ESSAYS OF THE POWER OF USER-GENERATED CONTENTS AND ONLINE COMMUNITIES

By

Yuan Zhang

DISSERTATION

Submitted in fulfillment of the requirements

for the degree of Doctor of Philosophy at

The University of Texas at Arlington

May, 2020

Arlington, Texas

Supervising Committee:

Dr. Jennifer Jie Zhang, Supervising Professor

Dr. Ritesh Saini

Dr. Sridhar Panchapakesan Nerur

Dr. Mahyar Sharif Vaghefi

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

LIST OF TABLES

CHAPTER TWO

CHAPTER THREE

CHAPTER FOUR

ABSTRACT

Essays of The Power of User-Generated Contents and Online Communities

Yuan Zhang

The University of Texas at Arlington, 2019

Supervising Professor: Jennifer Jie Zhang

How the internet and mobile technologies shape users' online content generation and how user-generated content on social media and online communities influence their subsequent behaviors are under-researched in the IS field. This dissertation contains three essays that examine three aspects of user-generated content (UGC) and online communities. The first study examines the spillover effects of location-based mobile applications on local businesses' performance through analyzing the entry and penetration of a location-based augmented reality application under a natural experiment setting. A rich-get-richer effect is identified in that the internalization of the reputation spillover varies with the current reputation of the local businesses and the time frames. The second study investigated how the viewership of user-generated videos on different platforms impact product sales and customer stickiness. A Panel-Vector-Auto- Regression model is used to show that live streaming video platforms can improve both the short-term and long-term product sales and customer stickiness, and pre-recorded video platforms can only enhance

long- term product sales and customer stickiness. Additionally, live streaming video platforms have a more significant and stronger predictive relationship than pre-recorded video platforms with the response level and explanatory power of product sales and customer stickiness. The media richness theory and the social presence theory are applied to explain the results. The third study is to examine the impacts of the group-level social comparison on group members' exercise participation. Our research context is set by a fitness app with the aim of improving users' exercise and usage participation through the group-level social and gamification function design. The group-level social interaction features are, for example, allowing users to join and exercise with other group members, hold and participate in offline meetup events, and group social activities. A conceptual framework based on the cognitive processes under the social comparison theory is built to explain the effects of the within-group and between-group social comparison on the group members' exercise participation. We also propose and verify that offline group social activities significantly moderate the main effects.

CHAPTER ONE

Introduction

In recent years, the Internet has evolved to various new artifacts equipped with multiple converged technologies such as the live streaming video platform, location-based mobile applications with augmented reality features, hobby-based community mobile applications with both the online and offline social interaction features. With these new user-generated content (UGC) and online communication artifacts, users are observed to have different subsequent behaviors and responses as well as the impacts on related businesses. Motivated to investigate the pattern and underlying mechanisms regarding the behaviors and impacts, this dissertation focus on providing logical mechanisms and conceptual frameworks, as well as both the theoretical and managerial implications for both academia and industry.

To study how the converged mobile technologies affect people's mobility patterns, daily lives, and local business, we select and observe the entry and penetration of a phenomenal location-based augmented (LBAR) reality mobile application. After the entry of this LBAR application, app users are immersed in the connection between the in-app virtual world and the real physical world. While users are encouraged to largely explore their vicinity areas and new places, local restaurants, and stores in users' proximity areas may experience the spillover effects in terms of new traffics and potential business opportunities. We use the online reputation as the proxy of the business performance of local restaurants. To examine the spillover effect (main effect) before and after the entry of the LBAR application between the spatially treated and control restaurants, we conduct a natural experiment through the difference-in-difference analysis on a multiple-source unique dataset. We further conduct a series of additional analyses, such as the difference-in-difference-in-difference analysis and treatment intensity analysis, to examine whether and how the internalization of the spillover effects varies with localized factors. Generally, we find that the entry of the LBAR technologies can significantly improve the treated restaurants' online reputation. However, the internalization of the main effects demonstrates a rich-get-richer pattern.

To investigate the emergence of another novel technology and its potentially substantial influences on other participants in the internet environment, in the second study, we select and analyze whether and how the consuming contents on the novel (live streaming video sharing platform) and traditional (pre-recorded video sharing platform) online content sharing platforms impact on users' purchase and retention. We rely on the media richness theory and the social presence theory to reason the underlying mechanism regarding the impacts. Later, we conduct a comprehensive dynamic panel vector auto-regression analyses on multiple-source, unique and big dataset, and a lab experiment through Amazon's MTurk to validate our reasoning and hypotheses. We find that watching videos containing product related information on both the live streaming and pre-record video sharing platforms can improve users' purchase volume and stickiness to the products. However, in terms of both the prediction powder and the response strengths, live streaming platform works more efficiently as an influencer marketing channel than the pre-recorded video platform to boost consumers' subsequent both purchase and stickiness.

Researching the underlying mechanism regarding how mobile users' group-level responses to both online and offline social comparison and interactions is complicated and gapping. To examine the social comparison effects and subsequences of offline social activates from the group perspective, in the third study, we select and manually record users' exercise data from a fitness application for outdoor runners. We comprehensively summarize key propositions and sub-category models under the social comparison theory and based on the literature, through analyzing and reasoning the cognitive processes of social comparison based on the underlying determinants, we build a theoretical framework to model the direct effects of between- and within-group social comparisons as well as the moderation effects of the offline social activities on the direct impacts. We validate our conceptual model through a series of empirical analyses. This study provides implications for designing and operating all similar online community platforms and applications. It also contributes to IS literature regarding how we understand collective level social comparison and the importance of the integration of offline social activities.

Coauthors help the author to polish this dissertation and improve its quality. Dr. Jie (Jennifer) Zhang helps to improve logical reasoning, theoretical and managerial contributions, writing polishing, etc., and the overall quality and chance of publication through her astonishing knowledge and rich experience. Dr. Liu Zilong and Dr. Xiaolong Song help the author manually record and collect the valuable dataset for study 3, and they also provide suggestions regarding the theoretical framework. Dr. Saini, Dr. Nerur, and Dr. Vaghefi, as the crucial dissertation members, also give the author valuable suggestions regarding methodologies, theoretical framework building, and the overall improvement of the research quality.

To sum up, this dissertation investigates the impacts of novel and influential technologies that emerged in recent years. The rest of the dissertation is organized as follows. In Chapter 2, essay 1 examines the spillover effects and the main effect internalization of the entry and penetration of LBAR applications on local businesses. In Chapter 3, essay 2 examines whether and how live streaming and pre-recorded video platforms shape users' purchase behaviors and their stickiness to related products. In Chapter 4, essay 3 studies the effects of between- and within-group social comparison on users' exercise participation and the moderation effects of offline social activities. In Chapter 5, the author summarizes the key findings and implications of this dissertation.

CHAPTER TWO

SPILLOVER EFFECTS OF LOCATION-BASED AUGMENTED REALITY MOBILE APPLICATIONS ON LOCAL BUSINESSES

Yuan Zhang

Jie (Jennifer) Zhang

Abstract

Mobile applications implemented with location-based services (LBS) and augmented reality (AR) technologies have become a new trend. They alter the users' app usage patterns, expand their mobility areas, and shape their daily lives. Consequently, local businesses experience the spillover effects in terms of store visits and business opportunities, as well as online reputation. As businesses located in proximity to the physical app portals are observed to piggyback the location-based augmented reality (LBAR) applications, a windfall gain is not always achieved. We leverage a natural experiment involving the launch of an LBAR mobile application to examine its business impacts on nearby restaurants' online reputation. We find that in general, restaurants in proximity to the physical app portals do benefit from the spillovers of the LBAR application in improving their online reputation. This spillover effect of LBAR applications varies over time and with reputation metrics. Moreover, internalization of the spillover effects depends on the restaurant features (e.g., competition intensity and other locational factors) and established reputation levels. Specifically, the spillover effect of the LBAR application on a restaurant is strengthened when it is surrounded by an agglomeration of restaurants and diminishes with the geographical distance from the app portals at an increasing rate. Additionally, local businesses in proximity to the app portals with a higher reputation or located in a more competitive marketplace will benefit more from the spillovers of LBAR applications in the long run, demonstrating a rich-get-richer pattern. The results are shown robust to a variety of validation models. This paper provides theoretical and practical insights into the applications of LBAR technologies on businesses.

Keywords: Augmented Reality (AR) Technologies, Location-Based (LB), Technologies, Spillover Effect, Rich-get-richer Effect, Online Reputation

1 Introduction

In recent years, mobile applications implemented with location-based-service (LBS) technologies, such as Waze, Foursquare and Curbside, have gained widespread popularity with billions of users. LBS technologies can push geographically personalized information to users, such as mobile advertising and SNS notification. The LBS applications gradually shape users' daily lives and online behaviors: they encourage users to explore new places in vicinity areas, expand their mobility areas, enrich their daily lives, and even stimulate their online activities on social media. For example, Foursquare and Yelp facilitate users to discover and explore interesting places in local areas, and they popularize the idea of checking in and sharing locations on social media. Moreover, some of these LBS mobile applications are converging with the Augmented Reality (AR) technology, which "describes the visual alignment of virtual content with real-world contexts and has been defined as a medium in which digital information is overlaid on the physical world that is in both spatial and temporal registration with the physical world and that is interactive in time" (Rauschenabel et al. 2017, Craig, 2013, p. 20). These location-based augmented reality (LBAR) applications such as Ingress, Harry Porter Wizards Unite, and Minecraft Earth encourage users' social interactions through geographical features by connecting the in-game portals to real locations and attract players to hang around in such locations (e.g. Figure A1 in Appendix A).

As the users of an LBAR app explore new places attracted by app contents or portals and share their findings on social media, businesses located in the surrounding areas may benefit from the app's spillover effects in the forms of store visits, business opportunities and online reputation. Editorial and consultant reports (Fulton 2016, Friedenthal 2016) encourage business owners and practitioners to utilize the spillovers of the LBAR apps to engage in proximity marketing and branding campaigns (e.g. Figure A2 in Appendix A). However, the penetration of the LBAR apps cannot always guarantee positive influences to the local businesses. For instance, Filloon (2016) and Zhu (2016) cast doubt on the spillover effect of Pokemon GO, the most popular LBAR app launched in 2016. They reported that not all restaurants close to the physical app portals experience more sales or more visits, and that players in some restaurant "came in but did not get anything to eat or drink". Thus, a practical puzzle is how much an LBAR app can effectively improve business performance by converting the attracted visitors into customers.

Driven by the above practical puzzle, this study aims to examine how the entry of the LBAR mobile applications impacts local business. More specifically, we study the spillover effects of a particular LBAR mobile app on local restaurants' online reputation. We formalize our objective into the following research questions: (1) How does the entry of the LBAR app impact restaurants' business performance measured by online reputation? In other words, what are the dynamic effects of the entry of LBAR app on online reputation of restaurants near physical app portals? (2) How is the relationship between the entry of the LBAR app and restaurants' online reputation influenced by restaurant, app, or market specific factors? We use online reputation metrics as the proxies of restaurants' business performance for the following two reasons: First, Literature (Luca 2016a, Xie et al. 2014, Qiu et al 2018, Luo et al. 2013) has revealed that online reputation is significantly positively associated with business performance. For example, Luca (2016) showed that one star increase in Yelp rating leads to a 5-9% increase in restaurant revenue. Second, online reviews are frequently generated and immediately disclosed. Thus, they can better capture the popular trends about the customer traffic and sentiments, and can provide timely signals about the changes in market demand. Therefore, when the real traffic and revenue data are not accessible, we can use online reputation metrics to measure restaurants' performance.

To address the above questions, we first build a panel dataset by merging the mobile app's geographic data and adoption data, Google trend index, restaurant profiles and characteristics, and restaurant review data. Then we conduct comprehensive analyses to examine the spillover effects of the LBAR app from various perspectives. After propensity score matching, we use a difference-in-difference (DID) analysis approach considering restaurants near physical app portals as the treatment group, and restaurants without a portal nearby as the control group to estimate the DID coefficients of each the review volume and rating metrics. We also compare the DID effects in both the short and the long terms. Further, we provide a series of additional analyses: we examine both the short and long term DID effects on additional reputation metric; we run DID on the partitioned samples to examine if the impacts of the LBAR application vary with heterogeneous restaurants; we examine the treatment intensity and boundary conditions through a series of Difference-in-Difference-in-Difference (DDD) analyses and a polynomial regression by incorporating the polynomial functions of the distance between a restaurant and its closest portal. We also conduct a series of robustness validations: we first discuss and eliminate the potential self-selection threats by summarizing and discussing all the developers' official posts and announcements regarding the portal application; we formally run a reversal regression of restaurants' online reputation metrics on the condition of whether or not being close to a portal to check the potential falsification issue and the reverse causality threat; we use the relative time approach to test the parallel assumption of DID estimation --- the holding of the assumption guarantees the counterfactual treatment groups, and the control groups are supposed to have the same time trends; we estimate a regression discontinuity model to restore the validity of the control group by comparing a pair of adjacent restaurants near the cutoff line of the forcing variable (distance); we also check and void several potential confounding concerns.

The results of our study suggest that LBAR applications do significantly impact restaurants' online reputation. In general, the spillover effect is largely verified as positive, in the sense that the entry of the LBAR app benefits restaurants near the physical app portals by improving their online reputation in both the short and long terms. In the short term, online reputation of restaurants with portals nearby is significantly improved by 5.7% in ratings and 6.5% in review volume. In the long term, online reputation of restaurants near the portals is significantly improved by 8% in rating and 1.5% in check-in. Moreover, we also find that the spillover effect of the LBAR application is strengthened when there is an agglomeration of restaurants, and diminishes with the geographical distance from the physical app portals at an increasing rate. The results of the partitioned subsample analyses and DDD analyses show that the beneficial spillover due to the entry of the LBAR app varies with restaurant heterogeneity and time frames. The treated restaurants with low established reputation can gain the positive spillover from the entry of the game, but only for the short term. In the long term, there is a rich-get-richer effect such that the positive influence only favors treated restaurants that have good established reputation. More specifically, only high-rating restaurants and/or high-review-volume restaurants can reap the benefits from such LBAR applications in the long term.

This study makes the following contributions. First, our study contributes to the LB and AR literature by providing empirical evidence to examine the business value of LBAR applications. This study uncovers the AR application's spillover through encouraging users' movement and their visits and lingering behaviors to the businesses in newly explored locations or locations close to physical app portals, which creates positive externality effect. Second, it contributes to related spillover studies by providing more refined findings. We discover a rich-get-richer phenomenon in the LBAR application's positive spillover effects in the long run. That is, not all restaurants in proximity to a physical app portal equally benefit from the LBAR app's launch. The internalization of the spillover effects varies with the restaurant's features and reputation level over time. Additionally, our findings indicate that the impacts of the LBS application are strengthened with a larger number of local competitors, and are negatively associated with geographical distance. We can also use our results to estimate the economic impact of the spillover effects due to the LBAR app's entry. Third, we provide managerial implications to local businesses and monetization strategies to LBAR application developers. For example, based on the findings of our main model and the treatment intensity analysis about the geographical distance, we recommend business owners to actively advertise their locational advantages of being close to physical app portals and to engage in proximity marketing to take advantage of the positive spillover effects. Our findings also suggest app developers to allow businesses to apply for a portal status and work with more diversified and heterogeneous restaurants rather than limiting the sponsorship program only to few big chains.

The rest of this paper is organized as follows. In Section 2, we review the related literature, summarize the context mechanism, and propose the hypotheses. Section 3 describes data sources, data processing and summary of statistics, and introduces the identification strategy. We present our model and findings in Section 4. We also provide a series of additional analyses in Section 5. Section 6 includes the robustness tests and validation checks. The theoretical and managerial implications are in Section 7. At last, we conclude the paper with limitations and future research.

2. Context, Literature and Hypotheses Development

We introduce related literature about location-based and augmented reality technologies. Meanwhile, we propose the hypotheses regarding the spillover effects of LBAR technologies on local businesses.

2.1 Context and the Awareness Mechanism

To examine the spillover effects of LBAR applications on local businesses, we choose the phenomenally popular LBAR mobile game (Pokemon GO) at its inception as our research context. Players follow the virtual-reality map to explore vicinity areas and capture Pokemon using one-time use virtual items, which can be restocked at physical app portals for free or through in-app purchases. The portals, called PokeStops, are real locations of interest predetermined by the developer, distributed in the real world and mapped with the same virtual world coordinates. Players move into close proximity of PokeStops to utilize the two important features: collecting virtual items and luring pokemon. Players can only collect free virtual items near PokeStops. Because each restocking session only supplies a limited quantity of the virtual items and takes several minutes to "cool down", players usually linger at the PokeStop and explore its vicinity areas in order to stock sufficient virtual items. Besides, PokeStops have a significantly higher Pokemon spawn rate than other places[1], plus players can only drop a "Lure Module" (a virtual item designed to increase Pokemon activity) in the surrounding area of a PokeStop. Thus players are also attracted by PokeStops to "lure" more and rarer Pokemon.

We illustrate the mechanism regarding Pokemon GO's spillover on local business in Figure 1. PokeStops (represented by the statue in Figure 1) are usually landmarks, statues or arts which

[1] https://pokemongohub.net/generation-2/researching-pokemon-go-spawn-mechanics/

conveying historical, cultural and educational values[2]. PokeStops attract players to move towards their vicinity areas by providing free in-game items, which are essential for players. During a PokeStop's "cooling down" time, players tend to stay within a walking distance to the PokeStop for the next round of restocking and meanwhile, players may explore new locations in the vicinity area and linger there (e.g. Figure A3 in Appendix A). The app design increases player awareness of surrounding businesses by encouraging movement and visits to/lingering behaviors in new locations. The players' lingering behaviors and exploration can result in new foot traffic and potential business opportunities to nearby restaurants as well as a new dining experience to the players themselves. In other words, the app design can increase users' likelihood to visit those businesses. Consequently, according to the economic patronage effect in Liu et al. (2018) and Duan et al. (2008), players, in turn, post reviews on social media. Based on the above awareness mechanism, we expect that there exists a significant spillover effect on the online reputation of the local restaurants that have a PokeStop nearby due to the entry of Pokemon GO.

[2] https://niantic.helpshift.com/a/pokemon-go/?s=PokeStops&f=submitting-a-PokeStop-nomination&l=en&p=web; https://niantic.helpshift.com/a/pokemon-go/?l=en&s=PokeStops&f=what-makes-a-high-quality-PokeStop&p=web

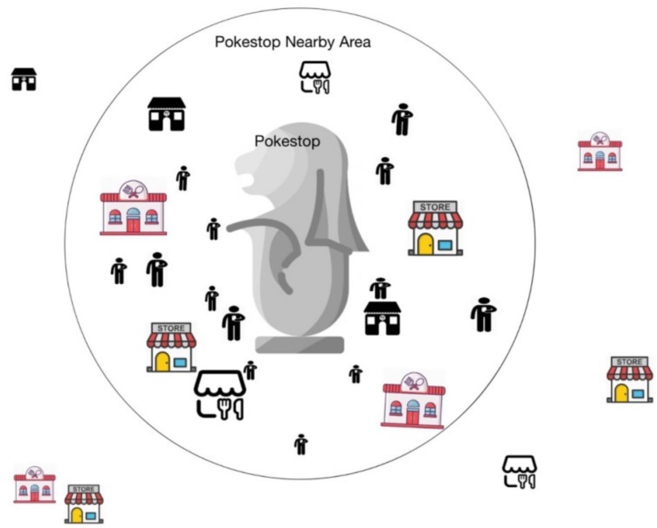


Figure 1. Illustration of the Awareness Mechanism of PokeStops' Attracting Players as Traffic to Nearby Stores

2.2 The AR Literature

By integrating layers of virtual information and virtual contents into users' perception of the real-world (Rauschnabel et al. 2017), augmented reality technologies can enable spatial, social, or real-time interactions between users and the environment or among users, and can provide users an enriched, immersive and interactive experience. Nowadays, AR technologies have been implemented as wearable devices such as Google glasses, and have been widely applied in such business areas as e-commerce, tourism and sightseeing, live broadcasting, and online education. The literature on AR technologies' business impacts consists of mainly conceptual studies (Bigham 2005, Bulearca and Tamarjan 2010, Kunkel et al. 2016) and surveys (Yuan and Wu 2008, Rauschnabel et al. 2017, 2019). Rauschnabel et al. (2019) and Kunkel et al. (2016) both proposed that AR app usage leads to users' positive perceived emotional gratification due to perceived augmentation quality in additional to the utilitarian and hedonic benefits; thus, it can improve a user's awareness, favorability and consideration of the brands reflected and projected by the AR app. Bigham (2005) and Bulearca and Tamarjan (2010) claimed that AR technologies can improve customer satisfaction by enhancing perceived experience value. Yuan and Wu (2008) also found that AR technologies can improve customers' entire experience throughout the purchase stages. These studies suggested that the adoption and use of AR technologies can boost consumers' brand awareness and purchase intentions. However, there lacks empirical evidence regarding how the penetration of AR technologies impacts business performance by shaping the users' behaviors and decisions.

Our study fills in the gap by examining the AR technologies' business impacts based on a natural experiment of the entry of an AR mobile app. Under our research context, when the app users explore the vicinity areas with their mobile devices, AR technologies first scan and project the real-world environment, then add layers of virtual contents (e.g. portals and items), and display the "augmented reality" on the users' mobile devices (Craig 2013). The augmented contents improve the awareness, favorability, and consideration of the reflected brands through the users' perceived emotional gratification, especially the hedonic benefits (Rauschnabel et al. 2019). In other words, activities enabled by AR technologies, such as collecting and consuming virtual items and capturing Pokemon, can improve a user's perceived hedonic benefits of the perceived augmented view of reality, and increase the positive attitude of the brands that appear on the AR camera mode. In addition, given the unique features of PokeStops (catching free virtual items and a higher likelihood of catching more and rarer Pokemon near PokeStops), those AR-enabled activities are more likely to occur in proximity to the physical app portal (PokeStops) than in other places. Thus, players are more inclined to play around PokeStops, where they also gain more hedonic benefits induced by AR technologies. Furthermore, local businesses' brands reflected by the AR mode in the LBAR app have a higher propensity to piggyback off users' gratification halo when they are close to the portals. Therefore, having portals nearby can be considered as an added feature of a local business, e.g. a restaurant. According to Susskind and Chan (2000) and Xiang et al. (2017), having diversified services and hospitality features can have a positive predictive relationship on users' evaluation like Yelp rating. Thus, we can consider the portals predetermined by the LBAR application as a positive external feature of a restaurant, and we expect that the being close to portals can have a positive association with users' evaluation of the restaurant in terms of review ratings.

Based on the above AR marketing theory, we formally propose Hypothesis 1:

Hypothesis 1: The entry of the LBAR application can significantly improve the rating of a restaurant with portals nearby.

2.3 The LBS and LBAR Literature

Location-based technologies (LBT) have been broadly adopted by mobile marketing (Ververidis and Polyzos 2002, Fang et al. 2015). The literature has documented the business and social impacts of LBT applications in the following two major ways.

First, LBT technologies are found to have significant impacts on users' daily activities (Cousins and Varshney 2001), and therefore support and improve the efficiency of proximity marketing strategies such as locational targeting promotions (Fang et al. 2015, Chan et al. 2014, and Andrews et al. 2015). By surveying app users, Cousins and Varshney (2001) showed that mobile advertising through LBT increases users' mobility and has a timely influence on users' purchase intentions. However, we still do not know the potential impacts of LBS apps on local businesses. Through randomized field experiments and surveying focal retailers, Fang et al. (2015) showed that given the real-time and location-sensitive features, LBS technologies can leverage the advantages of geo-marketing, and thus are popularly adopted by businesses as a novel marketing tool. Through field experiments, Chan et al. (2014), Fang et al. (2015), and Andrews et al. (2015) also demonstrated LBS technologies as an effective promotion tool due to their immediate and accumulated impacts on product sales. By constructing and estimating a structural model through data from an online restaurant review platform, Qiu et al. (2018) studied the check-in records of LBS social network users, and found that LBS technologies can diversify the users' and their friends' restaurant discovery and can consequently improve their observational learning outcome. Most of LBS literature investigated the firm usage of LBS technologies for their own marketing purposes, but did not focus on unintended spillover effects of the individual use of LBS apps on neighboring businesses. More importantly, we reveal the mechanism about how local businesses can piggyback off the spillovers of an LBS app. Moreover, though most of the literature on LBS technologies found that the business impact of LBS technologies is positive and effective, the practical puzzle revealed by Filloon (2016) and Zhu (2016) suggested that not all local business managers can rip the benefits of such LBS apps. This paper intends to verify the specific business impacts of an LBS mobile app and uncover the mechanism through comprehensive empirical evidence.

Second, LBT apps with other technologies like AR can directly influence consumer behaviors and subsequently, the performance of local businesses. Through a large-scale survey of Pokemon GO players in different locations, Colley et al. (2017) reported how the game changed players' movement patterns: about 60% of respondents visited at least one new location within a 3-km vicinity area, and about 10% players moved beyond 3 km while playing Pokemon GO. That report also demonstrated very heterogeneous types of newly-explored locations adjacent to PokeStops and Gyms, including city landmark buildings, parks, libraries, universities, museums, stores and restaurants. More importantly, 46% of interviewees reported that they had a consumption or dining experience at nearby venues due to Pokemon GO related activities. Another survey of Pokemon GO players in the US, Zach and Tussyadiah (2017) summarized similar respondent reports, i.e., besides improving players' daily functions and psychosocial functions, Pokemon GO can enhance players' sense of community, mobility and physical activity, as well as the unplanned consumption in retail (11% of respondents), restaurants (29%), services (13%) and travel (17%). The above literature summarizes rich first-hand observations that the LBAR application enlarges its users' mobility in nearby business areas and induces their unplanned consumption behaviors. However, there still lacks research about the business implication of adopting or piggybacking off these LBAR applications. The above literature gaps call for a comprehensive examination of LBAR applications' influence on local business. Motivated thus, we will explain and justify how the entry and penetration of an LBAR application shape the performance of local businesses.

