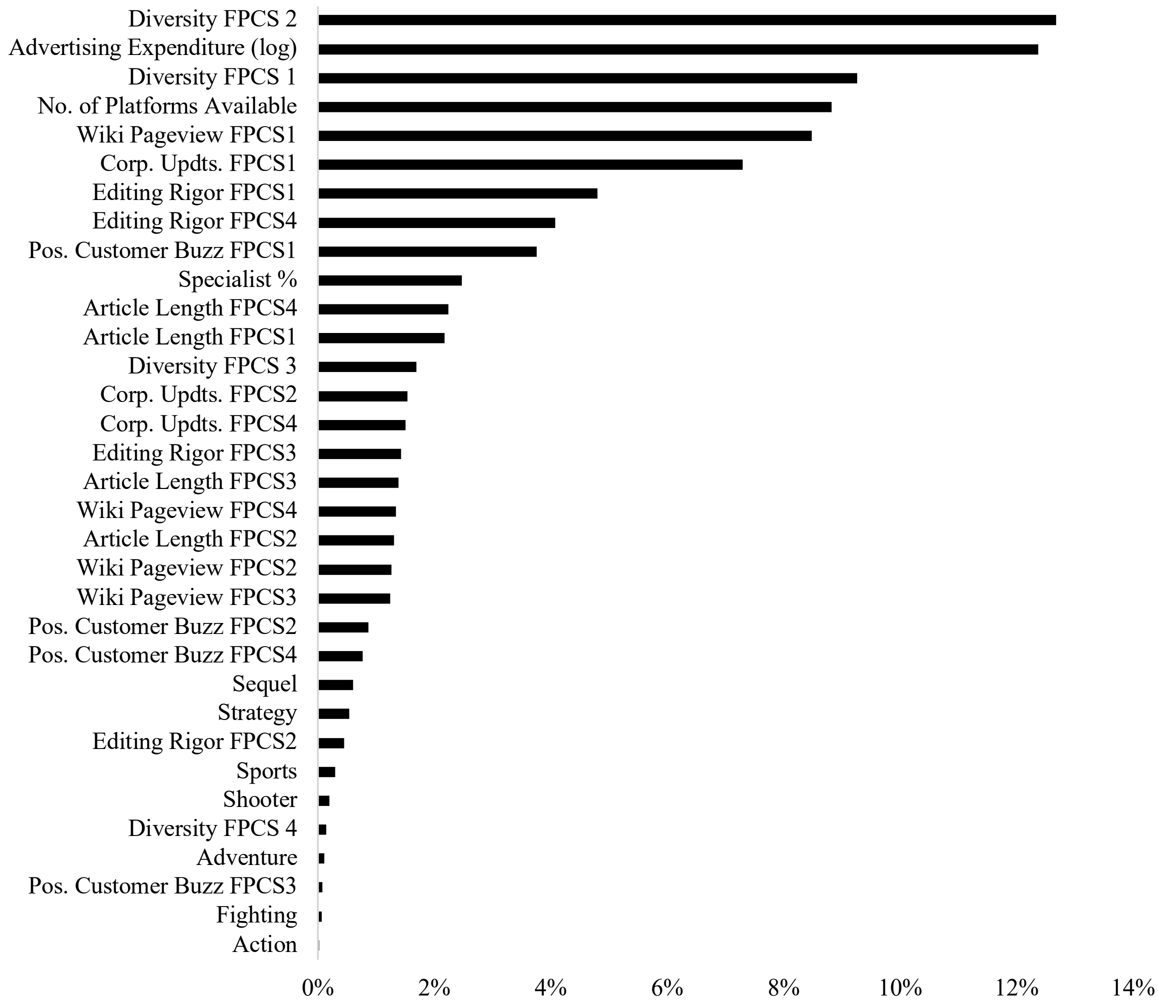


Figure 2.4
Significant Predictors Contributing to the Prerelease NPS Forecast Accuracy (MSE) based on the Random Forest (bagging) Results[16]

[16] Figure 2.4 is based on the estimated *permutation importance* of each predictor in the RF (bagging) models. Proposed by Breiman (2001) and Liaw and Wiener (2008), *permutation importance* is a conventional metric evaluating each predictor's relative contribution to the overall forecasting accuracy of the RF model. In essence, for each tree, the prediction error (i.e., mean squared error/MSE) based on the out-of-bag dataset is first evaluated. Then, with the permutation of each predictor, another MSE is evaluated. Following that, the difference between these two MSEs is calculated and averaged by the number of all trees in the process. Finally, the averaged value is further scaled (i.e., normalized) by the standard deviation of the differences of all trees. In the case when the standard deviation equals zero, the division is not defined, and the *permutation importance* of that predictor is set to zero.

REFERENCE

Akturk, M Serkan, Michael Ketzenberg, and Gregory R Heim (2018), “Assessing Impacts of Introducing Ship-to-Store Service on Sales and Returns in Omnichannel Retailing: A Data Analytics Study,” Journal of Operations Management, 61, 15-45.

Athey, Susan and Guido W Imbens (2019), “Machine Learning Methods That Economists Should Know About,” Annual Review of Economics, 11, 685-725.

Barasch, Alixandra and Jonah Berger (2014), “Broadcasting and Narrowcasting: How Audience Size Affects What People Share,” Journal of Marketing Research, 51 (3), 286-99.

Bargh, John A, Katelyn YA McKenna, and Grainne M Fitzsimons (2002), “Can You See the Real Me? Activation and Expression of the “True Self” on the Internet,” Journal of social issues, 58 (1), 33-48.

Bauer, Eric and Ron Kohavi (1999), “An Empirical Comparison of Voting Classification Algorithms: Bagging, Boosting, and Variants,” Machine learning, 36 (1), 105-39.

Bilgicer, Hasan (2014), “Driver and Consequences of Multichannel Shopping,” Columbia University.

Breiman, L, J Friedman, R Olshen, and C Stone (1984), “Classification and Regression Trees,” Wadsworth International Group, Belmont, California.

Breiman, Leo (2001), “Statistical Modeling: The Two Cultures (with Comments and a Rejoinder by the Author),” Statistical science, 16 (3), 199-231.

Breiman, L (2003), “Rf/Tools: A Class of Two-Eyed Algorithms,” in SIAM workshop.

Brown, John Seely and Paul Duguid (1991), “Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation,” Organization science, 2 (1), 40-57.

Brynjolfsson, Erik, Yu Hu, and Mohammad S Rahman (2009), “Battle of the Retail Channels: How Product Selection and Geography Drive Cross-Channel Competition,” Management Science, 55 (11), 1755-65.

Chakravarty, Anindita, Yong Liu, and Tridib Mazumdar (2010), “The Differential Effects of Online Word-of-Mouth and Critics' Reviews on Pre-Release Movie Evaluation,” Journal of Interactive Marketing, 24 (3), 185-97.

Chaney, Paul K, Timothy M Devinney, and Russell S Winer (1991), “The Impact of New Product Introductions on the Market Value of Firms,” Journal of business, 573-610.

Cheung, Christy MK and Dimple R Thadani (2012), “The Impact of Electronic Word-of-Mouth Communication: A Literature Analysis and Integrative Model,” Decision Support Systems, 54 (1), 461-70.

Chintagunta, Pradeep K, Shyam Gopinath, and Sriram Venkataraman (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), 944-57.

Chintagunta, Pradeep K and Jonathan Lee (2012), “A Pre-Diffusion Growth Model of Intentions and Purchase,” Journal of the academy of marketing Science, 40 (1), 137-54.

Coleman, Les (2011), “Losses from Failure of Stakeholder Sensitive Processes: Financial Consequences for Large Us Companies from Breakdowns in Product, Environmental, and Accounting Standards,” Journal of Business Ethics, 98 (2), 247-58.

Cucu, TC, G Varzaru, C Turcu, ND Codreanu, I Plotog, and R Fuica (2008), “1d and 2d Solutions for Traceability in an Electronic Manufacturing Services Company,” in 2008 31st International Spring Seminar on Electronics Technology: IEEE.

Cui, Dapeng and David Curry (2005), “Prediction in Marketing Using the Support Vector Machine,” Marketing Science, 24 (4), 595-615.

Davenport, Thomas H and Laurence Prusak (1998), Working Knowledge: How Organizations Manage What They Know: Harvard Business Press.

De Haan, Evert, PK Kannan, Peter C Verhoef, and Thorsten Wiesel (2018), “Device Switching in Online Purchasing: Examining the Strategic Contingencies,” Journal of Marketing, 82 (5), 1-19.

Derlega, Valerian J, Sandra Metts, Sandra Petronio, and Stephen T Margulis (1993), Self-Disclosure: Sage Publications, Inc.

Duan, Wenjing, Bin Gu, and Andrew B Whinston (2008), “Do Online Reviews Matter?—an Empirical Investigation of Panel Data,” Decision Support Systems, 45 (4), 1007-16.

Elgammal, Ahmed, David Harwood, and Larry Davis (2000), “Non-Parametric Model for Background Subtraction,” in European conference on computer vision: Springer.

Eliashberg, Jehoshua and Steven M Shugan (1997), “Film Critics: Influencers or Predictors?,” Journal of Marketing, 61 (2), 68-78.

Entertainment Software Association (2020), “2020 Essential Facts About the Video Game Industry”, Entertainment Software Association, (accessed March 23, 2021), [available at https://www.theesa.com/esa-research/2020-essential-facts-about-the-video-game-industry/].

Fan, Yingying, Natasha Foutz, Gareth M James, and Wolfgang Jank (2014), "Functional Response Additive Model Estimation with Online Virtual Stock Markets," The Annals of Applied Statistics, 8 (4), 2435-60.

Faraj, Samer, Sirkka L Jarvenpaa, and Ann Majchrzak (2011), "Knowledge Collaboration in Online Communities," Organization science, 22 (5), 1224-39.

Ferster, Charles B and Burrhus Frederic Skinner (1957), "Schedules of Reinforcement."

Fourt, Louis A and Joseph W Woodlock (1960), "Early Prediction of Market Success for New Grocery Products," Journal of Marketing, 25 (2), 31-38.

Foutz, Natasha Zhang and Wolfgang Jank (2010), "Research Note—Prerelease Demand Forecasting for Motion Pictures Using Functional Shape Analysis of Virtual Stock Markets," Marketing Science, 29 (3), 568-79.

Freund, Yoav and Robert E Schapire (1996), "Experiments with a New Boosting Algorithm," in icml Vol. 96: Citeseer.

Geladi, Paul and Bruce R Kowalski (1986), "Partial Least-Squares Regression: A Tutorial," Analytica chimica acta, 185, 1-17.

Ghose, Anindya and Panagiotis G Ipeirotis (2006), "Designing Ranking Systems for Consumer Reviews: The Impact of Review Subjectivity on Product Sales and Review Quality," in Proceedings of the 16th annual workshop on information technology and systems Vol. 303.

Gislason, Pall Oskar, Jon Atli Benediktsson, and Johannes R Sveinsson (2006), "Random Forests for Land Cover Classification," Pattern recognition letters, 27 (4), 294-300.

Go, Alec, Richa Bhayani, and Lei Huang (2009), "Twitter Sentiment Classification Using Distant Supervision," CS224N project report, Stanford, 1 (12), 2009.

Godes, David and Dina Mayzlin (2004), "Using Online Conversations to Study Word-of-Mouth Communication," Marketing Science, 23 (4), 545-60.

Greenstein, Shane, Grace Gu, and Feng Zhu (2020), "Ideology and Composition among an Online Crowd: Evidence from Wikipedians," Management Science.

Gruen, Thomas W, Talai Osmonbekov, and Andrew J Czaplewski (2006), "Ewom: The Impact of Customer-to-Customer Online Know-How Exchange on Customer Value and Loyalty," Journal of business research, 59 (4), 449-56.

Guide Jr, V Daniel R (2000), "Production Planning and Control for Remanufacturing: Industry Practice and Research Needs," Journal of Operations Management, 18 (4), 467-83.

Hamilton, James Douglas (2020), Time Series Analysis: Princeton university press.

Hansen, Morten T, Nitin Nohria, and Thomas Tierney (1999), "What's Your Strategy for Managing Knowledge," The knowledge management yearbook 2000–2001, 77 (2), 106-16.

Hars, A and S Ou (2001), "Work-Ing for Free? Motivations of Par-Ticipating in Open Source Projects []," in Theth Annual Hawaii In-ternational Conference on System Sciences.

Hodge, David C, Marcia B Baxter Magolda, and Carolyn A Haynes (2009), "Engaged Learning: Enabling Self-Authorship and Effective Practice," Liberal Education, 95 (4), 16-23.

Houston, Mark B, Ann-Kristin Kupfer, Thorsten Hennig-Thurau, and Martin Spann (2018), "Pre-Release Consumer Buzz," Journal of the academy of marketing Science, 46 (2), 338-60.

Hu, Kejia, Jason Acimovic, Francisco Erize, Douglas J Thomas, and Jan A Van Mieghem (2019), "Forecasting New Product Life Cycle Curves: Practical Approach and Empirical Analysis: Finalist–2017 M&Som Practice-Based Research Competition," Manufacturing & Service Operations Management, 21 (1), 66-85.

Hyde, V, E Moore, and A Hodge (2004), "Functional Pca for Exploring Bidding Activity Times for Online Auctions," Term paper, Research Interaction Team, VIGRE program, Univ. Maryland.

Isaksen, Scott G and Donald J Treffinger (1985), "Creative Problem Solving," The Basic Course. New York: Bearly Limited.

Jank, Wolfgang and Galit Shmueli (2006), "Functional Data Analysis in Electronic Commerce Research," Statistical science, 21 (2), 155-66.

Johnson, Benjamin K and Judith E Rosenbaum (2015), "Spoiler Alert: Consequences of Narrative Spoilers for Dimensions of Enjoyment, Appreciation, and Transportation," Communication Research, 42 (8), 1068-88.

Kahn, Kenneth B (2010), "New-Product Forecasting," Wiley International Encyclopedia of Marketing.

Kahn, Kenneth B (2014), "Solving the Problems of New Product Forecasting," Business Horizons, 57 (5), 607-15.

Kane, Gerald C, Robert G Fichman, John Gallaugher, and John Glaser (2009), "Community Relations 2.0," Harvard business review, 87 (11), 45-50, 132.

Kimmons, Royce M (2011), "Understanding Collaboration in Wikipedia," First Monday.

Komorita, Samuel S, CD Parks, and LG Hulbert (1992), "Reciprocity and the Induction of Cooperation in Social Dilemmas," Journal of Personality and Social Psychology, 62 (4), 607.

Lave, Jean and Etienne Wenger (1991), Situated Learning: Legitimate Peripheral Participation: Cambridge university press.

Lee, Jumin, Do-Hyung Park, and Ingoo Han (2008), "The Effect of Negative Online Consumer Reviews on Product Attitude: An Information Processing View," Electronic Commerce Research and Applications, 7 (3), 341-52.

Lemon, Katherine N and Peter C Verhoef (2016), "Understanding Customer Experience Throughout the Customer Journey," Journal of Marketing, 80 (6), 69-96.

Lewis, Roger J (2000), "An Introduction to Classification and Regression Tree (Cart) Analysis," in Annual meeting of the society for academic emergency medicine in San Francisco, California Vol. 14.

Lih, Andrew (2004), "Wikipedia as Participatory Journalism: Reliable Sources? Metrics for Evaluating Collaborative Media as a News Resource," Nature, 3 (1), 1-31.

Liu, Xiao, Param Vir Singh, and Kannan Srinivasan (2016), "A Structured Analysis of Unstructured Big Data by Leveraging Cloud Computing," Marketing Science, 35 (3), 363-88.

Lunetta, Kathryn L, L Brooke Hayward, Jonathan Segal, and Paul Van Eerdewegh (2004), "Screening Large-Scale Association Study Data: Exploiting Interactions Using Random Forests," BMC genetics, 5 (1), 1-13.

Maas, Andrew, Raymond E Daly, Peter T Pham, Dan Huang, Andrew Y Ng, and Christopher Potts (2011), "Learning Word Vectors for Sentiment Analysis," in Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies.

Martín-Perpiñá, Maria de las Mercedes, Ferran Viñas Poch, and Sara Malo Cerrato (2019), "Personality and Social Context Factors Associated to Self-Reported Excessive Use of Information and Communication Technology (Ict) on a Sample of Spanish Adolescents," Frontiers in psychology, 10, 436.

Melis, Kristina, Katia Campo, Els Breugelmans, and Lien Lamey (2015), "The Impact of the Multi-Channel Retail Mix on Online Store Choice: Does Online Experience Matter?," Journal of Retailing, 91 (2), 272-88.

Mousavi, Seyed Yaghoub, Renae Low, and John Sweller (1995), "Reducing Cognitive Load by Mixing Auditory and Visual Presentation Modes," Journal of educational psychology, 87 (2), 319.

Mudambi, Susan M and David Schuff (2010), "Research Note: What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon. Com," MIS quarterly, 185-200.

Mullainathan, Sendhil and Jann Spiess (2017), "Machine Learning: An Applied Econometric Approach," Journal of economic perspectives, 31 (2), 87-106.

Neelamegham, Ramya and Pradeep Chintagunta (1999), "A Bayesian Model to Forecast New Product Performance in Domestic and International Markets," Marketing Science, 18 (2), 115-36.

Nelson, Phillip (1970), "Information and Consumer Behavior," Journal of political economy, 78 (2), 311-29.

Norris, Pippa (2002), "The Bridging and Bonding Role of Online Communities," Sage Publications Sage CA: Thousand Oaks, CA.

Park, Sungsik, Woochoel Shin, and Jinhong Xie (2021), "The Fateful First Consumer Review," Marketing Science.

Prasarnphanich, Pattarawan and Christian Wagner (2009), "The Role of Wiki Technology and Altruism in Collaborative Knowledge Creation," Journal of Computer Information Systems, 49 (4), 33-41.

Ramsay, James O and CJ Dalzell (1991), "Some Tools for Functional Data Analysis," Journal of the Royal Statistical Society: Series B (Methodological), 53 (3), 539-61.

Ramsay, James O and Bernard W Silverman (2007), Applied Functional Data Analysis: Methods and Case Studies: Springer.

Ramsay, James O and Bernard W Silverman (2008), "Functional Data Analysis," İnternet Adresi: http.

Rodriguez-Galiano, Victor Francisco, Bardan Ghimire, John Rogan, Mario Chica-Olmo, and Juan Pedro Rigol-Sanchez (2012), "An Assessment of the Effectiveness of a Random Forest Classifier for Land-Cover Classification," ISPRS Journal of Photogrammetry and Remote Sensing, 67, 93-104.

Rosario, Ana Babić, Kristine de Valck, and Francesca Sotgiu (2020), "Conceptualizing the Electronic Word-of-Mouth Process: What We Know and Need to Know About Ewom Creation, Exposure, and Evaluation," Journal of the academy of marketing Science, 48 (3), 422-48.

Rust, Roland T, V Kumar, and Rajkumar Venkatesan (2011), "Will the Frog Change into a Prince? Predicting Future Customer Profitability," International Journal of Research in Marketing, 28 (4), 281-94.

Schapire, Robert E (1999), "A Brief Introduction to Boosting," in Ijcai Vol. 99: Citeseer.

Schapire, Robert E, Yoav Freund, Peter Bartlett, and Wee Sun Lee (1998), "Boosting the Margin: A New Explanation for the Effectiveness of Voting Methods," Annals of statistics, 26 (5), 1651-86.

Schwartz, Shalom H (1970), "Elicitation of Moral Obligation and Self-Sacrificing Behavior: An Experimental Study of Volunteering to Be a Bone Marrow Donor," Journal of Personality and Social Psychology, 15 (4), 283.

Segal, Mark R (2004), "Machine Learning Benchmarks and Random Forest Regression."

Seraj, Mina (2012), "We Create, We Connect, We Respect, Therefore We Are: Intellectual, Social, and Cultural Value in Online Communities," Journal of Interactive Marketing, 26 (4), 209-22.

Sethi, S Prakash, Emre A Veral, H Jack Shapiro, and Olga Emelianova (2011), "Mattel, Inc.: Global Manufacturing Principles (Gmp)–a Life-Cycle Analysis of a Company-Based Code of Conduct in the Toy Industry," Journal of Business Ethics, 99 (4), 483-517.

Shmueli, Galit and Otto R Koppius (2011), "Predictive Analytics in Information Systems Research," MIS quarterly, 553-72.

Shmueli, Galit, Ralph P Russo, and Wolfgang Jank (2007), "The Barista: A Model for Bid Arrivals in Online Auctions," The Annals of Applied Statistics, 1 (2), 412-41.

Sood, Ashish, Gareth M James, and Gerard J Tellis (2009), "Functional Regression: A New Model for Predicting Market Penetration of New Products," Marketing Science, 28 (1), 36-51.

Sproull, L. and M. Arriaga (2007). "Online Communities" in Handbook of Computer Networks: Distributed Networks, Network Planning, Control, Management, and New Trends and Applications (Vol. 3), H. Bidgoli (ed.), New York: John Wiley & Sons.

Stigler, George J (1961), "The Economics of Information," Journal of political economy, 69 (3), 213-25.

Strobl, Carolin, James Malley, and Gerhard Tutz (2009), "An Introduction to Recursive Partitioning: Rationale, Application, and Characteristics of Classification and Regression Trees, Bagging, and Random Forests," Psychological methods, 14 (4), 323.

Sweller, John (2011), “Cognitive Load Theory,” in Psychology of Learning and Motivation, Vol. 55: Elsevier.

Tirunillai, Seshadri and Gerard J Tellis (2012), “Does Chatter Really Matter? Dynamics of User-Generated Content and Stock Performance,” Marketing Science, 31 (2), 198-215.