Given the spillover awareness mechanism, related literature, and the important functions of PokeStops, we expect that, PokeStops have a higher likelihood of attracting players to linger around through LBS technologies. When players play at PokeStop areas, LBS technologies can further encourage movement, visits and exploration to new locations and can increase user awareness of surrounding businesses. This discovery process can increase users' likelihood of visiting businesses in the vicinity area. Hence, businesses such as restaurants located close to PokeStops are expected to benefit from the spillover effects of new traffic. In turn, users can post their experiences through reviews on social media. We expect to observe a larger review volume of those restaurants close to PokeStops. Combining the above literature of LBS/ LBAR technologies and summarizing the above reasoning, we propose Hypothesis 2.

Hypothesis 2: The entry of the LBAR application can significantly increase the review volume of a restaurant with portals nearby.

3. Identification Strategy and Data Processing

3.1 Data

We build a rich dataset by merging data from various sources: the Pokemon GO geographic data through PogoDev API[3], restaurant profiles and reviews from Yelp, the app downloads data from Prior Data[4], and Google Trend index of "Pokemon GO". The data were merged by geo-coordinates and date. We also develop new locational attributes such as geo-distance between a restaurant and PokeStops, and PokeStop density around a restaurant.

Following the literature (Xu et al 2016, Imbens and Wooldridge 2009, and Li 2016), we conduct the analyses on a monthly level. As validation, we also run the analyses on a weekly level controlling seasonal volatility in Section 6.6. The official launch date of Pokemon GO was on July 6, 2016. However, due to the large user base of Pokemon series games across different game platforms, and the prominent popularity during the global beta test since March 2016 prior to the US beta test, Pokemon GO has gained a large number of test users since the US beta test at the end of May 2016. Considering the potential influence of the beta versions, we chose June 2016 as the entry month.

3.1.1 Pokemon GO Geographic Data and Pokemon GO App Store Data

The Pokemon GO geographic data include the IDs and geo-coordinates of PokeStops and gyms in the Dallas-Fort Worth area of Texas.[5] This dataset is collected from Niantic via a third-

---

[3] https://docs.pogodev.org/
[4] https://prioridata.com/
[5] We included Gyms in our consideration of PokeStops for the following reasons: 1) Similar to PokeStops, Gyms can also attract visits to nearby locations, and players also need to stay around a Gym if they want to battle with other players. 2) The Pokemon

party API PGOAPI provided by PogoDev. More specifically, we first use the open source tool www.PokemonGOmap.info to visually select all PokeStops and gyms located in the DFW area on a Google map. Then via PGOAPI, we download the geo-coordinates of PokeStops and Gyms, based on which we calculate the locational variables such as PokeStop $Density$, $Competition$ among restaurants, and $Distance$ to nearest PokeStop for each restaurant based on the general form of the Haversine formula[6]

$$hav\left(\frac{d}{r}\right) = hav(\varphi_2 - \varphi_1) + cos(\varphi_1) * cost(\varphi_2) * hav(\lambda_2 - \lambda_1), \quad (1)$$

where hav is the Haversine function: $hav(\theta) = sin^2\left(\frac{\theta}{2}\right) = \frac{1-cos\,(\theta)}{2}$; d is the distance between two locations, r is the radius of the earth (3,959 miles or 6,371 kilometers); φ_1 and φ_2 are the latitudes of locations 1 and 2, respectively; and λ_1 and λ_2 are the longitudes of locations 1 and 2, respectively.

In addition, we collect the Pokemon GO App's monthly downloads (denoted as $Gamedownloads$) from the mobile app market intelligence provider Priori Data, and the Google trend index ($GoogleTrend$) of the key word "Pokemon GO" from https://trends.google.com as the macro level control variables for the game popularity trend.

3.1.2 Restaurants' Business Information and Reviews from Yelp

We choose to focus on restaurants in the Dallas-Fort Worth area of Texas. We obtain an exhaustive list of restaurants by a default organic search by location at Yelp.com. Through Yelp API, we collect the restaurants' profiles and reviews from Yelp.com. Restaurant attributes extracted from business information include restaurant name, address, zip code, price level,

GO in our study was in its first generation, where both Gyms and PokeStops are single-function portals. Thus they are homogeneous in the ways of attracting visits.

6 https://en.wikipedia.org/wiki/Haversine_formula, https://web.archive.org/web/19991010004728/http://www.geocities.com/ResearchTriangle/2363/trig02.html

cuisine text tags, overall ratings and total review volume. Review attributes include posting date, rating, check-ins, Yelper ID, and Yelper characteristics.

Based on the official definitions of chain restaurants[7], we manually examine each restaurant and assign it a $Chain$ dummy. To capture a restaurant's $CuisinesType$, we first classify the customer reported tags of all restaurants to extract 135 unique phrases with the NLP application spaCy, then cluster them into 21 cuisine categories based on Jaccard similarity, and manually label each category by the main cuisine type of the phrase group, such as "American", "Chinese", and "Italian" (See the full list in Table 2). Based on the address and zip code of each restaurant, we collect its location city from the database provided by federalgovernmentzipcodes.us and generate a $City$ dummy vector to control the city level fixed effect. The list includes 36 cities such as Dallas and Garland. Through the Google Map API, we obtain each restaurant's geo-coordinates by its address, through which we can calculate such control variables as $Competition$ among restaurants, $Distance$ to nearest PokeStop, and PokeStop $Density$ for each restaurant via the Haversine formula. We calculate the $Distance$ as the distance between a restaurant and any other one and measure the *Competition* variable by the count of restaurants within a radius of 46 meters[8]. We calculate the distances between a restaurant and any one of the PokeStops, and then sort them to obtain the shortest distance, which is denoted as *Distance* to nearest PokeStop. Since players may walk around the plaza to interact with other PokeStops, we control the PokeStop *Density* which measures the count of PokeStops within a radius of 46 meters. We use a restaurant's historical average rating to replace the missing values in monthly ratings of restaurants.

[7] https://en.m.wikipedia.org/wiki/Chain_store and http://www.nevobusiness.com/independentrestaurants/chain.html.
[8] According to https://en.wikipedia.org/wiki/Strip_mall#cite_note-1, the average size of strip malls in the US is 4880 square meters. Considering the shape of a mall may be irregular polygonal, we take the average (46 meters) between the radius of a round mall and the side of a square mall.

We use the search filter "PokeStop Nearby"[9] provided by the Yelp website and mobile App shortly after the launch of Pokemon GO. The validity of this filter is justified with a regression discontinuity design using *Distance* to nearest PokeStop as the forcing variable in Section 6.5. We use this filter to download the URLs of restaurants with PokeStops nearby as the treatment group, and the rest of the restaurants without PokeStops nearby as the control group. The parallel trend assumption validated in Section 6.3 ensures that in the absence of Pokemon GO, the difference between the treatment group and the control group does not vary over time.

3.2 Descriptive Statistics

After data cleaning, there are a total of 59,999 reviews for 1215 restaurants from January to December 2016. 77.5% of them have at least a PokeStop nearby. Table 1 provides the summary statistics of the key variables and controls.

3.3 Propensity Score Matching

To address potential endogeneity concerns of a restaurant's assignment to the treatment group, we conduct propensity score matching (PSM) before the difference in difference (DID) estimation, as recommended by Li (2016) and Xu et al. (2017). The PSM procedure ensures that the matched pairs of treated and control restaurants have a similar probability of receiving the treatment (having a PokeStop nearby). Thereby, all the processed observations are comparable, so that we can eliminate the self-selection bias and endogeneity concern. Matched observations can improve the propensity that our estimation meets the parallel trend assumption, which is a counterfactual condition that if there were no Pokemon GO entry or PokeStops, restaurants in the treated and the control groups would have identical time trends in terms of reputation metrics. In other words, if the estimated DID coefficients are significant, it is caused by the treatment

[9] https://blog.yelp.com/2016/07/catch-em-yelps-new-PokeStop-filter.

conditions but not by alternative factors. The parallel trend assumption is formally tested through dynamic DID estimations (a relative time approach) in Section 6.3.

The objective of PSM is to find matching pairs of treated (a restaurant with a PokeStop nearby) and control (one without) restaurants before the entry of Pokemon GO. Then we drop the unmatched observations and keep the matched ones for DID estimation. Observable restaurant characteristics are chosen as pre-treatment covariates, including a restaurant's overall rating, total review volume, chain or non-chain category, cuisine type, price level and local competition. Next we estimate the probability of a restaurant receiving the treatment as the function of the above six covariates through a probit regression (results in Table 2a). There are nineteen covariates significantly associated with a restaurant's receiving the treatment on the 0.05 significance level. Accordingly, we obtain the predictive propensity score for each control and treated restaurant.

To avoid the potential confounding concern caused by the entity order, we first randomly sort the dataset. Then, we match the treatment and control groups based on their predicted propensity scores, which are estimated by the above probit model (Table 2a). We adopt the Nearest Neighbor (NN) matching algorithm with replacement following Rishika et al. (2013) and Li (2016). According to Marco and Sabine (2005), one of the advantages of this algorithm is that it ensures high matching quality and can reduce matching bias when the propensity scores of the treatment and control groups are not close. Recommended by Smith and Todd (2005), we also impose the matching with a caliper (0.05), which is the tolerance level on the maximum propensity score distance. By doing so, we can avoid bad matches when the distance between the treated and control restaurants is large. A high tolerance level can ensure that the obtained matches have a high level of propensity similarity and a high matching quality. The PSM yields

273 matched untreated restaurants. To balance the size of the control and treatment groups, we keep equal number of treated restaurants that have the closest propensity scores to their untreated counterpart. We end up with 273 pairs, or 546 restaurants in our analyses. To evaluate the matching quality, we run t-tests on the means of matched treated and control restaurants. The results in Table 2b suggest that after matching, at a 0.05 significance level, a control restaurant and a treated one have the same propensity of receiving treatment. Thereby, the current matched control and treatment groups are comparable, and the balance matched observations allow us to use the restaurants without PokeStops nearby as the benchmark to examine the impact of the entry of Pokemon GO through a DID analysis.

4. Econometric Analyses

4.1 Econometric Model

In order to capture the causal impacts of the entry of Pokemon GO on the online reputation of restaurants with PokeStops nearby, we adopt the Difference-in-Difference (DID) model in our main analysis. Specifically, we examine the changes of the reputation metrics of restaurants with PokeStops nearby before and after the launch of Pokemon GO, relative to those of restaurants without PokeStops nearby over the same time period. For more results and robustness checks, we adopt additional models and conduct further analyses in Sections 5 and 6, e.g. DDD, polynomial regression, and reversal regression.

Inspired by Angrist and Pischke (2008) and Imbens and Wooldridge (2008), we have the following main econometric estimation models:

$$Volume_{i,t} = \alpha_0^V + \alpha_1^V PKG_i + \alpha_2^V d_t + \beta^V (PKG_i * d_t) + \Theta^V R_i + \Psi^V X_t + \eta_i + \gamma_t + \varepsilon_{i,t}^V \quad (2)$$

$$Rating_{i,t} = \alpha_0^R + \alpha_1^R PKG_i + \alpha_2^R d_t + \beta^R (PKG_i * d_t) + \Theta^R R_i + \Psi^R X_t + \eta_i + \gamma_t + \varepsilon_{i,t}^R \quad (3)$$

We use online reputation metrics review volume ($Volume_{i,t}$) and rating ($Rating_{i,t}$) as proxies for the foot traffic and user evaluation, respectively, of restaurant *i* at period *t* (Hu et al. 2017, Moe and Schweidel 2012, Luca 2016). In addition to these two metrics, customers' social check-ins are shown to have significant impacts on their peers' restaurant awareness, discovery and selection (Qiu et al. 2018), and Yelp Elite reviewers are also considered important for restaurants to maintain a good reputation (Luca 2016). Thus, check-in volume ($Checkin_{i,t}$) and Elite review volume ($Elite_{i,t}$) can also be business performance indicators. Due to the generality concern regarding these metrics, we examine the impacts of entry of the LBAR app on restaurants' check-in volume in Section 5.1, and on the Elite review volume in Appendix B.

In models (2) and (3), PKG_i is the treatment dummy: $PKG_i = 1$ when restaurant *i* is located near PokeStops, and 0 otherwise. d_t is the time switch dummy: $d_t = 1$ when it is after June 2016, and $d_t = 0$ otherwise. R_i is vector of the observable restaurant time-invariant characteristics: $R_i = (Density_i, Competition_i, Distance_i, PriceLevel_i, CuisinesType_i, City_i, Chain_i)$. $PriceLevel_i$ is a price level category vector, including four price level dummies P_1, P_2, P_3, and P_4, which represent the restaurant's price levels $ (under $10), $$ ($11 to $30), $$$ ($31 to $60), and $$$$ (above $61), respectively.[10] X_t is the vector of macro-level time-variant controls, including Google Trend monthly index ($GoogleTrend_t$), representing Pokemon GO's general popularity, and Pokemon GO's monthly downloads ($Gamedownloads_t$). η_i is the restaurant fixed effect, γ_t the time fixed effect, and $\varepsilon_{i,t}$ the residual. We take natural log for the dependent variables, $Density_i$, $Competition_i$, $Distance_i$, and $Gamedownloads_t$ to remove the scale effects.

---

[10] https://www.yelp.com/topic/san-diego-can-anyone-give-me-the-actual-dollar-range-for-the-dollar-sign-symbols

The two DID coefficients β^V and β^R estimate the average effects of the entry of Pokemon GO on the treated restaurants' review volume and rating, respectively. They are of key research interests. Considering the game effect may be saliently different between a short-term window and a long-term one, we conduct the DID estimations in both terms. Given June 2016 as the treatment launch time, we symmetrically select three months prior to and post this launch time (March to September 2016) as the short-term window, and select January to December 2016 as the long-term window.

4.2 Findings

The DID estimation results in both the short and the long terms are reported in Table 3. The short-term results validate the hypotheses that Pokemon GO positively impacts local restaurants in terms of their online reputation. They suggest that after the launch of Pokemon GO, being close to PokeStops consequently improves the restaurant's online review volume by 6.5% (p<0.05), and the average rating by 5.7% (p<0.01) in the short term. In the long term, only the DID coefficient of review rating is significantly improved by 8% (p<0.01), but the positive short-term DID effect on review volume diminishes to a nonsignificant level.

In sum, the launch of Pokemon GO significantly enhances restaurants' online rating in both time frames and review volume in the short run in our research setting. Thus both H1 and H2 are supported. Based on the AR marketing theory, as an LBAR app, Pokemon GO can have positive spillover on the local businesses that appear in the AR views during users' gameplay. This positive spillover effect enhances the value of those restaurants located in proximity of PokeStops. Thus, as a positive and desirable attribute, "PokeStop nearby" contributes to customers' perceived quality of a restaurant and thus improve their evaluations in terms of online ratings. The improvement in the ratings of the treated restaurants validates H1, and justifies and extends the current AR literature (Rauschnabel et al. 2019, Bulearca and Tamarjan 2010).

The significant short-term boost of treated restaurants' review volume supports H2. However, this impact tapers after the short term and diminishes to a non-significance level in the long term. Since a customer can only post one review per restaurant on Yelp[11], the review volume of a restaurant reflects the number of new customers but not repeated ones. After its entry, Pokemon GO has gained extreme popularity rapidly, resulting in a significant increase of players who move around and explore nearby areas hunting Pokemon. This LBAR app changes the users' online and offline social and physical activities within a very short time frame. When they find interesting landmarks or landscape and experience dining services from restaurants in the vicinity of PokeStops, these players may become customers and post their experiences on social media, which increase the review volume of the restaurants. However, as the phenomenal popularity of Pokemon GO dies down quickly, the monthly US downloads dropped from the peak of 103.4 million to 2.7 million by December 2016. As a result, the long-term review volume increment becomes insignificant due to the decrease of new traffic.

5. Additional Analyses

To cross validate our main results and also to identify richer patterns, we conduct additional analyses: (1) we check the impact of the entry of the LBAR app on another performance indicator, $Checkin$ volume, of the treated restaurants (Section 5.1); (2) we show the varying impacts of Pokemon GO on restaurants of heterogeneous established reputation levels (Section 5.2); (3) to verify the awareness mechanism proposed in Section 2.1, we inspect the treatment

[11] New experiences of the same customer on the same restaurant have to be posted as updates to the existing review and rating.

intensity from a geographical perspective by replacing the DID interaction term with the polynomial terms of the distance between a restaurant and its closest PokeStop and conducting a *polynomial regression* (Section 5.3); (4) to further examine how the internalization of the spillover effects depends on restaurants' established reputation levels, we conduct the *Difference in Difference in Difference (DDD) analyses* for the long-term (Section 5.4); (5) to further check potential *moderation effects*, we incorporate more interaction terms (e.g. between game downloads (Section 5.5), the chain/non chain dummy (Section 5.6), the competition level (Section 5.6) and the treatment condition) into the DID estimation models through a series of DDD analyses.

5.1 The DID Analysis on Check-ins

Yelp allows its users to "check in" by sharing their location information on social media when visiting a restaurant. By using the check-in feature, the customers broadcast to their followers on Yelp that they are at that restaurant. Accordingly, check-ins will potentially increase the awareness of the restaurant through observational learning and peer influence. Check-ins may not directly measure a restaurant's online reputation; however, it is a critical metric that business owners value (Qiu et al. 2018). We consider the volume of check-ins as another business performance indicator and conduct a DID estimation for both the short and long terms to examine the impacts of Pokemon GO on the treated and control restaurants (Equation 4).

$$Checkin_{i,t} = \alpha_0^C + \alpha_1^C PKG_i + \alpha_2^C d_t + \beta^C (PKG_i * d_t) + \Theta^C R_i + \Psi^C X_{i,t} + \eta_i + \gamma_t + \varepsilon_{i,t}^C \quad (4)$$

The results are reported in Table 4. After the entry of Pokemon GO, the check-in volume of restaurants with PokeStops nearby insignificantly increases in the short-term, and the increment accumulates to a significant level by 1.5% (p<0.05) in the long term. On Yelp, unlike posting reviews, a user can check-in to a restaurant multiple times. Thus check-in volume can reflect the traffic of both new and repeat customers to a restaurant. While review volume result (Table 3) suggests that the entry of Pokemon GO causes an immediate increase in new traffic to restaurants in proximity to PokeStops, even though that effect tapers off with the cooling down of the app downloads in the long run, the check-in result in Table 4 shows that Pokemon GO can still cause significant increase of repeated visits of the app users in the long term.

5.2 Findings of Partitioned Samples

Moreover, to uncover Pokemon GO's impacts on restaurants that are heterogeneous in established online reputation, we adopt the median split method used in Li (2016) and Demers and Lewellen (2003): we first partition the full sample into 4 subsamples by the median of overall review volume (MV) and the median of overall ratings (MR), then run DID estimation for each partitioned sample, and finally summarize and compare the results.

The full sample of restaurants is partitioned by the median of each restaurant's average annual review volume (32), and by the median of the average rating (3.9) separately. These two annually average metrics, calculated over the time periods from the inclusion year of the restaurant on Yelp up to 2015, describe a restaurant's *established* online reputation prior to the entry of Pokemon GO. We obtain four subsamples: low-volume restaurants (those with annual review volume lower than 32), high-volume ones (those otherwise); low-rating restaurants (those with an average overall rating lower than 3.9), and high-rating ones (those otherwise).

For each subsample, we conduct DID estimation in both short and long terms on review volume, rating and check-in volume and summarize the results in Table 5. From the short term to long term, the results of Pokemon GO's impacts on these reputation metrics are generally consistent with those of the full sample estimation. The results of the different partitioned samples also demonstrate significant variation. For the high-review-volume treated restaurants, the entry of Pokemon GO significantly improves review rating by 6.5% ($p<0.01$) in the short term and by 9.8% ($p<0.01$) in the long term; the DID effect of check-in volume is positive but insignificant in the short term but accumulates to a significant improvement of 1.4% ($p<0.05$) in the long term. For the low-review-volume restaurants with PokeStops nearby, the entry of Pokemon GO significantly improve their ratings by 6.3% ($p<0.01$) in the short run and by 7.4% ($p<0.01$) in the long run; check-in volume is insignificantly associated with the entry of the game in both terms. Estimation results of samples partitioned by established rating show a similar pattern. The review volume increments of treated restaurants with established high reputation are significantly improved for both the short and long terms by 5.2% and 4.3% ($p<0.05$), respectively. The increment of check-in volume of high-established rating restaurants is significant only in the long term, 1.1% ($p<0.05$). For treated restaurants with a low established rating, the entry of Pokemon GO significantly boosts the review volume by 4.5% ($p<0.1$) in the short term and the increase diminishes to a non-significant level in the long term; the check-in volume is insignificantly and positively associated with the entry of the game in both terms.

The above results about the DID effect demonstrates a *rich-get-richer* pattern, that is, the positive influence of the LBAR app only favors treated restaurants that have good established reputation. More specifically, in the long term, only restaurants of high established reputation (rating or review volume) can benefit from Pokemon GO in terms of more check-ins, only high-rating restaurants can significantly increase review volume, and high-review-volume restaurants incur a greater boost in their ratings than low-volume ones. This rich-get-richer phenomenon can be explained by consumers' consideration of online reputation when selecting restaurants in PokeStop vicinity areas. Hence, for reputed restaurants, having PokeStop nearby serves as a superior attribute that can attract even more new visits, turn players into customers or even repeat customers, and eventually further improve their online reputation over time.

5.3 Treatment Intensity Analysis - Polynomial Regression with the Distance Variable

In the previous analyses, the treatment condition of a restaurant is determined by whether it has a PokeStop nearby. However, it is unclear if the treatment effects increase or decrease for a business closer to or further away from PokeStops. To examine whether and how the treatment intensity shapes the effect size, more specifically, to examine if the geographical distance plays a primary role in determining the effect size of PokeStops on local businesses, we run the following polynomial regression on review volume and review rating for the long term by entering the distance to the closet PokeStop and its square term as indicators of treatment intensity.

$$Volume_{i,t} = \alpha_0^V + \alpha_1^V Distance_i + \alpha_2^V Distance_i^2 + \Theta^V R_i + \Psi^V X_{i,t} + \eta_i + \gamma_t + \varepsilon_{i,t}^V \tag{5}$$

$$Rating_{i,t} = \alpha_0^R + \alpha_1^R Distance_i + \alpha_2^R Distance_i^2 + \Theta^R R_i + \Psi^R X_{i,t} + \eta_i + \gamma_t + \varepsilon_{i,t}^R. \tag{6}$$

The results of the polynomial regression in summarized in Table 6. We find that two treatment intensity indicators, the distance and the area size, are significantly and negatively associated with both review volume and rating. More specifically, a 1% distance closer to the closest PokeStop can increase the restaurant's review volume by 2.20% (p<0.01), and increase the review rating by 0.15% (p<0.05). In addition, the significant and negative coefficients of the *Distance*^{2} term suggest that the spillover effects of the LBAR app on reputation metrics diminish with the distance to PokeStops at an increasing rate. Thus, the geographical distance has a significant direct relationship with the reputation of local businesses.

This finding strengthens our hypotheses and provides powerful support to our awareness mechanism. App users are attracted by PokeStops in order to interact with them for the many useful features. From there, they wonder around and linger in the adjacent shops and restaurants as new traffic. The result also implies that being closer to a PokeStop (a decrease in distance to PokeStops) is an attractive attribute of restaurants, and contributes to the positive spillover effects at an increasing rate.

5.4 Moderation Effects of Established Reputation through DDD Analyses

To examine the how the established online reputation moderates the DID effects, we conduct Difference in Difference in Difference (DDD) analyses on the potential moderation effects of the established review volume and rating. According to Gruber (1994) and Wooldridge (2007), the DDD analysis can recover all pairwise interactions, and demonstrate the moderation impacts of the third "difference" factor. Meanwhile, this method can control potential confounding effects due to the different trends across the treated and control groups. Thus, the results of the following DDD analyses can also further validate our previous DID estimation findings. The DDD models of established reputation metrics are:

$$Volume_{i,t} = \alpha_0^V + \alpha_1^V PKG_i + \alpha_2^V d_t + \beta_{DD}^V (PKG_i * d_t) + \rho^V HR_i + \beta_{DDD}^V (PKG_i * d_t * HR_i) + \Theta^V R_i + \Psi^V X_{i,t} + \eta_i + \gamma_t + \varepsilon_{i,t}^V \quad (7)$$

$$Rating_{i,t} = \alpha_0^R + \alpha_1^R PKG_i + \alpha_2^R d_t + \beta_{DD}^R (PKG_i * d_t) + \rho^R HV_i + \beta_{DDD}^R (PKG_i * d_t * HV_i) + \Theta^R R_i + \Psi^R X_{i,t} + \eta_i + \gamma_t + \varepsilon_{i,t}^R \quad (8)$$

$$Checkin_{i,t} = \alpha_0^C + \alpha_1^C PKG_i + \alpha_2^C d_t + \beta_{DD}^C (PKG_i * d_t) + \rho_V^C HV_i + \beta_{DDD,V}^C (PKG_i * d_t * HV_i) + \rho_R^C HR_i + \beta_{DDD,\mathrm{R}}^V (PKG_i * d_t * HR_i) + \Theta^C R_i + \Psi^C X_{i,t} + \eta_i + \gamma_t + \varepsilon_{i,t}^C \quad (9)$$

The dummy variables HV_i and HR_i are obtained by the partitioned sample in Section 5.2: HV_i equals 1 if the median of restaurant *i*'s established review volume is larger than or equal to 32, and HR_i equals 1 if the median of restaurant *i*'s established review ratings is greater than or equal to 3.9. β_{DDD}s are the DDD coefficients, and they estimate the average moderation effects of established online reputation metrics on Pokemon GO's DID effects. The DDD analysis is conducted on the matched pairs of observations for both the long and the short terms. The results are summarized in Table 7.

The coefficients of interaction terms illustrate how the established reputation moderates the DD effects. In the long run, if a treated restaurant has built a high reputation in the previous years, the established high rating significantly positively moderates the Pokemon GO's DID effects on review volume and check-ins, the established high review volume significantly positively moderates the Pokemon GO's DID effects on check-ins. Moreover, the significance level of the moderation effects increases from the short term to the long term. All of these phenomena are consistent with the rich-get-richer pattern in the partitioned sample analyses in Section 5.2. Thus "having PokeStop nearby" can be considered as long-lasting added value to the restaurants with an established reputation.

5.5 Moderation Effect of Game Download through DDD

To check potential moderation effects caused by game downloads on the treatment density, we incorporate the interaction term and re-run the DID estimation for the long term.

$$DV_{i,t} = \alpha_0^V + \alpha_1^V PKG_i + \alpha_2^V d_t + \beta_{DD}^V (PKG_i * d_t) + \rho_{Download}^V Download_t + \beta_{DDD,Download}^V (PKG_i * d_t * Download_t) + \Theta^V R_i + \Psi^V X_{i,t} + \eta_i + \gamma_t + \varepsilon_{i,t}^V, \quad (10)$$

where $DV_{i,t}$ represents review volume $Volume_{i,t}$, $Rating_{i,t}$, or $Checkin_{i,t}$. $Download_t$ measures the volume of the first-time downloads, excluding the volume of app updates. This metric can reflect the game's popularity on the app market among mobile users. Estimation of the long-term results are presented in Table 8.

The game download volume has a significantly positive direct effect on restaurants' reputation metrics. More importantly, game downloads significantly strengthen both the DID effects on review volume and check-in volume in the long run. This result validates that the app users contribute to the traffic and volume of social media posts of the restaurants in proximity to PokeStops.

5.6 Moderation Effects of Restaurants' Competition Level and Chain/Non-Chain Condition

To further examine whether a restaurants' competition level and its chain/ non-chain feature strengthen or suppress the treatment density after the entry of Pokemon GO, we incorporate the competition metric and chain/ non-chain dummy with their interaction terms to re-run the DID estimation for the long-term frame. We summarize the results in Table 9.

Results in Table 9a show that the competition level of a restaurant is positively associated with its review volume. This is consistent with the key finding of Liu et al. (2018) that geographic concentration, or agglomeration of restaurants can be a driver of the volume of online reviews. More importantly, we find that business agglomeration can also strengthen the spillover effects of Pokemon GO on restaurants' online reputation. That is, PokeStops attract the Pokemon GO app users to gather around, which creates the patronage effect for the nearby businesses. Then among the restaurants close to PokeStops, those located in a higher concentrated area will attract more visitors due to the agglomeration effect (e.g. more dining choices). And those visitors become potential customers and are expected to generate greater amount of reviews. In addition, restaurants that can thrive at a high-level competitive business environment provide certain unique features and experiences. The virtual contents in the LBAR app can augment customers' real experiences, improve the variety of services and features of the treated restaurants, thereby enhance the customers' evaluation.

Additionally, results in Table 9b suggest that chain restaurants benefit less of the positive spillover effect from the entry of the LBAR application. For chain restaurants, such as McDonalds and Chipotle, customers were already aware of the brands, and formed perceived expectation and evaluation before visiting the stores. However, for non-chain restaurants, the LBAR app can substitute the branding effect to a certain degree by promoting the awareness, consideration and favorability to the app users. Thus, independent or non-chain restaurants will benefit more piggyback effects from the LBAR app.

6. Robustness Validations

To validate the above findings, we discuss the potential issues and conduct a series of robustness tests. First, we prudently scrutinize and summarize all the official documents, updates and announcements regarding the portal nominations to exclude the potential self-selection and endogeneity issues (Section 6.1). Second, in order to remove the reversal causality concern, we run a falsification test by regressing reputation metrics back on the DID indicators (Section 6.2). Third, to check the parallel trend assumption, we use the relative time approach to examine the dynamic DID effects for each period (Section 6.3). In Section 6.4, we check the potential confounding issues. In Section 6.5, we design a regression discontinuity regression for the control group validation. At last, we validate our DID estimation results at a different time frequency--- the weekly level (Section 6.6).

6.1 Self-Selection Discussion Regarding Application Policy

The potential self-selection bias could jeopardize the randomness of the treatment and might threaten the validity of the DID estimation results. However, according to the Pokemon GO game introduction on the official video game media websites and the discussions of the app developers on online communities[12], in our research setting and during our data collection time period (the first generation Pokemon GO), restaurant owners cannot turn their business locations into PokeStops. In Pokemon GO's first generation, including client versions 0.29.0_1.00 to 0.51.0_1.21.0, all PokeStops and Gyms are imported from the portal database of Ingress, which is a previously launched AR mobile game by Niantic[13]. PokeStops or Ingress Portals are crowdsourced and are generally considered "interesting" locations by the public. Yet, players cannot create new portals or PokeStops in the game[14]. Instead, all candidate portals or PokeStops must be submitted to Niantic through an application process. Nonetheless, Niantic suspended the application process since September 2015 and started to accept new applications since September 2017[15]. From the time frame perspective, during our sample time span of the year 2016, there are no new PokeStops, and thus, restaurant owners have no opportunity to self-select their locations to become PokeStops. Additionally, all PokeStops' coordinate information is fixed at least during our research time frame.

Second, we search all related official documents to examine whether any features of PokeStops or Portals can potentially cause the alternative difference between the treated and control restaurants. The PokeStop/Portal nomination submission guidelines[16] by Niantic state that PokeStops or Portals are aimed to help players explore, discover, and enjoy vicinity areas. Eligible nominations include locations with historical, educational, or cultural value and significance, with pieces of art or unique architecture, and locations like hyper-local spots or

12 https://www.destructoid.com/review-pokemon-go-374630.phtml, https://venturebeat.com/2015/12/16/how-niantic-will-marry-animated-characters-with-mobile-location-data-in-pokemon-go/, https://mashable.com/2016/07/10/john-hanke-pokemon-go/, https://www.ibtimes.com/pokemon-go-map-updated-osm-google-maps-what-openstreetmap-2622624
13 https://en.wikipedia.org/wiki/Pok%C3%A9mon_Go, https://www.ibtimes.com/pokemon-go-map-updated-osm-google-maps-what-openstreetmap-2622624
14 https://ingress.fandom.com/wiki/Portal
15 http://ingressportal.com/research/portals/creating-new-portals/, https://pokemongohub.net/post/news/ingress-portal-submissions-reopened/
16 https://niantic.helpshift.com/a/pokemon-go/?s=Pokestops&f=submitting-a-Pokestop-nomination&l=en&p=web, https://niantic.helpshift.com/a/pokemon-go/?l=en&s=Pokestops&f=what-makes-a-high-quality-Pokestop&p=web

described as “hidden gem”, such as tourist or adventurous attractions, public parks, libraries or transportation hubs. Niantic also mentions locations that are ineligible to be PokeStops or Portals include but are not limited to schools, private properties, temporary and mobile locations, cemeteries, indoors, generic business locations, and natural landscapes. Based on the above official selection criteria, restaurant owners can hardly turn their business locations to be PokeStops and cannot intervene in the treatment conditions in our research setting.

6.2 Falsification Test through Reverse Regression

To validate the DID estimation results, formally remove the reversal causality concerns, and to invalidate the claim that online reputation metrics of businesses presiding in that location determine the presence of PokeStop in that area, we first perform the following reverse regression by regressing reputation metrics back on the DID indicators:

$$DID_{i,t} = \gamma_0^V + \gamma_1^V Volume_{i,t} + \gamma_2^V R_i + \Psi^V X_{i,t} + \eta_i + \gamma_t + \varepsilon_{i,t}^V \tag{11}$$

$$DID_{i,t} = \gamma_0^R + \gamma_1^R Rating_{i,t} + \gamma_2^R R_i + \Psi^R X_{i,t} + \eta_i + \gamma_t + \varepsilon_{i,t}^R \tag{12}$$

where $DID = PKG_i * d_t$.

Next, we incorporate the first order lag term of online reputation metrics $Volume_{i,t-1}$ and $Rating_{i,t-1}$ in the above regressions. The results are summarized in Table 10a and 10b. Neither reputation metrics nor their lag terms is significantly associated with the DID effects. Thus, the results illustrate that online reputation metrics do not reversely predict restaurants' treatment condition, or the DID effects. The reversal causality concern is addressed.

6.3 Relative Time Approach

We next test whether our DID analyses meet the parallel trend assumption, which guarantees that the counterfactual treatment group and the control group have the same time trends. If that assumption holds, there is no pre-treatment heterogeneity in trends between the control and the treatment groups. In other words, if the parallel trend assumption holds, without the entry of Pokemon GO, there is no difference in terms of the changes of online reputation metrics between the treated and control restaurants. Meeting this assumption can eliminate the potential threats caused by self-selection bias (Gao 2016, Chan et al. 2014).

A rigorous and widely-adopted method to verify the parallel trend assumption (Lu et al. 2017, Liu and Lu 2015) and to examine the validity of the DID estimation (Autor 2003 and Pischke 2005) is the relative time approach. Typical relative time estimation is conducted by reforming the general form of DID model and rewriting the time fixed effects into an additional set of time dummies to measure the distance between the current time and the time when treatments initiated. In our research setting, the Pokemon GO entry time is the same for all restaurants. We can rewrite the time treatment interaction for each period to estimate the dynamic DID effects based on the relative distances to the treatment entry time without having to generate the entity specific relative distance dummies. More specifically, our strategy is, first, to include interactions between the time fixed effect and the treatment dummy for the four pre-treatment months (January to April) ahead of the game entry month and to remove the interaction for the last pre-treatment month (May) given the dummy variable trap. Second, we rewrite the interactions related to the month prior to the treatment, May, which serves as the baseline. Thus, if our online reputation metrics satisfy the parallel trend assumption, the pre-DID coefficients (January to May) would all be insignificant while at least some of the post-DID coefficients (June to December) would be significant. According to Autor (2003) and Pischke (2005), the distinctive advantage of this method is that the interaction terms after treatment are shown in a dynamic way, which illustrates how the DID effects change over time. Specifically, the above method is to expand our main models (2) and (3) into generalized expressions and form the interaction such that:

$$Volume_{i,t} = \sum_j \beta_j^V * PriorPKG_{i,t}(j) + \delta^V * PKG_{i,t} + \sum_k \beta_k^V * PostPKG_{i,t}(k) + \theta^V R_i + \psi^V X_{i,t} + \eta_i + \gamma_t + \varepsilon_{i,t}^V \quad (13)$$

$$Rating_{i,t} = \sum_j \beta_j^R * PriorPKG_{i,t}(j) + \delta^R * PKG_{i,t} + \sum_k \beta_k^R * PostPKG_{i,t}(k) + \theta^R R_i + \psi^R X_{i,t} + \eta_i + \gamma_t + \varepsilon_{i,t}^R \quad (14)$$

According to Autor (2003) and Gong et al. (2017), the subscripts *j* and *k* are the chronical time distances. β_j describes the treatment's lag effect—effects before the event. Accordingly, β_k depicts the treatment's lead effect—effects after the event. If the parallel trend assumption holds, all lags (β_j) should be insignificant, meanwhile all or partial leads (β_k) should be significant. The dynamic effect results for both reputation metrics are summarized in Table 11.

The results in Table 11 show that all lag coefficients are insignificant and most of the lead coefficients are significant. Thus, the parallel trend assumption is verified under our research setting.

6.4 Confounding Check

Another potential threat might be caused by Yelp's promoting action—the launch of the "PokeStop nearby" filter after Pokemon GO's official launch. This potential confounding effect might affect the Yelp reviews of those restaurants that are close to PokeStops. If Yelp's action influences the treatment group significantly, the DID estimation could be upwards biased.

In order to examine and eliminate the confounding threat to our estimation results, we consider the launch of the Yelp filter as the treatment switch and operate the DID estimation on the reputation metrics. More specifically, through weekly data after the same propensity score matching process, we examine the changes of review volume and review rating before and after the launch of the Yelp's filter between restaurants with PokeStops nearby with those without PokeStops nearby. Yelp.com launched the "PokeStop nearby" filter and posted an announcement on their official blog on July 15, 2016. Thereby, we choose the third week of July as the time switch. We consider the first two weeks of July as the pre-period and the second two weeks as the post-period. The DID coefficient of review volume is 0.063 (Std. Err = 0.129) and the coefficient of rating is 0.003 (Std. Err = 0.067). All DID coefficients given by Yelp's reaction after Pokemon GO's launch are non-significant. Therefore, our results are free of the potential effect due to the launch of the Yelp filter, and our estimation is unbiased.

6.5 Control Group Validation through Regression Discontinuity

We have used the Yelp filter to assign restaurants into the treated and control groups in our previous analyses. To justify the validity of the control group, we adopt the Regression Discontinuity (RD) design and use the geo-distance between restaurants and the nearest PokeStop as the forcing variable. According to Anderson and Magruder (2012), the cutoff value c is determined by the intersection of kernel density distance distributions of filtered treated and control groups as:

$$Volume_{i,t} = \alpha_0^V + \beta^V * I(Distance_i \leq c) + \alpha_1^V * (c - Distance_i) + \alpha_2^V * (c - Distance_i) * I(Distance_i \geq c) + \Theta^V R_i + \Psi^V X_{i,t} \quad (15)$$

$$Rating_{i,t} = \alpha_0^R + \beta^R * I(Distance_i \leq c) + \alpha_1^R * (c - Distance_i) + \alpha_2^V * (c - Distance_i) * I(Distance_i \geq c) + \Theta^V R_i + \Psi^R X_{i,t} \quad (16)$$

where $I(Distance_i \leq c)$ is an indicator function, which equals 1, suggesting that restaurant i is determined as a treatment entity when $Distance_i \leq c$; and equals 0 (restaurant i as a control entity) otherwise. The indicator function intervenes the discontinuity in each dependent variable. The coefficient β^V or β^R estimates the causal effect of having a PokeStop nearby, or more specifically, having a shorter "distance to the nearest PokeStop" than the cutoff. As suggested by Li (2017), we include vectors of restaurant characteristics and macro level control covariates, including ($Density_i$, $Competition_i$, $PriceLevel_i$, $Chain_i$, $GameDownloads_t, GoogleTrend_t$).

The RD regression results are reported in Table 12. The results show that being close to a PokeStop can significantly improve the restaurant's review volume and rating by 41.9% (p<0.01) and 4.8% (p<0.05), respectively. These results are consistent with our DID estimation. Since the design of RD can guarantee local randomization of the treated and control restaurants near the cutoff line, the consistent results verify the control validity in the previous DID and DDD analyses.

6.6 Validation through a Weekly Frequency

For robustness, we also validate our DID estimation results at a different time frequency, i.e. on a weekly level. Given that seasonality and holiday may confound the estimation results, we incorporate time fixed effect and a holiday dummy[17] in the weekly data DID estimation on both review volume and rating. The primary results are reported in Table 13 and the full results in Appendix C. They are consistent with our previous monthly results in Table 3. More specifically, after the entry of Pokemon GO, review volume and rating of the treated restaurants increase by 2.1% (p<0.01) and 3% (p<0.01) respectively in the short term. The long-term review rating increases by 11.9% (p<0.01) and the review volume does not show significant increase in the long term.

7. Discussions and Implications

7.1 Theoretical Implications

[17] According to https://www.officeholidays.com/countries/usa/2016, weeks that include federal holidays are assigned 1 in the dummy matrix, and 0 otherwise.

The theoretical assessments and empirical findings in this paper advance our understanding regarding the business impacts of AR, LBS, and LBAR technologies. Through surveys and interviews, current AR literature (Bulearca and Tamarjan 2010, Rauschnabel et al. 2019) conceptually reasoned how the adoption and use of AR technologies can improve users' perceived emotional gratification and brand awareness, favorability and consideration, and can improve customers' entire experience throughout the purchase stages. However, in addition to user self-reported intentions and perceptions, very few studies have examined the impacts of AR on users and businesses through directly measuring users' brand evaluations as empirical evidence. In this sense, our study is among the first to examine the business impacts of AR technologies by providing direct, rich, and detailed empirical evidence to its positive influence on users' brand evaluation in terms of review valence. We find that the LBAR app can improve users' ratings of local restaurants. That is, the virtual content generated by the LBAR app can add value to users' experiences at a business, which is discovered through the app or while using it. Moreover, our research extends the AR marketing studies by examining the positive spillover effects of AR applications on consumers' evaluations on local businesses. Our finding uncovers the potential business value and positive externality of the LBAR app to local businesses.

LBS literature studied how LBS technologies impact on users' daily activities; however, very few studies focus on the potential impacts of LBS apps on local businesses. Moreover, most LBS literature only focuses on firms' applying the LBS technologies for their own marketing purposes, but did not study the unintended spillover effects of the individual use of LBS apps on neighboring businesses. To meet these two research gaps, we validate the LBS technologies' effectiveness in improving local businesses through the following perspectives:

First, we examine the LBS applications' positive spillover to location-related businesses and find that the entry of the LBAR app benefits restaurants near the physical app portals by improving their online reputation in both the short and long terms.

Second, to understand and explain the inconsistency between LBS literature and practical puzzles (Filloon 2016 and Zhu 2016) regarding the LBS technologies' proximity marketing efficiency, we investigate the internalization of the spillover effects for heterogeneous restaurants. Through a series of partitioned sample analyses and DDD analyses, our study reveals a rich-get-richer pattern of the LBAR app's positive spillover effects on restaurants with different features (e.g. competition intensity and other locational factors) and reputation levels. In addition, we also show the spillover effect of the LBAR app on a restaurant is strengthened when it is surrounded by an agglomeration of restaurants, and diminishes with the geographical distance from the portals at an increasing rate. Additionally, in the long run, local businesses with a higher reputation located in proximity to app portals or in a more concentrated business area will benefit more from the spillovers of the LBAR app.

At last, more importantly, we propose and verify the awareness mechanism about how local businesses can piggyback off the spillovers of an LBS app. The LBAR app can increase users' awareness of surrounding businesses by encouraging movement and visits to new locations and lingering behaviors, which can result in new foot traffic and potential business opportunities to nearby restaurants as well as new dining experiences to the users. Hence, the LBAR app can increase users' likelihood of visiting the nearby businesses, and in turn, post reviews on social media.

7.2 Managerial Implications for Local Businesses

Our findings suggest that "PokeStop nearby" can be considered as an attraction to users and validated as a positive feature of restaurants. In addition to the empirical results from our previous formal analyses (Sections 4-6), we also conduct text analyses of the Yelp reviews for more direct evidence. By parsing the review texts of the restaurants in our sample, we find that reviews with the keywords “Pokemon” or “PokeStops” have a higher proportion in 4 and 5-star ratings (Appendix D).

Our findings also show that the entry of this application can generally improve certain aspects of the online reputation of local businesses in both long and short terms. Based on those results, we can calibrate the economic significance of the spillover effects of the entry of LBAR applications in terms of restaurants’ revenue. According to Luca (2016), one star increase in a restaurant’s Yelp rating results in a 9% increase in its revenue. In addition, we obtain the average rating of the restaurants 3.81 from Table 1, and the DID results (0.057 in the short term and 0.080 in the long term) about the impact of Pokemon GO entry on restaurant ratings in Table 3. Combining these, we estimate that the entry of Pokemon GO results in around 1.96% (= 0.057*3.81*9%) increase in the monthly revenue of a PokeStop nearby restaurant in the short term, and 2.75% (=0.080*3.81*9%) increment in the long term.

PokeStops, a series of fixed locations connecting the virtual game world and the real world, turns to traffic access points by the LBS-featured application Pokemon GO. Thus, we recommend that business owners proactively adopt and advertise this for proximity marketing and branding to gain positive spillover effect. Specifically, for treated restaurants, managers can frequently advertise their exclusively nearby PokeStops by post tweets or announcement on social media such as Twitter or Facebook. Based on the dynamic effects estimated by the relative time approach, in order to sustain the traffic benefits, we recommend business owners to update their advertising posts (e.g. Appendix A) regarding PokeStops at least every three months, or right after the time when Niantic releases updates or initiates large Pokemon hunting events. Additionally, our treatment intensity analysis shows that a 1% distance closer to a PokeStop can lead to a 2.2% increase in review volume. Motivated thus, we suggest business owners, especially for restaurants having PokeStops within a walking distance, to broadcast their geographic advantages by claiming that, for example, "we have PokeStop in store." Given the promisingly considerable increase in traffic and revenue, this type of costless campaign through social media or posters is a very worthy investment.

Our findings also suggest that the spillover effects of location-based mobile applications cannot benefit all businesses for all time frames. Instead, it demonstrates a rich-get-richer pattern, that is, businesses with a higher established reputation benefit more from the entry of Pokemon GO in the long run, but those with a lower established reputation only obtain benefits temporarily. These findings suggest business owner not to merely passively wait for the spillover effect to take place, but to improve the fundamental product and service quality, and to establish and maintain a good reputation.

Our results also show that non-chain restaurants are more positively associated with the spillover effects on online reputation metrics like review volume and check-ins. Inspired by such, for restaurant without a rich reputation establishment in terms of below-average historical rating or, below-average review volume, beside improving their fundamental food and service quality, we suggest managers to also package and market their business with more specific and unique selling points such as seasonal or specific culture themed decoration, dining environment and arts. This strategy serves as the icing on the cake for restaurants with well-established reputation.

7.3 Managerial Implications for App Developer

Our study has demonstrated pronounced econometric significance of traffic and business opportunities brought by PokeStops during the time frame of this research. However, Niantic did not start working with some big-chains and marketing sponsored locations (aka "branded PokeStops") in the United States until early 2017, when Pokemon GO's second-generation was released. Till now, only Sprint and Starbucks but no individual small businesses successfully sponsored to become new PokeStops. Our study shows that Pokemon GO can generally improve restaurants' online reputation and especially for non-chain restaurants. Individual small local businesses have a higher motive to become PokeStops through sponsorship and to cooperate with Niantic. Though now Niantic has not adopted any form of monetization, in an interview by Brazil Globo[18], Niantic strategic VP is considering a cost-per-visit model by charging business partners at certain locations 0.5 dollar per visitor attracted by the game. 89% of the restaurants in our full sample are non-chain (independent) restaurants. According to our analysis, around 70% of these non-chain restaurants benefit from the entry of Pokemon GO. If the developer cooperates with small non-chain business as well, the shared revenue for Niantic will be substantial. Therefore, we recommend Niantic further open PokeStops application and work with more diversified and heterogeneous restaurants rather than limiting the sponsorship program only to big chains.

8. Conclusion

Our paper validates and extends the AR and LBS literature in the marketing and IS fields. We provide a comprehensive empirical analysis regarding the spillover effects of LBAR

[18] https://techcrunch.com/2017/05/31/pokemon-go-sponsorship-price/

application on local businesses. Pokemon GO as the most successful LBS-AR mobile application, indeed provides positive externality to local businesses, especially brings rich-get-richer effects to nearby restaurants with above-average established reputation. There are two directions for future studies. First, if sales or traffic data are available, LBAR apps' direct impact on local businesses can be examined. Second, till summer 2017, Niantic had updated this game and launched the third generation Pokemon GO. More LBAR games have also been launched. Future studies can continuously collect more data to verify whether and how this LBAR technological convergence application and local business would influence each other as an ecosystem in a longer time frame.

References

1. Andrews, M., Luo, X., Fang, Z. and Ghose, A., 2015. Mobile ad effectiveness: Hyper-contextual targeting with crowdedness. *Marketing Science*, 35(2), pp.218-233.
2. Anderson, Michael, and Jeremy Magruder. "Learning from the crowd: Regression discontinuity estimates of the effects of an online review database." *The Economic Journal* 122.563 (2012): 957-989.
3. Autor, David H. 2003 "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing." *Journal of Labor Economics* 21.1: 1-42.
4. Bigham, L., 2005. Experiential marketing, New Consumer Research.
5. Bulearca, M., & Tamarjan, D. (2010). Augmented reality: A sustainable marketing tool. *Global business and management research: An international journal*, 2(2), 237-252.
6. Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics*, 101(3), 442-451.
7. Chan, Jason, Anindya Ghose, and Robert Seamans. The Internet and Hate Crime: Offline Spillovers. Working Paper, 2014.
8. Cousins, K. and Varshney, U., 2001, July. A product location framework for mobile commerce environment. *Proceedings of the 1st international workshop on Mobile comm*erce (pp. 43-48). ACM.
9. Colley, A., J. Thebault-Spieker, A. Lin, D. Degraen, B. Fischman, J. Häkkilä, K. Kuehl, V. Nisi, N.J. Nunes, N. Wenig, D. Wenig, B. Hecht, J. Schöning, 2017 "The Geography of Pokemon GO: Beneficial and Problematic Effects on Places and Movement." *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM: 1179-1192.
10. Craig, Alan B. Understanding augmented reality: Concepts and applications. Newnes, 2013.
11. Demers, Elizabeth, and Katharina Lewellen. "The marketing role of IPOs: evidence from Internet stocks." *Journal of Financial Economics* 68.3 (2003): 413-437.
12. Duan, W., Gu, B., & Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *Journal of retailing*, 84(2), 233-242.
13. Fulton, Wil 2016 "Pokémon GO Is Going to Make Restaurants and Bars Tons of Money", Thrillist.
14. Fang, Z., Gu, B., Luo, X., & Xu, Y. (2015). Contemporaneous and delayed sales impact of location-based mobile promotions. *Information Systems Research*, 26(3), 552-564.
15. Friedenthal, Andrew. "How Pokémon GO and Augmented Reality Marketing Can Drive Your Sales", *Software Advice*, 2016.
16. Filloon, Whitney 2016 "How Restaurants Are Dealing with Pokémon GO Mania", *Eater*. Jul 11,
17. Gao, H., and Liu, H. (2014). Data Analysis on Location-Based Social Networks. In *Mobile Social Networking* (pp. 165-194). Springer New York.
18. Gao, Qiang, and Mingfeng Lin. 2016 "More than Just Money: Educational Impact of Online Charitable Crowdfunding", Working paper.
19. Ghose, A. and Han, S.P., 2014. Estimating demand for mobile applications in the new economy. *Management Science*, 60(6), pp.1470-1488.
20. Gong, J., B. Greenwood and Y. Song 2017, "Uber Might Buy Me a Mercedes Benz: An Empirical Investigation of Sharing Platforms and Durable Goods Purchase," Working paper.
21. Gruber, Jonathan. "The incidence of mandated maternity benefits." *The American economic review* (1994): 622-641.
22. Kunkel, N., Soechtig, S., Miniman, J., Stauch, C., 2016. Augmented and Virtual Reality Go to Work, Deloitte Report.