Von Krogh, Georg (1998), “Care in Knowledge Creation,” California management review, 40 (3), 133-53.

Wagner, Christian and Ann Majchrzak (2006), “Enabling Customer-Centricity Using Wikis and the Wiki Way,” Journal of management information systems, 23 (3), 17-43.

Wang, Shanshan, Wolfgang Jank, and Galit Shmueli (2008), “Explaining and Forecasting Online Auction Prices and Their Dynamics Using Functional Data Analysis,” Journal of Business & Economic Statistics, 26 (2), 144-60.

Wasko, M McLure and Samer Faraj (2000), ““It Is What One Does”: Why People Participate and Help Others in Electronic Communities of Practice,” The journal of strategic information systems, 9 (2-3), 155-73.

Wasko, Molly McLure and Samer Faraj (2005), “Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice,” MIS quarterly, 35-57.

Williams, Scott David (2002), “Self-Esteem and the Self-Censorship of Creative Ideas,” Personnel Review.

Woodman, Richard W, John E Sawyer, and Ricky W Griffin (1993), “Toward a Theory of Organizational Creativity,” Academy of management review, 18 (2), 293-321.

Wooldridge, Jeffrey M (2010), Econometric Analysis of Cross Section and Panel Data: MIT press.

Xiong, Guiyang and Sundar Bharadwaj (2014), “Prerelease Buzz Evolution Patterns and New Product Performance,” Marketing Science, 33 (3), 401-21.

Zhang, Jonathan Z and Chun-Wei Chang (2020), “Consumer Dynamics: Theories, Methods, and Emerging Directions,” Journal of the academy of marketing Science, 1-31.

Zhu, Feng and Xiaoquan Zhang (2010), “Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics,” Journal of Marketing, 74 (2), 133-48.

Zillmann, Dolf, T Alan Hay, and Jennings Bryant (1975), “The Effect of Suspense and Its Resolution on the Appreciation of Dramatic Presentations,” Journal of Research in Personality, 9 (4), 307-23.

CHAPTER 3

PRERELEASE ADVERTISING ADJUSTMENT AND NEW PRODUCT PERFORMANCE: THE ROLES OF CUSTOMER AND INVESTOR SENTIMENT[17]

[17] Zhang, P., A. Chakravarty and S. G. Bharadwaj. To be submitted to *Journal of Marketing*.

ABSTRACT

This essay investigates whether and how new product managers adjust their prerelease advertising deployments (both strategically and tactically) in response to customer and investor sentiment. Moreover, it examines the differential effects of such advertising adjustments in terms of new product sales and firm value. A three-step modeling analysis, comprised of panel vector autoregression model, Bayesian vector autoregression model, and Bayesian hierarchical model, shows that managers tend to increase both digital and traditional advertising expenditures for both new product promotional and brand marketing campaigns in response to the increased positivity of customer sentiment. However, they only increase traditional advertising expenditures for the brand marketing campaign in response to the increased investor bullish sentiment. Eventually, traditional advertising adjustments of new product promotional campaign in response to customer sentiment significantly and positively drive short term new product sales. Traditional advertising adjustments of both new product promotional and brand marketing campaigns in response to customer sentiment significantly and positively affect long term firm value.

INTRODUCTION

Prerelease period starting from the preannouncement[18] of a new product until its formal launch is considered a critical period. Marketing managers know a lot more about their product(s) and their marketing plans regarding consumer segments, market potential and marketing programs than external stakeholders such as consumers and investors. Thus, they attempt to alleviate the information asymmetry by employing advertising campaigns to communicate the benefits of the new product and its brand image so as to: 1) elicit positive emotion among prospective customers (Batra and Keller 2016; Moe and Fader 2002; Xiong and Bharadwaj 2014) and 2) enhance investors' confidence about the new product's contributing role to the launching firm's future cash flow generation as well as the firm's marketing commitment to the new product's success (Elberse and Anand 2007; Joshi and Hanssens 2009; Srinivasan and Hanssens 2009). The former is important generating demand and sales, the latter helps with the launching firm's financing ability to acquire more resources from the capital market for its future operation and innovation (Chan and Hameed 2006; James 1987).

However, there are uncertainties about the effectiveness and efficiency of these prerelease advertising campaigns in terms of their impact on customers' and investors' perceptions of the new product. Typical market research methods (e.g., focus group study, consumer interviews, or professional consulting), due to the product not yet being available for sale in the market and the lengthy execution time, are not always feasible to reach a reliable conclusion in a timely manner (Foutz and Jank 2010; Su and Rao 2010). Online chatter from

[18] A new product's preannouncement is the first formal marketing communication activity deliberately initiated by the launching firm. It may occur at any time throughout different stages of the new product development, including ideation, product design and engineering testing, trial and mass production, or new products introduction (Eliashberg and Robertson 1988). Accordingly, the length of prerelease period may vary from a few weeks to more than a year. In this study, the average prerelease period is 47 weeks for new video games launched in the U.S. market between 2012 and 2017.

prospective customers and investors, at this point, serve as an alternative information source for these managers to examine the sentiment toward the new product and its brand. The sentiment in turn provides guidance for managers to strategically adapt their advertising deployment and enhance the effectiveness and efficiency of the investment (Luan and Sudhir 2010; Juergens and Lindsey 2009; Suk, Chintagunta and Park 2019).

Houston et al. (2018) suggest that customer sentiment during the prerelease period (i.e., buzz) reflects a prospective customer's enthusiasm toward a forthcoming product, its features and the brand. An increasing positivity of customer sentiment foretells the customer's likelihood of owning the new product (You, Vadakkepatt and Joshi 2015; Xiong and Bharadwaj 2014). Investor sentiment, on the other hand, refers to investors' confidence (i.e., bullish expectation) about the strategic choices made by the firm (i.e., product launch), the target market potential of the new product, and the firm's marketing commitment to the new product success (see: Baker and Wurgler 2007; Luo, Jiang and Cai 2015). The more investors who are bullish (vs. bearish) about a firm's new product, the higher (vs. lower) the launching firm's stock price will be (see: Da, Engelberg and Gao 2011).

Facing the increased positivity of customer sentiment, managers are likely to adopt a value exploitation mindset by using available advertising campaigns (i.e., new product promotional and brand marketing campaigns) to quickly convert these customers who have already developed some interest in the new product (Arnold, Fang and Palmatier 2011). The increased bullish sentiment among investors challenges managers to explore all possible strategic choices (e.g., new product promotional and brand marketing ads) to expand the new product customer base (Vorhies, Orr and Bush 2011; Voss and Voss 2008). With the expected outcome being that the new product sales are likely to increase, managers are more likely to meet the bullish expectations among the investors.

Successful prerelease advertising deployment in response to customer and investor sentiment hinges on two pivotal factors: strategy and tactics. Strategy attends to what to communicate (see: Varadarajan 2010). Hereby, managers of new product have to determine to what extent they shall spend on new product promotional and brand marketing campaigns. New product promotional advertising is effective in communicating the key benefits of the product, so as to facilitate customers to better assess its value before making the purchase decision and signal investors the launching firm's confidence in product quality (Kapferer 2008; Kumar et al. 2010; Lane and Jacobson 1995; Sorescu and Shankar 2007). Alternatively, managers may also consider campaigns focusing on brand marketing. According to Balmer (2001) and Hatch and Schultz (1997), brand marketing advertisements promote a firm's vision and culture, both of which in turn serve as a distinct and competitive advantage differentiating the firm from its competitors (Keller and Aaker 1998). When a customer considers purchasing a new product, she is not only examining the functional benefits of the product, but also assessing other factors, such as the congruency between the firm's culture/image and her own social identity, or the quality of other similar products introduced by the firm historically (Brown and Dacin 1997; Muzellec and Lambkin 2009). For investors, brand marketing campaigns send signals that the launching firm has strong marketing commitment to supporting the new product success and driving sales, thus meeting their bullish expectations (Homburg, Theel and Hohenberg 2020).

Different from strategy, advertising tactics deals with issues of how to communicate (Holm 2006). While all advertising campaigns boost marketing communication, they play quite different roles in efficiently enhancing audiences' awareness and knowledge about a product, stimulating positive emotion, and improving purchase intent (Batra and Keller 2016; Dahlén and Edenius 2007; Van Doorn and Hoekstra 2013)[19]. Thanks to the targeting and retargeting technologies of digital advertising, advertisers are able to deliver specific ad content (either new product promotion or brand marketing) to their well-defined audiences and audience segments who share certain demographic characteristics or present certain online behaviors (Bruce, Murthi and Rao 2017). However, digital advertising has narrower audience reach (Sridhar et al. 2016). Moreover, advertisers have to be cautious about the inefficiency of digital advertising due to the ad effect measurement errors and audiences' inactivity (Gordon et al. 2021). Ad effect measurement errors directly affect targeting effectiveness. Audiences' inactivity (e.g., ad blocking or the reluctance to click paid ads) tremendously reduce the efficiency of digital advertising. On the contrary, traditional advertising (e.g., TV commercials, radio advertising, out-of-home advertising, etc.) is efficient reaching broad audiences through the national network of the media vehicles. Although the message delivered by this type of advertising is less specific and more generic (relative to digital advertising), it unobtrusively improves audiences' affinity with the brand and its product offerings, leading to sales and market value elevation (Batra and Keller 2016; Helm 2007). In short, both digital and traditional advertising are viable alternative tactics to promote new product and brand marketing. However, they also come with limitations that lead to managers' uncertainty about the ultimate outcome (e.g., new product sales).

New forms of programmatic planning allow firms to constantly adjust their advertising deployment based on the real time feedback from customers and investors, resulting in a more efficient execution of the advertising plan (Google Inc. 2021). For example, the Chicago Bulls, using an online social listening platform, adaptively manage its social media campaign to foster

[19] I use audience(s) to refer to the recipients of advertising message. In the context of this study, I specifically speak to two types of audiences: prospective customers and investors.

its relationship with fans all over the world (Sprout Social Inc. 2020). As a marketing manager of the basketball team explained, the success of an advertising campaign relies on managers' constant awareness of "what's happening, what the trends are, [and] what people are responding to", as well as their ability to adjust the campaigns appropriately in the real time based on these responses. This process of advertising adjustment echoes what Menon et al. (1999, p.36) have proposed in their seminal study that advertising deployment entails "managers [to] use information, when it is revealed to them, to update their beliefs about the responsiveness to their [advertising] actions. These beliefs, in turn, affect [advertising] budget decisions."

To the best of my knowledge, there is very little research providing insights into this area. Table 3.1 presents the literature review based on representative studies.

Against this backdrop, I seek to address the following two research questions within the context of the prerelease period. First, do managers of the launching firm respond to both customer and investor sentiment toward their forthcoming new product and adjust their strategic (i.e., new product promotional and brand marketing) and tactical (i.e., traditional and digital) advertising deployment during the prerelease period? Second, what are the consequences of such prerelease advertising adjustments undertaken by the launching firm, in terms of new product sales?

Empirically, I devise a three-step modeling estimation process incorporating panel vector autoregression (PVAR), Bayesian vector autoregression (BVAR), and Bayesian Hierarchical Model to address the research questions[20]. The proposed modeling process 1) successfully accounts for the heterogeneities from both firm and product category levels (i.e., game genre for

[20] I also examine the effect of such an advertising adjustment on firm value, measured by the difference of buy-and-hold abnormal return between the launching firm and its matched counterpart. The detailed description of this additional study is provided in the Additional Analysis section.

this study) and 2) controls the endogeneity concerns due to the dynamic feedback effects among customers and investors, the launching firm's advertising activities, competitors' advertising activities, and the media coverage during the prerelease period. I select video game industry as the research context, given its broad user base, and consistent large number of new video games introduction annually. I combine video game sales data from vgchartz.com, publisher companies' financial statistics from Compustat, and other data sources (e.g., Wikipedia, Reddit, STEAM, etc.). The final dataset consists of 3,175 product-week observations of 71 video games launched in the North American market during 2012 and 2017 by seven publishers with an average of 45 weeks of prerelease period.

The current study makes both academic and managerial contributions. First, it examines the role of customer and investor sentiment as an information source guiding managers' prerelease advertising deployment. While most past studies merely consider the use of advertising (at the aggregated level) to manage either customers' or investors' expectations, they ignore customers' and investors' participation in influencing the launching firms' prerelease advertising practices (e.g., Srinivasan et al. 2009; Joshi and Hanssens 2010). Among the few that attends to the influence of customer or investor sentiment on firm's advertising adjustment, they merely focus on one of the two (e.g., customers: Bass et al. 2007; Wies et al. 2019; investors: Luo, Jiang and Cai 2015; Mian, Sharma and Gul 2018). Furthermore, none of the above studies examine the context of new product introduction during the prerelease period.

Second, it captures advertising adjustment at both strategic and tactical levels in response to customer and investor sentiment during the new product prerelease period. Third, the empirically devised three-step approach accounts for the dynamism embedded in the advertising adjustment process, and quantifies the ultimate outcome of such dynamic adjustment, in terms of new product sales, accounting for both endogeneity and heterogeneity concerns. Finally, the current study offers insightful guidelines for managers to more effectively and efficiently optimize their advertising campaigns during the critical new product release period. While both types of sentiment serve as critical information sources motivating managers to undertake advertising adjustment, only those changes based on customer sentiment exert significant and immediate effect on new product sales. Moreover, product promotional (c.f., brand marketing) advertising delivered through traditional media vehicles performs better than other options.

For the rest of this chapter, I first review the representative studies in this area, followed by the elaboration on the development of the conceptual framework and hypotheses. Next, I discuss the research context, variable measurement, and methodology. Following that, I will present the model results and discuss implications. I conclude with research limitation and paths for further studies.

LITERATURE REVIEW, CONCEPTUAL FRAMEWORK AND HYPOTHESES DEVELOPMENT

Literature Review

Most earlier studies in this area examine managers' proactive use of advertising to influence customers' or investors' behaviors, so as to improve firms' business performance. For example, Lovett and Staelin (2016) find support that firms' paid advertising campaigns positively influence online customer buzz, leading to customers' purchase behavior. Xiong and Bharadwaj (2013) has explored the moderating effect of advertising on firm value in face of negative or positive media coverage. According to the authors, when there are positive news reports about a firm's product or service, increase in advertising expenditure amplifies the positive effect by drawing investors' attention to the good news and magnifying the firm's capability to generate greater cash flow in the future. In the context of new product introduction, Srinivasan et al. (2009) have documented that investors are more favorable for firms that demonstrated substantial advertising support for their new product launching, resulting in the elevation of these firms' stock market value. Among the few studies recognizing the influence of customers or investors on firms' advertising strategy adjustment, most of them have focused on the managerial myopic effect. For example, Chakravarty and Grewal (2011, 2016) propose and empirically demonstrate that managers confronted with the risks of missing earning expectations tend to strategically decrease their advertising expenditure. Based on the accounting rule, cut in advertising spends artificially bumps up firms' earnings, mitigating the market volatility (Dechow, Sloan and Sweeney 1995). Park, Chintagunta and Suk (2019) contend that managers may strategically use the favorable feedback of the stock market (i.e., unexpected positive abnormal returns of a new drug's FDA approval) to inform their decision to increase advertising expenditures. Such an upward adjustment eventually enhances the advertising effectiveness in terms of new drug sales. However, these researchers did not examine the potential of investor sentiment in molding the launching firm managers' advertising strategies and tactics.

As discussed earlier, both customers and investors express emotion and expectations about a new product during the prerelease period, with different foci. Managers must adjust their prerelease advertising deployment based on information extracted from both customer and investor sentiment, in order to boost the new product performance. Moreover, advertising adjustment is not merely an increase or decrease in expenditure. It is a reflection of launching firm managers' understanding of advertising strategy and tactics in terms of their effectiveness and efficiency in driving new product performance. Finally, both customer and investor sentiment toward a forthcoming product is constantly changing, as a result of individual-level life stage changes (Zhang and Chang 2021) as well as the iterative interactions (e.g., advertising) between the launching firm and the individual. Each interaction leads to the individual's updating her attitude and knowledge about the brand and the product, eventually the purchase or investment decision (Da, Engelberg and Gao 2011; Lemon and Verhoef 2016; Peterson and Kumar 2009). Therefore, managers when crafting their prerelease advertising adjustment must constantly monitor and react to the fluctuation of customer and investor sentiment, to ensure the effectiveness of the ad campaigns.

To the best of the author's knowledge, however, fewer studies have simultaneously examined the influences from customers and investors. Neither do they decompose the advertising adjustment at both strategic and tactical levels and empirically investigate managers' prerelease advertising adjustment in response to customer and investor. Table 3.1 presents the literature review based on representative studies.

Conceptual Framework Development

Drawing on literature in advertising, the integrated marketing communication (IMC) and financial economics, I develop a conceptual framework mapping out how managers of the launching firm adjust their prerelease advertising deployment in response to customer and investor sentiment respectively and discuss whether the prerelease advertising adjustment will ultimately influence the launching firm's new product sales. Regarding advertising adjustment, I categorize it on the basis of two dimensions. The first attends to the strategic advertising planning, where managers of the launching firm need to determine their spend on the new product promotion and the brand marketing strategies. The second is related to advertising tactics, where the managers must determine through which media vehicles to deploy those advertisements (i.e., choices being digital and traditional), so as to achieve new product success.

For the rest of this section, I first discuss the differences between advertising strategy and tactics. Next, I elaborate on why managers adjust their prerelease advertising deployment (both strategically and tactically) in response to customer and investor sentiment. Finally, I explain how advertising adjustment may occur at both strategic and tactical levels in response to customer and investor sentiment respectively.

Advertising Strategy vs. Advertising Tactics

According to Varadarajan (2010, p.120), marketing strategy is “an organization’s integrated pattern of decisions that specify its crucial choices concerning products, markets, marketing activities and marketing resources in the creation, communication and/or delivery of products that offer value to customers in exchanges with the organization and thereby enables the organization to achieve specific objectives.” In contrast, marketing tactics focus on the realization of a strategy based on the efficient execution of the marketing mix (i.e., product, price, place, and promotion/communication; Slevin and Pinto 1987; Webster 1992).

In the context of this study, I propose that prerelease advertising strategy reflects a launching firm’s decision to bolster its new product’ post-release sales, in response to the increased positivity of customer and investor sentiment about the new product during the prerelease period, through efficiently communicating the benefits and market potential of the new product and the launching firm’s marketing commitment to supporting the new product success. Tactically, managers of the launching firm need to determine how to use both digital and traditional advertising to achieve their strategic goal.

Prerelease Advertising Adjustment – A Process Catering to the Evolving Customer and Investor Sentiment

According to Zhang and Chang (2021), an individual’s sentiment and behaviors toward a brand and its product is ever-changing due to multiple factors, including the evolution of one’s life stage, the influence from her social circle, and the interactions with the firm (e.g., advertising contact). For example, as an individual ages, she tends to be more risk averse and less willing to switch brands or products (Du and Kamakura 2006). Peer influence from one's social network drives homophily within the group, leading to changes of this person's attitude toward the brand and the product as well (e.g., Zhang 2009; Nitzan and Libai 2011).

In the context of the current study, both customer and investor sentiment about a forthcoming product is also bound to changes because of their interactions with the launching firm through the prerelease advertising campaigns. Each advertising exposure leads to an update in one's knowledge and experience with the brand and its product, triggering her to adjust "weighting of relative importance of the brand image and product attributes" for either purchase or investment decisions (Erdem and Keane 1996; Li et al. 2011; Zhang and Chang 2021, p.168). Such changes, in turn, entails launching firm managers to update their advertising strategy and tactics to cater to people's updated focus on brand and the new product. For the sake of brevity, I will not repetitively articulate the necessity of such advertising adjustment in the rest of this section. Instead, the main focus is on *how* managers may make different advertising adjustments (both strategically and tactically) in response to customer and investor sentiment respectively.

Prerelease New Product Promotional Advertising Adjustment in Response to the Increased Positivity of Customer Sentiment.

Prerelease customer sentiment refers to an aggregated emotional expression toward a forthcoming new product (Houston et al. 2018). It reflects prospective customers' enthusiasm to own the new product. Myriad of past research has shown that the positivity of customer sentiment is associated with the new product sales (e.g., Babić Rosario et al. 2016; Houston et al. 2018; Xiong and Bharadwaj 2014). Thus, it requires the launching firm to closely monitor and promptly react to the positivity of customer sentiment, by deploying new product promotional advertising and communicating the unique attributes of the new product during the prerelease period.

Tactically, the firms may choose to increase both digital and traditional advertising for the new product promotion. As customer sentiment is increasingly positive, it is critical for managers to reach these customers and quickly drive them to convert. In line with prospective customers' varied stages of path to purchase, prospective customers can be categorized as top of funnel and bottom of funnel customers. Digital advertising allows managers to target specific customers and well-defined customer segments who are already interested in the new product (i.e., bottom-of-funnel customers) to quickly convert. The targeting characteristics of digital advertising ensure the specific and personalized ad content delivery. It makes these customers more knowledgeable, facilitating them to make the ultimate purchase decision (Ruth 2001). New product promotion ads through traditional media vehicles speak to top of funnel customers who may be aware of the new product due to the general positive sentiment but still need to be motivated to seek more information and develop interest in the new product. Hereby, traditional advertising with broader reach and generalized messaging is helpful, by stimulating these customers' intention to undertake more detailed information search about the new product (Sridhar et al. 2016). Eventually, it promotes customers' knowledge about the new product and improve their purchase intent ((Bartels 1988; Batra and Keller 2016; Kenny and Rice 1994). Thus, I hypothesize

H1: During the prerelease period, launching firm adjusts their new product promotional spending by increasing both a) digital advertising and b) traditional advertising in response to the increased positivity of customer sentiment.