23. Li, X. 2016 "Could Deal Promotion Improve Merchants' Online Reputations? The Moderating Role of Prior Reviews." *Journal of Management Information Systems* 33.1: 171-201.
24. Li, X. 2018 "Impact of Average Rating on Social Media Endorsement: The Moderating Role of Rating Dispersion and Discount Threshold." *Information Systems Research*, 29(3): 525-777.
25. Liu, Angela Xia, Jan-Benedict EM Steenkamp, and Jurui Zhang. "Agglomeration as a Driver of the Volume of Electronic Word of Mouth in the Restaurant Industry." *Journal of Marketing Research* 55.4 (2018): 507-523.
26. Liu, Q. and Lu, Y. 2015 "Firm Investment and Exporting: Evidence from China's Value-added Tax Reform." *Journal of International Economics* 97.2: 392-403.
27. Lu, Y., Tao, Z. and Zhu, L. 2017 "Identifying FDI Spillovers." *Journal of International Economics* 107: 75-90.
28. Luca, M 2016. "Reviews, Reputation, and Revenue: The Case of Yelp.com." Working paper, Harvard Business School.
29. Luo, X., Zhang, J., & Duan, W. (2013). Social media and firm equity value. *Information Systems Research*, 24(1), 146-163.
30. Caliendo, Marco, and Sabine Kopeinig. "Some practical guidance for the implementation of propensity score matching." *Journal of economic surveys* 22.1 (2008): 31-72.
31. Moe, Wendy W., and David A. Schweidel. "Online product opinions: Incidence, evaluation, and evolution." *Marketing Science* 31.3 (2012): 372-386.
32. Pischke, Jörn-Steffen. 2005 "Labor Market Institutions, Wages, and Investment: Review and Implications." *CESifo Economic Studies* 51.1: 47-75.
33. Qiu, Liangfei, Zhan Shi, and Andrew B. Whinston. "Learning from your friends' check-ins: An empirical study of location-based social networks." *Information Systems Research* 29.4 (2018): 1044-1061.
34. Rauschnabel, Philipp A., Alexander Rossmann, and M. Claudia tom Dieck. "An adoption framework for mobile augmented reality games: The case of Pokemon GO." *Computers in Human Behavior* 76 (2017): 276-286.
35. Rauschnabel, Philipp A., Reto Felix, and Chris Hinsch. "Augmented reality marketing: How mobile AR-apps can improve brands through inspiration." *Journal of Retailing and Consumer Services* 49 (2019): 43-53.
36. Rishika, R., Kumar, A., Janakiraman, R., & Bezawada, R. (2013). The effect of customers' social media participation on customer visit frequency and profitability: an empirical investigation. *Information systems research*, 24(1), 108-127.
37. Smith, Jeffrey A., and Petra E. Todd. "Does matching overcome LaLonde's critique of nonexperimental estimators?" *Journal of econometrics* 125.1-2 (2005): 305-353.
38. Susskind, Alex M., and Edwin K. Chan. "How Restaurant Features Affect Check Averages: A Study of the Toronto Restaurant Market." *Cornell Hotel and Restaurant Administration Quarterly* 41.6 (2000): 56-63.
39. Ververidis, C. and Polyzos, G., 2002, July. Mobile marketing using a location based service. *Proceedings of the First International Conference on Mobile Business*, Athens, Greece.
40. Imbens, Guido W., and Jeffrey M. Wooldridge. 2009 "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47.1: 5-86.
41. Zhu, Y. 2016 "5 Brilliant Ways Restaurants Are Using Pokémon GO to Level Up Sales". *Forbes*.
42. Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management*, 58, 51-65.

43. Xie, K. L., Zhang, Z., & Zhang, Z. (2014). The business value of online consumer reviews and management response to hotel performance. *International Journal of Hospitality Management*, 43, 1-12.
44. Xu, K., Chan, J., Ghose, A., & Han, S. P. (2016). Battle of the channels: The impact of tablets on digital commerce. *Management Science*, 63(5), 1469-1492.
45. Yuan, Y., Wu, C., 2008. Relationships among Experiential Marketing, Experiential Value and Customer Satisfaction, *Journal of Hospitality & Tourism Research*, 32 (3), 387-410.
46. Zach, Florian J., and Iis P. Tussyadiah. 2017 "To Catch Them All—The (Un)intended Consequences of Pokemon GO on Mobility, Consumption, and Wellbeing." *Information and Communication Technologies in Tourism* 2017. Springer, Cham, 21.

Tables

Table 1. Descriptive Statistics

Variables	Min	Max	Mean	S.E
Review Volume	0	108	3.58	0.22
Rating	1	5	3.81	0.01
Checkin Volume	0	38	0.99	0.07
Competition	0	5.42	3.53	5.36
Density	0	5.65	3.30	0.47
Distance (m)	0	389.33	34.94	39.55
Game Downloads (Million)	0	103.44	22.76	40.24
Google Trend Index	0	263	34.65	0.59

Table 2a. Probit Regression of Receiving Treatment Table 2b. t-Test Result on the Matching

Variables	Coef.	Std. Err.	Mean treated	Mean control	Mean diff.	t-stat.
Price Level $	0.345	0.240	0.346	0.346	0.001	0.10
Price Level $$	0.894***	0.239	0.608	0.604	0.004	0.50
Price Level $$$	-0.966***	0.255	0.040	0.044	-0.004	-1.31
Competition	-0.007***	0.001	40.820	42.256	-1.436	-1.71
Chain	0.214***	0.047	0.405	0.404	0.001	0.16
Overall Rating	-0.099	0.075	3.780	3.796	-0.016	-1.62
Total Volume	0.001	0.001	179.840	184.180	-4.340	-1.49
American	-0.324***	0.058	0.275	0.286	-0.011	-1.61
Chinese	0.976***	0.153	0.044	0.042	0.001	0.45
Japanese	0.565***	0.108	0.074	0.070	0.004	1.13
Korean	0.136	0.138	0.028	0.029	-0.001	-0.37
Indian	2.480***	0.304	0.003	0.003	0.000	0.56
Other Asian Fusion	1.528***	0.112	0.101	0.098	0.003	0.72
French	0.013	0.024	0.026	0.025	0.000	0.10
Italian	0.529***	0.084	0.098	0.099	-0.001	-0.15
Other European	0.487***	0.138	0.034	0.032	0.002	0.60
Mexican	0.444***	0.073	0.158	0.157	0.000	0.04
Other Latin American	-0.068	0.102	0.042	0.043	0.000	-0.15
Mediterranean	0.144	0.147	0.033	0.031	0.002	0.70
Bakeries and Dessert	-0.090	0.097	0.056	0.057	-0.001	-0.27
Bars, Beers, Wine and Liquor	0.302***	0.056	0.256	0.253	0.004	0.56
Breakfast and Brunch	0.891***	0.085	0.105	0.102	0.002	0.50
Coffee & Tea	0.931***	0.132	0.045	0.043	0.003	0.90

Deli	-0.495***	0.146	0.021	0.022	-0.001	-0.43
Fast Food	0.843***	0.063	0.222	0.226	-0.004	-0.63
Juice & Frozen Desserts	0.329*	0.192	0.019	0.018	0.001	0.46
Vegan & Vegetarian	0.006	0.096	0.057	0.057	0.000	0.00
Other Cuisines and Services	0.553***	0.141	0.033	0.031	0.002	0.79

Table 3. DID Estimation Results for Full Model

	Review Volume		Review Rating	
Variable(s)	Short-term	Long-term	Short-term	Long-term
Diff-in-Diff	0.065** (0.033)	0.025 (0.026)	0.057*** (0.013)	0.080*** (0.016)
Price Level $	-0.038 (0.035)	-0.117* (0.067)	-0.019 (0.015)	0.007 (0.044)
Price Level $$	-0.039 (0.034)	-0.126* (0.067)	-0.011 (0.015)	0.0120 (0.044)
Price Level $$$	0.067 (0.071)	0.011 (0.014)	0.086** (0.039)	0.009 (0.047)
Competition	0.011* (0.006)	0.016*** (0.005)	0.001 (0.001)	0.002 (0.003)
Density	0.002 (0.009)	0.010 (0.008)	0.002 (0.002)	0.002 (0.005)
Distance	-0.005 (0.011)	-0.011* (0.006)	-0.002 (0.002)	-0.009* (0.006)
Game Download	0.045*** (0.015)	0.046*** (0.015)	0.004 (0.003)	0.037*** (0.009)
Chain	0.021 (0.014)	-0.025** (0.011)	0.005 (0.004)	0.008 (0.007)
Google Trend	0.001 (0.001)	0.007** (0.004)	0.002** (0.001)	0.006*** (0.003)
R-square	0.52	0.51	0.61	0.62

Note: Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 4. DID Estimation on Checkin Volume Results for Full Model

Variable(s)	Short-term	Long-term
Diff-in-Diff	0.028 (0.068)	0.015** (0.009)
Price Level $	0.164** (0.073)	-0.001 (0.012)
Price Level $$	0.342*** (0.072)	0.011 (0.013)
Price Level $$$	-0.074 (0.185)	0.002 (0.003)
Competition	-0.001* (0.001)	0.002 (0.005)
Density	0.002*** (0.001)	-0.008 (0.006)
Distance	-0.011 (0.008)	-0.037*** (0.009)
Game Download	0.002*** (0.001)	0.023*** (0.008)
Chain	-0.052 (0.04)	-0.003*** (0.001)
Google Trend	0.001* (0.001)	0.01** (0.004)
R-square	0.64	0.62

Note: Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 5. Summary of Partitioned Sample Results

Reputation Metrics	Short-term		Long-term	
	High-volume	Low-volume	High-volume	Low-volume
Review Rating	0.065***(0.019)	0.063***(0.018)	0.098***(0.026)	0.074*** (0.022)

Check-in	0.019 (0.035)	0.038 (0.031)	0.014** (0.007)	0.054 (0.056)
	High-rating	Low-rating	High-rating	Low-rating
Review Volume	0.052** (0.023)	0.079* (0.042)	0.043** (0.018)	0.056 (0.044)
Check-in	0.053 (0.078)	0.003 (0.011)	0.011** (0.006)	0.013 (0.010)

Note: Diff-in-Diff Coef. (Standard errors in parentheses). * p < 0.1, ** p < 0.05, *** p < 0.01

Table 6. Results of Treatment Intensity Identity- Polynomial Regression

	Review Volume	Rating
Distance	-2.203*** (0.515)	-0.151** (0.06)
Distance^2	-0.268*** (0.076)	-0.099** (0.039)
Price Level $	-7.923** (4.038)	-0.017 (0.014)
Price Level $$	-6.174* (3.678)	-0.039*** (0.013)
Price Level $$$	-5.361** (2.114)	0.0823**(0.035)
Competition	0.301*** (0.0692)	0.045*** (0.009)
Density	0.237*** (0.057)	0.037*** (0.008)
Game Download	0.02346 (0.023)	0.021 (0.096)
Google Trend	0.011*** (0.007)	0.036 (0.295)
Chain	-6.429** (3.07)	-0.011* (0.005)
R-square	0.61	0.51

Note: Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 7 Summary of DDD Estimation Results of Overall Reputation

Reputation Metrics	DD	HighVolume	HighRating	HighVolume x DD	HighRating x DD
	Long-term				
Volume	0.017 (0.021)	-	0.040*** (0.005)	-	0.054*** (0.016)
Rating	0.078*** (0.018)	0.007** (0.003)	-	0.005 (0.014)	-
Check-in	0.013** (0.005)	0.07** (0.025)	0.040* (0.023)	0.086** (0.037)	0.027** (0.012)
	Short-term				
Reputation Metrics	DD	HighVolume	HighRating	HighVolume x DD	HighRating x DD
Volume	0.088*** (0.035)	-	0.011*** (0.004)	-	-0.047*(0.026)
Rating	0.057*** (0.015)	0.004* (0.002)	-	0.012* (0.007)	-
Check-in	0.012 (0.007)	0.040* (0.023)	0.013 (0.037)	0.001 (0.003)	0.087* (0.052)

Note: DDD (Coef. Standard errors in parentheses). * p < 0.1, ** p < 0.05, *** p < 0.01

Table 8 Summary of DDD Estimation Results of Game Trends

Reputation Metrics	DD	Game Download	Game Download x DD
Volume	0.022 (0.033)	0.024** (0.012)	0.038*** (0.012)
Rating	0.077*** (0.018)	0.021* (0.008)	0.004 (0.006)
Check-in	0.011* (0.006)	0.003* (0.002)	0.006** (0.003)

Note: DDD Coef. (Standard errors in parentheses). * p < 0.1, ** p < 0.05, *** p < 0.01

Table 9a Summary of DDD Estimation Results of Competition

Reputation Metrics	DD	Competition	Competition x DID
Volume	0.024 (0.027)	0.003** (0.001)	0.002*** (0.001)
Rating	0.078*** (0.017)	0.001* (0.001)	0.002* (0.001)
Check-in	0.011** (0.005)	-0.001* (0.001)	0.001* (0.001)

Table 9b Summary of DDD Estimation Results of Chain/Non-Chain

Reputation Metrics	DD	Chain	Chain x DID
Volume	0.040 (0.023)	-0.005* (0.003)	-0.012*(0.007)
Rating	0.076*** (0.014)	-0.001 (0.002)	0.003 (0.005)
Check-in	0.012** (0.005)	-0.002** (0.001)	-0.007** (0.003)

Note: DDD Coef. Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 10a. Reverse Regression, Rating on DID

Diff-in-Diff	Without Lag Term	With Lag Term
Rating	-0.006 (0.005)	-0.006 (0.005)
Rating_AR1	-	-0.007 (0.005)
Price Level $	-0.005 (0.127)	-0.015 (0.127)
Price Level $$	-0.003 (0.127)	-0.013 (0.127)
Price Level $$$	-0.017 (0.135)	-0.024 (0.135)
Competition	-0.001 (0.001)	-0.001 (0.001)
Density	0.001 (0.001)	0.001 (0.001)
Distance	-0.001 (0.001)	0.001 (0.001)
Game Download	0.001*** (0.001)	0.002 (0.001)
Google Trend	0.001*** (0.001)	0.001 (0.001)

Table 10b. Reverse Regression, Volume on DID

Diff-in-Diff	Without Lag Term	With Lag Term
Volume	0.009 (0.006)	0.008 (0.006)
Volume_AR1	-	-0.005 (0.009)
Price Level $	-0.001 (0.127)	-0.006 (0.127)
Price Level $$	0.003 (0.127)	-0.001 (0.127)
Price Level $$$	-0.012 (0.135)	-0.014 (0.135)
Competition	-0.001 (0.001)	-0.001 (0.001)
Density	0.001 (0.001)	0.001 (0.001)
Distance	-0.001 (0.001)	-0.001 (0.001)
Game Download	0.002*** (0.001)	0.002 (0.001)
Google Trend	0.001*** (0.001)	0.001 (0.001)

Note: Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 11. Dynamic Effects of Relative Time Approach DID Estimation

	Volume	Rating
5 Month Prior	0.051 (0.032)	0.013 (0.009)
4 Month Prior	0.014 (0.023)	0.012 (0.009)
3 Month Prior	-0.025 (0.023)	0.008 (0.009)
2 Month Prior	-0.011 (0.023)	0.013 (0.009)
1 Month Prior	-0.007 (0.031)	0.015 (0.012)
Current	0.088 (0.029)	0.029*** (0.010)
1 Month Post	0.066*** (0.023)	0.092*** (0.009)
2 Month Post	0.085*** (0.028)	0.071*** (0.009)
3 Month Post	0.039* (0.023)	0.016*** (0.008)
4 Month Post	0.036** (0.018)	0.011*** (0.007)
5 Month Post	0.024 (0.016)	0.057*** (0.006)
6 Month Post	0.005 (0.021)	0.015*** (0.009)

Note: Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 12. Regression Discontinuity Regression Results

	Review Volume	Review Rating
$I(Distance_i \leq c)$	0.419 *** (0.134)	0.048** (0.027)
$Distance_i \leq c$	0.017* (0.010)	0.014** (0.007)
$(Distance_i \leq c) * I(Distance_i \leq c)$	0.077 (0.067)	0.046 (0.076)
Price Level $	-0.379** (0.193)	0.052 (0.067)
Price Level $$	-0.352* (0.189)	0.087 (0.067)
Price Level $$$	0.395** (0.219)	0.097 (0.070)
Competition	0.025 (0.026)	-0.012 (0.017)
Density	0.062 (0.044)	0.026 (0.019)
Game Download	0.137* (0.083)	0.020 (0.031)
Chain	-0.075 (0.05)	-0.022** (0.012)
Google Trend	0.002* (0.001)	0.002* (0.001)
Entity Level VCE Cluster	Yes	Yes

Note: Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 13 Summary of Weekly Estimation Results

(Restaurant cuisines factor, city location dummy and holiday dummy are included)

Reputation Metrics	Short-term	Long-term
Volume	0.030** (0.012)	0.020 (0.02)
Rating	0.022*** (0.006)	0.115*** (0.019)

Note: DD Coef. Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

Appendix A. More Evidence about the Pokemom GO App, the PokeStops, the Marketing Initiatives and the Mixed Outcomes

We illustrate how Augmented Reality (AR) (Figure A1a) and location based (LBS) (Figure A1b) are implemented in the Pokemon GO app with two examples in Figure A1. In Figure A1b, the virtual PokeStop on the virtual map are corresponding to the real physical location of the app user.

Figure A1a. An AR Example of Pokémon GO's Spawn Screen

Figure A1b. An LBS Example of a PokéStop

Figure A1.

With the entry of the popular app Pokemon GO, a massive influx of users gathered and lingered in the neighboring area of PokeStops (Figure A3).

Figure A3. Many Pokémon GO Players Gathered at a PokéStop near Santa Monica Pier in August 2016.

Examples of restaurant managers taking advantage of having Pokestops nearby and the mixed marketing outcomes

Recommended by many marketing consultants (Friedenthal 2016, Filloon 2016 and Zhu 2016), managers of local restaurants attempted to take advantage of the penetration of this LBAR app and their nearby Pokestops to attract traffic and improve sales by posting game related advertisements, promotions, and special offers (Figure A3). For example, by Zhu (2016)'s editorial, a restaurant owner created a large Facebook event 'Pokémon GO: Battle for New York City!' for players to meet at their restaurants, which caught 6,200 people's interests and received 1,200 commitments to attend.

Most restaurants saw positive effects of this LBAR app: some restaurant owners also reported a surge in new foot traffic. after they purchased and used the 'Lure Module'[19][20]. A pizza bar in the Long Island City[21] reported that the PokéStop in the plaza drew so many players that the shops' business went up by 75% after the launch of Pokémon GO. Some chain restaurants offered gift cards to customers who post a photo of themselves with a Pokémon and check-in and at one of the company's locations on social media[22]. However, there were also restaurant managers complaining that Pokémon GO failed as a marketing tool: some restaurant owners did not experience more sales or more visits after they dropped lures because "players came in but did not get anything to eat or drink" (Filloon 2016 and Zhu 2016).

[19] https://www.reviewtrackers.com/use-pokemon-go-local-marketing
[20] https://www.forbes.com/sites/megykarydes/2016/07/14/how-pokemon-go-shots-are-driving- restaurant-sales/#506af3a41d26
[21] https://nypost.com/2016/07/12/pokemania-runs-wild- through-city-causing-crime-accidents/
[22] https://www.mic.com/articles/148439/pokemon-go-boosting-restaurant-sales-pokestop-lure-module

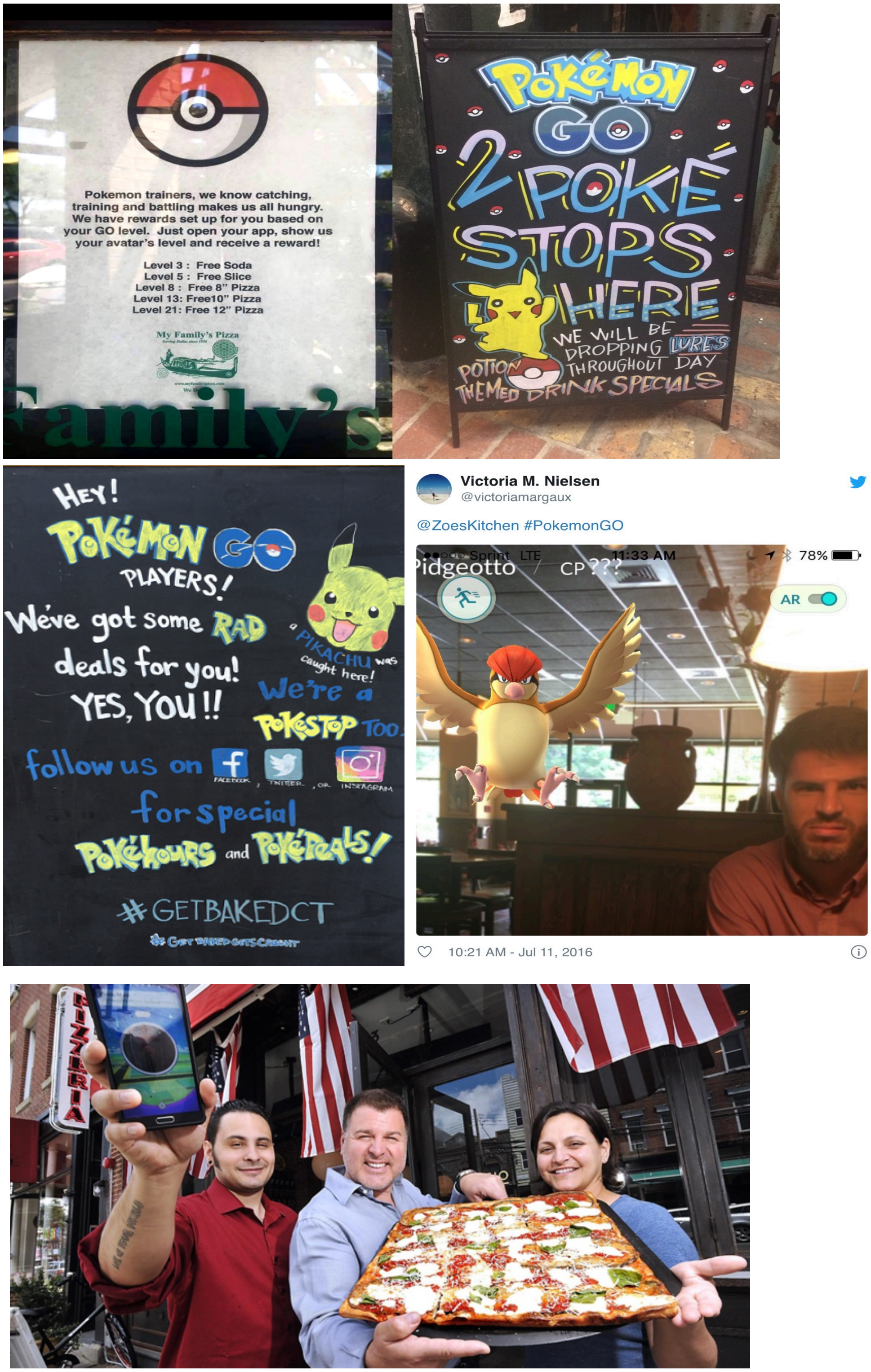


Figure A2. Examples of Restaurant Managers Taking Advantage of Having Pokestops Nearby

Appendix B. The DID Estimation on Elite Review Volume

We conduct a DID estimation of restaurant's Elite review volume data with the following model: $Elite_{i,t} = \alpha_0^E + \alpha_1^E PKG_i + \alpha_2^E d_t + \beta^E(PKG_i * d_t) + \Theta^E R_i + \Psi^E X_{i,t} + \eta_i + \gamma_t + \varepsilon_{i,t}^E$, and the results are reported in Table A. The entry of Pokemon GO weakly significantly improves elite volume of restaurants with PokeStops nearby by 1.2% in the short-term, and 0.9% (p<0.05) in the long term.

Table B. DID Estimation Results of Elite Review Volume for Full Model

Restaurant cuisines factor city location dummy are included.

Variable(s)	Short-term	Long-term
Diff-in-Diff	0.012*(0.007)	0.009* (0.005)
p1	-0.011 (0.008)	-0.032** (0.014)
p2	-0.012* (0.007)	-0.029** (0.014)
p3	0.002 (0.019)	0.003 (0.015)
Competition	0.003***(0.001)	0.002* (0.001)
Density	0.002 (0.002)	0.001 (0.002)
Distance	-0.002** (0.001)	-0.002** (0.001)
Game Download	0.003 (0.002)	0.003 (0.003)
Chain	-0.002 (0.001)	-0.001 (0.001)
Google Trend	0.001 (0.001)	0.001 (0.001)
R-square	0.64	0.67

Note: Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

Appendix C. The Full DID Estimation Based on Weekly Data

Table C. The Full DID Estimation Results Based on Weekly Data

Restaurant cuisines factor city location dummy are included.

	Review Volume		Review Rating	
Variable(s)	Short-term	Long-term	Short-term	Long-term
Diff-in-Diff	0.030** (0.012)	0.020 (0.02)	0.022*** (0.006)	0.115*** (0.019)
p1	-0.082*** (0.014)	-0.035 (0.023)	-0.004 (0.007)	0.002 (0.01)
p2	-0.029** (0.013)	-0.018 (0.023)	-0.002 (0.006)	-0.009 (0.01)
p3	0.081*** (0.028)	0.116** (0.049)	0.029* (0.015)	0.014 (0.021)
Competition	0.003*** (0.001)	0.010*** (0.004)	0.001 (0.001)	0.002 (0.002)
Density	0.018* (0.010)	0.011* (0.006)	0.003 (0.002)	0.001 (0.003)
Distance	-0.003*** (0.001)	-0.021*** (0.008)	-0.001 (0.001)	-0.005***(0.002)
Game Download	0.001 (0.001)	0.002*** (0.001)	0.002* (0.001)	0.002** (0.001)
Chain	-0.009*** (0.005)	-0.023 (0.009)	-0.004 (0.003)	-0.022*** (0.004)
Google Trend	0.001 (0.001)	0.001 (0.001)	0.002* (0.001)	0.004** (0.001)
R-square	0.49	0.63	0.69	0.69

Note: Standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01

Appendix D. Validation from Customer Perceptive, Text Analyses of Yelp Reviews

To further explore how Pokemon GO affects restaurants' online reputation, we run a text analysis in this session. We parse all review texts with the keywords "Pokemon" or "PokeStop" in July and August 2016 along with individual rating and customers' information on Yelp. There are 40 reviews (0.278% of all reviews) from these two months in our sample, including Pokemon GO as keyword. Thirty-eight of them are from the treatment group-- restaurants with PokeStops nearby and two reviews are from the control group-- restaurants without PokeStops nearby. In July, in the treatment group, the frequency of "Pokemon" is 19, and the frequency of "PokeStop" is 11. In August, in the treatment group, the frequency of "Pokemon" is 13, and the frequency of "PokeStop" is 9. In July, in the control group, the frequency of "Pokemon" is one, and the frequency of "PokeStop" is 1. In August, in the control group, the frequency of "Pokemon" is 0, and the frequency of "PokeStop" is 0.

We manually read each of the identified reviews and found most of these reviews were about sharing information related to Pokemon species or PokeStops in nearby areas. Through a customer's perspective, review texts show three patterns regarding Pokemon GO's impact on restaurants' online reputation.

First, PokeStops can draw the players' attention and lead players to visit nearby restaurants. Some yelpers point out that for restaurants that serve similar cuisine, those with PokeStops nearby will be highly preferred.

Second, for the yelpers who play Pokemon GO, having PokeStops nearby is shown as a positive attribute to them (an example in Figure D1a), and these yelpers usually leave higher ratings and sentiment scores. Besides being stated as an experience enhancement, being close to

PokeStops or being inhabited by rare Pokemon also improves customers' experiences and utilities while waiting for seats or services and prevents negative emotions during waiting.

Third, ways of how restaurant owners take advantage of the game and restaurant's attitudes to players affect customer's ratings and sentiment. Several of the restaurants provide Pokemon GO related check-in promotional offers or special menus, and customers leave feedbacks on Yelp regarding their dinging and playing experience. For customers, given average food and service quality, Pokemon GO is considered as a plus for "not-bad" restaurants. However, there are also negative reviews (an example in Figure D1b) with low-rating in texts in the parsed texts. Few owners might consider Pokemon GO an annoying fad. Nevertheless, restaurant managers and servers' bad manners and attitudes to customers who love Pokemon GO could hurt customers' feelings and restaurants' online reputation.

Meanwhile, we do not observe text showing that having a crowded atmosphere due to the Pokemon GO heat decreases yelper's rating and sentiment, either. In Figure D2a, we have an overall view of the different rating proportions of parsed Pokemon GO related texts. Generally, the impact of Pokemon GO on yelpers' attitudes to restaurants is positive, and as we expect in the previous session, Pokemon GO is verified to be an attractive feature to restaurants. Accordingly, we parse the texts of reviews without keywords like "Pokemon GO" and "PokeStop" and summarizing their rating distribution in the pie chart Figure D2b. By comparing the two figures, we observe an evident difference between the two text samples' rating distribution. The 5-point rating proportion of the sample without keywords is 18% less than that of the sample with keywords. The 1-point, 2-point, and 3-point rating proportions of the sample without keywords are larger than their counterparts of the sample with keywords by 5%, 8%, and

7%, respectively. The 4-point rating proportion of the sample without keywords is slightly lower than that of the sample with keywords.

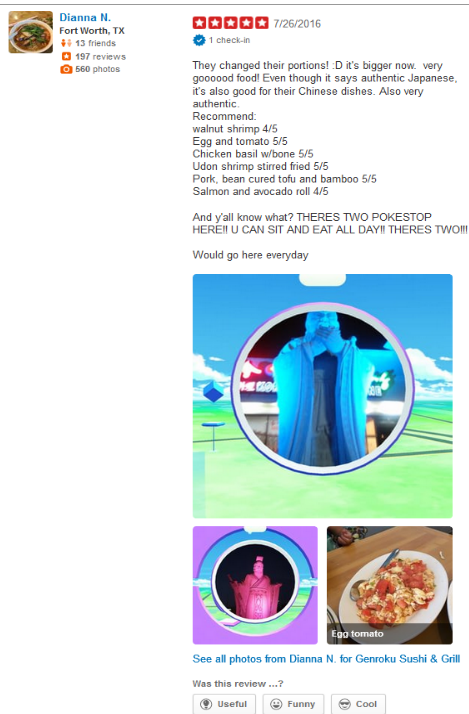
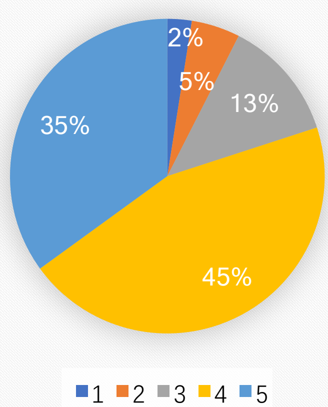


Figure D1a. Positive 64

Figure D1b. Negative Review Sample

Figures D2a. With Pokemon GO Related Keywords



Figures D2b. Without Pokemon GO Related Keywords

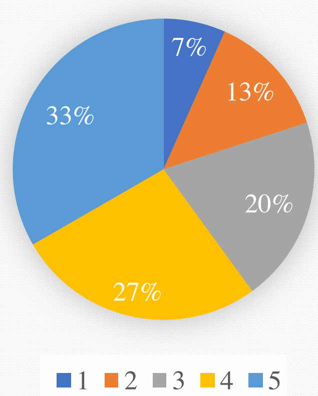


Figure D2. Contrast of Rating Distributions of Reviews with and without Pokemon GO Related Keywords

Further, we compare the means of two samples. The mean rating of the sample with Pokemon GO related keywords is significantly higher (*p*= 0.035) than that of the sample without those keywords. This salient difference further validates our expectation that Pokemon GO is an attractive feature for both restaurants and customers by enhancing restaurant differentiation and improving customers' dining experience and perceived quality.

CHAPTER TWO

"LIVE" TO WIN: THE IMPACTS OF DIFFERENT VIDEO PLATFORMS ON PRODUCT SALES PERFORMANCE- A PVAR ANALYSIS

Yuan Zhang

Jie (Jennifer) Zhang

Abstract

User-generated-video platforms are growing rapidly, has engendered massive proliferation, and is evolving to various forms, which includes more and more interpersonal interaction and socialization features. This study focuses on live streaming video platforms and pre-recorded video platforms. The objective is to examine (1) how effectively do live streaming video platform and pre-recorded video platforms play in influencer marketing, in terms of product sales and customer stickiness? (2) Do platforms perform equally in converting viewers to shoppers and why? Through Panel-Vector-Auto-Regression analysis, we find that (1) Live streaming video platforms can improve both the short-term and long-term product sales and customer stickiness. (2) Pre-recorded video platforms can only improve the long-term product sales and customer stickiness. (3) Live streaming video platforms have a more significant and stronger predictive relationship than pre-recorded video platforms with the response level and explanatory power of product sales and customer stickiness. We used the media richness theory and the social presence theory to explain the results. Collectively, this research adds the contributions to marketing, IS and communication literature by extending the propositions of media richness theory and social presence theory and by providing managerial implications to managers who conduct content and influencer marketing.

Keywords: live streaming video platforms, pre-recorded video platforms, content marketing, UGC, Panel VAR analysis, media richness theory, social presence theory

A picture is worth a thousand words and a moving picture is worth a million people.

(Huffpost 2014)[23]

1. Introduction

A vast of studies have verified that contents on user-generated-content (UGC) platforms are popular earned media (Lovett and Staelin 2016) and have significant impacts on product sales (Luca (2016), Chen, et al. 2015). For example, by studying radio play and blog buzz, Dewan and Ramaprasad (2014) found that audio UGC, as free online sampling music, is positively related to both song and album level of music sales while blog buzz is not associated with album sales and negatively associated with song sales. UGCs used as earned media are found to be more influential than paid and owned media given the same exposure in marketing and branding (Lovett and Staelin 2016). Luo and colleagues (2012) found that web blogs have a significant predictive relationship with firm equity value. However, most of UGC genres that were researched are text or pictures. There is few study on the influence of user-generated videos. Videos, as one of the most information-rich UGC types, can attract more attention and customer engagement. For example, after Twitter added inline videos in addition to pictures and texts in Nov. 2013, Tweets were 94% more likely to be Retweeted.[24] Hubspot's digital marketing report[25] showed that video advertising is more acceptable to content viewers than other content formats such as blogs articles and pictures (Figures 1). As a result, companies and practitioners across various industries utilize user-generated video platforms such as YouTube, Periscope and Twitch for content and influencer marketing. According to Forrester's new Video Advertising Forecast[26], video ad spending is expected to grow from $91 billion in 2018 to $103 billion by

[23] https://www.huffingtonpost.com/scott-macfarland/if-a-picture-video-production_b_4996655.html
[24] www.higher-education-marketing.com
[25] https://blog.hubspot.com/marketing/video-marketing
[26] https://www.marketingdive.com/news/forrester-video-ad-spending-will-hit-103b-by-2023/530657/

2023.WordStream[27] reported that by 2019 86% marketers using user-generated video for marketing and branding. Besides posting or live streaming such as How-to, advertising and branding campaign, and expert explanation videos as owned media, marketers can also take advantage of user-generated videos as earned media by either inserting advertising in user-generated videos or targeting and sponsoring influential video generators (aka influencer marketing).

Despite support from marketing managers, there lacks evidence in the literature to show how effectively online video platforms can shape product sales or enhance branding. In response to the prevalent practices and literature gap regarding the influence of user-generated videos, this paper formally examines the effectiveness of user-generated videos on different platforms as influencer marketing media.

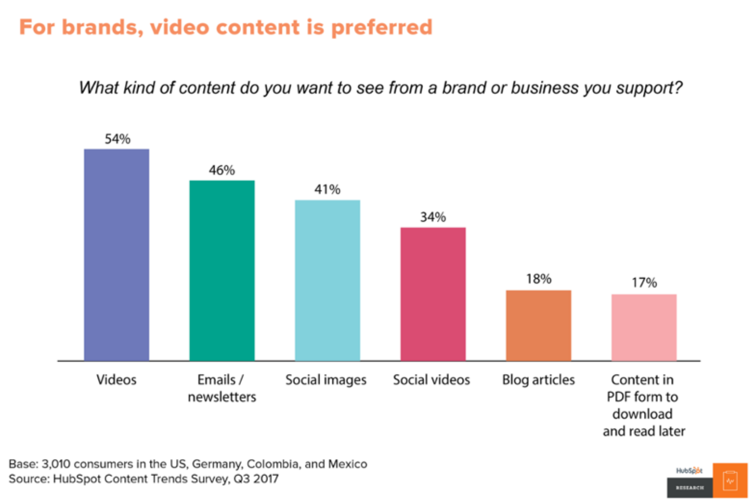


Figure 1. Branding tactics comparison

There are two types of video platforms: pre-recorded and live streaming video platforms. On pre-recorded video platforms, such as YouTube's core service, content generators have sufficient time to create and launch their contents. Since Vlogging (video blogging) became a dominant

27 https://www.wordstream.com/blog/ws/2019/03/12/video-advertising-trends

online content category[28], platforms that provide pre-recorded video sharing service are largely used as UGC marketing channels. Live streaming refers to online streaming media simultaneously recorded and broadcast in real-time to the viewers. Given its "live" attribute, live streaming videos enable social interaction features during the streaming. For example, viewers can send real-time texts and GIF graphs publicly or privately to chat with streamers, leave time-spot comments (called "barrages") in the chatting rooms, and send real-time virtual gifts or donate to their favorite streamers. At the same time, streamers can respond to their viewers by audio chatting and other means while streaming. Many social networking companies launch their own live streaming services, such as Twitter Periscope and Facebook Live, for users to document and share daily lives. As a new form of social media, live streaming platforms have been widely used for content and influencer marketing in such industries as cosmetics, food, video game, music, and movie[29]. Practitioners have identified several unique advantages of live-streaming platforms in marketing[30], such as encouraging trust and transparency, cost-effectiveness, easy sharing content, and the active social interaction and high user participation.

The similarity between the two platforms is that these online video platforms are both crowd-sourced and diversified in terms of content genre. More importantly, they own massive traffic and user base and thus are both largely used and recommended as media for influencer and content marketing. For example, HubSpot recommends YouTube as the most efficient inbound marketing channel, while Steam Lab reported Twitch increased its concurrent users on Steam.com by 67 percent in the third quarter of 2017.

---

28 http://mediakix.com/2018/08/vloggers-on-youtube-influencer-marketing/#gs.ftBxFbW1
29 https://www.forbes.com/sites/kathleenchaykowski/2017/10/18/digital-video-marketing-is-a-135-billion-industry-in-the-u-s-alone-study-finds/#d12ed58d4dda
30 https://www.salesforce.com/blog/2017/07/live-streaming-for-business.html

The basic distinction between the two types of platforms is their ways of streaming. On pre-recorded video platforms such as YouTube's core service, Vimeo and Dailymotion, content generators have sufficient time to film and edit their contents, and then they decide when to share their content. Yet on live streaming video platforms such as Periscope, Facebook Live and Twitch, content generators need to create and share content "live." That is, content generation and consumption are simultaneous. In addition, the two types of platforms also differ in their social interaction features. On pre-recorded video platforms, there is a delay for content generators to provide immediate feedback for their viewers and thus they cannot instantly change video content to meet the viewers' desire. Live streaming video platforms, on the other hand, given its "live "or even simulated "face-to-face "live nature, have more social interaction gadgets, such as real-time chat, gifting, and sponsorship, barrage comments, etc. Hence, the streaming contents can include more timely communication and immediate feedback between content generators and consumers. We summarize the key differences between these two platforms in Table 1. We also introduce two examples– Twitch.tv representing live streaming video platforms and YouTube as the representation of pre-recorded video platforms (Figure A1 in the Appendix).

Table 1. Platforms difference

Pre-recorded video platform	Live streaming video platform
• Linear streaming • Pre-recorded and edited • Filtered content/ campaigns • Social interaction features: -. static comments	• Non-linear streaming • Live and on-demand • Unfiltered content/ campaigns • Social interaction features: -. Real-time text and audio chatting -. Gifting and sponsorship -. Time-spot pop-up comments ("Barrage")

In addition to the literature gap regarding the impacts of online video platforms on business performance, no previous studies compare the relative marketing efficiency of the two platforms. Thus, in response to the literature gap and practical puzzles, our research questions are:

(1) How effective are live streaming video platforms and pre-recorded video platforms in influencer marketing, in terms of product sales and stickiness (time to use the product)?
(2) Which platform is relative more effective in improving product sales and customer stickiness?

Based on studies and propositions regarding the media richness theory and the social presence theory, platforms have different capabilities to conveying social cues, which are hints to guide and facilitate communications. Given the volume and diversity of conveying cues, platforms then have different levels of media richness. A higher level of media richness can provide platform users (in this research, video viewers) more social presence, which interprets intimacy, immediacy, and efficiency in interpersonal communication and interaction. The user's social presence level brought by rich media can positively influence on users' purchase intention of related products and their intention to use the product. Thereby, we propose that: Both live streaming (hypothesis 1) and pre-recorded video (hypothesis 2) platforms have a significant predictive relationship with product sales and customer stickiness. Live streaming video platforms have a stronger predictive relationship with product sales and stickiness than pre-recorded video platforms (hypothesis 3).

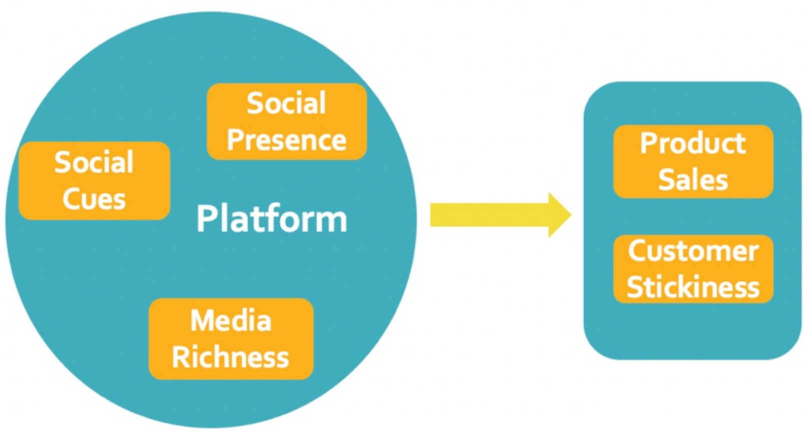
To examine the dynamic impacts and verify hypotheses, we conduct two studies- study 1, a PVAR analysis and study 2, an online experiment. In study 1, we first build a unique and comprehensive dataset by merging three sources: products' sales and profiles data, product-related video data from a live streaming video platform and product-related video data from a pre-recorded video platform. The merged dataset includes 1,023,342 daily observations of 1029 products from August 19, 2015, to March 15, 2018. To examine the dynamic impacts of video platform metrics on product sales metrics, we conduct a Panel VAR analysis. We find the following dynamic patterns. Given the difference of conveying cues, media richness, and social presence: (1) Live streaming video platforms can improve immediate and accumulated product sales and stickiness. (2) Pre-recorded video platforms can only improve accumulated product sales and stickiness. (3) Live streaming video platforms have a higher predictive performance than pre-recorded video platforms in terms of both response level and explanatory power. The patterns are explained and supported by the media richness theory and the social presence theory. To further validate the hypothesis 3 and to exclude potential self-selection and confounding issues, in study 2, we conduct an online experiment through Amazon MTurk and Qualtrics. After the experiment, we summarize valid responses from 238 highly responsible participants and we find that first, people's preference to either pre-recorded and live streaming video platforms has no significant difference in purchase intention. Second, regardless the condition of self-selecting or being randomly-assigned, watching video on live streaming video platform rather than on pre-recorded video platforms is associated with higher purchase intention.

This paper contributes to the literature in several ways. First, our research quantifies the marketing effectiveness of online video platforms by showing that both live streaming and pre-recorded video platforms can improve accumulated product sales and customer stickiness. Second, in both IS and marketing fields, our study is the first to provide rich empirical findings and reveal the influence of live streaming video platforms, and also the first to compare the marketing efficiency of live streaming video platforms with traditional pre-recorded video platforms. Third, most of the media richness theory studies (Dennis and Kinney 1998, Treviño et al. 2000, Rice 1992) focus on the consequence of how organizations choose and use media to fulfill group tasks. We extend and verify the application of media richness theory to the individual's choice and use of communication technologies. Most of the social presence theory literature (Rice 1992, Hajli et al. 2017, Hajli 2015, Liang et al., 2011 and Zhang et al., 2014) uses intention as the outcome variable to capture and analyze the impacts of social presence We validate studies of the social presence theory with a more direct measurement- the purchase outcome (sales) rather than the purchase intention.

In the remainder of this paper, we first summarize related backgrounds, literature and hypotheses in §2. §3 introduces the data sample and measurements. §4 describes the Panel VAR analysis. §5 presents all findings. The last section includes implications, contributions, and conclusions.

2. Literature and Hypothesis Development

We propose the following logic flow (Figure 2) that links online video platforms with product sales performance. Platforms have different capabilities to convey cues and have different levels of media richness. Platforms then provide viewers different levels of social presence and consequently bring different impacts on product sales and stickiness. Next, we



illustrate each building blocks with related theories and propositions.

Figure 2. Logic flow

2.1 Social Cues and Media Richness

Social cues are hints for communications and social interactions. Cues can reflect communicators' attitudes and can be expressed as verbal or non-verbal signals such as facial expression, tones, body gestures, proximity, etc. On the pre-recorded video platforms, communication cues can be conveyed through comments between video generators and their audiences. In a live streaming video environment, cues can be conveyed through various interaction gadgets or extensions between streamers and viewers such as real-time comments, gifting, chatting, etc. The cues can be tones, emoji sentiments in observable chatting, facial expression and eye contacts of streamers, etc. Martin and Postmes (2003) found that the ability to enable the multiplicity of cues and immediacy of feedback can guarantee the richness level of communication media.

Media richness theory (or information richness theory) (Daft and Lengel 1986) describes the ability of communication technology to diffuse, reproduce and convey the information sent over it. By studying traditional communication media like face-to-face chatting or live chatting, Martin Tanis and Tom Postmes (2003) found that given task equivocality, higher media richness contributes to improved task performance, social fulfillment and more efficient communication for organizations. However, these findings regarding media richness (Mendes-Filho and Tan 2009) are mostly supported under the traditional media context such as meetings and letters. Studies focusing on new media (i.e., Internet-based and computer-mediated communication tools) have not reached consensus to support the media richness theory. For example, Dennis and Kinney (1998) tested the media richness theory and found that given new communication technologies, selecting media richness level to match the level of task equivocality did not enhance the task performance. On the contrary, Flanagin and Metzger (2001) proposed more factors to measure the richness level for new media. These factors include assessment of need fulfillment, social norms and peer evaluations of media, etc. They found support for the media richness theory applying to the new media. Daft and Lengel (1986) asserted four factors that influence and shape the media richness- the ability to transmit multiple cues (e.g., vocal inflection, gestures)- the immediacy of feedback, language variety, and the personal focus of the media. However, most of these studies evaluated and measured the choice and use of media on the organizational level rather than on the individual level. It is unclear whether the central proposition of media richness theory applies to the context of individual use.

Based on the above factors that influence new media proposed by the recent literature, it is obvious that video platforms including both the pre-recorded and the live streaming platforms can convey more information and social cues than traditional UGC forms and platforms such as word-picture blogs and audio podcasts. According to the use and gratification theory, technology users actively seek specific media to satisfy particular needs and to fulfill certain objectives, such as seeking product information, strengthening consideration for purchase intention, engaging social interactions. Many previous studies like Luca (2015), Luo et al. (2013), Dewan and Ramaprasad (2014) and Li (2016) have verified that contents on UGC platforms have significant impacts on task outcomes, such as product branding, product sales performance, and business performance. Thereby, combining propositions of media richness theory and conclusions from prior studies on the UGC's impacts, we expect that individual choice and use of video platforms have salient impacts on the objective fulfillment of the individuals.

2.2 Social Presence Theory

Another related concept is Social Presence, which is described by Gunawardena (1995) as "the degree of salience of the other person in interaction or the degree to which a person is perceived as a real person." Social presence assesses emotional connections, group identification, social orientation and psychological distance between communicators. Short et al. (2010)

revealed that social presence explains primitive responses to social cues and thus can shape communicators' group perception and behaviors.

Many studies, such as Freeth et al. (2013), Kumar and Benbasat (2006), Miranda and Saunders (2003), and Ned (2004), agreed that richer media provides higher social presence to its users, and different levels of social presence have various consequences on media users. Hajli et al. (2017), and Chen et al. (2017) found that by improving the consideration and trust in the purchase funnel, a high-level social presence can improve the platform users' purchase intention. By studying group shopping and coupling shopping behaviors respectively, Liang et al. (2011) and Zhang et al. (2014) found that higher perceived social presence can improve the social fulfillment of group users and thus can improve group purchase likelihood. Zhu et al. 2010 and Li et al. 2006 found that by improving users' trust, social presence can improve users' stickiness to products or brands in terms of higher re-use intention, longer using time and lower churn rate. Moreover, Choi et al. (2011) found a moderating effect that social presence's impact is more significant if the product is hedonic than utilitarian. Yet, most of these studies focused on users' intention rather than actions and more observable and objective outcome variables like sales. It remains largely untested: can the media that provides its users with higher social presence improve users' purchase and ultimately improving the product sales?

Thereby, further combining propositions regarding social presence theory, we can specify the individual use objective to product sales and customer stickiness and propose

H1: Live streaming video platforms have a significant predictive relationship with (H1a) product sales and (H1b) customer stickiness.

H2: Pre-recorded video platforms have a significant predictive relationship with (H2a) product sales and (H2b) customer stickiness.

The platform comparison in Table 1 suggests that live streaming video platforms can convey more social cues, have a higher level of media richness and provide visitors better social presence than pre-recorded video platforms. Especially compared with live streaming video platforms, pre-recorded video platforms have limited cues like static comments and can be considered as “lean” media (Rice 1992). According to Dennis and Kinney (1998), limited capacity to convey social cues has negative consequences for task performance. We assume that as influencer marketing media, live streaming video platforms perform better than pre-recorded video platforms. Formally we propose H3 below.

H3: Live streaming video platforms have a stronger predictive relationship with (H3a) product sales and (H3b) stickiness than pre-recorded video platforms.

3. Data and Measures

To examine the impacts of the two platforms, we select Twitch.tv as the representative for live streaming platforms, YouTube[31] as the representative for pre-record video platforms and gather data regarding product sales and customer stickiness. Compared with pictures and texts, videos are more informative to convey experience attributes (information about the experience of consuming the product) in addition to search attributes. Thus, video platforms are largely adopted for marketing and branding experience goods. We select PC games as the focal product because they are one of the most profitable online retail experience goods. According to the report by SuperData[32], PC games aggregately generated 32.3 billion US dollars by 2017. We collect product sales data from Steam because it is the dominant digital PC game selling platform. PC games sales on Steam contributed to a rough gross revenue of 4.3 billion dollars in

31 YouTube also provides live streaming service but in this study, we exclude its live streaming contents and only consider its pre-record contents.
32 https://www.dsogaming.com/news/pc-games-sales-in-2017-are-almost-as-big-all-console-sales-combined/

the first season of 2017. Steam owned over 150 million accounts with a peak of 18.5 million concurrent users online[33]. This platform can provide sufficient product sales data for our analysis. We select Twitch as the live-streaming platform in our study because among all the applications of live streaming technologies, Twitch.tv is the leading and most profitable live streaming UGC platform in the world[34]. By the end of February 2018, Twitch had over 2 million streamers (broadcasters) monthly and over 15 million daily active users. Similarly, we select YouTube as the pre-recorded video platform because it is the largest with 1.5 billion monthly active users and 300 hours of video are uploaded every minute[35].

We build a unique and comprehensive dataset by merging data from three sources: (1) Game profiles and sales data on Steam from Valve's API and Steamspy[36]. More specifically, games' profile and owners' data is extracted from the HTML file of source page of each game by a Python package. (2) From YouTube API, we gathered pre-recorded video viewership data for each game excluding game studio official contents and live streaming recordings. (3) Gaming video viewership data on the live streaming video platforms is gathered from Twitch.tv. We merged the datasets by game ID and date. After data processing, the merged dataset forms a panel including 1,023,342 observations of 1,029 games on Steam from August 19, 2015, to March 15, 2018. The time-frequency is daily. For each game, we measure the following metrics summarized in Table 2.