Prerelease Brand marketing Advertising Adjustment in Response to the Increased Positivity of Customer Sentiment.

In the above subsection, I have discussed the role of new product promotional advertising campaign in accelerating bottom-of-funnel customers' new product conversion (through digital media vehicles) and driving top-of-funnel customers' new product interest (through traditional media vehicles). As a matter of fact, purchase decision of the new product is also influenced by the brand image of the launching firm. In this regard, brand image is considered as the projection of a customer's own social identity (Bhattacharya and Sen 2003; Chiu et al. 2014; Overby and Lee 2006; Tajfel and Turner 1979; Voss, Spangenberg and Grohmann 2003). According to Lam et al. (2010), customers develop a relationship with a brand based on the belief that the brand and themselves share the same attributes. Such a relationship improves customers' positive sentiment toward the new product and promotes their loyalty. Anecdotally, there have been many discussions about consumers' identification as Mac users versus PC users (Nations 2020). In addition, a brand marketing campaign also reminds the customer of other products historically launched by the firm and assure her that the new product is trustworthy (Atuahene-Gima 2005; Olsen 2002).

Tactically, both digital and traditional advertising in response to customer sentiment enable the managers to bolster customers' decision to convert, but work through different routes. On the one hand, digital advertising can be used to directly target customers who have already developed a relationship with the launching firm and its brand (i.e., bottom of funnel customers). A digital ad focused on brand marketing efficiently reminds these people of their intimacy with the launching firm as well as their favorable user experience with the products launched by the firm historically. Both in turn exerts a call to action effect, driving those customers' intention to purchase the new product. On the other hand, brand marketing advertising through traditional media vehicles articulates the brand value to broader audiences and enables the public acceptance of the firm's brand image (Goldfarb and Tucker 2011). Eventually, such publicly accepted brand image reassures top of funnel customers of their belief in the brand's value

(Tajfel and Turner 1979; Lam et al. 2010). In other words, increased brand marketing advertising campaign in response to positive consumer sentiment helps promote customers' affinity with the brand and enabling increased purchase likelihood. Hence, I hypothesize

H2: During the prerelease period, launching firm adjusts their brand marketing spending by increasing both a) digital advertising and b) traditional advertising in response to the increased positivity of customer sentiment.

Prerelease New Product Promotional Advertising Adjustment in Response to the Increased Bullish Investor Sentiment.

Different from prospective customers whose ultimate objective is focused on new product purchase, investors tend to connect the new product release with the overall firm performance, which further guides their investment decisions (Pauwels et al. 2004; Rappaport 1999). Formally, investor sentiment refers to the varied degrees of investors' confidence (i.e., bullish vs. bearish) about the new product in terms of its impact on firm value (Baker and Wurgler 2007; Luo, Jiang and Cai 2015). The influence of investor sentiment on a firm's stock price, according to Malcolm and Wurgler (2007 p. 132), originates from the information asymmetry between the firm and the investors. According to these authors, investors are prone to speculate a firm's value when they have limited and uncertain information to determine the firm's profitability. In the context of the current study, when a new product is preannounced to the market, information about the product attributes and its success is rather limited. Thus, investors will have to rely on their and others' sentiment about the new product to make investment decisions.

An increased bullish sentiment among investors toward the new product represents their confidence about the new product's capability to enhance the launching firm's future cash flow. Here, managers of the launching firm have to utilize all marketing means to guarantee the new product success (e.g., sales growth), so as to meet such a high expectation held by investors

(Peltier, Zahav and Lehmann 2013; Voss and Voss 2008). Prior literature has recognized that advertising's role in promoting audience's product awareness and need recognition (Percy and Rossiter 1992). The information conveyed by the new product promotional ads (e.g., price, key attributes, quality, etc.) supports audience in better evaluating the new product benefits and stimulates their purchase decision (Devos et al. 2015; Nelson 1970, 1974; Stigler 1961). In this vein, managers of the launching firm, by employing product promotional advertising adjustment in response to investor sentiment, are able to expand the pool of prospective customers, improve the new product sales, and meet investors' bullish expectation.

Tactically, digital advertising is an efficient tool signaling the strength of the new product by allowing managers to customize and deliver the detailed messages to those bottom-of-funnel customers (Gordon et al. 2021; Sridhar et al. 2016). The more detailed information about the benefits of the new product customers can receive, the better chance they will make the purchase decision, leading to the increased new product sales (Lemon and Verhoef 2016). Compared with it, traditional advertising, with its simple ad content delivery is able to increase top-of-funnel customers' new product awareness and need recognition, resulting in new product sales subsequently (Ganesh, Arnold and Reynolds 2000; Morimoto and Chang 2006). Hence,

H3: During the prerelease period, launching firm adjusts their new product promotion spending by increasing both a) digital advertising and b) traditional advertising in response to investors' bullish sentiment.

Prerelease Brand marketing Advertising Adjustment in Response to the Increased Bullish Investor Sentiment.

Brand marketing advertising campaigns in response to investor sentiment signal investors the launching firm's marketing commitment to supporting new product success and the determination to meet the bullish expectation (Rubera and Kirca 2012; Tellis 2003). Specifically, the emotional appeal of brand marketing advertising campaign affectively influences audiences' affinity toward the brand and its products (Silverman, Sprott and Pascal 1999). In addition, it reminds audiences of the successful products introduced by the firm historically, which plays a critical role in molding these people's interest and confidence for the current one (Garber et al. 2004; Warren and Sorescu 2017). For example, when Tesla launched its Model 3 electric vehicle, the company did not deploy any advertising campaigns relevant to the new car. Instead, it deployed an advertising campaign inspiring audience's emotional attachment to the company as an industry icon and instilling in prospective consumers the notion that Tesla (as a brand) represents the future of the auto industry (Isidore 2019). The introduction of Model 3 led analysts to raise their price target from $368 to $411. Admittedly, the outstanding performance in the stock market hinges on investors' confidence in the Model 3's market potential. However, it is also relevant to Tesla's leading position, brand image, as well as the phenomenal success of its previous electric vehicle launches, which promote the consideration of Tesla among the general public.

Tactically, both digital and traditional advertising is likely to be increased in response to the bullish sentiment among investors. For digital advertising, its efficient targeting characteristics allow managers to identify audiences who are already aware of the new product and/or the brand. The personalized ad content promoting the launching firm's brand image is better able to remind these people of the launching firm's capability to keep their promise in terms of new product quality. It also reinforces their intimacy with the launching firm, as aforementioned. Thus, it drives new product sales by accelerating prospective customers adoption time (Chalil, Dahana and Baumann 2020; Yang and Ghose 2010). Brand marketing campaigns through traditional media vehicles are able to publicize the brand image of the launching firm among broader audiences, which is relatively more critical for people who are unfamiliar with the product or its brand yet. The general and easy-to-process brand message delivered through traditional media vehicle improves these people's awareness, which potentially expands the customer base (Argenti and Druckenmiller 2004). Thus, I hypothesize:

H4: During the prerelease period, launching firm adjusts their brand marketing spending by increasing both a) digital advertising and b) traditional advertising in response to investors' bullish sentiment.

DATA AND MEASUREMENT

The research context of this study is the video game industry well-known for its huge customer base and steady growth in the past few years. Based on the annual report from the Entertainment Software Association (2020), 75% of American families own at least one video game device in 2020. Over 64% of the adults and 70% of young people under the age of 18 claim themselves as regular video game players. In 2019, U.S. customers spent approximately $35.4 billion purchasing video games, which is a 2% growth over 2018. In addition, video games, similar to other entertainment products, usually have shorter life cycle. A large majority of the sales are realized in the first few weeks after the official release[21]. Therefore, managers of the launching firms are more motivated to closely monitor the market reactions after new product preannouncements and adjust their prerelease advertising deployment to maximize opening week sales.

For this study, I concatenated data from multiple sources, aiming at addressing the research questions. First, using vgchartz.com I identified the list of new video games and their weekly sales information and the names of the launching firms[22]. I used Compustat to locate

---

[21] This is further supported by my empirical data. Figure 3.1 presents the order volume of the 71 new video games launched between 2012 and 2017 in the North American market. On average, 21% of the order volume was realized within the first week since their official launch and 88% were achieved within the first year since launch.

[22] vgchartz.com is a website that tracks and records launching information and sales data of all video games worldwide. Extant prior literature (e.g., Marchand and Hennig-Thurau 2013; Xiong and Bharadwaj 2014) has used information from this site for

public companies among all those launching firms and acquired their financial information. Third, I used Factiva to identify the dates when these new video games were preannounced by the launching firms. Therefore, I could calculate the prerelease period which is the number of weeks elapsed between the preannouncement date and the official release date. Next, I used Kantar Ad$pender to acquire advertising expenditure information of these new video games during their prerelease periods. Ultimately, I collected 3,175 product-week observations of 71 video games launched in the North American market during 2012 and 2017 by seven public companies with an average of 45 weeks of prerelease period. In terms of genres, 19 (27%) of them are action, adventure and strategy, 18 (25%) are sports games, 17 (24%) are shooter games, and the rest of the games (24%) have genres, such as action, fighting, role-playing, etc. Figure 3.2 (Panel A and B) shows the distribution of all these video games by year and their launching companies respectively. In the rest of this section, I will share more details on the source and measurement of each variable of interest to this study.

Dependent Variable

I aim to explore the impact of managers' prerelease advertising adjustment in response to customer and investor sentiment in terms of new product sales. Specifically, I use the new product's opening week order volume (*log-transformed*) acquired from the vgchartz.com. As a top line measure, the use of opening week sales suits the objective of this study in two folds. First, it directly examines the effectiveness of the advertising adjustment in the sense that both the increase or decrease of advertising expenditure directly influence customers' cognitive, affective, and eventually behavioral activities toward the new product along their paths to purchase (Lemon and Verhoef 2016; Srinivasan, Rutz and Pauwels 2015). Second, given the

research. To ensure the data accuracy, I also cross-checked information about these new video games using other sources, including Wikipedia, STEAM, and Reddit.

relatively shorter life cycle of video games, managers are more eager to know whether their new video game will be a market hit during the opening week rather than waiting for a few months or years (Golder and Tellis 2004; Xiong and Bharadwaj 2014)[23].

Independent Variables – Launching Firms' Advertising Adjustment

Launching Firms' Advertising Deployment for New Product Promotion.

I used Kantar Media's Ad$pender to access the weekly advertising expenditure data of the launching firms at the product level. In general, Ad$pender tracks firms' advertising expenditures on four channels at both national and local levels: TV (i.e., cable, network, Hispanic, spot and syndicated), radio (local, network, and national), print (magazines and newspapers), digital (website display and search), and outdoor. Following the practice of Sridhar et al. (2016), I manually entered the query of each new video game's name on Ad$pender to extract its weekly advertising spends information. Then, I further grouped these weekly advertising spends into two categories: traditional ads and digital ads[24].

Launching Firms' Advertising Deployment for Brand marketing.

I first undertook an exploratory data extraction for brand marketing advertising expenditure made by the seven launching firms based on the 'corporate promotion' label provided by Ad$pender. However, the data I extracted are extremely scarce (48 observations for all 71 video games during their prerelease periods). This could be related to the fact that most video game products are published by a few big companies (e.g., Microsoft Xbox Studio, Sony

[23] In an additional analysis, I also examined the effect of the prerelease advertising adjustment on the launching firm's long term stock market value, measured by the buy-and-hold abnormal return difference between the launching firm and its matched counterpart. More detailed description on the variable measurement and the modeling results have been provided in the Additional Analysis section.

[24] Based on Ad$pender data, less than 8% of the sampled firms had regional/local advertising campaigns for their new video games during the prerelease period. Thus, I eventually classified the advertising spend into traditional and digital, rather than national, regional, and online suggested by Sridhar et al. (2016).

Entertainment Inc.), some specialist companies (e.g., Take-Two Interactive Software Inc., Electronic Arts Inc.), or their subdivisions (Zackariasson and Wilson 2012). In this oligopoly market, customers and investors are fairly familiar with those brand names *per se*. Therefore, the launching firms are more reliant on advertising of their incumbent products (same genre products of the launching firm that are actively sold in the market at the time of preannouncement of the new product) to differentiate themselves from the competitors, remind customers and investors of the network effects of their products, and boost their brand image (Hamilton 2009; Lee and O'Connor 2003). Accordingly, I measured the brand marketing advertising as the launching firms' advertising expenditures for their incumbent products. Based on this criterion, I searched vgchartz.com and finalized the list of incumbent products owned by the launching firms during their new video games' prerelease period. Next, I followed the same procedure to acquire the weekly advertising spend data of all these incumbent products for each launching firm from both traditional and digital media vehicles.

Independent Variables – Customer and Investor Sentiment

Customer Sentiment.

I use customer buzz measure as a proxy of prospective customers' reactions toward the new product introduction during the prerelease period. In line with prior literature, I extracted the textual data of customer buzz from an online gaming forum NeoGaf. Founded in 2006, NeoGaf is at present one of the largest online forums dedicated to video game products. As of January 2021, this forum has had over 150,000 registered members generating almost one million threads and over 131 million posts. Users of this platform engage in chatters for not only the existing video games but also those to be released to the market.

More specifically, I manually entered query of each sampled new video game in the search box of the website and recorded each post within the prerelease period and the posting time. Next, I used a machine-learning-based Gated Recurrent Unit (GRU) model to evaluate the valence of each post[25]. Based on the GRU evaluation, each textual post is assigned a valence score ranging from 0 to 1, with 0 as the least positive and 1 as the most positive. Then, I calculated the mean value of the valence scores of all posts pertaining to one video game on the weekly basis to capture the prospective customers' overall sentiment toward the new product.

Investor Sentiment

I selected StockTwits.com to extract investors' online chatter about those new video games and use it as a proxy of the investors' reactions toward the new product during the prerelease period. StockTwits.com, founded in 2009, is an online social media platform (similar to Twitter) where investors may exchange investing information, voice opinions about investment ideas, and vent their sentiment about companies' profitability. So far, there have been more than two million registered users worldwide according to Investopedia.com. For the purpose of this study, I manually searched for investor posts relevant to each sampled new video games during the prerelease period and recorded their opinions about the stock price of the launching firm due to the new product introduction (i.e., bullish or not). I then calculated the percentage of bullish posts on the weekly basis[26]. A higher value of this measure implies that there are more investors who are bullish about the new product success and its positive impact on the firm value.

---

[25] The GRU model of this chapter has the same configuration as those described in Exhibition 2.4 of Appendix 1.

[26] On StockTwits.com, there is an option for users to label their opinions about a certain stock as bullish, bearish, or neutral. I collected this information and recoded them as 1 (bullish) and 0 (not bullish). In addition, I also trained a GRU model (the same as that for customer buzz data) and evaluated the valence of investors' sentiment. The GRU estimated results are very close to the bullish/bearish/neutral labels self-reported by the investors themselves (> 90%). To more accurately reflect investors' opinions, I eventually determined to use those labels directly.

Control Variables

Besides of the independent variables aforementioned, I recognize there could be a dynamic system where both customer and investor sentiment and launching firm's advertising adjustment are interdependent on other factors (e.g., media coverage of the new product and the competitors' reactions). Both Mizik and Jacobson (2003) and Srinivasan et al. (2009) have shown that firms' marketing activities directly influence investors' perception of firms' profit potential, eventually their stock returns. On the customers' side, a large body of research has empirically documented the feedback effect of firms' marketing communication (e.g., advertising) on both the volume and valence of customer word-of-mouth (e.g., Hewett et al. 2016; Srinivasan, Rutz and Pauwels 2016; De Vries, Gensler and Leeflang 2017). Rinallo and Basuroy (2009) have discussed changes in advertising expenditure may increase or decrease media coverage of the firm. Numerous studies (e.g., Bowman and Gatignon 1995; Dube and Manchanda 2005) have also disclosed that competitors promptly respond to a firm's new product introduction via their advertising expenditure adjustment. Finally, there could be interactions among customers, investors, media and competitors. For example, customers after seeing so much media coverage of the new product may become more active in expressing their emotion or feedback on social media platforms (see: Lemon and Verhoef 2016). Competitors' advertising expenditure adjustment in reaction to the launching firm's new product introduction may inform investors of market competition or remind them of the superiority of the new product, rendering the changes with their expectations on the new product (Robertson, Eliashberg and Rymon 1995). Hence, I include the following control variables in the modeling system.

Media Coverage of the New Product

Extant research in both marketing and business strategy disciplines has articulated the essential role of media coverage in shaping customers' perception of a firm's brand image and product quality and eventually their purchase behavior. To capture media coverage measure, I used *Factiva* to extract media coverage records related to each new video game during the prerelease period. Next, I calculated the total number of media coverage records[27] on the weekly basis and include it into the modeling analysis.

Competitors' Reaction to the Upcoming New Product Launch.

I measure the competing firms' reaction with the use of the total advertising expenditure on their incumbent products within the same genre as the new video game. Following the same procedure, I acquired advertising expenditure data from Ad$pender by manually entering the names of those incumbent video games and their affiliated firms.[28] Next, I grouped the weekly advertising spends across different media vehicles into two categories: traditional and digital advertising.

Other Control Variables

Consistent with prior studies, for the new product sales model, I also control for the carry-over effect of the franchised video games (e.g., Madden NFL video game series, NBA series, etc.). In this regard, sales of a new version of the franchised video game may be influenced by the popularity of the previous ones. Following Bayer et al. (2020) and Danaher, Bonfrer and Dhar (2008), I calculated the one-year ad stock variable of all franchised video games in my dataset ($N = 17$) to proxy the carry-over effect and put it into the sales model as

[27] The final measure is the natural logarithm transformed value of the actual count of media coverage plus one, due to the non-negative distribution of the raw data.

[28] The list of incumbent video games owned by competing firms is created based on information from vgchartz.com and cross-checked based on information from Wikipedia.org, STEAM.com, and Reddit.com.

control[29]. Specifically, I set $Adstock_{it} = \psi\ Adstock_{i,t-1} + (1 - \psi)\ advertising\ spending_{it}$, where $Adstock_{it}$ refers to the advertising stock value of the franchised video game *i* at time *t*. ψ is a smoothing parameter ranging between 0 and 1, whose value is determined based on the grid search algorithm proposed by Dinner et al. (2014) and Bayer et al. (2020).

Besides, I also control for the total number of pre-orders (*log-transformed*) of the new product, holiday season fixed effect, number of video game console availability, total number of firm PR releases, and game genres (see: Chintagunta, Gopinath and Venkataraman 2010; Xiong and Bharadwaj 2014). For the firm long-term value model (an additional study explained below), I account for effects from the launching firm's advertising expenditure and R&D intensity, firm size, leverage, book-to-market ratio, total PR releases since new product preannouncement till the end of its first anniversary of the official launch, and the number of historical new product releases of the launching firm (see: Saboo, Chakravarty and Grewal 2016; Sorescu and Shankar 2007; Warren and Sorescu 2018). Table 3.2 below summarizes the detailed information of variable measurement and data source.

MODELING PROCEDURE

Three-step Modeling Procedure

The main objective of this study is two-fold. First, I would like to investigate whether managers dynamically adjust their prerelease advertising campaign (both strategically and tactically) in response to customer and investor sentiment. Second, are these adjustments beneficial for launching firm, in terms of new product sales and firm value (additional study)? Addressing these two questions demands an approach that: 1) examines the validity of the prerelease advertising adjustment across all firms and their new products; and 2) specifically

[29] The list of the franchised video games is cross-checked based on information from Wikipedia.org and vgchartz.com.

ascertain how varied magnitudes and directions of such an adjustment for each new product may lead to different outcomes.

Hereby, I devised a three-step modeling procedure. Specifically, I first utilized the panel VAR (PVAR) model to analyze the panel dataset and explore the evidence of managers' prerelease advertising adjustment in response to customer and investor sentiment aggregated across all firms and all new products. In addition, the use of PVAR allows me to account for the dynamic feedback effects not only between the launching firms and its customers and investors, but also among other market participants, including media and competitors.

As the second step, I exploited Bayesian VAR (BVAR) model to ascertain the varied directions and magnitudes of the prerelease advertising adjustment at the product level and extracted the impulse response function coefficients (IRFs) for each significant prerelease advertising adjustment (empirically supported based on the PVAR results) for each new product. These IRFs serve as the proxies of the launching firms' prerelease advertising adjustment in response to customer and investor sentiment for each new product. Moreover, due to the extremely small number of observations for some new video games (e.g., five observations) and the relatively large number of independent variables (i.e., 37 advertising adjustment schemes), the use of BVAR allows me to set up informative priors and estimate the product-level IRF coefficients with the 'borrowed' information from firm and genre level models.

At the third step, I used the Bayesian Hierarchical Model to investigate how different values of the IRFs, may impact the ultimate outcome (i.e., new product sales and the firm value in the additional analysis). The use of Bayesian Hierarchical Model deals with the heterogeneity concerns at both firm and genre levels. In this regard, the effectiveness of a launching firm's advertising campaign is often constrained by the availability of its resources, the capabilities to flexibly make advertising adjustment, and the constraint of the product category (i.e., game genre for this study). While large firms (relative to smaller one) usually have bigger R&D team, more resources, and affluent marketing investment for the new product design and promotion (Ettlie and Rubenstein 1987; Garcia and Calantone 2002), smaller firms are often more agile detecting and responding to customer needs for new product introduction (Acs and Audretsch 1990). In terms of genres, same advertising campaign may lead to quite different video game sales performance due to varied customer bases for each game genres. Based on the 2018 data from statista.com[30], 38.5% of total sales volume were generated by video games with genres of action, adventure, and strategy, followed by the shooter games (20.9%). In contrast, racing games merely contributed 5.8% of the total sales volume.