Table 2. Metric measurements

Key covariates	Game sales	$gamesales_{it} = \Delta(Aggregatedgamesales_{it})$
	Play time	$playtime_{it} = \Delta$ (*The total play time per player on game i till time t*)
	Twitch views	$TwitchViews_{it}$ = Number of views of the game related channels on Twitch at *t*

33 https://store.steampowered.com/about/
34 https://www.wowza.com/blog/the-best-interactive-video-platforms-with-user-generated-content
35 https://www.socialmediatoday.com/social-networks/youtube-reports-updated-user-count-vr-180-and-vertical-video-support
36 www.steamspy.com. Steamspy uses the developer API of the Steam software distribution service that is owned by Valve Corporation to estimate the number of sales of game software titles offered on the service.

	YouTube views	$YouTubeViews_{it}$= Number of views of the game related videos on YouTube at t
Control variables	Price	$price_{it}$
	Review rating	$NetRating_{it} = SumPosRating_{it} - \text{abs}(SumNegRating_{it})$

According to Steam API, the sales of each game are calculated by the daily change of the number of owners. Owners show the number of users having a game on their Steam accounts, and thus the daily change of owners indicates the number of users who get and add a game on their account by a day. Playtime represents till time t, the change of total play time per user spent on game i. The $TwitchViews$ and $YouTubeViews$ are defined as the daily viewership for all videos or channels related to the game on the Twitch or YouTube platforms, respectively. We classify videos and channels as related to the game if their titles or descriptions contain the game's name.

Review ratings of games on Steam are generated by Steam users based on a 7-point scale[37]: "-3 for overwhelmingly negative, -2 for mostly negative, -1 for negative, 0 for neutral, +1 for positive, +2 for mostly positive, and +3 for overwhelmingly positive". Steam grouped the ratings based on the positive and negative signs, and sum them up by game *i* and time *t* into two metrics: $SumPosRating_{it}$ (positive number) and $SumNegRating_{it}$ (negative number). We use a net review rating $NetRating_{it}$, which is the sum of the above two metrics to represent a game's reputation.

The descriptive statistics for all covariates and control variables are summarized in Table 3.

Table 3. Descriptive statistics

	Metrics	Mean	Max	Min	Std. Dev	Median
Key covariates	$gamesales$	6893.969	1502514	1	10889.630	4373.5
	$playtime$	292.876	8183.207	15.26	489.257	181.557

[37] https://store.steampowered.com/reviews/

	$TwitchViews$	417.056	1137953	0	7353.362	391.23
	$YouTubeViews$	13459.140	5357960	0	72785.170	33.786
Control variables	$NetRating$	9.006	116	-2403	34738	11.563
	$price$	13.533	99.990	0	10.569	9.99

4. Study 1- Panel VAR Analysis

To decompose and capture the causes and effects (Lof and Malinen 2014, Fort et al 2013, Abrigo and Love 2015) in the research questions, we use a Panel Vector Auto Regression model with GMM estimation to describe the dynamic relations between four endogenous covariates-product sales, customer stickiness, the viewership of gaming video contents on the live streaming video platform and the viewership on the pre-recorded platform. We choose Panel VAR model with GMM estimation for the following reasons: First, Panel VAR can present us both the static and dynamic interdependencies between covariates (Adomavicius et al. 2012, Canova & Ciccarelli 2013). Second, according to Love & Zicchino 2006 and Canova & Ciccarelli 2013, Vector Auto Regression with panel-data approach allows for unobserved individual heterogeneity and accounts for cross-sectional dynamic heterogeneities. Third, according to Abrigo and Love 2015, GMM estimators have been proposed to calculate consistent estimates of equations, especially for a long time span and large panel size. Moreover, according to Newey & Rosen 1988, Holtz-Eakin 1998 and Abrigo and Love 2015, GMM estimation can improve the overall efficiency of the model.

4.1 Model Specification

The Panel VAR model specification is:

$$\boldsymbol{Y_{it}} = \sum_{\boldsymbol{k=1}}^{\boldsymbol{p}} \boldsymbol{A_k} \boldsymbol{Y_{it-k}} + \boldsymbol{B X_{it}} + \boldsymbol{e_{it}},$$

where $\boldsymbol{Y_{it}} = \begin{pmatrix} gamesales_{it}, \\ playtime_{it} \\ TwitchViews_{it} \\ YouTubeViews_{it} \end{pmatrix}$, $i \in \{1,2, \dots N\}, t \in \{1,2, \dots T\}$. N (=1029) is the number of games, T (=1061) is the time span. e_{it} is the vector of the idiosyncratic errors. $\boldsymbol{Y_{it}}$ is a vector of dependent variables. $\boldsymbol{X_{it}}$ is a vector of control variables. The matrix A and matrix B are parameters to be estimated through Panel VAR model. Here the idiosyncratic errors e_{it} are assumed to have stable trend s.t. $E[e_{it}] = 0$, $E[e'_{it}e_{it}] = \sum$, and $E[e'_{it}e_{is}] = 0,\ \forall\ t \neq s$. The stationarity checks will be shown later.

4.2 Analysis Procedure and Results

Based on Love (2015), Hayakawa (2015), Canova & Ciccarelli (2013), and Holtz-Eakin et al. (1988), we take Helmert transformation for all covariates to avoid heteroskedasticity and take natural log to remove scale effect. Inspired by Panel VAR analysis like Dewan and Ramaprasad (2014), Abrigo and Love (2015), Chen et al. 2015, and Fort et al. (2013) and classic cumulative VAR analysis like Luo et al. (2013), Luo (2009) and Srinivasan et al. (2010), our analyses follow six steps illustrated below.

4.2.1 Step 1. Pre-Estimation Unit Root Rest for Stability

Before the estimation, we first check the stationarity for all covariates to guarantee the Panel VAR assumption. Stationarity illustrates that the endogenous variable is mean-reverting and all salient fluctuation or deviation will dissolve eventually. Thus, under our research setting, the deterministic pattern stationary can be estimated unbiasedly and consistently and can be free of a permanent shift. The variances of stationary covariates are finite and time-invariant. For VAR family models, the typical way to test stationarity is the unit root test. Here, the pre-estimation stationarity check is conducted through the ADF unit root test. According to Whitehead (2002), the stationarity test first performs a unit root test on each panel's series separately, and then a combined p-value for the overall panel will be used to determine whether the panel series contains a unit root. The test results are presented in Table 4, all covariates pass the ADF test and are stable.

Table 4. ADF unit root test

Covariate	$gamesales_{it}$	$playtime_{it}$	$TwitchViews_{it}$	$YouTubeViews_{it}$
Z(t)	13.367	77.595	84.638	32.154
p-value	0	0	0	0
Stationary	√	√	√	√

Note: 1% critical value=3.43, 5% critical value=3.43, 10% critical value=2.57

4.2.2 Step 2. Select Optimal Lag Terms

Another preliminary procedure is to select the optimal highest order of all covariates for the PVAR estimation (Love & Zicchino 2006, Andrews and Lu 2001). The selection criteria are summarized in Table 5. Based on the maximal coefficient determination and minimal information criteria, we consider previous four lags for all covariates in the PVAR estimation.

Table 5. Model selection

Lag	Coefficient Determination	MAIC	MQIC	MBIC
1	0.9846717	69.875	69.883	69.895
2	0.9847247	68.087	68.124	68.184
3	0.9849711	67.906	67.973	68.080
4	**0.9855173***	**67.747***	**67.901***	**68.070***
5	0.9603621	67.819	67.915	68.115

4.2.3 Step 3. PVAR GMM Estimation

Given the optimal coefficient determination and information criteria, we fit the Panel VAR model with first fourth lags for all covariates for the estimation. The coefficient estimation of

$gamesales_{it}$ and $playtime_{it}$ as the dependent variable is summarized in Table 6 and the estimates are used to build the impulsive response functions in the later procedure.

Table 6. PVAR Coefficient estimation

EQ3: dep.var : h_playtime			
	b_GMM	se_GMM	t_GMM
L.h_YouTubeViews	1.745	0.206	8.463
L.h_TwitchViews	2.038	1.160	1.757
L2.h_YouTubeViews	1.215	0.155	7.849
L2.h_TwitchViews	0.993	0.114	8.741
L3.h_YouTubeViews	0.961	0.148	6.511
L3.h_TwitchViews	0.267	0.107	2.503
L4.h_YouTubeViews	1.362	0.158	8.632
L4.h_TwitchViews	1.620	0.110	14.701
EQ5: dep.var : h_ln_gamesales			
	b_GMM	se_GMM	t_GMM
L.h_YouTubeViews	0.187	0.011	16.746
L.h_TwitchViews	1.768	0.123	14.409
L2.h_YouTubeViews	0.115	0.011	10.661
L2.h_TwitchViews	0.156	0.012	12.668
L3.h_YouTubeViews	0.111	0.011	10.335
L3.h_TwitchViews	1.485	0.123	12.044
L4.h_YouTubeViews	0.153	0.011	14.220
L4.h_TwitchViews	0.159	0.012	12.926

4.2.4 Step 4. Granger Causality Test

Before the forecasting procedures, we examine the significance of the Granger causality between explanatory covariates and dependent covariates. Test results are presented in Table 7. On the aggregated level, $TwitchViews_{it}$ Granger causes $gamesales_{it}$ and $playtime_{it}$ on a 99% significance level. $YouTubeViews_{it}$ Granger causes $gamesales_{it}$ and $playtime_{it}$ on a 90% significance level.

Table 7. Panel VAR-Granger Causality Wald Test

Ha: Excluded variable Granger-causes Equation variable Ho: Excluded variable does not Granger-cause Equation variable		
Chi^2 (Significant level)		
Response to	$gamesales_{it}$	$playtime_{it}$
$TwitchViews_{it}$	49.35(***)	37.64(***)
$YouTubeViews_{it}$	8.73(*)	8.17(*)

Note: ***=0.01 sig level. **=0.05 sig level, *=0.1 sig level

4.2.5 Step 5. Impulse Response Function (IRF)

In the next step, given the estimated parameters of the Panel VAR model, we generate the impulse response functions (IRF). The impulse response function Φ_i can be captured by the re-shaping the reduced form Panel VAR model to infinite vector moving average form and the VMA parameters $\Theta_i = \begin{cases} I_k & , i = 0 \\ \sum_{j=1}^{i} \varphi_{t-j} A_j , & i = 1,2 \ldots \end{cases}$. The IRF with estimated coefficients can gauge the net effects of one percent of unexpected change in UGC metric *i* on $gamesales_{it}$ at time *t*. Standard errors are generated by Monte Carlo simulation with 500 repetition and coefficients' significance is tested by 0.95 confidence interval. We summarize the combined IRF graphs of key covariates in Figure 3. Based on innovation simulation, IRF illustrates how many percentage y-axis metric changes given one percent change of x axis metric. The dotted lines form a cone, which represents the significance level. Here, the range includes two standard deviation. In Table 8, we summarize the immediate predictive elasticity and the accumulated value that combines all significant effects across the forecasting periods from each IRF.

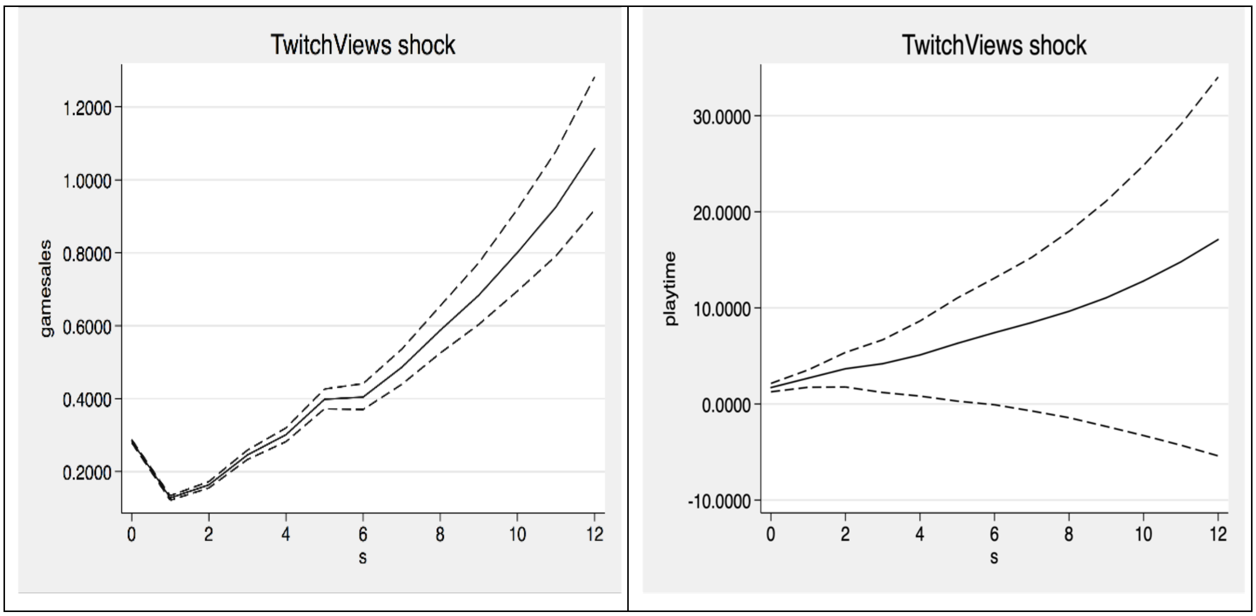
Table 8. Summary of IRF patterns

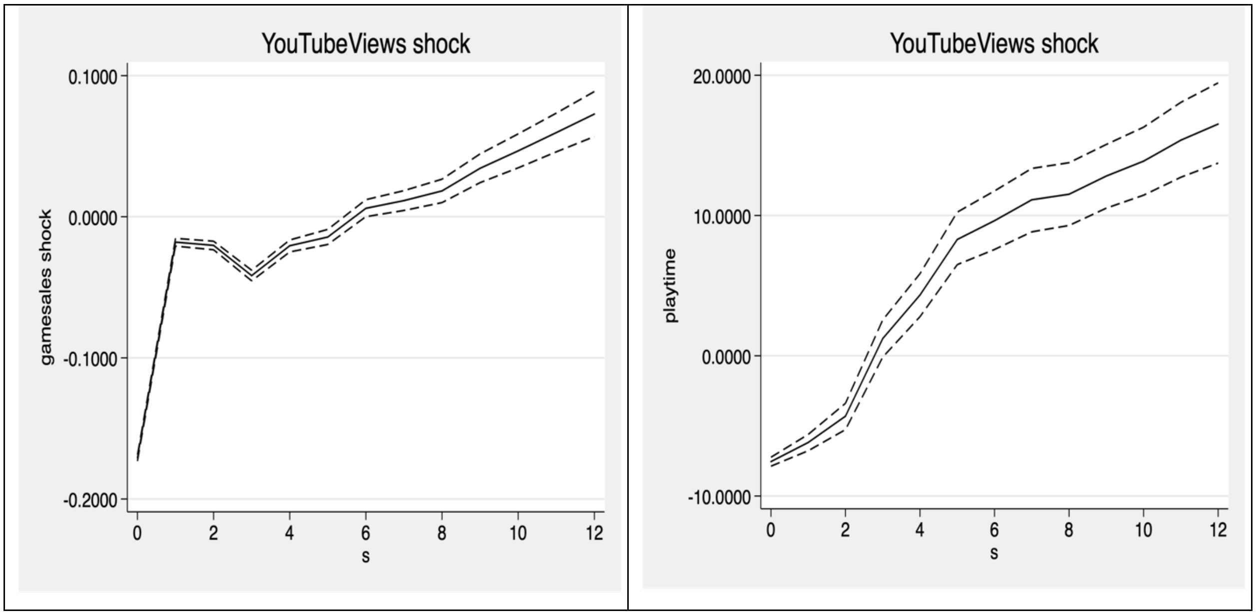
	$gamesales_{it}$		$playtime_{it}$	
	Immediate	Accumulated	Immediate	Accumulated
$TwitchViews_{it}$	0.284***	6.496***	1.706**	104.937**
$YouTubeViews_{it}$	-0.171*	0.804**	-7.547*	86.647**
$TwitchViews$ > $YouTubeViews$				
Kruskal-Wallis Statistics	18.778***		16.131***	
F Statistics	4.39***		24.587***	

The IRF results illustrate that $TwitchViews$ improves immediate and accumulated $gamesales$ and $playtime$. $YouTubeViews$ only improves accumulated $gamesales$ and $playtime$. More specifically, one percent $TwitchView$ can increase $gamesales$ by 6.5% and

$playtime$ by 105% in the long term; one percent $YouTubeView$ can increase $gamesales$ by 0.8% and $playtime$ by 86% in the long term. Consistent with the previous study (Tanis and Postmes 2003), pre-recorded video platforms as a lean media with limited capacity conveying cues and may bring immediate negative consequence. Such video content can be considered as a pure immediate substitute for games and thus could delay the purchase. Comparing the IRF coefficients for each period, the impact of $TwitchViews$ is significantly greater than that of $YouTubeViews$.

Figure 3. Impulse Responses for Panel VAR





4.2.6 Step 6. Forecast Error Variance Decomposition (FEVD)

Based on the estimated PVAR parameters, we derive the forecast- error- variance-decomposition (FEVD) estimates to isolate the contribution of viewership from two platforms to game sales and consumer stickiness. Similar to the determination coefficient, FEVD illustrates the relative predictive power of each covariate in explaining the variance of dependent covariates over time without assuming a causal ordering. Based on Abrigo and Love 2015 and Canova and Cirrarelli 2013, the h-step ahead forecast-error is described as $gamesales_{i,t+h} - E\left[gamesales_{i,t+h}\right] = \sum_{i=0}^{h-1} e_{i,(t+h-i)}\Theta_i$ and $playtime_{i,t+h} - E\left[playtime_{i,t+h}\right] = \sum_{i=0}^{h-1} e_{i,(t+h-i)}\Theta_i$. Based on Lutkepohl (2005), percentages representing the variance contributions in Table 9 are normalized relative to the 10-period ahead forecast-error variance of covariates. Thus, the FEVD results in Table 9a and 9b can identify the relative predictive value of $Twitchviews$ and $YouTubeViews$ to $gamesales$ and $playtime$.

Table 9a. Forecast-Error Variance Decomposition

	s	$playtime$	$gamesales$
$Twitchviews$	2	6.590%	6.875%
$Twitchviews$	4	4.195%	5.257%

$Twitchviews$	6	1.029%	9.339%
$Twitchviews$	8	3.643%	4.190%
$Twitchviews$	10	0.388%	0.785%
$YouTubeViews$	2	3.389%	5.219%
$YouTubeViews$	4	3.808%	3.175%
$YouTubeViews$	6	2.648%	5.308%
$YouTubeViews$	8	1.420%	2.307%
$YouTubeViews$	10	0.263%	0.348%

Table 9b. Average FEVD

	$gamesales$	$playtime$
$Twitchviews$	6.303%	5.168%
$YouTubeViews$	2.871%	1.844%
$Twitchviews > YouTubeViews$		
Kruskal-Wallis Statistics	5.771**	6.818**

4.3 Findings

Table 7 presents the immediate and cumulative impulsive response elasticities. The magnitude of the elasticities reflects how many percentage changes of $gamesales$ and $playtime$ in response to one percentage unexpected change of $Twitchviews$ and $YouTubeViews$. Hypotheses verification is shown in Table 10, and all of the hypotheses are supported on 95 significance level.

Table 10. Hypotheses verification

	Game sales	Stickiness
Live streaming video platform	H1a-supported	H1b-supported
Pre-recorded video platform	H2a-supported	H2b-supported
Live streaming video platform >Prerecorded video platform	H3a-supported	H3b-supported

4.3.1 Short- and Long-Term Predictive Values of Online Video Platforms

- **Live Streaming Video Platforms**

As shown in Table 8 and Figure 3, $TwitchViews$ has significantly positive predictive relationships with both the immediate and accumulated $gamesales$ and $playtime$. More specifically, one percent increase of $TwitchViews$ improves 0.284 (p<0.01) and 1.706 (p<0.05) percentage of immediate $gamesales$ and $playtime$. One percentage increase of

$Twitchviews$ improves 6.496 ($p<0.01$) and 104.937 ($p<0.05$) percentage of accumulated $gamesales$ and $playtime$. Therefore, the results support both H1a and H1b. The findings are consistent with Haili et al. 2017 and Chen et al. 2017 that viewing live streaming contents can improve viewers' consideration in purchasing related products during the purchase funnel and hence increase the sales.

- **Pre-Record Video Platforms**

In Table 8 and Figure 3, $YouTubeViews$ has weakly- significant negative predictive relationship with $gamesales$ and $playtime$ on the immediate level and has a significantly positive predictive relationship with the dependent covariates on the accumulated level. More specifically, one percent increase of $YouTubeViews$ decrease 0.171 ($p<0.1$) and 7.547 ($p<0.1$) percent of immediate $gamesales$ and $playtime$. One percent increase of $YouTubeViews$ improves 0.804 ($p<0.05$) and 86.647 ($p<0.05$) percentage of accumulated $gamesales$ and $playtime$. Consistent with Dennis and Kinney (1998), pre-recorded video platform, as a lean media with limited capacity conveying cues, may bring immediate negative consequence such as task performance and purchase intention. We considered such video content on lean media as a pure immediate substitute for games and thus can delay the purchase.

4.3.2 Relative Strength of the Predictive Value of Live Streaming Video Platforms versus Pre-Recorded Video Platforms

In the IRF procedure, we compare the coefficients of impulsive responses to viewership on two platforms for each period, the impacts of $TwitchViews$ are significantly larger than that of $YouTubeViews$. Thereby, both H3a and H3b are supported ($p<0.01$) such that live streaming video platforms do have a stronger predictive relationship with product sales and stickiness than pre-recorded video platforms.

In the FEVD procedure, Table 9a and 9b is constructed through Monte Carlo simulation with 500 repetitions for every 2th day up to 10 days and the significance level is 0.95. Specifically, $Twitchviews$ can explain 6.3% of $gamesales$ and 5.2% of $playtime$. $YouTubeViews$ can explain 2.9% of $gamesales$ and 1.8% of $playtime$. According to the Kruskal-Wallis test, $Twitchviews$'s explanatory power is also significantly better than that of $YouTubeViews$.

4.4 Validation and Robustness Check

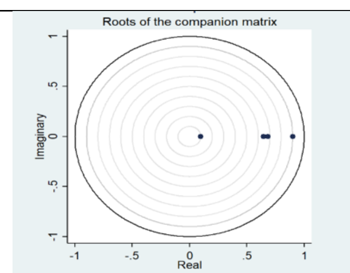
We conduct several additional tests to validate the robustness of the results.

4.4.1 Stability of the Estimated PVAR Check through the Eigenvalue Test

Based on Abrigo and Love 2015, we need to conduct the eigenvalue test to check the stability conditions of IRF and FEVD estimates. Covariates estimates are stable only if all the eigenvalues lie inside the unit circle. According to test result in Table 11, the estimated PVAR satisfies stationarity condition and the IRF and FEVD estimates are stable and consistent.

Table 11. Eigenvalue Stationarity Test

	Eigenvalue		Modulus
	Real	Imaginary	
$gamesales_{it}$	0.902	0	0.902
$playtime_{it}$	0.686	0	0.686
$TwitchViews_{it}$	0.648	0	0.648
$YouTubeViews_{it}$	0.098	0	0.098



4.4.2 Reverse Effect Check through Granger Causality Test

In Table 12. In the long run, the reverse effects of $gamesales_{it}$ and $playtime_{it}$ are either very weak or non-significant. Neither $gamesales_{it}$ or $playtime_{it}$ Granger causes $TwitchViews_{it}$ or$YouTubeViews_{it}$ on at least 95-significance level. Thus, there is no reverse causal effect.

Table 12. Panel VAR-Reverse Effect Check

Ha: Excluded variable Granger-causes Equation variable Ho: Excluded variable does not Granger-cause Equation variable		
Chi^2 (Significant level)		
Response to	$TwitchViews_{it}$	$YouTubeViews_{it}$
$gamesales_{it}$	9.03(*)	7.83(*)
$playtime_{it}$	4.32	7.67

5. Study 2- Online Experiments

5.1 Experiment Design

The objective of study 2 is to further compare the marketing efficiency between two platforms, validate hypothesis three, and to rule out the correlation between customer's purchase intention and preference of selecting platforms. More specifically we conduct a randomized experiment on Amazon MTurk through Qualtrics. We send 260 surveys to MTurk workers through Qualtrics and receive 238 valid responses. All recruited workers are located in the US with Human Intelligence Task (HIT) approval rate over 95 % and with Master Qualifications. According to MTurk, "workers who have demonstrated excellence across a wide range of HITs are awarded the Masters Qualification." Workers' languages are English. In this experiment, we use YouTube to represent for pre-recorded video platforms and Twitch to represent for live streaming video platforms. For each response, we provide 0.5 dollars as rewards.

The experiment flow includes the following procedures. First, we briefly introduce the study, live streaming video platforms and pre-recorded video platforms to all participants through a short information consent. Participants will then read a description and picture of a video game. After introducing the game, half of the randomly selected participants are surveyed their purchase intentions (pre-watch intention) to this game, and the other half are not. There are two purpose of this randomized pre-treatment design. The first is to rule out potential contagious issue due to the question order. The second is to exclude potential confounding issue. If the outcome variables of pre-questions and non-pre-questions do not have significant difference, the contagious issue will be excluded[38] and the following treatment effects are not caused by confounding factors. Next, in the treatment procedure, half of the randomly selected participants are told to self-select a video from YouTube or Twitch. The other half will be randomly assigned a video either from YouTube or Twitch. Then all participants will watch the self-select or randomly-assigned video. After watching the video, all participants will be asked about their purchase intentions (post-watch intention). At last, participants will be surveyed by demographic questions. Participants will also be surveyed about their previous game playing, the game purchasing experience, and experience about using two online video platforms (YouTube and Twitch). The demographic statistics are listed in the Appendix.

Based on the above experiment flow, we can have the following 4 subsamples. M_1 includes participants who self-select Twitch to watch the video. M_2 includes participants who self-select YouTube to watch the video. M_3 includes participants are randomly assigned the Twitch video to watch and M_4 includes participants are randomly assigned the YouTube video to watch. If there is significant difference between the post-watch intention of participants who randomly-assigned Twitch and that of participants who randomly-assigned YouTube, we can validate the marketing efficiency difference between two platforms. If there is self-selection bias, whether subjects can self-select video will affect their purchase intention before or after watching the video, and there will be a significant difference between the pre-and the post-watch purchase intentions of self-select and randomly assigned subjects.

38 The comparison of post-video-watching purchase intention between pre-question and no-pre-question subjects shows that there is no significant difference (p=0.337) and thus the study is free of the contagious issue.

More specifically, we will first examine the post-video watching purchase intention and that is the self-selection bias can be eliminated if the purchase intention of M_1 is no significantly different from that of M_3 and the purchase intention of M_2 is no significantly different from that of M_4. Moreover, the empirical analysis results can be further validated if the purchase intention of M_1 is significantly larger than that of M_2 and the purchase intention of M_3 is significantly larger than that of M_4. The pre-watch purchase intention of different subjects will be examined in later.

5.2 Findings

Among all participants, 53.54% are female, and 46.46% are male. Age ranges from 18 to 69 (mean 34.19). The main results are in Table 13. Based on Table 13, whether people self-select or are randomly assigned a platform to watch the video is not associated with their purchase intention. Therefore, this study is free of the self-selection issue. For both self-selection or the condition of being randomly assigned, watching video on live streaming video platform rather than on pre-recorded video platforms is associated with higher purchase intention. This result is consistent with and validates our empirical results.

To examine the re-watch intention and to rule out potential contagious and confounding issues for further validating the causality in our empirical result, we divide responses and analyze the following subsamples. We parse the responses that are asked the purchase intentions before watching the video also into the four groups. M_1' are participants who self-select Twitch to watch the video, M_2' are participants who self-select YouTube to watch the video, M_3' participants are randomly assigned the Twitch video to watch and M_4' includes participants are randomly assigned the YouTube video to watch. Comparison results are listed in Table 14. There is no significant difference in the pre-watch intentions for self-selected Twitch, self-selected

YouTube. Thus, we can consider people's preference to either Twitch or YouTube has no significant difference in purchase intention. This means the study has no self-selection bias, potential contagious and confounding issues.

Table 14 shows that there is no significant difference between the pre-watch intention of people who prefer and select Twitch, people who prefer and select YouTube, people who are randomly assigned Twitch and people who are randomly assigned YouTube. This cross-subsample indifference rules out the influence of confounding factors. Thus, this result guarantees that the different post-watch intentions of people who watching video on different platforms are not caused by confounding factors. This also strengthens the internal validity of our empirical results.

Table 13. Post-Watch Purchase Intention

	M1(Self-Select Twitch)	***M2(Self-Select YouTube)***	***M3(Random-Assign-Twitch)***	***M4(Random-Assign-YouTube)***
Mean	3.463	2.5	3.125	2.719
Variance	0.855	1.296	1.529	1.169
Observations	45	73	60	60
	Self-Selection bias rule-out		**Empirical results validation**	
Ho	*M1=M3*	*M2=M4*	*M1=M2*	*M3=M4*
t Stat	1.542	-1.117	4.888	1.855
P(T<=t) two-tail	0.126	0.266	0.000	0.033

Table 14. Pre-Watch Intention

	M1'(Self-Select Twitch)	***M2'(Self-Select YouTube)***	***M3'(Random-Assign-Twitch)***	***M4'(Random-Assign-YouTube)***
Mean	3.095	2.657	2.786	3.100
Variance	0.790	1.173	0.989	1.197
Observations	31	35	30	30
	Self-Selection bias rule-out		**Empirical results validation**	
Ho	*M1'=M3'*	*M2'=M4'*	*M1'=M2'*	*M3'=M4'*
t Stat	1.127	-1.636	1.674	1.673
P(T<=t) two-tail	0.265	0.107	0.124	0.258

6 Discussion

The first objective of this study is to examine how effectively do live streaming video platforms and pre-recorded video platforms play in influencer marketing, in terms of product sales and customer stickiness. Second, it is to compare their efficiency. The results suggest that given the difference of conveying cues, media richness, and social presence, live streaming video platforms can improve immediate and accumulated product sales and stickiness. Pre-recorded video platforms can only improve accumulated product sales and stickiness. Moreover, live streaming video platforms have a higher predictive performance than pre-recorded video platforms in terms of both response level and explanatory power. These findings are robust through Granger causality test, Eigenvalue test, and the online experiment. These findings provide unique and important implications for the theory and practice of UGC platforms.

6.1 Theoretical Implications

This study contributes to the literature across IS, marketing and communication fields. User-generated video platforms are valuable influencer marketing media, information diffusion channels across social media, and are timely and cues-rich communication technologies. Business editorials' comment "A picture is worth a thousand words, and a moving picture is worth a million"[39] may lack validation. However, product related videos do have a higher level of media richness than blogs or broadcast type of UGCs and do bring significant immediate and accumulated impact on product sales performance. In this sense, our research adds to the literature on digital content marketing. Specifically, our results indicate that online video platforms can improve accumulated product sales and customer stickiness. These findings provide empirical evidence for the marketing efficiency of different types of user-generated

[39] https://www.huffingtonpost.com/scott-macfarland/if-a-picture-video-production_b_4996655.html

videos. Thus, managers should allocate certain marketing budgets and resources across different online video platforms according to the significant predictive power of these rich media.

Furthermore, our study is the first to reveal the influence of live streaming video platforms and is the first to compare the marketing efficiency of live streaming video platforms with traditional pre-recorded video platforms by providing rich empirical findings in both IS and marketing fields. While both two type platforms can improve product sales and customer stickiness, only live streaming video platforms can give an immediate boost in product sales performance in the long term. Interestingly, we find that the immediate marketing efficiency of pre-recorded video platforms is negative in terms of both product sales and customer stickiness. This surprising finding supports Dennis and Kinney (1998)'s conclusion that limited richness of a media has negative consequence regarding task performance. Thus, for firms, especially for those to sell hedonic products like video games, conducting influencer marketing on live streaming platforms can bring better and immediate efficiency.

Also, through surveys and experiments, prior communication studies focusing on media richness theory and social presence theory have demonstrated the predictive power of media's richness level and their ability to show social presence on user's task performance such as purchase and re-use intention on the organization level. Our findings support their propositions and extend these streams of research in the following directions: (1) We apply the media richness theory on the individual level under the new communication technology setting and validate its main proposition through comprehensive empirical analysis. Specifically, live streaming video platforms, given its higher ability to convey more social cues and a higher level of media richness than pre-recorded video platforms, can provide users better social fulfillment, more efficient communication and better task performance. For example, rich cues can enhance individual's consideration through the purchase funnel and use process, and thereby improve the purchase and re-use intention, which, according to the use and gratification theory, are objectives of using such platforms. (2) We provide empirical evidence and validate the proposition of social presence theory by quantifying purchase and re-use intentions to sales and customer stickiness metrics.

Finally, we conduct a Panel VAR analysis, which demonstrates the dynamics influence and causality directions among covariates, as well as both immediate and accumulated impacts in terms of both elasticity and prediction power. Our model prevents neglecting the enduring effects of user-generated video platforms and avoids being confounded by the reverse causality, which may threaten the internal validity of findings.

6.2 Managerial Implications

Our research also informs managers in several ways. Both live streaming video platforms and pre-recorded video platforms can serve to be an effective marketing channel in the long term to improve sales performance. Still, allocating marketing budgets and efficiently utilizing between online video platforms are confounded by not knowing their dynamic impacts and their essential difference. More and more companies plunge in and enlarge their advertising budgets to content marketing but many of them suffer from uncertain investment returns and low marketing efficiency. However, our results illustrate the distinct efficiency of different online video platforms for marketing in the short and long terms.

Analyzing the short-term and immediate IRF between live streaming and pre-record video platforms will demonstrate managers the instant social media investment return on each platform and will also show them wear-in periods for positive returns and wear-out periods for negative returns. As reported in Table 8 and Figure 3, the immediate returns of investing YouTube in terms of sales and customer stickiness are both negative, and the accumulated returns will turn positive after 2 to 6 days. Given the limited capacity to provide rich social interaction gadgets, content generators on YouTube have fewer means to communicate with their viewers. Accordingly, their videos lack feedbacks and interactions and are more concentrated just on gameplay, which makes the content a pure substitute for games. To mitigate the immediate negative effect, managers who have input plenty of resources need to, given the wear-out periods, add more online campaigns (For example, launching short live webinars, voting campaigns and product-related discussion posts on YouTube along with other social media like Instagram) to encourage interactions with customers and to improve their sense of social presence. Such additional online activities can mitigate the limited ability of lean media to convey cues.

Comparing the long-term IRF can show managers the aggregated returns of investing on different platforms. Based on our results, we generally suggest managers, especially those who run business in the entertainment and digital product industry, (1) to create and share vlog type of videos embedding with products and branding information, or (2) to sponsor and cooperate with influencers on such platforms for branding and inbound marketing. Moreover, based on our findings, we recommend managers, whose goal is to gain immediate benefits, to conduct content marketing or influencer marketing on live streaming video platforms.

7 Conclusion and Future Research

This study reveals the salient influence and the predicting power of live streaming video platforms. Moreover, it sheds lights on the importance of this revolutionary UGC platform type and social media. By conduct a comprehensive PVAR analysis and experimental study, we quantify the efficiency of online video platforms as influencer marketing channels. Moreover, we compare the efficiencies of the live streaming video platforms and pre-recorded video platforms and discover the dynamic patterns of their different impacts. By combining the findings in Media Richness Theory, Social Presence theory and Use and Gratification Theory as well as the findings of our online experiments, we explain the different impacts of two types of platforms. Additionally, we extend the application of the above theories to the contexts of the live streaming online environment, which is an under-researched façade of the internet.

This research has several limitations, which can serve to be the future research directions. First, if viewers' data, and more specifically, the viewers' click-through data is available, the future research can track users' actual purchase behaviors after consuming online contents and specify and model the attributes of online video platforms that cause the purchase. Second, if text blog contents and broadcast contents for each game are available, future research can further compare the marketing efficiency across different types of UGC. Third, if there will be a chance to conduct field experiment under the setting of utilitarian products rather than only hedonic products like music and video gaming, the future research can expand our findings to a more general level and improve the external validity.

Reference

Abrigo, Michael RM, and Inessa Love. "Estimation of panel vector auto regression in Stata: A package of programs." manuscript, Feb 2015.

Andrews, Donald WK, and Biao Lu. "Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models." *Journal of Econometrics* 101.1 (2001): 123-164.

Adomavicius G, Bockstedt J, Gupta A (2012) Modeling supply-side dynamics of IT components, products, and infrastructure: An empirical analysis using vector auto regression. *Information Systems Research*. 23(2):397–417.

Blundell, Richard, and Stephen Bond. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of econometrics* 87.1 (1998): 115-143.

Canova, Fabio, and Matteo Ciccarelli. "Panel Vector Autoregressive Models: A Survey" VAR Models in Macroeconomics–Sims. Emerald Group Publishing Limited, 2013. 205-246.

Chen H, De P, Hu Y J. IT-enabled broadcasting in social media: An empirical study of artists' activities and music sales. *Information Systems Research*, 2015, 26(3): 513-531.

Choi J, Lee H J, Kim Y C. The influence of social presence on customer intention to reuse online recommender systems: The roles of personalization and product type. *International Journal of Electronic Commerce*, 2011, 16(1): 129-154.

Daft R L, Lengel R H. Organizational information requirements, media richness and structural design. *Management science*, 1986, 32(5): 554-571.

Dennis A R, Kinney S T. Testing media richness theory in the new media: The effects of cues, feedback, and task equivocality. *Information Systems Research*, 1998, 9(3): 256-274.

Dewan S, Ramaprasad J. Social Media, Traditional Media, and Music Sales: A Panel VAR Approach. *MIS Quarterly*, 2014, 38(1).

Flanagin A J, Metzger M J. Internet use in the contemporary media environment. *Human communication research*, 2001, 27(1): 153-181.

Fort, Teresa C., et al. "How firms respond to business cycles: The role of firm age and firm size." IMF Economic Review 61.3 (2013): 520-559.

Freeth M, Foulsham T, Kingstone A. What affects social attention? Social presence, eye contact and autistic traits. *PloS one*, 2013, 8(1): e53286.

Gunawardena C N. Social presence theory and implications for interaction and collaborative learning in computer conferences. *International Journal of Educational Telecommunications*, 1995, 1(2): 147-166.

Love, Inessa, and Lea Zicchino. "Financial development and dynamic investment behavior: Evidence from panel VAR." *The Quarterly Review of Economics and Finance* 46.2 (2006): 190-210.

Lof M, Malinen T. Does sovereign debt weaken economic growth? A panel VAR analysis. *Economics Letters*, 2014, 122(3): 403-407.

Hajli N. Social commerce constructs and consumer's intention to buy. *International Journal of Information Management*, 2015, 35(2): 183-191.

Hajli N, Sims J, Zadeh A H, et al. A social commerce investigation of the role of trust in a social networking site on purchase intentions. *Journal of Business Research*, 2017, 71: 133-141.

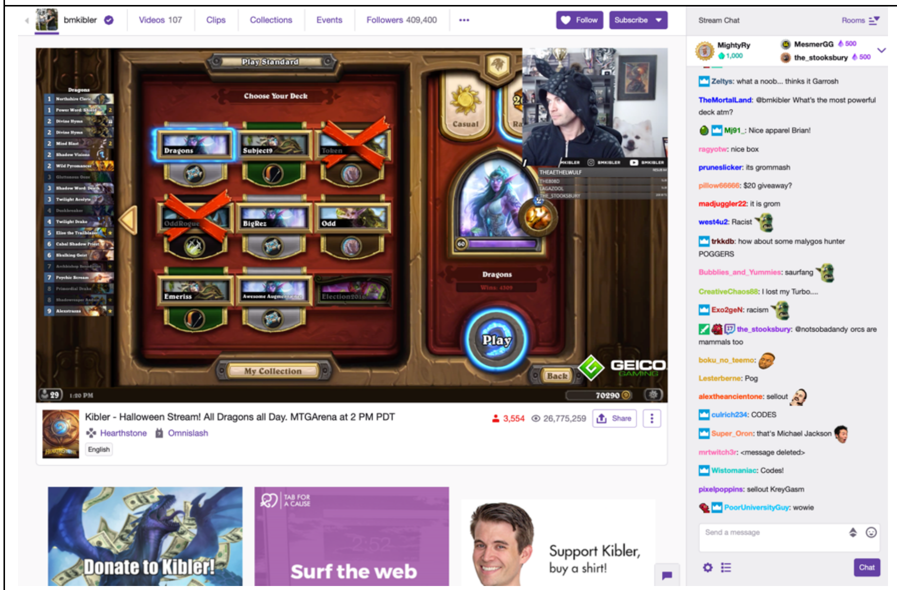
Hayakawa K. Improved GMM estimation of panel VAR models. Computational Statistics & Data Analysis, 2016, 100: 240-264.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen. "Estimating vector auto regressions with panel data." *Econometrica*: Journal of the Econometric Society (1988): 1371-1395.
Kumar N, Benbasat I. Research note: the influence of recommendations and consumer reviews on evaluations of websites. *Information Systems Research*, 2006, 17(4): 425-439.
Kock, Ned. "The psychobiological model: Towards a new theory of computer-mediated communication based on Darwinian evolution." *Organization science* 15.3 (2004): 327-348.
Li, Xitong. "Could deal promotion improve merchants' online reputations? The moderating role of prior reviews." Journal of Management Information Systems 33.1 (2016): 171-201.
Li D, Browne G J, Wetherbe J C. Why do internet users stick with a specific web site? A relationship perspective. International journal of electronic commerce, 2006, 10(4): 105-141.
Liang T P, Ho Y T, Li Y W, et al. What drives social commerce: The role of social support and relationship quality. International Journal of Electronic Commerce, 2011, 16(2): 69-90.
Lof, Matthijs, and Tuomas Malinen. "Does sovereign debt weaken economic growth? A panel VAR analysis." Economics Letters 122.3 (2014): 403-407.
Lowenthal P R. Social presence//Social computing: Concepts, methodologies, tools, and applications. IGI Global, 2010: 129-136.
Luca, Michael. "Reviews, reputation, and revenue: The case of Yelp.com." Com (March 15, 2016). Harvard Business School NOM Unit Working Paper 12-016 (2016).
Luo, Xueming, Jie Zhang, and Wenjing Duan. "Social media and firm equity value." Information Systems Research 24.1 (2013): 146-163.
Luo X (2009) Quantifying the long-term impact of negative word of mouth on cash flows and stock prices. Marketing Sci. 28(1):148–165.
Lütkepohl, Helmut. New introduction to multiple time series analysis. Springer Science & Business Media, 2005.
Mendes-Filho L, Tan F B. User-generated content and consumer empowerment in the travel industry: A uses & gratifications and dual-process conceptualization. PACIS 2009 Proceedings, 2009: 28.
Miranda S M, Saunders C S. The social construction of meaning: An alternative perspective on information sharing. Information systems research, 2003, 14(1): 87-106.
Tanis M, Postmes T. Social cues and impression formation in CMC. Journal of Communication, 2003, 53(4): 676-693.
O'Neill, Mark, et al. "Condensing Steam: Distilling the Diversity of Gamer Behavior." Proceedings of the 2016 Internet Measurement Conference. ACM, 2016.
Rice R E. Task analyzability, use of new media, and effectiveness: A multi-site exploration of media richness. Organization science, 1992, 3(4): 475-500.
Srinivasan, Shuba, Marc Vanhuele, and Koen Pauwels. "Mind-set metrics in market response models: An integrative approach." Journal of Marketing Research 47.4 (2010): 672-684.
Sjöblom, Max, and Juho Hamari. "Why do people watch others play video games? An empirical study on the motivations of Twitch users." Computers in Human Behavior 75 (2017): 985-996.
Treviño L K, Webster J, Stein E W. Making connections: Complementary influences on communication media choices, attitudes, and use. Organization Science, 2000, 11(2): 163-182.
Whitehead, A. Meta-Analysis of Controlled Clinical Trials. Chichester, UK: Wiley 2002.
Yee, Nick, Nicolas Ducheneaut, and Les Nelson. "Online gaming motivations scale: development and validation." Proceedings of the SIGCHI conference on human factors in computing systems. ACM, 2012.

Yee, Nick. "The demographics, motivations, and derived experiences of users of massively multi-user online graphical environments." Presence: Teleoperators and virtual environments 15.3 (2006): 309-329
Zhang H, Lu Y, Gupta S, et al. What motivates customers to participate in social commerce? The impact of technological environments and virtual customer experiences. Information & Management, 2014, 51(8): 1017-1030.
Zhu L, Benbasat I, Jiang Z. Let's shop online together: An empirical investigation of collaborative online shopping support. Information Systems Research, 2010, 21(4): 872-891.

Appendix

Figure a1. Twitch- An example of live streaming video platforms



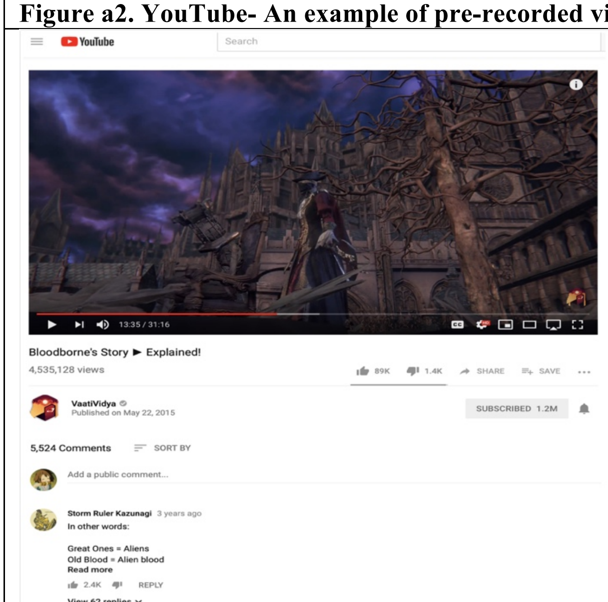
Interaction:

- Immediate feedback
- Direct communication

Gadgets:

1. Live chat
2. Customizable gifting
3. Customizable emoji
4. Conspicuous donation
5. Real-time barrage comments
6. Customizable streamer extensions

Figure a2. YouTube- An example of pre-recorded video platforms



Interaction

- Indirect communication
- Delayed feedback

Gadgets:

1. Comment

b. Experiment Demographic Statistics

Demographic statistics			
Annual Income	**Frequency (ratio)**	**Employment Status**	**Frequency (ratio)**
Less than 20,000	47 (20.79%)	Employed	113 (50.00%)
20,000 ~ 34,999	62 (27.43%)	Student	32 (14.16%)
35,000 ~ 49,999	50 (22.12%)	Homemaker	18 (7.96%)

50,000 ~ 74,999	52 (23.01%)
75,000 ~ 99,999	6 (2.65%)
Over 99,999	9 (3.98%)

Self-employed	55 (24.34%)
Unemployed	8 (3.54%)

Education level	Frequency (ratio)
Less than high school	4 (1.77%)
High school	65 (28.76%)
College	136 (60.18%)
Graduate school	21 (9.29%)

Gender	Frequency (ratio)
Female	53.43%
Male	46.46%

Marriage Status	Frequency (ratio)
Single	144 (63.72%)
Married	82 (36.28%)

c. Survey flow

Survey Flow

Standard: Information Consent (3 Questions)

Branch: New Branch

If

If Attention Check Question Which platform will not be considered in this study? Twitter Is Not Selected

EndSurvey: Advanced

Standard: Questions about gaming experiences (5 Questions)

Standard: Questions about platform experiences (7 Questions)

BlockRandomizer: 1 - Evenly Present Elements

BlockRandomizer: 1 - Evenly Present Elements

Standard: Experiment - Random Allocation (PreQ) (8 Questions)

Standard: Experiment - Random Allocation (No PreQ) (7 Questions)

BlockRandomizer: 1 - Evenly Present Elements

Standard: Experiment - Self Select (PreQ) (10 Questions)

Standard: Experiment - Self Select (No PreQ) (9 Questions)

Standard: Demographical questions (7 Questions)

WebService: GET - http://reporting.qualtrics.com/projects/randomNumGen.php - Fire and Forget

Standard: Show Survey Code (1 Question)

Page Break

d. Survey questions

Start of Block: Information Consent

Information Consent My name is Yuan Zhang, and I am requesting your participation in a UT Arlington research study. The purpose of this study is to examine how user-generated video platforms impact product sales performance. The procedures that you will follow as a research subject are 1). Reading informed consent. 2). Answering survey questions to provide your experience about using YouTube, Twitch purchasing and playing video games before and after being introduced a game and its related video. 3). Entering demographic data (NO identifiable data), such as age, gender, game preference and experience etc., and it should take about 3 to 5 minutes. There are no perceived risks in this study. You might be able to learn a new platform and service, live streaming. There are no alternatives to this research project, but you may quit at any time. You must be at least 18 years old to participate. You will receive compensation from Amazon MTurks for participating in this research study. No identifiable information will be collected, and all records will be kept confidential with access limited to the research team. We may publish, present, or share the results, but your name will not be used. If you have questions about the study, you can contact me at Yuan Zhang, yuan.zhang@uta.edu. For questions or concerns, contact the UTA Research Office at 817-272-3723 or regulatoryservices@uta.edu. By clicking on the arrow button to the next page, you confirm that you are 18 years of age or older, and have read or had this document read to you. You have been informed about this study's purpose, procedures, possible benefits, and risks.

Introduction about research

This survey is aimed to study the how people's choice of online video platforms influence their purchase and gaming behavior. We will compare Twitch (the live streaming, "live" video) with YouTube (pre-recorded videos). Twitch is an online service for watching and streaming digital video broadcasts and Twitch is equipped with a live chat feature that allow users to interact with other audience and broadcasters immediately. YouTube is a video sharing service where users can watch, like, share, comment and upload their own videos, and on YouTube YouTuber and their audience can communicate through leaving comments. The survey will take about 3 to 5 minutes

---

Attention Check question

Which platform will **not** be considered in this study?

○ YouTube (1)

○ Twitter (2)

○ Twitch (4)

End of Block: Information Consent

Start of Block: Questions about gaming experiences

Questions about gaming experiences

1. Have you purchased video games before (PC games, mobile games or console games)?

○ Yes (1)

○ No (2)

Skip To: End of Block If 1. Have you purchased video games before (PC games, mobile games or console games)? = No

2. On average, how much do you spend in purchasing games and gaming related products (peripherals or equipment) **per month**?

- ○ 1-25 dollars (1)
- ○ 26-50 dollars (2)
- ○ 51-75 dollars (3)
- ○ 76-100 dollars (6)
- ○ more than 101 dollars (5)

3. Which genres of games do you purchase? (You can select multiple options)?

- ☐ Action (1)
- ☐ Strategy (2)
- ☐ RPG (Role-Playing Game) (3)
- ☐ Indie (4)
- ☐ Adventures (5)
- ☐ Sports (6)
- ☐ Simulation (7)
- ☐ MOBA (Multiplayer Online Battle Arena) (8)
- ☐ Others (9)

4. On average, how much time do you spend in playing video games **per day**?

- ○ I don't play video games (1)
- ○ 1-30 minutes (2)
- ○ 31-60 minutes (3)
- ○ 61-90 minutes (4)
- ○ 91-120 minutes (5)
- ○ More than 121 minutes (6)

End of Block: Questions about gaming experiences

Start of Block: Questions about platform experiences

Questions about platform experiences

5. Do you watch YouTube?

- ○ Yes (1)
- ○ No (2)

Skip To: Q16 If 5. Do you watch YouTube? = No

6. Are you a YouTuber who publishing videos on YouTube?

- I am a YouTuber (1)
- I only watch videos on YouTube as an audience (2)
- Both (3)

7. On average, how much time do you spend on YouTube **per day**?

- 1-30 minutes (1)
- 31-60 minutes (2)
- 61-90 minutes (3)
- 91-120 minutes (5)
- More than 121 minutes (4)

8. Do you watch Twitch?

- Yes (1)
- No (2)

Skip To: End of Block If 8. Do you watch Twitch? = No

9. Are you a Twitch streamer who broadcast channel on Twitch?

- ○ I am a Twitch streamer (1)
- ○ I only watch videos on Twitch as an audience (2)
- ○ Both (3)

10. On average, how long time do you spend in visiting Twitch **per day**?

- ○ 1-30 minutes (1)
- ○ 31-60 minutes (2)
- ○ 61-90 minutes (3)
- ○ 91-120 minutes (5)
- ○ More than 121 minutes (4)

End of Block: Questions about platform experiences

Start of Block: Experiment - Random Allocation (PreQ)

Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Please read the following description of a game: PLAYERUNKNOWN'S BATTLEGROUNDS is a battle royale shooter that pits 100 players against each other in a struggle for survival. Gather supplies and outwit your opponents to become the last person standing.