For the rest of this section, I will further discuss the necessity of such a three-step modeling procedure setup and offer more detailed description on the implementation of each step.

Step One: Panel Autoregression Model (panel VAR)

At this step, I employed a panel vector autoregression model (panel VAR) to holistically evaluate whether managers of the launching firm make dynamic advertising adjustment in response to customer and investor sentiment. It is a fairly standard method that has been applied in many prior studies. For example, Dewan and Ramaprasad (2014) revealed the interdependence among customer buzz, radio play, and music sales, based on the evaluation of panel VAR model. Chen, De and Hu (2015) used panel VAR model to find that there is a significant interrelationship between social media promotions and music sales. In particular, user-generated social media messages plays a more positive role in driving sales than firm-

[30] https://www.statista.com/statistics/189592/breakdown-of-us-video-game-sales-2009-by-genre/

generated ones on the platform. In finance, Love and Zicchino (2006) employed panel VAR model to uncover the dynamicity between a country's financial advancement and its investment behaviors. In terms of this study, the use of panel VAR model is notably appropriate for the research objective, as it does not assume any *a priori* exogeneity of the independent variables. Instead, it treats all variables in the system as jointly endogenous and evaluate how changes with one variable at one period may lead to the changes with other variables and itself in the following period (Adomavicius, Bockstedt and Gupta 2012). Furthermore, the traditional VAR model can merely evaluate one cross-sectional unit at a time and requires a relatively large number of time series observations of that unit (Holtz-Eakin, Newey and Rosen 1988). In contrast, the panel VAR model, taking advantage of the panel data nature, allows me to exploit instrument variables within the model system (e.g., lagged dependent variables in GMM) to estimate the cross-sectional effect of all endogenous variables, while accounting for the unobserved individual heterogeneities. Thus, it relaxes the requirement of long time series observations within each cross-sectional unit (Abrigo and Love 2016; Holtz-Eakin, Newey and Rosen 1988; Love and Zicchino 2006). For this study, in particular, I collected a dataset of 3,175 product-week observations. While a few video games had a fairly long prerelease period (>100 weeks), some others have very short prerelease period (5 weeks). It is more efficient for me to use panel VAR model to complete the estimation.

Model Specification.

I specify the panel VAR model as follows:

$$\begin{bmatrix} DigAd_{i.t.f}^{prod\ promo} \\ TraAd_{i.t.f}^{prod\ promo} \\ DigAd_{i.t.f}^{corp\ brand} \\ TraAd_{i.t.f}^{corp\ brand} \\ Buzz_{i.t}^{cust} \\ Bull_{i.t}^{inv} \\ \vdots \end{bmatrix} = \begin{bmatrix} \alpha_{01} \\ \alpha_{02} \\ \alpha_{03} \\ \alpha_{04} \\ \alpha_{05} \\ \alpha_{06} \\ \alpha_{07} \\ \alpha_{08} \\ \alpha_{09} \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{19} \\ \vdots & \ddots & \vdots \\ \alpha_{91} & \cdots & \alpha_{99} \end{bmatrix} \sum_{p=1}^{P} \begin{bmatrix} DigAd_{i.p.f}^{prod\ promo} \\ TraAd_{i.p.f}^{prod\ promo} \\ DigAd_{i.p.f}^{corp\ brand} \\ TraAd_{i.p.f}^{corp\ brand} \\ Buzz_{i.p}^{cust} \\ Bull_{i.p}^{inv} \\ \vdots \end{bmatrix} + \boldsymbol{\varphi Exog_{i.t.f}} + \boldsymbol{\mu_{i.t.f}}$$

[Equation 1]

, where $DigAd_{i.t.k}^{prod\ promo}$ and $TraAd_{i.t.k}^{prod\ promo}$ respectively denote the launching firm *f*'s prerelease digital and traditional new product promotional advertising expenditures for new product *i* in week *t*, $DigAd_{i.t.k}^{corp\ brand}$ and $TraAd_{i.t.k}^{corp\ brand}$ respectively refer to the launching firm *f*'s digital and traditional prerelease brand marketing advertising expenditures for new product *i* in week *t*, $Buzz_{i.t.k}^{cust}$ refers to customer positive sentiment for new product *i* in week *t* during the prerelease period, and $Bull_{i.t}^{inv}$ is the investors' bullish sentiment for the new product *i* in week *t* during the prerelease period. For the sake of brevity, non-focal endogenous variables (i.e., weekly media coverage for the new product, competitors' weekly digital and traditional advertising expenditures for their incumbent products) are not listed in Equation 1. $\boldsymbol{\alpha}$ is the matrix of coefficients to be estimated. p refers to the lag period of the PVAR model. At this point, I set $p = 1$ based on the GMM-based J statistic. $\boldsymbol{Exog_{i.t.k}}$ is a vector of exogenous variables controlled in the PVAR system (i.e., weekly preorder volume and the number of historical new product releases made by the launching firm as of week *t*). $\mu_{i,k}$ is the idiosyncratic error term.

The panel VAR model in this study is estimated based on the standard GMM method documented by Abrigo and Love (2016). Before the estimation, I first undertook statistical tests to examine the issue of stationarity and determine the number of lags. For the test of stationarity, I applied augmented Dickey Fuller Fisher tests (Choi 2001) to the nine endogenous variables. The results were insignificant, which implies there could be unit root issue with the raw dataset. Therefore, I follow the previous studies (e.g., Abrigo and Love 2016; Enders and Sandler 1993) to first difference these endogenous variables and take the test again. The final results show that all those first-differenced variables are stationary. Next, I took a series of statistical measures to determine the optimal number of lags in the panel VAR system. At this point, I used the GMM-based J statistic of overidentifying restriction test developed by Andrews and Lu (2001). Similar to the lag selection criteria used for regular VAR model (i.e., AIC, BIC, and HQIC), the J statistic based MAIC, MBIC, and HQIC metrics evaluate the optimal number of lags in the panel VAR system. Moreover, it allows the overidentification situation when the number of instrumental parameters exceeds that of observations, which is quite common for panel VAR models (For detailed description of the method, see: Andrews and Lu 2001; Hansen 1982). I eventually selected lag one for the model evaluation (Equation 1), as the J statistic based MAIC, MBIC, and HQIC metrics are the smallest (i.e., optimal) compared with those estimated values based on more than one lag for each of the endogenous variables and that for all variables (see: Table 3.3). After the model estimation, I took the stability test by computing the modulus of each fitted model's eigenvalue. According to Lutkepohl (2005), all the eigenvalues of the moduli should be less than one for the panel VAR model to be stable. Based on the results, the panel VAR model is stable (see Figure 3.3).

Step Two: Product-level Bayesian Vector Autoregression Model (BVAR)

I now move on to the product level and examine whether such dynamic prerelease advertising adjustment affects new product performance[31]. Due to the complex dynamicity and

---

[31] I only consider quantifying prerelease advertising adjustment schemes that are significant based on Step One results.

multicollinearity of the endogenous variables in the dynamic feedback system, I cannot directly use the estimated coefficients to proxy the prerelease advertising adjustment schemes (Ramos 1996; Sims 1980). In this regard, the Impulse Response Function (IRF) allows me to estimate the *sheer* reactions of all other endogenous variables in the system over time, by assigning one endogenous variable as a shock (Litterman 1984; Bronnenber and Mahajan 2000). In other words, the IRFs are exogenous by nature, which makes the Step Three estimation easier to execute.

I used the impulse response functions (IRFs) of Bayesian Vector Autoregression (BVAR) model at the product level as the measure of prerelease advertising adjustment in response to respective customer and investor sentiment. Using IRFs of PVAR model is not feasible, as they represent the average level of prerelease advertising adjustment in response to customer and investor sentiment across all launching firms and their new products. I might be able to run separate vector autoregression (VAR) models for each product to extract IRFs at the product level. However, regular VAR model requires a large number of observations within a time series to make efficient estimation (Holtz-Eakin Newey and Rosen 1988; Chakravarty and Grewal 2011). The use of BVAR allows me to analyze a dataset with small number of observations but large number of parameters (i.e., over-parameterization) by imposing either non-informative or informative prior distributions to the parameters, which results in the shrinkage of the parameter space and consistent estimation of the parameters (Doan, Litterman and Sims 1984; Leeper et al. 1996).

In this regard for the product-level IRF estimations, I used informative priors based on the posterior estimates of both firm- and genre-level BVAR models[32]. The use of informative (relative to non-informative) priors for product-level IRF estimation has both methodological and practical advantages. First, it avoids the consequence that the estimates of the most parameters are shrunk to zeros. This is because most non-informative priors follow the normal distribution $N \sim (0, \sigma^2)$. Second, as discussed earlier, different launching firms possess different levels of resources and capabilities. New products from different categories (i.e., genres) have quite different customer bases and potential sales volumes. Following this logic, it is reasonable to believe that managers from the same launching firm promoting new products with the same product category (i.e., genre) may respond to customer and investor sentiment in a similar manner, which is unobservable to the researcher. Therefore, I first ran BVAR models at the firm- and genre-levels to evaluate new product managers' different prerelease advertising adjustment behaviors (in response to the customer and investor sentiment). Next, I converted the hyperparameters estimated based on the posterior distributions of both models as the informative priors of the product-level BVAR model. By so doing, I not only overcome the over-parameterization issue but also account for the potential heterogeneity concerns at the firm- and genre-levels.

Specifically, I set the three BVAR models (firm level, genre level, and product level), as follows:

$$Y_{t.f} = B_0 + B_1 y_{t-1,f} + \cdots B_9 Y_{t-m,f} + \epsilon_{t,f}, \qquad \epsilon_{t,f} \sim N(0, \Sigma) \qquad \text{[Equation 2]}$$

[32] It could be feasible to set up a hierarchical BVAR model to account for the heterogeneity concerns at both firm and genre levels. However, it does not resolve the small number of observation issue for some new video games. Thus, the setup of informative priors allows me to 'borrow' information at the firm and genre levels and avoid the estimated values of those IRFs to be shrunk to zero for those new products with fewer number of observations.

, where $Y_{t.f}$ denotes an $n \times 1$ vector of the nine endogenous variables (the same as those in Equation 2) in week *t* for firm *f*, $\epsilon_{t,f}$ is the $n \times 1$ vector of the error terms for firm *f*, where the value of n depends on the actual number of observations for each firm. $B_0, B_1, \ldots, B_9$ as well as the Σ refer to the matrices of unknown parameters to be estimated in the BVAR model. I set $m = 1$ based on Step One results showing lag 1 is optimal. The genre level and the product level model follow the same set-up. The only difference is the observation level. At the firm level, I have $Y_{t,f}$, $Y_{t-m,f}$, and $\epsilon_{t,f}$. At the genre and product levels, these vectors are replaced with $Y_{t,g}$, $Y_{t-m,g}$, $\epsilon_{t,g}$ and $Y_{t,p}$, $Y_{t-m,p}$, $\epsilon_{t,p}$ respectively.

For the firm- and genre-level BVAR models, I allow the distribution of the vector of endogenous variables to be multivariate normal. This is consistent with prior literature in both marketing and econometrics (e.g., Chakravarty and Grewal 2011; Del Negro and Schorfheide 2011; Koop and Korobilis 2010; Zeithammer and Lenk 2006). With respect to the prior distribution, I set the Minnesota priors of the three hyperparameters (i.e., hyperpriors), λ, SOC, and SUR, using Gamma densities with mode equal to 0.2, 1, and 1 and standard deviations equal to 0.4, 1, 1 respectively. This is recommended by Sims and Zha (1998) as a default option. At this point, the essential part of the Minnesota prior is the hyperparameter λ. It controls the degree of all variances and covariances within the BVAR system, eventually leading to the variation in the tightness of the Minnesota prior. In addition to λ, Doan, Litterman and Sims (1984) and Sims (1994) respectively propose the use of two additional hyperparameters: sum-of-coefficients (SOC) and single-unit-root (SUR) to collectively account for the unit root and cointegration concerns[33].

[33] For more detailed description and mathematical deduction of the three hyperparameters, see Giannone, Lenza and Primiceri (2012).

At the end of the BVAR estimations of the firm- and genre-level models, I extracted two sets of the optimal posteriors (i.e., firm-level and genre-level) of those hyperparameters. Finally, I took the mean values of the estimated firm- and genre-level posteriors for each new product and used them as the informative priors of the BVAR model for the product-level IRFs estimation.

Step Three: Bayesian Hierarchical Model

At the end of Step Two, I eventually configurated a dataset comprised of all significant prerelease advertising adjustment schemes at the product level (i.e., IRFs), along with all other control variables discussed in the previous section. The final step is to evaluate how managers' varied prerelease advertising adjustments (in response to customer and investor sentiment) affect the new product sales. To address this question, I need account for the heterogeneity issue in terms of advertising effectiveness estimation, where the same advertising deployment may present quite different effects in terms of sales due to the genre-level heterogeneity (see: Sethuraman, Tellis, and Briesch 2011)[34]. Second, the final sample I have is a cross-sectional dataset with 71 distinct new video games, along with 37 significant advertising adjustment schemes (i.e., IRFs from Step Two) that I would like to evaluate their disparate impacts on launching firms' business performance, and other control variables. Such a set-up results in an over-parameterization issue where the number of independent variables (both IRFs and control variables) is quite large, more than 65% of the observations in the dataset. In addition, the final dataset is also rather sparse. Not all managers of all new games respond to customer and investor sentiment at all times. Hence, there are many zero values for those advertising adjustment schemes (i.e., IRFs) in the observation. Typical asymptotic methods are efficient in coefficient

[34] I also examine the firm-level heterogeneity. However, the results are not significant. Thus, I only focus on the genre-level heterogeneity in Step Three.

estimation, only if there are enough observations (e.g., 1,000 observations per parameter for choice modeling; see: McCulloch and Rossi 1994; Rossi and Allenby 2003). The Bayesian Hierarchical Model, at this point, outperforms its frequentist counterparts in two folds. First, it is natural and very straightforward to set up a hierarchical structure in parameter estimation where one may estimate a subset of parameters at the first stage, then move on to the second stage for the estimation of the remaining parameters. In this regard, the number of parameters to be estimated at each step is tremendously reduced and the over-parameterization issue can be solved. Second, the Bayesian Hierarchical Model is capable of dealing with the large number of parameters based on a relatively smaller number of observations. By setting up a non-informative normal prior per parameter, the analyst actually imposes a phenomenon called 'shrinkage' to the posterior estimation, where the estimates of outliers are pushed toward the prior mean (i.e., zero for non-informative prior) rather than the unit-level maximum likelihood (Rossi and Allenby 1993; 2003). Hence, the number of parameters to be estimated will tremendously decrease, leaving the situation more manageable.

Specifically, I established the Bayesian Hierarchical Model for this study as follows:

Level 1- Product Level:

$$y_i | \mu_i, \sigma^2 \sim N(\mu_i, \sigma^2) \quad \text{Equation [3]}$$

$$\mu_i = \beta_0 + \sum_{p=1}^{P} \beta_p X_{ip} + \boldsymbol{Ctrl_i} \quad \text{Equation [4]}$$

, where y_i is the opening week sales volume of each new video game. $p = 1, 2, \ldots 37$ denotes the number of advertising adjustment schemes, β_p is the estimated coefficient of effectiveness of adjusting advertising based on customer and investor sentiment, and X_{ip} is the vector of product-level IRFs of each scheme (i.e., capturing the effects of customer and investor sentiment on strategic and tactical advertising deployment). $\boldsymbol{Ctrl_i}$ refers to a vector of control variables to be accounted. For the new product sales model, I account for the carry-over effect of franchised video games, holiday season fix effect, the number of console platform available for new video game, corporate PR releases for the new product during the prerelease period, and total number of pre-orders (*log*).

Level 2 – Genre Level:

$$\beta_p | \nu_g, \tau^2 \sim N(\nu_g,\ \tau^2) \qquad \text{Equation [5]}$$

$$\nu_g = \sum_{g=1}^{G} \gamma_g Genre_g \qquad \text{Equation [6]}$$

At Level 2, I am interested in examining the genre-level heterogeneity in terms of advertising effectiveness. Therefore, I measure γ_g – the unknown coefficient evaluating how different genres (i.e., $g = 1, 2, 3, 4$) may lead to differential advertising effectiveness captured by ν_g. Following prior literature in both marketing and statistics (e.g., Gelman 2006; Hoff 2009; Rossi and Allenby 2003; Suk, Chintagunta and Park 2019), I set up non-informative priors of product-level parameters ($\boldsymbol{\beta}$) and genre-specific parameters ($\boldsymbol{\gamma}$) to follow a multivariate normal distribution $\beta_0, \beta_1, \dots\ \beta_P \sim MVN(0, \kappa^2)\ for\ some\ \kappa > 0$ and $\gamma_1, \dots\ \gamma_G \sim MVN(0, \partial^2)\ for\ some\ \partial > 0$. The priors of the variances of $\boldsymbol{\beta}$ (κ) and $\boldsymbol{\gamma}$ (∂) are inverse gamma distributed as $\kappa \sim igamma(shape = 0.01, slope = 0.01)$ and $\partial \sim igamma(shape = 0.01, slope = 0.01)$. Finally, priors for all control variable parameters are also set to be normally distributed and non-informative.

For the actual estimation, I employed MCMC methods with three chains and 50,000 iterations. The first 20,000 were discarded as burn-in, while the rest 30,000 were used for approximation. Convergences were achieved based on a series of diagnostics measures, including Monte Carlo standard errors and Geweke diagnostics (Suk, Chintagunta and Park 2019).

RESULTS

Prerelease Advertising Adjustment in Response to Customer and investor sentiment.

As a reminder, I used PVAR model to holistically investigate firms' prerelease advertising adjustment in response to customer and investor sentiment, while the BVAR model is set up at the product level for the sake of IRF extraction. Based on the PVAR results (see: Table 3.4), both lagged customer sentiment and investor sentiment serve as significant information sources for launching firm managers' prerelease advertising adjustment. For new product promotion advertising campaigns, managers tend to increase both digital ($\alpha = 4.624, p < .001$) and traditional ($\alpha = 4.624, p < .001$) advertising in response to the increased positivity of customer sentiment in the previous week. Thus, H1 is supported. Regarding brand marketing advertising, managers also tend to increase both digital ($\alpha = 0.097, p < .01$) and traditional ($\alpha = 0.653, p < .001$) ads in response to the increased positivity of customer sentiment in the previous week. Therefore, H2 is supported. On the investors' side, the increased investor bullish sentiment from the previous week leads to the decreased expenditures on new product promotional advertising campaigns for both digital ($\alpha = -3.017, p < .001$) and traditional ($\alpha = -0.951, p < .001$) deployments. H3 is rejected. I will elaborate on this results in the discussion section. Finally, the PVAR finds that managers of the launching firm tend to decrease their digital advertising expenditure ($\alpha = -0.158, p < .001$) for brand marketing campaigns (in response to increased investor bullish sentiment), while increasing spends for brand marketing campaigns deployed through traditional media vehicles ($\alpha = 0.177, p < .001$). H4 is partially supported, and I will discuss the implication in the following section. All these results were bolstered by the Granger Wald Test, whose results were presented in Table 3.5.

The Impact of Prerelease Advertising Adjustments on New Product Sales

In terms of new product opening week sales, the results of the Bayesian Hierarchical Model (see: Table 3.6 Panel A) shows that only the increase in advertising expenditure for new product promotional campaigns through the traditional media vehicles exerts significant and positive effect on the new product's opening week sales. No other prerelease advertising adjustments in response to customer and investor sentiment significantly affect new product's opening week sales. Furthermore, such significant effects are only observed for video games with sports genre ($\beta = 22.916$) and those from fighting, racing, music genres (i.e., genres labeled as 'others'; $\beta = 17.099$).

ADDITIONAL ANALYSIS

The Effect of Prerelease Advertising Adjustment on Long-term Firm Value

As an additional analysis, I am interested in whether such a prerelease advertising adjustment in response to customer and investor sentiment may usher into any impact on the launching firm's long-term stock market value. I used the buy-and-hold abnormal return differences between the launching companies and their matched counterparts (*hereinafter* BHAR-Diff) as the firm value measure. Similar to other stock market measures, BHAR-Diff evaluates the present value of the expected future cash flows of a firm after an event occurs. Thus, it is forward looking (Srivastava, Shervani and Fahey 1999) and has been widely used in the marketing discipline in the past to proxy firms' longer-term values (e.g., Mizik and Jacobson 2007; Shivakumar 2000; Saboo, Chakravarty and Grewal 2016).

Specifically, I calculated the BHAR-Diff by extracting the difference between the one-year buy-hold abnormal returns of the launching companies due to the new product release and that of matched companies during the same time period. The data source is Compustat. In line with prior literature (Barber and Lyon 1997), I establish the following selection criteria for matched companies: 1). Both the launching and matched companies are from the same industry (i.e., the same two-digit Standard Industrial Classification / SIC code); 2). The matched companies shall have market values within 70% to 130% of that of the launching companies; 3). The launching and the matched companies shall have the closest book-to-market values. (Exhibition 3.1 of Appendix 2 lists all matched companies accompanied with the launching companies, based on the three criteria.) The one-year estimation window reflects the launching companies' new product introduction cycle. Based on my empirical data observation, over 88% of sales volumes of these new video games were realized within the first year since official release (see Figure 3.1). Moreover, many of these new video games were replaced by their sequels on a yearly basis (e.g., NBA 2K12, NBA 2K13, Gears of War, Gears of War 2, 3, 4 etc.). Eventually, the one-year window aligns with investors' anticipation of the discounted net present value of the cash flows generated by these new video games since their official launch date. Mathematically, I define the BHAR-Diff as

$$BHAR - Diff_{it+P} = \prod_{p=1}^{P} 1 + R_{ip} - \prod_{p=1}^{P} 1 + R_{matched,\ p} \qquad \text{Equation [7]}$$

, where R_{ip} denotes the holding returns for company i during the time period P, $R_{matched,\ p}$ is the holding returns of the matched company for company i during period P. As discussed earlier, I set $P = 1$ to align with the typical life cycle of new video games.

While the set-up of the Bayesian Hierarchical Model is the same as the opening week sales model in Step Three, the list of control variables is different. Consistent with Saboo, Chakravarty and Grewal (2016); Sorescu and Shankar (2007), Warren and Sorescu (2017), I controlled the launching firm's advertising and R&D intensities, leverage, book-to-market ratio, one-year advertising expenditure (*log-transformed*) since new product launch, corporate PR releases for both prerelease period and the following one year since new product launch, total number of pre-orders (*log-transformed*).

The modeling results show that prerelease advertising adjustment for both new product promotional and brand marketing campaigns significantly and positively drive the launching firms' BHAR-Diff (Table 6 Panel B). Same as the opening week sale model, these significant results are only observed for advertising adjustment in response to customer sentiment. Specifically, the increased traditional advertising expenditure for the new product promotional campaigns enhances the launching firm's BHAR-Diff when the new video game's genre is sports ($\beta = 73.889$) or fighting, racing, and music ($\beta = 55.409$). In the similar vein, the increased traditional advertising expenditure for brand marketing campaigns enhances the launching firm's BHAR-Diff when the new video game's genre is sports ($\beta = 46.470$) or fighting, racing, and music ($\beta = 34.830$).

DISCUSSION

In this study, I investigate and find support that managers adjust their advertising deployment (both strategically and tactically) in response to customer and investor sentiment during the prerelease period of their new product launch. These adjustments originate from managers' uncertainty about customers' and investors' perception of the new product value and its potential impact on the firm's future cash flow generation. In this regard, customer sentiment represents prospective customers' enthusiasm to possess the new product. It positively correlates with the new product post-release sales, thus motivating the launching firm managers to increase both their new product promotional and brand marketing advertising campaigns through both digital and traditional media vehicles. Such that, customers holding positive sentiment toward the new product will more quickly convert.

Investor sentiment (in terms of a new product launch) signifies their confidence (i.e., bullish belief) in the new product's contribution to the firm value and the launching firm's marketing commitment to supporting the new product's success. It motivates managers to increase their prerelease new product promotional and brand marketing ads to expand their customer base, eventually enhancing the probability to meet investors' expectation.

The modeling results show, however, that launching firm managers decrease their digital advertising for both new product promotional and brand marketing campaigns, in response to the increased investor bullish sentiment. This could be related to digital ads audiences' inactivity. One common billing practice of digital advertising is pay-per-click (i.e., PPC), which prescribes publishers to charge the advertiser only if the targeted audience clicks a paid ad on a website or the search engine result page (SERP). It leads to a seemingly paradoxical situation. On the one hand, the advertiser (i.e., launching firm) wants to spend more on its digital campaign for new product promotion. On the other hand, the reached audience rejects to interact with the advertiser, resulting in the decrease of digital advertising expenditure. In fact, such an 'under-spending' situation is quite common in the real world, when managers are faced with pressures to spend money (Keehler 2020).

In terms of traditional advertising, managers of the launching firm decrease their spends on new product promotional campaign, while increasing expenditure on brand marketing campaigns. One possible reason could be the budget constraint which is unnecessarily dealt with in the current study. Faced with limited advertising budget, managers are more likely to spend media dollars on campaigns that are capable of driving overall sales. At this point, brand marketing campaigns through traditional media vehicle not only supports audiences' affinity with the brand for the purchase of the new product, but also bolster sales of other product owned by the brand. Thus, it is more likely to improve the launching firm's overall performance and meet investors' bullish expectation.

Regarding the impact of the prerelease advertising adjustment on new product opening week sales, customer sentiment is the only significant and effective factor guiding managers' prerelease advertising adjustment. Customers are the end-buyers of the product. Prerelease advertising adjustment based on their sentiment directly drives sales. It is worth noting, however, only the adjustment of those new product promotional ads through the traditional media vehicles are drivers of the new product opening week sales, while digital ads are not significant. This result, to some degree, supports the findings of many recent studies about the inefficiency of digital advertising due to ad frauds, ad blocking, and measurement errors (see: Deighton, Mela and Moorman 2021; Gordon et al. 2021). With respect to the link between the brand marketing ads and the new product sales, adjustment in neither digital nor traditional campaigns exert any significant effects. Brand marketing ads (relative to new product promotional ads; Bass et al. 2007) takes longer time to affect customers' purchase behavior. Considering the short lifecycle of video games, its effect comes too late and too slow (as shown by significant effects on long term firm value).

In terms of the impact on long-term firm value, the result implies that both new product promotional and brand marketing campaigns through traditional media vehicles drive the increase in launching firms' buy-and-hold abnormal returns relative to their matched counterparts (BHAR-Diff). The former drives product sales, the latter builds up the brand equity and improves customers' positive emotion, leading to improved customer loyalty and steady future cash flows. Both of them together positively influence the enhancement of the launching firms' stock market value. At this point, digital advertising (either new product promotional or brand marketing) is direct response driven but simultaneously short lived (Sridhar et al. 2016). In addition, its limited audience reach and the inefficiency due to the aforementioned reasons hamper its influence on the long-term firm value.

The current study offers theoretical implications for future research in the literature as follows. First, it illustrates the differential roles of customer and investor sentiment in guiding managers' prerelease advertising adjustment. It offers a novel perspective for future research by treating advertising adjustment as an endogenous factor influenced by not only advertisers but also marketplace participants (i.e., customers and investors). Second, it shows the impact of such prerelease advertising adjustment in terms of both short-term new product sales and long-term firm value. In either case, traditional advertising (relative to digital advertising) is more effective. This is quite interesting and resonates one research priority proposed by the MSI between 2020 and 2022, in terms of the inefficiency of digital advertising. It is well accepted digital advertising is better able to reach finely defined audience and audience segments due to the targeting technology. However, it is highly bounded by the accuracy of ad effect measurement, audience's willingness to interact with the advertiser, and ad fraud (see Gordon et al. 2021). Thus, future studies may shed light on to what extent various inefficiencies aforementioned may influence the advertising effectiveness and their contingencies.

Practically, this study decomposes advertising deployment at both strategic and tactical levels. It well aligns with managers' advertising planning practice, when these managers must determine how to deploy both new product promotional and brand marketing campaigns through both digital and traditional media vehicles. Varied combinations of these components usher into different advertising effects in terms of new product sales and firm value. It implies that new product managers when executing their advertising campaigns should have a clear vision in terms of their emphasis on either short-term sales, long-term firm value, or both.

LIMITATION AND FURTHER STUDIES

The current study contributes to the literature in three folds. First, it recognizes the differential values of customer and investor sentiment as an information source driving managers' prerelease advertising adjustment. Second, it advances our understanding of the effectiveness of prerelease advertising adjustment by decomposing it at both strategic and tactical levels and collectively examining their influences on new product sales and firm value. Third, it offers actionable guidelines for managers to optimize their advertising campaigns during the critical new product release period.

However, there are some limitations that I would explore in further studies. First, the current study does not empirically measure the managerial uncertainty, which serves as the driver of launching firm managers' prelease advertising adjustment in response to customer and investor sentiment. It would be of great value to further understand how varied uncertainties may usher into different advertising adjustment behaviors which lead to the changes in new product performance. As the next step, I will use the Hidden Markov Model (HMM) to tackle this issue.

Second, while the current study shows that managers respond to customer and investor sentiment and make advertising adjustment as soon as one week. This is based on the panel VAR results, which captures the big picture across all sampled firms and their new products. Needless to say, there may exist some individual level heterogeneity when some firms respond to customers or investors much faster than others (i.e., velocity). According to Palmatier et al. (2013), firms' varied response velocities to customers' needs influence their relationship with the customers, resulting in the rise or decline of firm performances. In the context of this study, it is reasonable to assume the varied response velocities (in the form of prerelease advertising adjustment) to different customer and investor sentiment may also influence the new product performance. Thus, I aim to address this question in further studies.

Third, many recent studies in our discipline have expressed concerns about the inefficiency of the digital advertising. In the current study, I use the inefficiency of advertiser-audience interactivity to explain the decreased new product promotional advertising spend through digital media vehicles in response to the increased investor bullish sentiment. However, due to the data availability constraint, I am unable to empirically ascertain this assumption. In further studies, I strive to collect more data and investigate whether there exists such an inefficiency and how this inefficiency affects the advertising effectiveness in terms of new product sales and firm value.

TABLES

Table 3.1
Representative Studies Related to Advertising, Customer and Investor Activities, and Firm Performance

Study	**New Product Introduction as Research Context**	**Recognition of Customers' and/or Investors' Influence on Firms' Advertising Adjustment**	**Examination of Customer Sentiment**	**Examination of Investor Sentiment**	**Empirically Controlling for the Dynamic Feedback Effect of Advertising Adjustment**	**Decomposing Advertising at both Strategic and Tactical Levels**
Chakravarty and Grewal (2011)	No	Yes	No	No	Yes	No
Chakravarty and Grewal (2016)	No	Yes	No	No	Yes	No
Hewett et al. (2016)	No	Yes	Yes	No	No	No
Joshi and Hanssens (2010)	No	No	No	No	No	No
Lovett and Staelin (2016)	No	No	Yes	No	Yes	No
Markovitch and Golder (2018)	No	No	No	No	Yes	No
Mian, Sharma and Gul (2018)	No	Yes	No	Yes	No	No
Mizik (2010)	No	Yes	No	No	No	No
Park, Chintagunta and Suk (2019)	Yes	Yes	No	No	Yes	No
Pauwels et al. (2004)	Yes	No	No	No	No	No
Robertson, Eliashberg and Rymon (1995)	Yes	No	No	No	No	No
Srinivasan et al. (2009)	Yes	No	No	No	No	No
Srinivasan, Rutz and Pauwels (2016)	No	No	Yes	No	Yes	No
Wies et al. (2019)	No	No	No	No	No	No
Xiong and Bharadwaj (2013)	No	No	No	No	No	No
This Study	***Yes***	***Yes***	***Yes***	***Yes***	***Yes***	***Yes***

Table 3.2
Variable Measurement

Variable Name	Measurement	Data Source
Dependent Variables		
Opening Week Sales Volume	Order volume (log-transformed) of a new video game during the firm week of its official release	vgchartz.com
Buy-and-hold Abnormal Return Difference	Difference between the launching firm's buy-and-hold abnormal returns (12 months since the new product preannouncement date) and that of its matched firm based on the three matching criteria: 1) the same two-digit SIC code; 2). market values within 70% to 130% of that of the launching firm; 3). the closest book-to-market values.	Compustat
Independent Variables		
Customer Sentiment (prerelease)	Mean value (log-transformed) of the sentiment scores of those textual online forum posts on the weekly basis, relevant to the new product on the weekly basis. The sentiment score is estimated based on Gated-Recurrent-Unit model with 1 being the most positive and 0 being the least positive.	NewGaf.com
Investor Sentiment (prerelease)	Percentage of bullish posts over total number of posts made by investors on the weekly basis.	Stocktwits.com
Digital New Product Promotional Advertising (prerelease).	Weekly advertising expenditure (log transformed) made by each launching firm for the new product through online media vehicles.	Ad$pender
Traditional New Product Promotional Advertising (prerelease).	Weekly advertising expenditure (log transformed) made by each launching firm for the new product through all offline media vehicles (e.g., national and local TV networks, OOH, radio, print, etc.)	Ad$pender
Digital Brand marketing Advertising (prerelease)	Weekly online advertising expenditure (log transformed) made by each launching firm for other incumbent products with the same genre as the to be released new product.	Ad$pender
Traditional Brand marketing Advertising (prerelease)	Weekly offline advertising expenditure (log transformed) by each launching firm for other incumbent products with the same genre as the to be released new product.	Ad$pender
Control Variables		
Launching Firm's PR Release (prerelease)	Weekly number of PR releases (log-transformed) made by each launching firm on the weekly basis	Factiva

Variable Name	Measurement	Data Source
Control Variables (cont'd)		
Carry-over Effect of Franchised Games	Advertising stock values of video game *i* at time *t* calculated based on Bayer et al (2020).	Ad$pender
Weekly Media Coverage	Weekly number of media reports (log-transformed) covering the new product during the prerelease period.	Factiva
Competitors' Reactions (digital advertising)	Weekly online advertising expenditure (log transformed) made by all competing firms that owns incumbent products with the same genre as the to be released new product owned by the launching firm.	Ad$pender
Competitors' Reactions (traditional advertising)	Weekly offline advertising expenditure (log transformed) made by all competing firms that owns incumbent products with the same genre as the to be released new product owned by the launching firm.	Ad$pender
Launching Firm's Advertising Expenditure Intensity	Ratio of a firm's advertising expenditures to its assets	Compustat
Launching Firm's R&D Expenditure Intensity	Ratio of a firm's R&D expenditures to its assets	Compustat
Leverage	Ratio of the firm's long-term book debt to its total assets	Compustat
Book-to-market Ratio	Ratio of book value of equity to market value of equity	Compustat
Launching firm's advertising Expenditure (post-release)	Total amount of advertising expenditure (log-transformed) made by each launching firm 12 months after the new product official release	Compustat
Launching Firm's PR Release (post-release)	Total number of PR releases (log-transformed) made by each launching firm 12 months after the new product official release	Compustat
Holiday Season	Dummy variable with 1 coded as the preannouncement of a new product made during the period between November 15 and January 15; 0 otherwise.	Factiva
Genre Group	Categorical variable. 1) Action and Adventure; 2) Shooter; 3) Sports; 4) All other genres	vgchartz.com

Table 3.3
Estimation of Number of Lags for Panel VAR Estimation

lag	CD	J		MBIC	MAIC	MQIC
1	0.86	758.88	***	-2662.12	-105.12	-1028.96
2	0.93	514.29	***	-1766.39	-61.71	-677.61
3	0.95	258.79	***	-881.55	-29.21	-337.16

Note: *** $p < .001$; ** $p < .01$; * $p < .05$

Table 3.4
Results of Prerelease Advertising Adjustment in Response to Customer and Investor sentiment

Panel A: Advertising for New Production Promotion

Information Source	Advertising for New Product Promotion			
	Digital Advertising		Traditional Advertising	
Customer Sentiment	4.624	***	1.938	***
Investor Sentiment	-3.017	***	-0.951	***

Note: *** $p < .001$; ** $p < .01$; * $p < .05$

Panel B: Advertising for Brand marketing

Information Source	Advertising for Brand marketing			
	Digital Advertising		Traditional Advertising	
Customer Sentiment	0.097	**	0.653	***
Investor Sentiment	-0.158	***	0.177	***

Note: *** $p < .001$; ** $p < .01$; * $p < .05$

Table 3.5
Granger Wald Test Results of Panel VAR Model

Information Source	Advertising for New Product Promotion				Advertising for Brand marketing			
	Digital Advertising		Traditional Advertising		Digital Advertising		Traditional Advertising	
Customer Sentiment	201.313	***	209.563	***	10.255	**	66.994	***
Investor Sentiment	161.356	***	145.02	***	103.28	***	64.334	***

Note: *** $p < .001$; ** $p < .01$; * $p < .05$

Table 3.6
Significant Impacts of Prerelease Advertising Adjustment

Panel A: Impact on Opening Week Sales of the New Video Games

Dependent Variable	**Information Source**	**Prerelease Advertising Adjustment**	**Increase vs. Decrease Ad Expenditures**	**Impact on DV**	**Game Genre**	**Coefficient**	**SD**
Opening Week Sales	customer sentiment	New Product Promotion (Traditional Advertising)	Increase	helps	sports	22.916	13.758
			Increase	helps	others	17.099	11.382
	sig2					1.682	0.507
	sig2.alpha					80.246	174.282

Panel B: Significant Impacts on Market Value of Launching Firms (BHAR-Diff)

Dependent Variable	**Information Source**	**Prerelease Advertising Adjustment**	**Increase vs. Decrease Ad Expenditures**	**Impact on DV**	**Game Genre**	**Coefficient**	**SD**
BHAR-Diff	customer sentiment	New Product Promotion (Traditional Advertising)	Increase	helps	sports	73.889	44.290
			Increase	helps	others	55.409	34.606
		Brand marketing (Traditional Advertising)	Increase	helps	sports	46.470	29.286
			Increase	helps	others	34.830	23.956
	sig2					42.097	11.316
	sig2.alpha					299.914	871.037

Note: All reported estimated coefficients are significant at the 95% credibility interval.

FIGURES

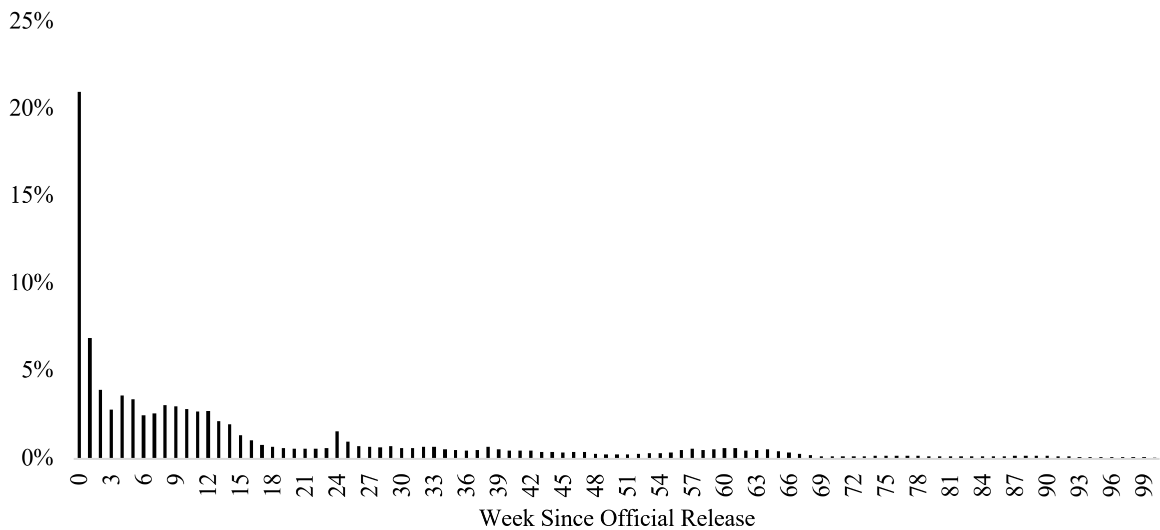


Figure 3.1
Mean Sales Distribution of Sampled New Video Games

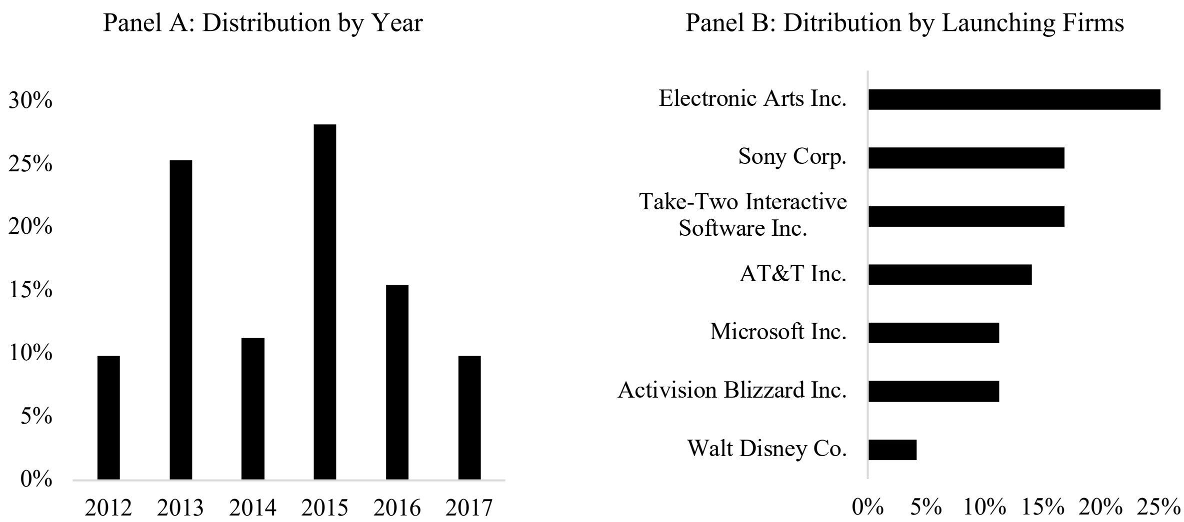


Figure 3.2:
Distribution of Video Games by Year and the Launching Firms

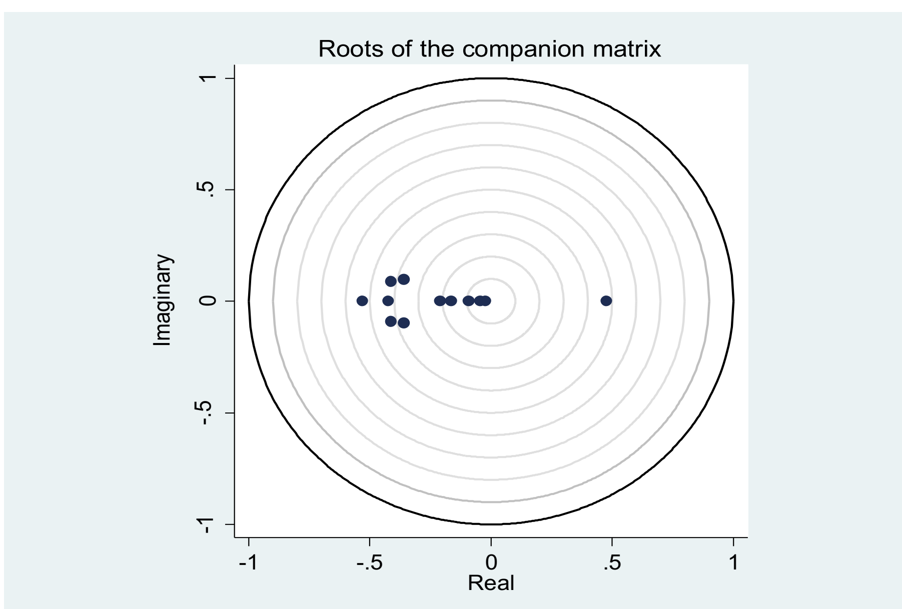


Figure 3.3
Stability Test Results of Panel VAR Model

REFERENCE

Abrigo, Michael RM and Inessa Love (2016), "Estimation of Panel Vector Autoregression in Stata," The Stata Journal, 16 (3), 778-804.

Acs, Zoltan J and David B Audretsch (1990), Innovation and Small Firms: Mit Press.

Adomavicius, Gediminas, Jesse Bockstedt, and Alok Gupta (2012), "Modeling Supply-Side Dynamics of It Components, Products, and Infrastructure: An Empirical Analysis Using Vector Autoregression," Information Systems Research, 23 (2), 397-417.

Andrews, Donald WK and Biao Lu (2001), "Consistent Model and Moment Selection Procedures for Gmm Estimation with Application to Dynamic Panel Data Models," Journal of econometrics, 101 (1), 123-64.

Argenti, Paul A and Bob Druckenmiller (2004), "Reputation and the Corporate Brand," Corporate Reputation Review, 6 (4), 368-74.

Arnold, Todd J, Eric Er Fang, and Robert W Palmatier (2011), "The Effects of Customer Acquisition and Retention Orientations on a Firm's Radical and Incremental Innovation Performance," Journal of the academy of marketing Science, 39 (2), 234-51.

Atuahene-Gima, Kwaku (2005), "Resolving the Capability–Rigidity Paradox in New Product Innovation," Journal of Marketing, 69 (4), 61-83.

Babić Rosario, Ana, Francesca Sotgiu, Kristine De Valck, and Tammo HA Bijmolt (2016), "The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors," Journal of Marketing Research, 53 (3), 297-318.

Baker, Malcolm and Jeffrey Wurgler (2007), "Investor Sentiment in the Stock Market," Journal of economic perspectives, 21 (2), 129-52.

Balmer, John MT (2001), "Corporate Identity, Brand marketing and Corporate Marketing-Seeing through the Fog," European Journal of marketing.

Barber, Brad M and John D Lyon (1997), "Detecting Long-Run Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics," Journal of Financial Economics, 43 (3), 341-72.

Bartels, Robert (1988), "The History of Marketing Thought."

Bass, Frank M, Norris Bruce, Sumit Majumdar, and BPS Murthi (2007), "Wearout Effects of Different Advertising Themes: A Dynamic Bayesian Model of the Advertising-Sales Relationship," Marketing Science, 26 (2), 179-95.

Batra, Rajeev and Kevin Lane Keller (2016), “Integrating Marketing Communications: New Findings, New Lessons, and New Ideas,” Journal of Marketing, 80 (6), 122-45.

Bhattacharya, Chitrabhan B and Sankar Sen (2003), “Consumer–Company Identification: A Framework for Understanding Consumers’ Relationships with Companies,” Journal of Marketing, 67 (2), 76-88.

Bowman, Douglas and Hubert Gatignon (1995), “Determinants of Competitor Response Time to a New Product Introduction,” Journal of Marketing Research, 32 (1), 42-53.

Bronnenberg, Bart J, Vijay Mahajan, and Wilfried R Vanhonacker (2000), “The Emergence of Market Structure in New Repeat-Purchase Categories: The Interplay of Market Share and Retailer Distribution,” Journal of Marketing Research, 37 (1), 16-31.

Brown, Tom J and Peter A Dacin (1997), “The Company and the Product: Corporate Associations and Consumer Product Responses,” Journal of Marketing, 61 (1), 68-84.

Bruce, Norris I, BPS Murthi, and Ram C Rao (2017), “A Dynamic Model for Digital Advertising: The Effects of Creative Format, Message Content, and Targeting on Engagement,” Journal of Marketing Research, 54 (2), 202-18.

Chakravarty, Anindita and Rajdeep Grewal (2011), “The Stock Market in the Driver's Seat! Implications for R&D and Marketing,” Management Science, 57 (9), 1594-609.

Chan, Kalok and Allaudeen Hameed (2006), “Stock Price Synchronicity and Analyst Coverage in Emerging Markets,” Journal of Financial Economics, 80 (1), 115-47.

Chen, Hailiang, Prabuddha De, and Yu Jeffrey Hu (2015), “It-Enabled Broadcasting in Social Media: An Empirical Study of Artists’ Activities and Music Sales,” Information Systems Research, 26 (3), 513-31.

Chintagunta, Pradeep K, Shyam Gopinath, and Sriram Venkataraman (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), 944-57.

Chiu, Chao-Min, Hsin-Yi Huang, Hsiang-Lan Cheng, and Pei-Chen Sun (2015), “Understanding Online Community Citizenship Behaviors through Social Support and Social Identity,” International journal of information management, 35 (4), 504-19.

Choi, In (2006), “Combination Unit Root Tests for Cross-Sectionally Correlated Panels,” Econometric Theory and Practice: Frontiers of Analysis and Applied Research: Essays in Honor of Peter CB Phillips. Cambridge University Press, Chapt, 11, 311-33.

Da, Zhi, Joseph Engelberg, and Pengjie Gao (2011), “In Search of Attention,” The Journal of Finance, 66 (5), 1461-99.

Dahlén, Micael and Mats Edenius (2007), “When Is Advertising Advertising? Comparing Responses to Non-Traditional and Traditional Advertising Media,” Journal of Current Issues & Research in Advertising, 29 (1), 33-42.

De Vries, Lisette, Sonja Gensler, and Peter SH Leeflang (2017), “Effects of Traditional Advertising and Social Messages on Brand-Building Metrics and Customer Acquisition,” Journal of Marketing, 81 (5), 1-15.

Dechow, Patricia M, Richard G Sloan, and Amy P Sweeney (1995), “Detecting Earnings Management,” Accounting review, 193-225.

Deighton, John A, Carl F Mela, and Christine Moorman (2021), “Marketing Thinking and Doing,” SAGE Publications Sage CA: Los Angeles, CA.

Del Negro, Marco and Frank Schorfheide (2011), “Bayesian Macroeconometrics,” Citeseer.

Devos, Erik, Wei Hao, Andrew K Prevost, and Udomsak Wongchoti (2015), “Stock Return Synchronicity and the Market Response to Analyst Recommendation Revisions,” Journal of Banking & Finance, 58, 376-89.

Dewan, Sanjeev and Jui Ramaprasad (2014), “Social Media, Traditional Media, and Music Sales,” MIS quarterly, 38 (1), 101-22.

Dhaoui, Abderrazak and Sami Bacha (2017), “Investor Emotional Biases and Trading Volume’s Asymmetric Response: A Non-Linear Ardl Approach Tested in S&P500 Stock Market,” Cogent economics & finance, 5 (1), 1274225.

Doan, Thomas, Robert Litterman, and Christopher Sims (1984), “Forecasting and Conditional Projection Using Realistic Prior Distributions,” Econometric reviews, 3 (1), 1-100.

Dubé, Jean-Pierre, Günter J Hitsch, and Puneet Manchanda (2005), “An Empirical Model of Advertising Dynamics,” Quantitative marketing and economics, 3 (2), 107-44.

Elberse, Anita and Bharat Anand (2007), “The Effectiveness of Pre-Release Advertising for Motion Pictures: An Empirical Investigation Using a Simulated Market,” Information economics and policy, 19 (3-4), 319-43.

Eliashberg, Jehoshua and Thomas S Robertson (1988), “New Product Preannouncing Behavior: A Market Signaling Study,” Journal of Marketing Research, 25 (3), 282-92.

Enders, Walter and Todd Sandler (1993), “The Effectiveness of Antiterrorism Policies: A Vector-Autoregression-Intervention Analysis,” American Political Science Review, 829-44.

Entertainment Software Association (2020), “2020 Essential Facts About the Video Game Industry”, Entertainment Software Association, (accessed March 23, 2021), [available at https://www.theesa.com/esa-research/2020-essential-facts-about-the-video-game-industry/].

Ettlie, John E and Albert H Rubenstein (1987), "Firm Size and Product Innovation," Journal of Product Innovation Management, 4 (2), 89-108.

Fehle, Frank, Sergey Tsyplakov, and Vladimir Zdorovtsov (2005), "Can Companies Influence Investor Behaviour through Advertising? Super Bowl Commercials and Stock Returns," European Financial Management, 11 (5), 625-47.

Foutz, Natasha Zhang and Wolfgang Jank (2010), "Research Note—Prerelease Demand Forecasting for Motion Pictures Using Functional Shape Analysis of Virtual Stock Markets," Marketing Science, 29 (3), 568-79.

Ganesh, Jaishankar, Mark J Arnold, and Kristy E Reynolds (2000), "Understanding the Customer Base of Service Providers: An Examination of the Differences between Switchers and Stayers," Journal of Marketing, 64 (3), 65-87.

Garber, Tal, Jacob Goldenberg, Barak Libai, and Eitan Muller (2004), "From Density to Destiny: Using Spatial Dimension of Sales Data for Early Prediction of New Product Success," Marketing Science, 23 (3), 419-28.

Garcia, Rosanna and Roger Calantone (2002), "A Critical Look at Technological Innovation Typology and Innovativeness Terminology: A Literature Review," Journal of Product Innovation Management: An international publication of the product development & management association, 19 (2), 110-32.

Garretson, Judith A, Dan Fisher, and Scot Burton (2002), "Antecedents of Private Label Attitude and National Brand Promotion Attitude: Similarities and Differences," Journal of Retailing, 78 (2), 91-99.

Golder, Peter N and Gerard J Tellis (2004), "Growing, Growing, Gone: Cascades, Diffusion, and Turning Points in the Product Life Cycle," Marketing Science, 23 (2), 207-18.

Goldfarb, Avi and Catherine Tucker (2011), "Online Display Advertising: Targeting and Obtrusiveness," Marketing Science, 30 (3), 389-404.

Google Inc. (2021), "Programmatic Ad Buying For Every Format", Google Inc., (accessed March 23, 2021), [available at https://admanager.google.com/home/capabilities/formats/].

Gordon, Brett R, Kinshuk Jerath, Zsolt Katona, Sridhar Narayanan, Jiwoong Shin, and Kenneth C Wilbur (2021), "Inefficiencies in Digital Advertising Markets," Journal of Marketing, 0022242920913236.

Hamilton, Stephen F (2009), "Informative Advertising in Differentiated Oligopoly Markets," International Journal of Industrial Organization, 27 (1), 60-69.

Hansen, Lars Peter (1982), “Large Sample Properties of Generalized Method of Moments Estimators,” Econometrica: Journal of the econometric society, 1029-54.

Hatch, Mary Jo and Majken Schultz (1997), “Relations between Organizational Culture, Identity and Image,” European Journal of marketing.

Hewett, Kelly, William Rand, Roland T Rust, and Harald J Van Heerde (2016), “Brand Buzz in the Echoverse,” Journal of Marketing, 80 (3), 1-24.

Holm, Olof (2006), “Integrated Marketing Communication: From Tactics to Strategy,” Corporate Communications: An International Journal.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey S Rosen (1988), “Estimating Vector Autoregressions with Panel Data,” Econometrica: Journal of the econometric society, 1371-95.

Homburg, Christian, Marcus Theel, and Sebastian Hohenberg (2020), “Marketing Excellence: Nature, Measurement, and Investor Valuations,” Journal of Marketing, 84 (4), 1-22.

Houston, Mark B, Ann-Kristin Kupfer, Thorsten Hennig-Thurau, and Martin Spann (2018), “Pre-Release Consumer Buzz,” Journal of the academy of marketing Science, 46 (2), 338-60.

Isidore, Chris (2019), “Tesla has never needed to advertise, but that might have to change”, cnn.com, (accessed March 23, 2021), [available at https://www.cnn.com/2019/05/31/tech/tesla-advertising].

James, Christopher (1987), “Some Evidence on the Uniqueness of Bank Loans,” Journal of Financial Economics, 19 (2), 217-35.

Joshi, Amit and Dominique M Hanssens (2010), “The Direct and Indirect Effects of Advertising Spending on Firm Value,” Journal of Marketing, 74 (1), 20-33.

Joshi, Amit M and Dominique M Hanssens (2009), “Movie Advertising and the Stock Market Valuation of Studios: A Case of “Great Expectations?”,” Marketing Science, 28 (2), 239-50.

Juergens, Jennifer L and Laura Lindsey (2009), “Getting out Early: An Analysis of Market Making Activity at the Recommending Analyst's Firm,” The Journal of Finance, 64 (5), 2327-59.

Kapferer, Jean-Noel (2008), The New Strategic Brand Management: Creating and Sustaining Brand Equity Long Term: Kogan Page Publishers.

Keehler, John (2020), “Three Ways To Determine Your Digital Marketing Budget”, forbes.com, (accessed March 23, 2021), [available at https://www.forbes.com/sites/forbesagencycouncil/2020/09/11/three-ways-to-determine-your-digital-marketing-budget/?sh=1af9972b47ea].

Keller, Kevin Lane and David A Aaker (1998), "The Impact of Corporate Marketing on a Company's Brand Extensions," Corporate Reputation Review, 1 (4), 356-78.

Kenney, Patrick J and Tom W Rice (1994), "The Psychology of Political Momentum," Political Research Quarterly, 47 (4), 923-38.

Koop, Gary and Dimitris Korobilis (2010), Bayesian Multivariate Time Series Methods for Empirical Macroeconomics: Now Publishers Inc.

Kumar, Vita, Lerzan Aksoy, Bas Donkers, Rajkumar Venkatesan, Thorsten Wiesel, and Sebastian Tillmanns (2010), "Undervalued or Overvalued Customers: Capturing Total Customer Engagement Value," Journal of service research, 13 (3), 297-310.

Lam, Son K, Michael Ahearne, Ye Hu, and Niels Schillewaert (2010), "Resistance to Brand Switching When a Radically New Brand Is Introduced: A Social Identity Theory Perspective," Journal of Marketing, 74 (6), 128-46.

Lee, Yikuan and Gina Colarelli O'Connor (2003), "New Product Launch Strategy for Network Effects Products," Journal of the academy of marketing Science, 31 (3), 241-55.

Leeper, Eric M, Christopher A Sims, Tao Zha, Robert E Hall, and Ben S Bernanke (1996), "What Does Monetary Policy Do?," Brookings papers on economic activity, 1996 (2), 1-78.

Litterman, Robert (1984), "Forecasting and Policy Analysis with Bayesian Vector Autoregression Models," Quarterly Review (fall).

Love, Inessa and Lea Zicchino (2006), "Financial Development and Dynamic Investment Behavior: Evidence from Panel Var," The Quarterly Review of Economics and Finance, 46 (2), 190-210.

Luan, Y Jackie and K Sudhir (2010), "Forecasting Marketing-Mix Responsiveness for New Products," Journal of Marketing Research, 47 (3), 444-57.

Luo, Miao, Daniel X Jiang, and Jun Cai (2014), "Investor Sentiment, Product Features, and Advertising Investment Sensitivities," Asia-Pacific Journal of Financial Studies, 43 (6), 798-837.

Lütkepohl, Helmut (2005), New Introduction to Multiple Time Series Analysis: Springer Science & Business Media.

Marchand, André and Thorsten Hennig-Thurau (2013), "Value Creation in the Video Game Industry: Industry Economics, Consumer Benefits, and Research Opportunities," Journal of Interactive Marketing, 27 (3), 141-57.

McCulloch, Robert and Peter E Rossi (1994), "An Exact Likelihood Analysis of the Multinomial Probit Model," Journal of econometrics, 64 (1-2), 207-40.

Menon, Anil, Sundar G Bharadwaj, Phani Tej Adidam, and Steven W Edison (1999), "Antecedents and Consequences of Marketing Strategy Making: A Model and a Test," Journal of Marketing, 63 (2), 18-40.

Mian, G Mujtaba, Piyush Sharma, and Ferdinand A Gul (2018), "Investor Sentiment and Advertising Expenditure," International Journal of Research in Marketing, 35 (4), 611-27.

Mizik, Natalie and Robert Jacobson (2003), "Trading Off between Value Creation and Value Appropriation: The Financial Implications of Shifts in Strategic Emphasis," Journal of Marketing, 67 (1), 63-76.

Moe, Wendy W and Peter S Fader (2002), "Fast-Track: Article Using Advance Purchase Orders to Forecast New Product Sales," Marketing Science, 21 (3), 347-64.

Moliner, Miguel Angel, Diego Monferrer-Tirado, and Marta Estrada-Guillén (2018), "Consequences of Customer Engagement and Customer Self-Brand Connection," Journal of Services Marketing.

Muzellec, Laurent and Mary C Lambkin (2009), "Brand marketing and Brand Architecture: A Conceptual Framework," Marketing Theory, 9 (1), 39-54.

Nations, Daniel (2021), "What's the Difference Between a Mac and a PC?", Lifewire.com, (accessed March 23, 2021), [available at https://www.lifewire.com/what-is-a-mac-4155662].

Nations, Daniel (2021), "What's the Difference Between a Mac and a PC?", Lifewire.com, (accessed March 23, 2021), [available at https://www.lifewire.com/what-is-a-mac-4155662].

Nelson, Phillip (1970), "Information and Consumer Behavior," Journal of political economy, 78 (2), 311-29.

Nelson, Phillip---- (1974), "Advertising as Information," Journal of political economy, 82 (4), 729-54.

Olsen, Svein Ottar (2002), "Comparative Evaluation and the Relationship between Quality, Satisfaction, and Repurchase Loyalty," Journal of the academy of marketing Science, 30 (3), 240-49.

Overby, Jeffrey W and Eun-Ju Lee (2006), “The Effects of Utilitarian and Hedonic Online Shopping Value on Consumer Preference and Intentions,” Journal of business research, 59 (10-11), 1160-66.

Palmatier, Robert W, Mark B Houston, Rajiv P Dant, and Dhruv Grewal (2013), “Relationship Velocity: Toward a Theory of Relationship Dynamics,” Journal of Marketing, 77 (1), 13-30.

Park, Kyung M, Pradeep K Chintagunta, and Inho Suk (2019), “Capital Market Returns to New Product Development Success: Informational Effects on Product Market Advertising,” Journal of Marketing Research, 56 (1), 37-56.

Pauwels, Koen, Jorge Silva-Risso, Shuba Srinivasan, and Dominique M Hanssens (2004), “New Products, Sales Promotions, and Firm Value: The Case of the Automobile Industry,” Journal of Marketing, 68 (4), 142-56.

Peltier, James W, Debra Zahay, and Donald R Lehmann (2013), “Organizational Learning and Crm Success: A Model for Linking Organizational Practices, Customer Data Quality, and Performance,” Journal of Interactive Marketing, 27 (1), 1-13.

Percy, Larry and John R Rossiter (1992), “Advertising Stimulus Effects: A Review,” Journal of Current Issues & Research in Advertising, 14 (1), 75-90.

Petersen, J Andrew and V Kumar (2009), “Are Product Returns a Necessary Evil? Antecedents and Consequences,” Journal of Marketing, 73 (3), 35-51.

Rappaport, Alfred (1999), Creating Shareholder Value: A Guide for Managers and Investors: Simon and Schuster.

Rinallo, Diego and Suman Basuroy (2009), “Does Advertising Spending Influence Media Coverage of the Advertiser?,” Journal of Marketing, 73 (6), 33-46.

Robertson, Thomas S, Jehoshua Eliashberg, and Talia Rymon (1995), “New Product Announcement Signals and Incumbent Reactions,” Journal of Marketing, 59 (3), 1-15.

Rossi, Peter E and Greg M Allenby (1993), “A Bayesian Approach to Estimating Household Parameters,” Journal of Marketing Research, 30 (2), 171-82.

Rossi, Peter E and Greg M Allenby ---- (2003), “Bayesian Statistics and Marketing,” Marketing Science, 22 (3), 304-28.

Rubera, Gaia and Ahmet H Kirca (2012), “Firm Innovativeness and Its Performance Outcomes: A Meta-Analytic Review and Theoretical Integration,” Journal of Marketing, 76 (3), 130-47.

Ruth, Julie A (2001), "Promoting a Brand's Emotion Benefits: The Influence of Emotion Categorization Processes on Consumer Evaluations," Journal of Consumer Psychology, 11 (2), 99-113.

Saboo, Alok R, Anindita Chakravarty, and Rajdeep Grewal (2016), "Organizational Debut on the Public Stage: Marketing Myopia and Initial Public Offerings," Marketing Science, 35 (4), 656-75.

Sethuraman, Raj, Gerard J Tellis, and Richard A Briesch (2011), "How Well Does Advertising Work? Generalizations from Meta-Analysis of Brand Advertising Elasticities," Journal of Marketing Research, 48 (3), 457-71.

Shankar, Venkatesh (1999), "New Product Introduction and Incumbent Response Strategies: Their Interrelationship and the Role of Multimarket Contact," Journal of Marketing Research, 36 (3), 327-44.

Shivakumar, Lakshmanan (2000), "Do Firms Mislead Investors by Overstating Earnings before Seasoned Equity Offerings?," Journal of Accounting and Economics, 29 (3), 339-71.

Silverman, Steven N, David E Sprott, and Vincent J Pascal (1999), "Relating Consumer-Based Sources of Brand Equity to Market Outcomes," ACR North American Advances.

Sims, Christopher A and Tao Zha (1998), "Bayesian Methods for Dynamic Multivariate Models," International Economic Review, 949-68.

Slevin, Dennis P and Jeffrey K Pinto (1987), "Balancing Strategy and Tactics in Project Implementation," Sloan management review, 29 (1), 33-41.

Smith, Ken G, Christopher J Collins, and Kevin D Clark (2005), "Existing Knowledge, Knowledge Creation Capability, and the Rate of New Product Introduction in High-Technology Firms," Academy of management Journal, 48 (2), 346-57.

Sorescu, Alina, Venkatesh Shankar, and Tarun Kushwaha (2007), "New Product Preannouncements and Shareholder Value: Don't Make Promises You Can't Keep," Journal of Marketing Research, 44 (3), 468-89.

Spence, Michael (1974), "Competitive and Optimal Responses to Signals: An Analysis of Efficiency and Distribution," Journal of Economic theory, 7 (3), 296-332.

Sprout Social Inc. (2021), "How Sprout Social helps the Chicago Bulls take its social media game to the next level", Sprout Social Inc., (accessed March 23, 2021), [available at https://sproutsocial.com/insights/case-studies/chicago-bulls/].

Sridhar, Shrihari, Frank Germann, Charles Kang, and Rajdeep Grewal (2016), "Relating Online, Regional, and National Advertising to Firm Value," Journal of Marketing, 80 (4), 39-55.

Srinivasan, Shuba and Dominique M Hanssens (2009), "Marketing and Firm Value: Metrics, Methods, Findings, and Future Directions," Journal of Marketing Research, 46 (3), 293-312.

Srinivasan, Shuba, Koen Pauwels, Jorge Silva-Risso, and Dominique M Hanssens (2009), "Product Innovations, Advertising, and Stock Returns," Journal of Marketing, 73 (1), 24-43.

Srinivasan, Shuba, Oliver J Rutz, and Koen Pauwels (2016), "Paths to and Off Purchase: Quantifying the Impact of Traditional Marketing and Online Consumer Activity," Journal of the academy of marketing Science, 44 (4), 440-53.

Srivastava, Rajendra K, Tasadduq A Shervani, and Liam Fahey (1999), "Marketing, Business Processes, and Shareholder Value: An Organizationally Embedded View of Marketing Activities and the Discipline of Marketing," Journal of Marketing, 63 (4_suppl1), 168-79.

Stigler, George J (1961), "The Economics of Information," Journal of political economy, 69 (3), 213-25.

Su, Meng and Vithala R Rao (2010), "New Product Preannouncement as a Signaling Strategy: An Audience-Specific Review and Analysis," Journal of Product Innovation Management, 27 (5), 658-72.

Tellis, Gerard J (2003), Effective Advertising: Understanding When, How, and Why Advertising Works: Sage Publications.

Turner, John C, Rupert J Brown, and Henri Tajfel (1979), "Social Comparison and Group Interest in Ingroup Favouritism," European journal of social psychology, 9 (2), 187-204.

Van Doorn, Jenny and Janny C Hoekstra (2013), "Customization of Online Advertising: The Role of Intrusiveness," Marketing Letters, 24 (4), 339-51.

Varadarajan, Rajan (2010), "Strategic Marketing and Marketing Strategy: Domain, Definition, Fundamental Issues and Foundational Premises," Journal of the academy of marketing Science, 38 (2), 119-40.

Vorhies, Douglas W, Linda M Orr, and Victoria D Bush (2011), "Improving Customer-Focused Marketing Capabilities and Firm Financial Performance Via Marketing Exploration and Exploitation," Journal of the academy of marketing Science, 39 (5), 736-56.

Voss, Glenn B and Zannie Giraud Voss (2008), "Competitive Density and the Customer Acquisition–Retention Trade-Off," Journal of Marketing, 72 (6), 3-18.

Voss, Kevin E, Eric R Spangenberg, and Bianca Grohmann (2003), "Measuring the Hedonic and Utilitarian Dimensions of Consumer Attitude," Journal of Marketing Research, 40 (3), 310-20.

Warren, Nooshin L and Alina Sorescu (2017), “Interpreting the Stock Returns to New Product Announcements: How the Past Shapes Investors’ Expectations of the Future,” Journal of Marketing Research, 54 (5), 799-815.

Webster Jr, Frederick E (1992), “The Changing Role of Marketing in the Corporation,” Journal of Marketing, 56 (4), 1-17.

Wies, Simone, Arvid Oskar Ivar Hoffmann, Jaakko Aspara, and Joost ME Pennings (2019), “Can Advertising Investments Counter the Negative Impact of Shareholder Complaints on Firm Value?,” Journal of Marketing, 83 (4), 58-80.

Xiong, Guiyang and Sundar Bharadwaj (2014), “Prerelease Buzz Evolution Patterns and New Product Performance,” Marketing Science, 33 (3), 401-21.

You, Ya, Gautham G Vadakkepatt, and Amit M Joshi (2015), “A Meta-Analysis of Electronic Word-of-Mouth Elasticity,” Journal of Marketing, 79 (2), 19-39.

Zackariasson, Peter and Timothy L Wilson (2012), The Video Game Industry: Formation, Present State, and Future: Routledge.

Zeithammer, Robert and Peter Lenk (2006), “Bayesian Estimation of Multivariate-Normal Models When Dimensions Are Absent,” Quantitative marketing and economics, 4 (3), 241-65.

APPENDICES

Appendix 1

Exhibition 2.1 Sales Distribution of the 300 New Video Games

As discussed in the study, most of the entertainment products (e.g., video games, movies, etc.) have relatively shorter lifetime and most of their sales are generated in the first few weeks or months after the initial launch. This is further backed by my empirical data. Figure 2.5 below shows the average sales distribution of 300 new video games historically launched in the North American market between 2011 and 2017.

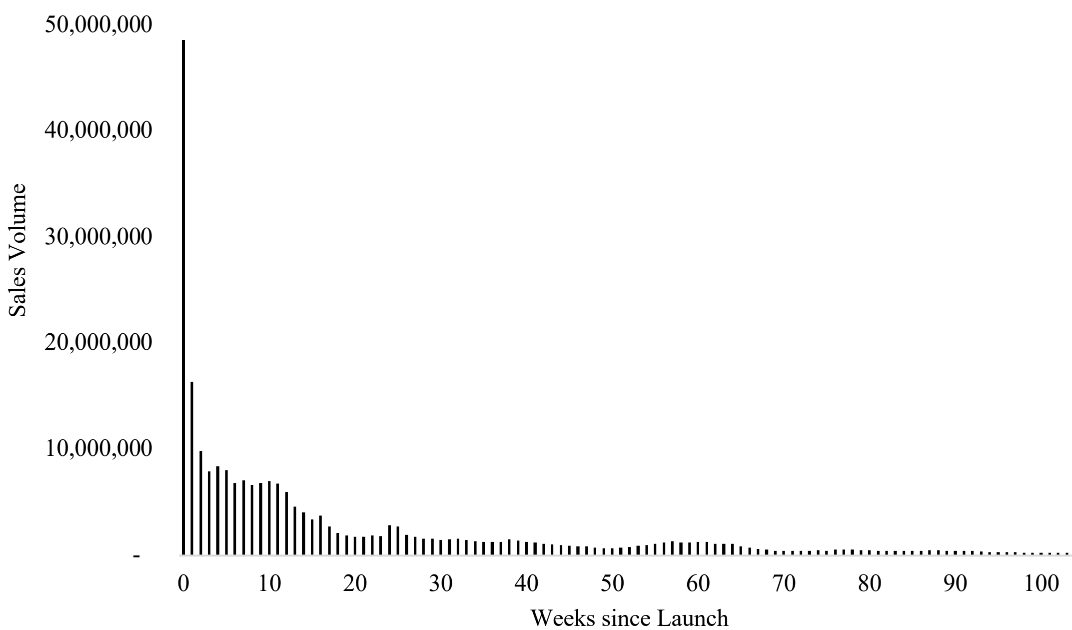


Figure 2.5

Sales Distribution of 300 Video Games Launched Between 2011 and 2017

Exhibition 2.2 Average Pageview Evolution of 300 New Video Games

To better illustrate how regular customers may proactively come to Wikipedia.org to examine relevant information about new video games prior to their official launch, I collected Wikipedia pageview data from https://dumps.wikimedia.org/other/pagecounts-ez/, where one can extract all pageview stats of existing Wikipedia pages starting from 2007. Particularly for the purpose of this study, I collected pageview data for the 300 new video games from the 58th week to the last week prior to their official launches.

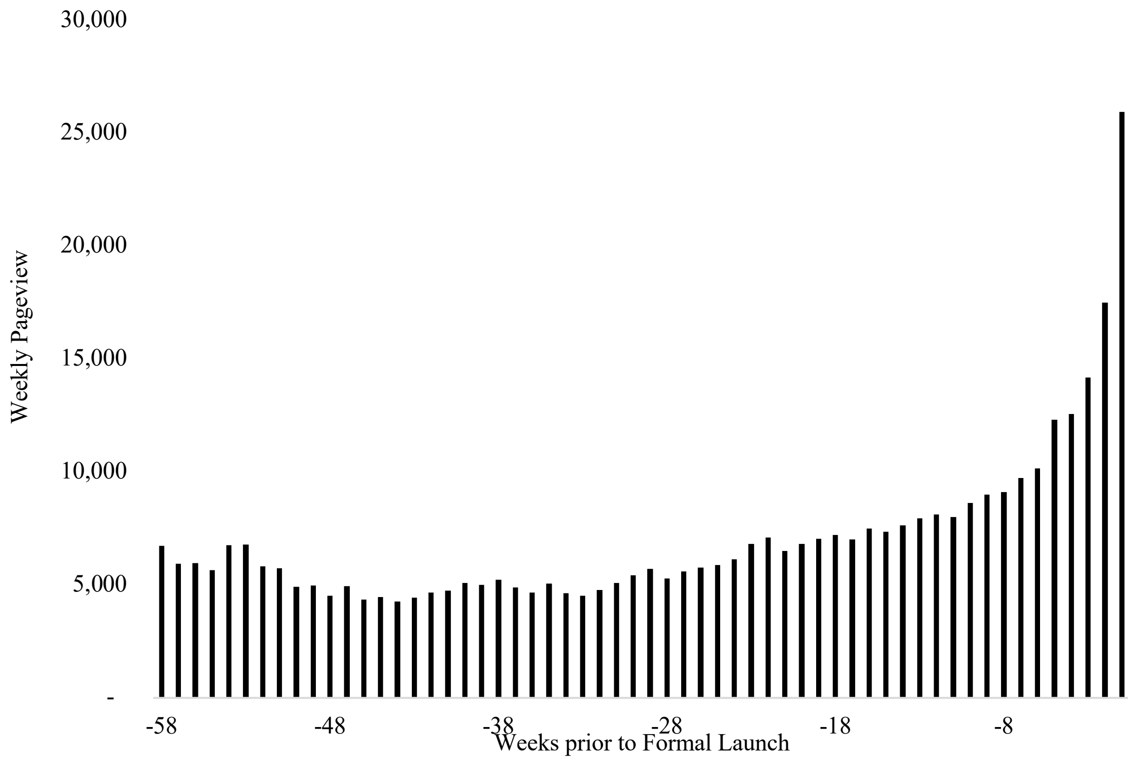


Figure 2.6
Average Wikipedia Pageviews of 300 New Video Games during Prerelease Period

Exhibition 2.3 Information Sources of Wikipedia Pages of the 300 New Video Games

To further investigate how knowledge contributors acquired essential information to co-create the Wikipedia pages of those 300 new video games during the prerelease period, the authors painstakingly read through each new video game's historical Wikipedia pages (prior to official launch) and collected the information sources based on the reference list of each page.

In general, there are four major sources that knowledge contributors utilized to garner information and co-created Wikipedia pages, including 1) official website or PR news release of the launching firms; 2) launching firms' official public pages on social media platforms, such as Facebook, Twitter, etc.; 3) professional reviews; and 4) other news coverages. Figure 2.7 below illustrates the average composition of the information sources used by knowledge contributors when co-creating those 300 new video games' Wikipedia pages.

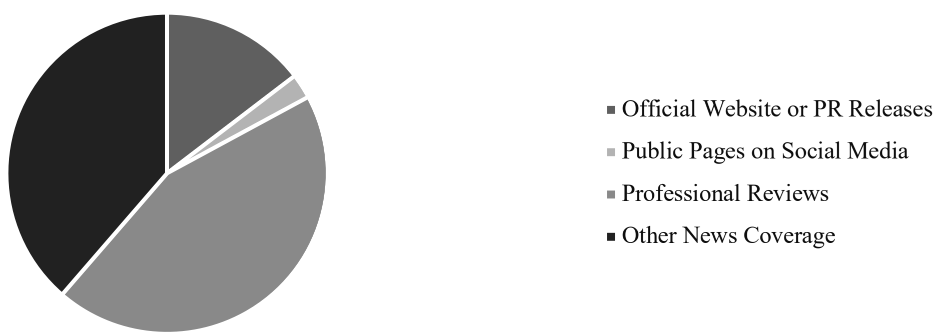


Figure 2.7

Information Source Composition across All 300 New Video Games

Exhibition 2.4 Recurrent Neural Network based Gated Recurrent Unit (RNN-GRU) Model and Its Application to Sentiment Analysis of New Video Game Customer Buzz

I chose the RNN-GRU method, a sequential modeling technique, for the sake of its high efficiency in natural language processing (relative to the 'bag-of-words' approach). One of the most salient advantages of the RNN model lies in the fact that it takes into account the sequential connection across words in a document. For example, a person may comment on a video game by saying "I don't think it is a good one." The word '*good*' can be considered as an indicator of a positive sentiment. However, the phrase "*don't*" ahead of it infers that this person actually does not like it (For a more detailed review, see Hu and Liu 2004; Pang, Lee, and Vaithyanathan 2002). In this case, a RNN model is able to detect the 'context' and assess the valence of the document efficiently.

The RNN model resolves the sequential issue by setting up a recurrent hidden state $\tilde{h}^t$, whose activation depends on the state h^{t-1} at previous time and the current information input (x^t). Formally, let's have a sequence $x = (x^1, x^2, \dots, x^t)$ and assume that

$$p(x^1, x^2, \dots, x^t) = p(x^1)p(x^2|x^1)p(x^3|x^1, x^2) \dots p(x^T|x^1, \dots, x^{t-1}) \qquad \text{Equation [1]}$$

Then, the recurrent hidden state ($\tilde{h}^t$) is updated based on:

$$\tilde{h}^t = \begin{cases} 0, & t = 0 \\ tanh(Wh^{t-1} + Ux^t + b), & otherwise \end{cases} \qquad \text{Equation [2]}$$

, where $tanh(\cdot)$ is a smooth and bounded hyperbolic tangent function, W and U are the weights, and b is the bias term.

As Bengio, Simard, and Frasconi (1994) and Chung et al. (2014) pointed out, such a vanilla RNN model often suffers from gradient vanishing or explosion issue in the backpropagation process, where the effect of long-term dependencies become exponentially small or big with respect to the sequence length. Therefore, it is challenging to train RNN models to capture long-term dependencies.

Based on the traditional RNN, Gated Recurrent Unit model (GRU)[35] deals with the vanishing gradient problems by pruning unnecessary connections and adaptively capturing the dependencies of different temporal scales within a sequence. Proposed by Cho et al. (2014), GRU is formally defined as:

$$u^t = \sigma(W_u h^{t-1} + U_u x^t + b_u) \quad \text{Equation [3]}$$

$$r^t = \sigma(W_r h^{t-1} + U_r x^t + b_r) \quad \text{Equation [4]}$$

$$\tilde{h}^t = \tanh\,(W_h(r^t \odot h^{t-1}) + U_h x^t + b_h) \quad \text{Equation [5]}$$

$$h^t = (1 - u^t) \odot h^{t-1} + u^t \odot \tilde{h}^t \quad \text{Equation [6]}$$

Different from the regular RNN model, the RNN-GRU model incorporates two new components which collectively adapt and prune the dependencies, u^t and r^t. u^t in Equation [3] is called the *update gate*, the 'key' adaptive part dependent upon the value of the previous state h^{t-1} and the new information available at the current time x^t. It prescribes how much the content gets updated. W_u and U_u are the weights, and b_u is the bias term. $\sigma\,(\cdot)$ is the activation function (a sigmoid function in my research context). r^t in Equation [4] is called the *reset gate* that prunes away some of *useless* memories. When $r_t = 0$, the GRU model will merely consider the current state and *forget* all past events.

In Equation [5], one has $\tilde{h}^t$, the hidden state determined by both current information (x^t) and the previous condition (h_{t-1}) whose influence is pruned by the reset gate r^t. $\odot$ denotes

[35] I am also aware of the long-short term memory model (LSTM), another special case of RNN. There are some differences regarding the modeling architecture of the two methods. However, based on prior literature both GRU and LSTM are fairly comparable in terms of modeling accuracy (see Chung et al. 2014; Fu, Zhang, and Li 2016). In addition, some studies have found GRU runs faster than LSTM (e.g., Rana 2016; Chung et al. 2014). However, comparison between both models is beyond my scope.

element-wise multiplication. Eventually, in Equation [6], the current state h^t is estimated based on the both the previous state h^{t-1} and the current hidden state $\tilde{h}^t$, whose respective influences are determined by the *update gate* u^t and $1 - u^t$.

Following the prior literature (Karpathy, Johnson and Fei-Fei 2015; Saxe, McClelland and Ganguli 2014; Zeiler 2012), I configure the GRU model architecture by setting three layers, using the Adam algorithm as the optimizer to control the learning rate. While the first layer deals with the fundamental 'meaning' of each word within a sequence, the second and the third layers focus on a more complex and high-level picture of each post. Usually, a three-layer GRU model is good enough. Too many layers may lead to the over-fitting issue (see: Zaremba, Sutskever and Vinyals 2014). The Adam algorithm is more advantageous compared with conventional gradient descent method in the sense that it enables a more efficient stepwise optimization and facilitates faster convergence (see more details from Kingma and Ba 2014; and Pascanu, Mikolov, and Bengio 2013).

Once setting up the GRU configuration, I utilize a pre-labeled movie review database from IMDb for the initial model training[36] [37]. This pre-labeled dataset contains 25,000 highly polarized movie reviews as the training set (i.e., either positive = 1 or negative =1), and another

[36] Given the large number of my posts, a 10% sample is 19,393 posts. If I use human raters for coding, it will extremely slow down the efficiency. However, I do conduct a robustness check for the accuracy of my sentiment analysis based on the pre-labeled IMDb dataset. The detailed process and result will be discussed shortly.

[37] I also notice that there is a Sentiment 140 dataset containing 1.6 million pre-labeled Twitter tweets (positive versus negative) provided by Stanford researchers Go, Bhayani, and Huang (2009) on http://help.sentiment140.com/home. I initially used that dataset for the GRU model training. However, the result is not satisfactory compared with the one based on IMDb dataset. I assume this is due to the length of Twitter tweet, which is constrained to 140 characters. Relative to it, the average number of words of my collected posts is 236.83. Thus, it may contain more complex idea. At this point, the movie reviews from the IMDb dataset are on average longer and contains more complex information for the model training.

25,000 highly polarized reviews as test set. The number of the positive and that of the negative reviews are equal[38]. There are no neutral reviews in the dataset.

It was first used by researchers Maas and colleagues (2011) and offered on their website (http://ai.stanford.edu/~amaas/data/sentiment/). In the past, multiple research studies in sentiment analysis have been conducted based on this pre-labeled dataset and have achieved very accurate results (e.g., Liu 2012; Mikolov et al. 2013). Given the fact that my goal at this stage is to identify which posts are positive, this dataset is a good fit. In the actual modeling process, I allocate another 25% of the original training data from the IMDb dataset as the evaluation set to further fine-tune the model before applying it to the testing dataset. After the model training, I applied it to the test dataset and evaluate the sentiment of each movie reviews. Then, I compared the results based on the model and the pre-labeled result. 87% of them are matched. Finally, I applied the trained GRU model to my 193,930 raw NeoGaf posts to identify which are positive. Specifically, if the GRU-estimated sentiment score is equal to or greater than 0.7, I then coded that post as positive and save it for further FDA analysis. If the sentiment score is smaller than 0.7, I then disregarded it. Based on this criterion, 115,044 posts (59.3%) are estimated by my RNN-GRU model as positive.

As a robustness check, I randomly sampled 500 RNN-GRU estimated posts and recruited two human raters to code the sentiment (positive or not) independently. The inter-rater agreement is 87%, the inter-rater reliability based on Cohen's kappa is 0.79 and Krippendorff's alpha is 0.80. Next, I compare the human raters' coding result with the GRU model result for those 500 posts. 90% of alignment is achieved. As of then, I completed all the pre-processing work to extract the positive customer buzz toward those 300 new video games.

[38] More details about the dataset description and the actual dataset can be retrieved from https://github.com/SrinidhiRaghavan/AI-Sentiment-Analysis-on-IMDB-Dataset

Exhibition 2.5 Regression Tree Analysis

In a classical linear regression set-up, the expected value of the dependent variable (Y), where $Y \in \mathbb{R}$, is determined by a linear function of the predictors ($\boldsymbol{X}$) and the random error (ϵ). Thus,

$$Y = \beta_0 + \boldsymbol{\beta X} + \epsilon \qquad \text{Equation [7]}$$

The estimation of Y in Equation (4) may also allow the potential interactions among predictors. However, there are two limitations. First, the analyst will need to determine which two or more predictors may interact with one another *ex-ante*. Second, as the number of the predictors grows, their potential interaction dimension may become higher (e.g., three- or four-way interactions, etc.x), leading to over-parameterization.

Breiman et al. (1984) proposed an alternative and nonparametric method – regression tree. Based on this approach, one can divide (i.e., partition) the global predictor dataspace ($\boldsymbol{X}$) into k small regions (R_k), where both the number of predictors and their interactions are more manageable to evaluate. Thus, the value of the dependent variable (Y) based on predictors ($\boldsymbol{X}$) within each region (R_k) can be written as $E(Y|\boldsymbol{X} \in R_k)$). The final predicted value of $\hat{Y}$ is based on the mean value of all the observed dependent variable (Y's) across all k regions. More visually illustrated in Figure 2.8, a tree grows through a sequence of binary partitioning based on different predictors. In this process, a node may grow at the end of a split with the introduction of a new predictor (e.g., *Diversity* FPCS1), resulting in an interaction with the former predictor (e.g., *Article Length* FPCS1). In other cases, a node can grow with the same predictor, hence examining the predictor-DV relationship in greater details. Such a splitting process continues as the residual sum of squares (SSR) keep decreasing and stops once the SSR of each ending node is minimized. Hereby, the residual sum of squares is defined as

$$SSR = \sum_i (y_i - \mu)^2 \qquad \text{Equation [8]}$$

, where y_i represents varied estimated values of the DV in a node and μ is the mean value of those $y_i's$ in the node

To take this study for example, one predictor I am interested in is the dynamic knowledge collaborative activities measured by the FPC scores of *Article Length*, the other is the FPC scores of *Diversity*. As illustrated in Figure 2.8, the tree may grow starting from the first FPC score of *Article Length* (i.e., the root node). Following that, a binary partition is made based on the value of the root node (e.g., Is the *Article Length* FPCS1 greater than 0?). One interior node is formed at one end of the split through the introduction of the second variable *Diversity* (Is the *Diversity* FPCS1 greater than 0?). At the other end of the split, however, two terminal nodes (or leaves) arise with the binary partition based on the criterion that "Is the *Article Length* FPCS1 greater than 5?". Eventually, the final estimation of $\hat{Y}$ is the average of the Y's within corresponding terminal nodes. For example, the estimated value of Y_1 on Figure 2.8 is estimated based on the average values of all Y's in a scenario when *Article Length* is greater than 5. Y_3 is the mean of all Y's in a scenario when both the *Article Length* FPCS1 is less than or equal to 0 and the *Diversity* FPSC1 greater than 0.

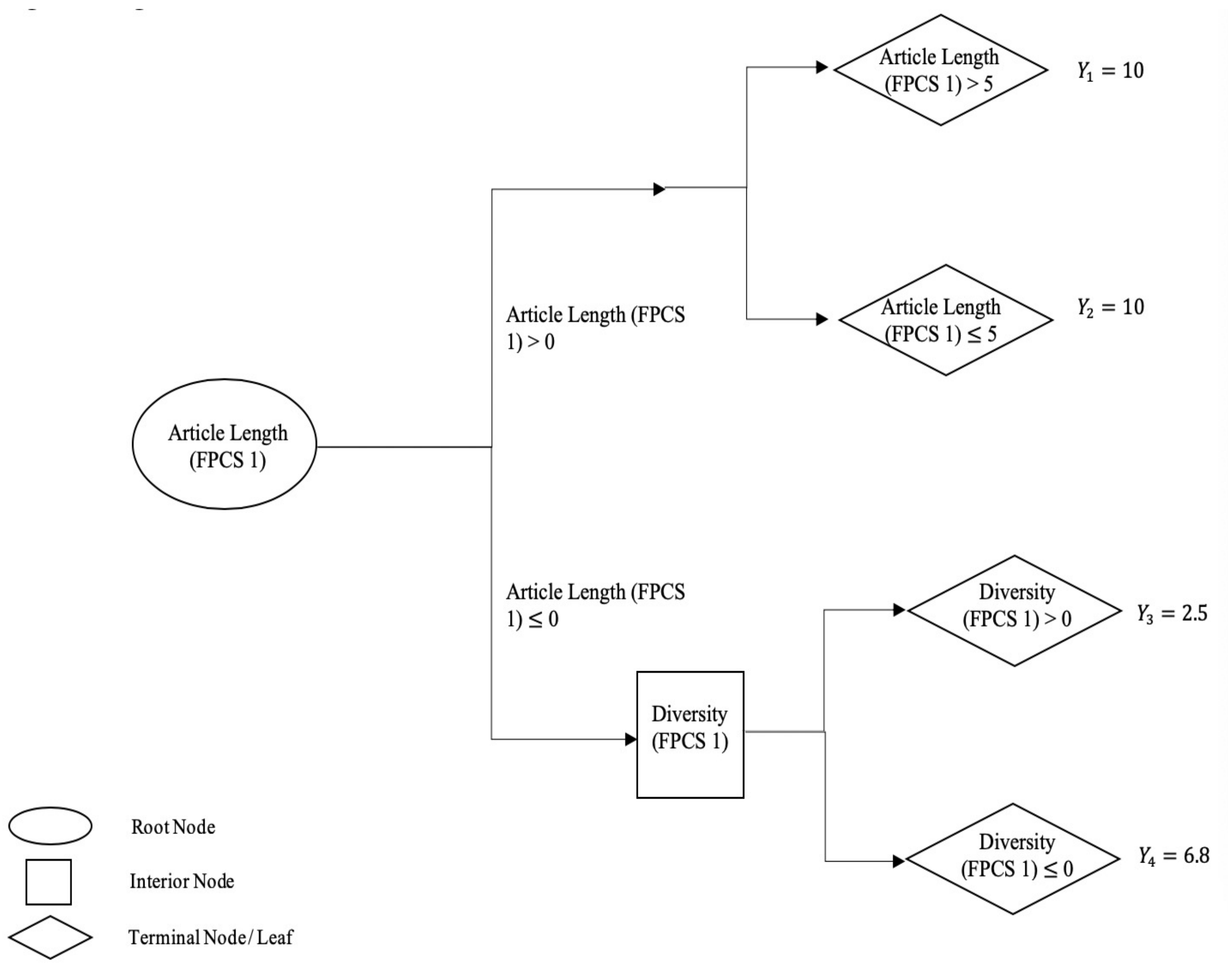


Figure 2.8
Regression Tree Illustration

Appendix 2

Exhibition 3.1 List of Information about the Sampled New Video Games and their Launching Firm's 12-month BHAR-Diff

Video Game	Preannouncement Date	Launching Firm	12-month BHAR (Launching Firm)	Matched Firm	12-month BHAR (Matched Firm)	BHAR-Diff
007 Legends	4/18/2012	Activision Blizzard Inc.	26.5%	NetEase Inc.	-36.8%	0.63
Borderlands 2	8/3/2011	Take-Two Interactive Software Inc.	63.1%	CACI Information Technology Co.	3.0%	0.60
Dance Central 3	6/4/2012	Microsoft Inc.	-28.3%	Google Inc.	-38.0%	0.10
FIFA 13	4/23/2012	Electronic Arts Inc.	159.0%	Fleetcor Technologies Inc.	129.0%	0.30
Lego Batman 2: DC Super Heroes	3/9/2012	AT&T Inc.	-16.6%	China Mobile Corp.	10.2%	-0.27
Madden NFL 13	4/10/2012	Electronic Arts Inc.	161.7%	Fleetcor Technologies Inc.	140.4%	0.21
World of Warcraft: Mists of Pandaria	10/21/2011	Activision Blizzard Inc.	46.2%	NetEase Inc.	-744.0%	7.90
Army of Two: The Devil's Cartel	8/2/2012	Electronic Arts Inc.	81.7%	Fleetcor Technologies Inc.	-5.4%	0.87

Video Game	Preannouncement Date	Launching Firm	12-month BHAR (Launching Firm)	Matched Firm	12-month BHAR (Matched Firm)	BHAR-Diff
Batman: Arkham Origins	4/9/2013	AT&T Inc.	-6.0%	China Mobile Corp.	-8.9%	0.03
Batman: Arkham Origins Blackgate	4/9/2013	Microsoft Inc.	54.7%	Google Inc.	-93.5%	1.48
Battlefield 4	7/17/2012	Electronic Arts Inc.	-339.9%	Fleetcor Technologies Inc.	-217.9%	-1.22
BioShock Infinite	8/12/2010	Take-Two Interactive Software Inc.	43.4%	CACI Information Technology Co.	15.8%	0.28
Doki-Doki Universe	6/10/2013	Sony Corporation	85.9%	Eaton Corp. PLC	-118.8%	2.05
FIFA Soccer 14	4/22/2013	Electronic Arts Inc.	-392.9%	Fleetcor Technologies Inc.	-288.2%	-1.05
God of War: Ascension	4/22/2012	Sony Corporation	48.5%	Eaton Corp. PLC	-20.0%	0.69
Gran Turismo 6	2/28/2013	Sony Corporation	85.9%	Eaton Corp. PLC	-118.8%	2.05
Injustice: Gods Among Us	5/31/2012	AT&T Inc.	-25.5%	China Mobile Corp.	-27.2%	0.02

Video Game	Preannouncement Date	Launching Firm	12-month BHAR (Launching Firm)	Matched Firm	12-month BHAR (Matched Firm)	BHAR-Diff
Killzone: Mercenary	1/31/2013	Sony Corporation	25.2%	Eaton Corp. PLC	-183.4%	2.09
Knack	2/11/2013	Sony Corporation	110.9%	Eaton Corp. PLC	-90.9%	2.02
LEGO Marvel Super Heroes	1/8/2013	AT&T Inc.	-6.0%	China Mobile Corp.	-8.9%	0.03
MLB 13: The Show	12/14/2012	Take-Two Interactive Software Inc.	43.4%	CACI Information Technology Co.	15.8%	0.28
NBA Live 14	5/21/2013	Electronic Arts Inc.	-258.0%	Fleetcor Technologies Inc.	-128.2%	-1.30
NHL 14	4/22/2013	Electronic Arts Inc.	-392.9%	Fleetcor Technologies Inc.	-288.2%	-1.05
Sly Cooper: Thieves in Time	6/7/2011	Sony Corporation	56.6%	Eaton Corp. PLC	-31.3%	0.88
The Last of Us	12/10/2011	Sony Corporation	-24.5%	Eaton Corp. PLC	-17.0%	-0.07
Borderlands: The Pre-Sequel!	4/11/2014	Take-Two Interactive Software Inc.	101.2%	CACI Information Technology Co.	52.0%	0.49
Destiny	5/7/2012	Microsoft Inc.	-123.5%	Google Inc.	-30.1%	-0.93

Video Game	Preannouncement Date	Launching Firm	12-month BHAR (Launching Firm)	Matched Firm	12-month BHAR (Matched Firm)	BHAR-Diff
Disney Infinity 2.0: Marvel Super Heroes	4/30/2014	Walt Disney Co.	21.6%	Vodafone Telecommunication Co.	-4.4%	0.26
Driveclub	2/20/2013	Sony Corporation	107.0%	Eaton Corp. PLC	20.9%	0.86
Forza Horizon 2	6/8/2014	Microsoft Inc.	-123.5%	Google Inc.	-30.1%	-0.93
Lego Batman 3: Beyond Gotham	5/27/2014	AT&T Inc.	-61.0%	China Mobile Corp.	48.6%	-1.10
MLB 14: The Show	2/10/2014	Sony Corporation	50.7%	Eaton Corp. PLC	-15.4%	0.66
The Sims 4	5/6/2013	Electronic Arts Inc.	94.3%	Fleetcor Technologies Inc.	70.8%	0.24
Batman: Arkham Knight	3/4/2014	AT&T Inc.	23.0%	China Mobile Corp.	-8.4%	0.31
Battlefield: Hardline	6/1/2014	Electronic Arts Inc.	-18.0%	Fleetcor Technologies Inc.	33.9%	-0.52
Bloodborne	6/5/2014	Sony Corporation	-90.8%	Eaton Corp. PLC	45.9%	-1.37

Video Game	Preannouncement Date	Launching Firm	12-month BHAR (Launching Firm)	Matched Firm	12-month BHAR (Matched Firm)	BHAR-Diff
Borderlands: The Handsome Collection	2/3/2015	Take-Two Interactive Software Inc.	-44.9%	CACI Information Technology Co.	41.2%	-0.86
Call of Duty: Black Ops 3	4/27/2015	Activision Blizzard Inc.	26.6%	NetEase Inc.	-40.3%	0.67
Disney Infinity 3.0	9/29/2014	Walt Disney Co.	-40.2%	Vodafone Telecommunication Co.	-1.7%	-0.39
Dying Light	6/13/2013	AT&T Inc.	32.6%	China Mobile Corp.	18.9%	0.14
Evolve	1/24/2013	Take-Two Interactive Software Inc.	-84.2%	CACI Information Technology Co.	16.9%	-1.01
FIFA 16	5/28/2015	Electronic Arts Inc.	-86.0%	Fleetcor Technologies Inc.	-3.5%	-0.83
Guitar Hero Live	6/13/2013	Activision Blizzard Inc.	37.6%	NetEase Inc.	101.4%	-0.64
Halo 5: Guardians	5/16/2014	Microsoft Inc.	5.4%	Google Inc.	33.2%	-0.28
LEGO Dimensions	4/9/2015	AT&T Inc.	8.7%	China Mobile Corp.	-30.2%	0.39
LEGO Jurassic World	1/30/2015	AT&T Inc.	23.0%	China Mobile Corp.	-8.4%	0.31

Video Game	Preannouncement Date	Launching Firm	12-month BHAR (Launching Firm)	Matched Firm	12-month BHAR (Matched Firm)	BHAR-Diff
NBA 2K16	3/20/2015	Take-Two Interactive Software Inc.	-6.1%	CACI Information Technology Co.	9.7%	-0.16
NBA Live 16	5/7/2015	Electronic Arts Inc.	-86.0%	Fleetcor Technologies Inc.	-3.5%	-0.83
NHL 16	5/20/2015	Electronic Arts Inc.	-86.0%	Fleetcor Technologies Inc.	-3.5%	-0.83
Rise of the Tomb Raider	6/9/2014	Microsoft Inc.	-46.7%	Google Inc.	10.9%	-0.58
Star Wars Battlefront (2015)	1/29/2015	Walt Disney Co.	-61.4%	Vodafone Telecommunication Co.	12.0%	-0.73
StarCraft II: Legacy of the Void	5/11/2015	Activision Blizzard Inc.	26.6%	NetEase Inc.	-40.3%	0.67
WWE 2K16	5/18/2015	Take-Two Interactive Software Inc.	64.6%	CACI Information Technology Co.	25.5%	0.39
Batman: A Telltale Game Series	6/23/2016	AT&T Inc.	-33.6%	China Mobile Corp.	-1.5%	-0.32

Video Game	Preannouncement Date	Launching Firm	12-month BHAR (Launching Firm)	Matched Firm	12-month BHAR (Matched Firm)	BHAR-Diff
Battleborn	7/10/2014	Take-Two Interactive Software Inc.	109.1%	CACI Information Technology Co.	-27.8%	1.37
Battlefield 1	5/6/2016	Electronic Arts Inc.	13.9%	Fleetcor Technologies Inc.	-45.6%	0.59
BioShock The Collection	7/1/2016	Take-Two Interactive Software Inc.	45.7%	CACI Information Technology Co.	-85.5%	1.31
EA Sports UFC 2	9/15/2013	Electronic Arts Inc.	-7.3%	Fleetcor Technologies Inc.	-26.5%	0.19
FIFA 17	5/10/2016	Electronic Arts Inc.	11.7%	Fleetcor Technologies Inc.	-34.2%	0.46
Gears of War 4	6/15/2015	Microsoft Inc.	17.1%	Google Inc.	-21.7%	0.39
NHL 17	5/10/2016	Electronic Arts Inc.	11.7%	Fleetcor Technologies Inc.	-34.2%	0.46
Teenage Mutant Ninja Turtles: Mutants in Manhattan	1/28/2016	Activision Blizzard Inc.	-9.0%	NetEase Inc.	24.5%	-0.34

Video Game	Preannouncement Date	Launching Firm	12-month BHAR (Launching Firm)	Matched Firm	12-month BHAR (Matched Firm)	BHAR-Diff
World of Warcraft: Legion	8/6/2015	Activision Blizzard Inc.	-52.3%	NetEase Inc.	-95.0%	0.43
XCOM 2	6/2/2015	Take-Two Interactive Software Inc.	85.7%	CACI Information Technology Co.	57.9%	0.28
Crash Bandicoot N. Sane Trilogy	12/5/2016	Activision Blizzard Inc.	-17.8%	NetEase Inc.	-183.8%	1.66
Gravity Rush 2	5/26/2016	Sony Corporation	13.8%	Eaton Corp. PLC	9.2%	0.05
Halo Wars 2	8/4/2015	Microsoft Inc.	-25.0%	Google Inc.	-15.0%	-0.10
Horizon Zero Dawn	6/16/2015	Sony Corporation	-16.0%	Eaton Corp. PLC	17.4%	-0.33
NBA 2K18	1/12/2017	Take-Two Interactive Software Inc.	-49.6%	CACI Information Technology Co.	36.9%	-0.86
NBA Live 18	5/29/2017	Electronic Arts Inc.	-102.7%	Fleetcor Technologies Inc.	54.0%	-1.57
NHL 18	1/31/2017	Electronic Arts Inc.	-102.7%	Fleetcor Technologies Inc.	54.0%	-1.57

REFERENCE

Bengio, Yoshua, Patrice Simard, and Paolo Frasconi (1994), "Learning Long-Term Dependencies with Gradient Descent Is Difficult," IEEE transactions on neural networks, 5 (2), 157-66.

Breiman, L, J Friedman, R Olshen, and C Stone (1984), "Classification and Regression Trees," Wadsworth International Group, Belmont, California.

Chung, Junyoung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio (2014), "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," arXiv preprint arXiv:1412.3555.

Fu, Rui, Zuo Zhang, and Li Li (2016), "Using Lstm and Gru Neural Network Methods for Traffic Flow Prediction," in 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC): IEEE.

Go, Alec, Richa Bhayani, and Lei Huang (2009), "Twitter Sentiment Classification Using Distant Supervision," CS224N project report, Stanford, 1 (12), 2009.

Liu, Bing (2012), "Sentiment Analysis and Opinion Mining," Synthesis lectures on human language technologies, 5 (1), 1-167.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean (2013), "Efficient Estimation of Word Representations in Vector Space," arXiv preprint arXiv:1301.3781.

Rana, Rajib (2016), "Gated Recurrent Unit (Gru) for Emotion Classification from Noisy Speech," arXiv preprint arXiv:1612.07778.

CHAPTER 4

CONCLUSION

The main objectives of this dissertation research are: 1) to examine the validity of the dynamic online knowledge collaboration activities in terms of its added value to the prerelease NPS forecast and 2) to investigate whether managers of the launching firm may dynamically undertake prerelease advertising adjustment (both strategically and tactically) in response to customer and investor sentiment, and what consequences such adjustments will lead to.

In my first essay, I reason the validity of employing online knowledge collaborative activities for the prerelease NPS forecast in the sense that the content (relevant to a forthcoming new product) co-created through such a dynamic process benefits prospective customers' continuous information search during the prerelease period and facilitates them to make the ultimate purchase decision. Empirically, the proposed method based on the FDA technique efficiently captures such dynamics. The results based on the random forest model, outperforming any regression-based approaches, show that the inclusion of the dynamic knowledge collaboration variables significantly improves the accuracy of the prerelease NPS forecast. Moreover, the superiority of the inclusion of the dynamic knowledge collaboration variables with the proposed method still holds for early forecast based on data 10 weeks prior to the final release.

However, there are some limitations to be addressed in further studies. First, the first essay due to its main focus does not necessarily investigate the causality between the online knowledge collaboration and the launching firm's marketing activities. Second, the research context of the study is the video game industry. It would be interesting to explore how such a dynamic online knowledge collaboration may contribute to prerelease NPS forecast for products of other categories embodying more utilitarian values (e.g., foods). Third, it would be beneficial to explore other available online knowledge collaboration platforms other than Wikipedia.org for data extraction.

My second essay finds support that managers of new product make advertising adjustments in response to both customer and investor sentiment during the prerelease period. Such adjustments apply to not only advertising strategies (i.e., new product promotion and corporate branding), but also the tactics (i.e., digital and traditional advertising). However, only the dynamic adjustment of new product promotional ads through the traditional media vehicles exerts significant and positive impact on the new product opening week sales. In the additional study, the model results imply that both new product promotional and corporate branding campaigns through traditional media vehicles drive the increase in launching firms' buy-and-hold abnormal returns relative to their matched counterparts (BHAR-Diff).

There are limitations of the second essay that I would explore in further studies. First, it does not empirically measure the managerial uncertainty, which serves as the driver of launching firm managers' dynamic prelease advertising adjustment in response to stakeholders' sentiment. Second, there may exist some individual level heterogeneity when some firms respond to one (or both) type of stakeholders much faster than others (i.e., velocity). Thus, it is reasonable to assume the varied response velocities (in the form of prerelease advertising adjustment) to different stakeholders' sentiment may also influence the new product performance.

In summary, the two essays of my dissertation offer novel perspectives for the prerelease NPS forecast and the launching firms' prereleases advertising adjustment. The findings are of essential values for both academia and the managerial practices.