17. After reading the above description about the game, are you interested in buying this game?

- Strongly Disagree (1)
- Disagree (2)
- Neutral (3)
- Agree (4)
- Strongly Agree (5)

Now please enjoy a short video on Twitch, and complete the last question afterward.
<u>Please do not skip watching the video, or you may not receive the reward.</u>
Now please enjoy a short video on YouTube, and complete the last question afterward.
<u>Please do not skip watching the video, or you may not receive the reward.</u>

19. Now are you interested in buying this game?

- Strongly Disagree (1)
- Disagree (2)
- Neutral (3)
- Agree (4)
- Strongly Agree (5)

20. After viewing the video, do you want to chat or interact with the video generator?

- Yes (1)
- No (2)

21. Regardless whether you owned this game, are you interested in playing this game?

- Yes (1)
- No (2)

End of Block: Experiment - Random Allocation (PreQ)

Start of Block: Experiment - Random Allocation (No PreQ)

Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Please read the following description of a game: PLAYERUNKNOWN'S BATTLEGROUNDS is a battle royale shooter that pits 100 players against each other in a struggle for survival. Gather supplies and outwit your opponents to become the last person standing.

Now please enjoy a short video on Twitch, and complete the last question afterward.

Please do not skip watching the video, or you may not receive the reward.

Now please enjoy a short video on YouTube, and complete the last question afterward.

Please do not skip watching the video, or you may not receive the reward.

19. Now are you interested in buying this game?

- ○ Strongly Disagree (1)
- ○ Disagree (2)
- ○ Neutral (3)
- ○ Agree (4)
- ○ Strongly Agree (5)

20. After viewing the video, do you want to chat or interact with the video generator?

- ○ Yes (1)
- ○ No (2)

21. Regardless whether you owned this game, are you interested in playing this game?

- ○ Yes (1)
- ○ No (2)

End of Block: Experiment - Random Allocation (No PreQ)

Start of Block: Experiment - Self Select (PreQ)

Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)
Please read the following description of a game: PLAYERUNKNOWN'S BATTLEGROUNDS is a battle royale shooter that pits 100 players against each other in a struggle for survival. Gather supplies and outwit your opponents to become the last person standing.

17. After reading the above description about the game, are you interested in buying this game?

- Strongly Disagree (1)
- Disagree (2)
- Neutral (3)
- Agree (4)
- Strongly Agree (5)

Now we have a pre-recorded video from YouTube and a live streaming video from Twitch and they are all about this game.
Twitch **YouTube**
18. Which platform will you choose to view the video?

- Twitch (1)
- YouTube (2)

Skip To: Q30 If 18. Which platform will you choose to view the video? = Twitch

Skip To: Q31 If 18. Which platform will you choose to view the video? = YouTube

Now please enjoy a short video on Twitch, and complete the last question afterward.
Please do not skip watching the video, or you may not receive the reward.

Skip To: Q33 If Now please enjoy a short video on Twitch, and complete the last question afterward. Please do not...() Is Displayed

Now please enjoy a short video on YouTube, and complete the last question afterward.
Please do not skip watching the video, or you may not receive the reward.

Skip To: Q33 If Now please enjoy a short video on YouTube, and complete the last question afterward. Please do no...() Is Displayed

19. Now are you interested in buying this game?

- Strongly Disagree (1)
- Disagree (2)
- Neutral (3)
- Agree (4)
- Strongly Agree (5)

20. After viewing the video, do you want to chat or interact with the video generator?

- Yes (1)
- No (2)

21. Regardless whether you owned this game, are you interested in playing this game?

- Yes (1)
- No (2)

End of Block: Experiment - Self Select (PreQ)

Start of Block: Experiment - Self Select (No PreQ)

Timing
First Click (1)
Last Click (2)
Page Submit (3)
Click Count (4)

Please read the following description of a game: PLAYERUNKNOWN'S BATTLEGROUNDS is a battle royale shooter that pits 100 players against each other in a struggle for survival. Gather supplies and outwit your opponents to become the last person standing.

Now we have a pre-recorded video from YouTube and a live streaming video from Twitch and they are all about this game.

Twitch **YouTube**

18. Which platform will you choose to view the video?

- Twitch (1)
- YouTube (2)

Skip To: Q62 If 18. Which platform will you choose to view the video? = Twitch

Skip To: Q63 If 18. Which platform will you choose to view the video? = YouTube

Now please enjoy a short video on Twitch, and complete the last question afterward.
Please do not skip watching the video, or you may not receive the reward.

Skip To: Q64 If Now please enjoy a short video on Twitch, and complete the last question afterward. Please do not...() Is Displayed

Now please enjoy a short video on YouTube, and complete the last question afterward.
Please do not skip watching the video, or you may not receive the reward.

Skip To: Q64 If Now please enjoy a short video on YouTube, and complete the last question afterward. Please do no...() Is Displayed

19. Now are you interested in buying this game?

- Strongly Disagree (1)
- Disagree (2)
- Neutral (3)
- Agree (4)
- Strongly Agree (5)

20. After viewing the video, do you want to chat or interact with the video generator?

- Yes (1)
- No (2)

21. Regardless whether you owned this game, are you interested in playing this game?

- Yes (1)
- No (2)

End of Block: Experiment - Self Select (No PreQ)

Start of Block: Demographical questions

Demographic questions
11. What is your age?

____

12. What is your annual income?

- Less than $20,000 (1)
- $20,000 to $34,999 (2)
- $35,000 to $49,999 (3)
- $50,000 to $74,999 (4)
- $75,000 to $99,999 (5)
- Over $100,000 (6)

13. What is your gender?

- Male (1)
- Female (2)

14. What is your current employment status?

- Employed (1)
- Student (2)
- Homemaker (3)
- Self-employed (4)
- Unemployed (5)

15. What is the highest degree or level of school you have completed?

- Less than high school (1)
- High school (2)
- College (3)
- Graduate school (4)

16. What is your marriage status?

- Single (1)
- Married (2)

End of Block: Demographical questions

Start of Block: Show Survey Code

Thank you for your participation. Please provide the completion code below on MTurk in order to receive your payment.
${e://Field/random}

Please be sure to click on >>> to complete the survey, so that your responses are recorded. We will not be able to pay you if your answers are not recorded.

End of Block: Show Survey Code

CHAPTER FOUR

RUN FOR THE GROUP: THE EFFECTS OF WITHIN- AND BETWEEN-GROUP SOCIAL COMPARISON AND OFFLINE SOCIAL ACTIVITIES ON GROUP USERS' EXERCISE PARTICIPATION - EVIDENCE FROM A MOBILE FITNESS APP

Yuan Zhang

Jie (Jennifer) Zhang

Zilong Liu

Xiaolong Song

Abstract

To encourage users to exercise more and to improve the retention, fitness application developers build apps with more social interaction features on the collective level, such as allowing users to join and work out with groups and hold offline team-building events and social activities for group members. However, little is known regarding the impacts of the within-group and between-group social comparison on the group members' exercise participation. Motivated thus, we build a conceptual framework to examine the effects based on the social comparison theory. Through analyzing the underlying determinants (comparison dimension reference and similarity) of three social comparison cognitive processes (contrast, assimilation and reflection), we propose that both the within- and between-group social comparisons can significantly improve group members' exercise participation. We also propose that offline group social activities can significantly moderate the effects of the effects of within-group and between-group social comparison on group members' exercise participation. We manually record and collect group level users' data from a mobile fitness app and conduct a series of comprehensive empirical analyses to test and validate the main and moderating effects. Our findings validate the positive effects of within-group and between-group social comparisons and reveal that the number of offline activities moderates the main effects in opposite directions. Our findings help fitness app developers to better understand the impacts of offline social activities on the participation of the online virtual groups, and further, we provide implications regarding how to make online community policies and design incentive mechanisms to stimulate and promote offline social activities.

Keywords: offline social activities, fitness app, group level social comparison, user participation

1. Introduction

Fitness applications facilitate users to record exercise activities and self-regulate health conditions and are found to help motivate people to work out more and live healthier (Zhou et al 2016). However, app developers found that only the recording function is not sufficient to engage and attract users and the retention rate drops off while the initial passion fades out (Sonders and Ana L. 2016). Accordingly, fitness apps like Runkeeper, Keep and Nike+, etc. are designed with social interactions on the individual user level with the aim of improving users' participation in both the exercise and the app use. While most prior research (Zhou et al 2016, Cavallo et al 2012 and Richardson et al 2010) focused the effects of peer influence and social comparison on the individual level, literature examining these effects conclude with mixed results. For example, through analyzing environmental factors in social comparison among individuals, Wu et al 2015, Zhou et al 2016 and Munson and Consolvo 2012 found that the social interaction among individual users are able to further improve users' activity level and overall health behaviors. However, some social network analysis literature reveal the deficiency of social interaction's effects on users' physical activities. For example, Aral and Nicolaides (2017) pointed that exercise is socially contagious, and less active runners influence more active runners, but not the reverse. Following these mixed findings, upward comparison benchmarks, like top performance runners on the leaderboard, might not efficiently contribute to motivating users to exercise more. It is unclear for practitioners either whether it is effective and efficient to adopt social interaction features to promote user participation. The effects of social comparison and peer influence on the individual calls for reasoning and examination for the underlying mechanisms based on key determinants.

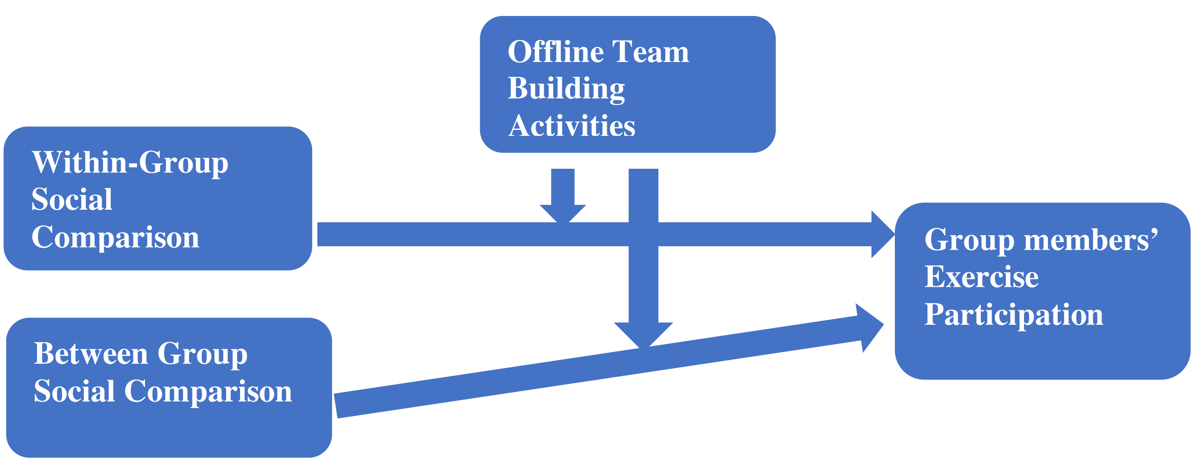
Given the uncertain efficiency of interpersonal integration features and the objective to improve users' exercise frequency and their retention. Developers recently equip their fitness app with group level social interaction and gamification functions. Besides the integration with social media, many apps also allow users to create and join groups, communicate through online group pages as well as initiate, participate and record offline group meetups, upload and share meetup pictures, etc. Group members can observe and might be encouraged to exercise more by the activity records of top runners, the activity records of group organizers, pictures and documents of group offline meetup events, the rank of the group and the overall activity records and performance of the group from the group board information. However, app developers know little about whether and how this" group" concept and the offline social interactions can effectively promote users to exercise more active and use their fitness app more often. Fewer researchers have studied on the group identification contours the effects of social comparison on users' participation and group behaviors. Moreover, little is known regarding the role of offline group member social interactions, such as teambuilding events and causal social activities, on the social comparison effects on users' exercise participation. In response to the above practical concerns and literature gaps, we aim to examine the effects of fitness apps with group level social interactions on group user's physical activities and propose two research questions. (1). How do within-group and between-group social comparisons affect group members' exercise performance? (2) How do offline social activities moderate on these effects?

To address these research questions, we build a conceptual model in Figure 1 mainly based on the social comparison theory. Given the consequences of upward assimilation due to within-group comparison, top performance members can be considered as an athlete role model and upward comparison benchmarks by other group members. The out-performed records can stimulate other members to exercise more active and more often. Since group organizers participate more frequently in group events and their records and information are more often exposed to other group members. As influencers, the exercise performance record of group organizers shall also have a significant impact in encouraging others to participate workout more. In general, the within-group social comparison affects group members' activity levels and can shape the group behavior (Kelman 1958, Mussweiler et al. 2004). This theory also demonstrates that environmental factors like competitive climate can have contingency impact on the effect of between-group comparison and consequently on the group level physical activity behavior (Brown et al. 2007 and We et al. 2005). Such competitive climates between groups can enlarge members' sense of "group" and amplify their focus on group task performance. Through examining the influence of team building activities on organization’s task performance, Tuckman (1965), reveals that casual social activities can improve social support and social ties and decrease social distance among colleagues who work in the same group. Team building or casual social events can thus enhance the group cohesion and motive co-workers to have better group task performance. We apply the findings of these organization studies to the online hobby-group setting and combine their logic to our theoretical reasoning based on the social comparison theory. We expect that the offline group social activities can moderate the effects of within-group and between-group social comparisons on group members’ exercise participation.

We test our model with a unique dataset manually recorded from an outdoor running app through a fixed-effect baseline analysis, the moderation analysis, and dynamic Panel VAR analysis. We find that both the within-group and the between-group social comparisons can significantly improve the group members’ exercise participation. The number of offline activities has a significant positively moderation impact on the relationship between the between-group competitive climate metrics and group participation. However, the interactions between offline team building and within-group social comparison are negative. The robustness and validation of analysis results are checked through different tests, and we also conduct several additional analyses to examine the dynamic impacts of group level social comparison further.

First, we extend the social comparison theory to the group level and examines the role of "group" by differentiating the social comparison within group members and the comparison between different groups. Second, in previous literature, group identity is examined as an environmental factor and measured mainly through surveying people's perceived belongingness. Our study provides direct empirical evidence to quantify the distinct consequences of with-in and between group social comparisons. Third, there are few previous research studying the user's offline social interactions and their impacts on user's online retention and exercise participation. By combining the Teamwork theory, we explain distinctive moderating effects of offline team building activities on the effects of within-group and between-group social comparison. Fourth, we extend the generalization of the Teamwork theory from a working setting in the human resource field to an online-hobby-community setting in the IS field by verifying its central propositions through rich empirical evidence.

Figure 1. Conceptual Framework

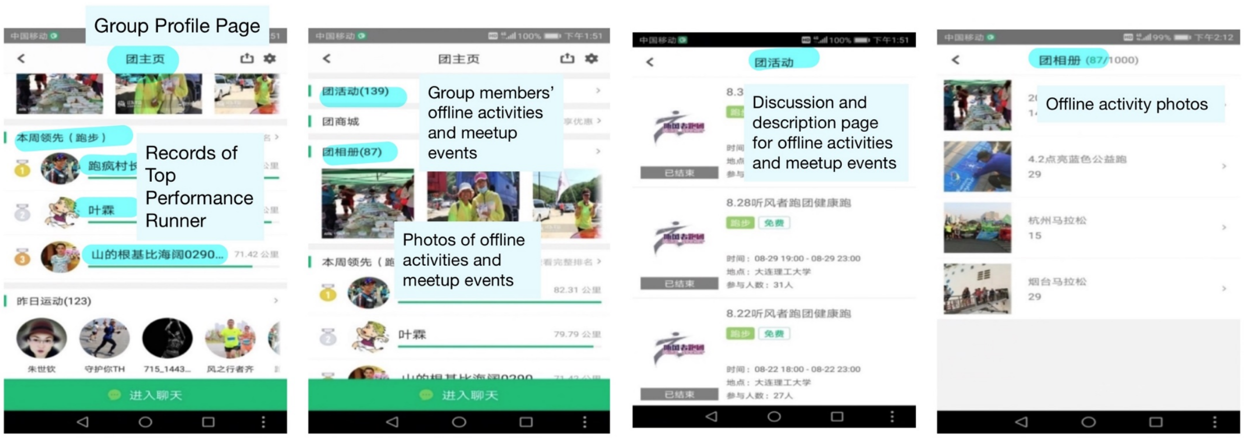


2. Theory Foundation and Literature Review

2. 1 Research Context and Conceptual Framework overview

To examine the group level social comparison effects and influence of offline activities, we select a fitness application as the research context of this study and manually observed and recorded data every week for about a year long. This app is one of the most popular fitness mobile applications for outdoor runners. The application enables users to record and monitor their exercise activities and health conditions, as well as helps users to create and improve exercise routines. Besides the essential functions, and similar to other fitness and healthcare application, this app is designed with social interaction features to improve users' exercise, engagement, and retention. However, unlike other apps, it is the social interaction function, and gamification features are designed on the group level. In this app, users can create and join virtual groups (Figure 2). After the user joins any running group, he/she can observe the records of both top performance members and that of the group organizer. On the group page, members can check the overall group performance and group ranks. Moreover, members can communicate with group members through online group pages as well as initiate offline group meetups events or social activities, and update information and pictures of these events. We choose this app based on the following reason. First, by the end of 2016, this app had owned about 80,000,000 active users. This large user base can provide us sufficient observations and guarantee the sample representativeness and external validity of our research. Second, its group-level socialization and gamification design provide us an ideal setting to examine the effects of group-level social interaction and comparison. Third, this app is among the first to allow and promote users' offline social activities with online community participation. This offline feature provides us an appropriate setting to examine the role and importance of offline social activities and its potential influence on the consequences of social comparison.

Figure 2. Research Context Snapshots



From the users' perspective, most users select and join a virtual group on the online hobby community platforms and apps to seek information, emotional, and psychological supports. More specifically, for example, on weight-loss or body-building communities, users, based on their objectives, can join a group and seek information for specific topics or successful experiences shared by other professional group members. Besides the drive for seeking information, users can also receive encouragement and empathy through communicating and socially connecting with other group members with similar goals. Moreover, as a user achieving goals with and being companioned by a group of like-minded people, he/she will also build social bonding and a sense of belongings with such a group. Joining and growing with a group can provide users the various type of social support, logically can secure users' objective to adopt these platforms or apps, and improve user participation and retention. Subsequently, developers bringing the group concept into the app design and enabling all essential function and social interaction on the group-level rather than on the individual-level mainly with the aim of, further motivating, encouraging and improving users' exercise frequency and participation through interpersonal comparison, interaction, and bonding among group members and competition between different groups.

The group's identity can logically distinguish users' social comparison and behaviors and, eventually, their physical activities in and out of the group. We depict an overall conceptual framework based on the social comparison theory in Figure 1. We can expect that both within and between-group social comparisons can have a significant impact on group members' exercise participation. However, it is unclear how offline social activities can moderate the above social comparison effects. In the section of literature review and hypotheses development, we will build this framework with more detail and rigorous reasoning.

2. 2 Social Comparison Theory

When users exercise with a group, they can observe and compare their running records with those of their group members, the overall group performance, as well as of other groups. Naturally, users involve mentally accounting the difference of their performance with those of others and form a social comparison process. We will use the social comparison theory as the core foundation for our logic reasoning and mechanism explanation. Festinger 1954 first proposed the social comparison theory, and its central premise is that people recurrently rely on similar others' ability, performance, and opinion to evaluate and define their own. The observable information of similar others (comparison targets) is perceived as a precise and stable basis for comparison and thus are considered as benchmark or references. People also tend to compare themselves with people who have similar attributes of interests, though there might be differences in other aspects (Goethals and Darley 1977). Sharp et al. 2011 demonstrated that especially when people are under high stress or competition, novelty or change, they are inclined to make comparisons. Thereby, being exposed to leaderboard information or being informed with the change of group performance rank can foster social comparison and its perception and behavioral consequences. With the exposure of others' higher or lower performance, people will form upward or downward comparison, and according to many classic research studies (Collins 1996, Wood 1989, Gordijin & Stapel 2006), both these two direction comparisons may call self-evaluation appraisal with positive, proactive coping behaviors and improved performance or the deterioration of self-esteem with negative, passive behaviors and impaired performance. As Hogg 2006's review, most of the classic psychological studies only conceptualized the comparison consequences under the lab experiment conditions and have not considered real-case environmental factors, such as competition level, comparison dimension relevance, and group identification, etc. There might be more covariates underlying the comparison process and shape the direction of the comparison consequences. In our research setting, with access to real cases and observations, we expect to reveal the mechanisms of the direction and consequences of social comparison by considering more scenario-specific factors.

Researchers summarized two cognition processes of social comparison- assimilation and contrast. In short, assimilation is the process when one perceives him/herself as similar to others and contrasts as different from others (Biernat at al 1997). Collin 1996, Buunk et al. 2001 and Carmona et al. 2006 demonstrated that given comparison dimension relevance and similarity between a person and his/her target, upward comparison can generate the contrast (upward contrast) and the assimilation (upward assimilation) process and downward comparison is inclined to generate contrast process (downward contrast). There is no convincing evidence that there exists downward assimilation. In our context, we can use records of top runners, the record of group organizer, belonged average group record, and other groups' records as comparison targets for group members. However, without detail explanations in the above literature, it is unclear that when there are both upward and downward targets, which process will be constructed or which process will dominate the other. Moreover, it is still unclear that how comparison dimension relevance and interpersonal similarity change with context and how the dynamics of these two determinants will shape the cognition processes. With the context of interactive groups, we expect to uncover further how in and out-group membership identification can change the above two determinants and the consequences of the two processes.

Besides the comparison cognition process, we use records of top runners, the record of group organizer, groups' average performance as our primary covariates mainly because these variables can be considered as optimal proximity to a standard. According to Garcia and Tor 2007, the social comparison will more likely to happen if comparison references are approximated to a standard, such as the top 1 ranking or the average score. Literature also listed several situational covariates that might affect the degree of social comparison, and based on the following literature, we accordingly select the following control variables for our later analysis. Garcia and Tor 2009 and Garcia 2010 verified that the intensity of social comparison is negatively associated with the comparison size. To control the decreasing scaling effect, we will incorporate group size as a control variable in our analysis. Analogous to social comparison, temporal comparison logically has a similar potential effect that, as the time length of group creation increases, the consequences of temporal comparison are inclined to dilute. Thereby, our later analysis will also include group age. Moreover, according to the local dominance effect in Zell and Alicke 2010, the consequences of social comparison are more salient when the comparisons are localized. For example, if a person want to assess the body fat, he/her might compare the fat ratio with that of his/her one or a group of good friends, colleagues, or even the mean ratio of people living in the same town. To control the potential local dominance effect, we will incorporate weather index and the shortest distance between a group's creation location to its nearest sports ground.

2. 3 Online Individual Social Interactions on App Engagement

Besides the social comparison theory literature as the foundation of our framework, empirical research of organizational and fitness application studies based on individual-level behaviors also provide us some insights. Most of the fitness app studies built their theoretical mechanism on the individual social comparison and observations learning, and social network analysis, and reached the consistent conclusion that social interaction among individual users can improve their activity levels and the overall health behaviors. Mainly built on the identification-contrast model by Buunk et al. 2013 and the model's criticism by Brown et al. 2017, studies like Wu et al. 2015 and Zhou et al. 2016 demonstrated that the competitive environment, such as being exposed to the public leaderboard or self-set goals, can amplify users' contrast processes, subsequent exercise behaviors or their adaptive defensive coping behaviors. More specifically, through structural equational modeling, the author verified that the perceived competitive climate could strengthen the social comparison's consequences in terms of user's view about exercising behaviors expedited by the running app, and the willingness of exercising behaviors. Under the competitive environment, the contrast process might dominate the assimilation process of individuals. However, it is still possible that users might consider the upward targets as a role-model or someone they would like to become, and it is also possible that when observing upward target on the leaderboard, users can form an upward assimilation process and response positively. Fitness app studies based on social network literature showed that observing or being exposed to other users' performance, users might will have observational learning behaviors but might not arouse positive or significant associations. For example, Aral and Nicolaides 2017 incorporated exogenous environment factors as instruments to detect the social contagion effects on users of a running app. They found that exercise behavior is socially contagious, and less active users can affect more active users but not reverse. This means that observing the down targets will have more salient social comparison effects than observing the upward targets. This could be explained by a downward assimilation process among individual users. However, the downward assimilation is considered as a non-evident scenario in the previous social comparison studies. How we capture the environmental factors and how the leaderboard gamification is designed by the app developers might affect the twisting cognition process caused by either upward or downward comparison among individual users. In our study, given the within-group upward target only setting and the between-group bi-directional target setting, we expect to clarify users' cognition process and subsequent behaviors due to different directional comparisons. Besides the mixed finding in the IS research, psychology works like Wood et al. 2018 revealed that social influence effects do not sustain. Motivated such, these twisting findings call for our thorough reasoning and analysis regarding users' social comparison effects and their subsequent behaviors. Moreover, most of the previous fitness app studies only focus on individual-level rather than the group level comparison. However, joining and interact within and between-group can form users' perceived categorization and the identification process, and may affect users' comparison dimension relevance and interpersonal similarities, which will shape the comparison direction and the underlying cognition processes and consequences. Our research setting naturally classifies users' social comparison into two categories- the within-group comparison and between-group comparison. In the hypotheses development section, we will reason the underlying mechanisms regarding the direction, cognitive process and subsequent behavior under within and between-group comparisons to examine our research questions and to meet the above research gaps.

3. Hypotheses Development

3. 1 Within-Group Social Comparison

The central idea of the social comparison theory (Festinger 1954) is that the underlying drive for people to observe and compare with similar others is to more accurately evaluate, maintain and improve self-esteem, ability, and performance. Meanwhile, the similarity in terms of individual characteristics and ability and the relevance of performance dimension (Goethals and Darley 1977) are two essential determinants for interpersonal social comparison. In the context of fitness applications, users' general objectives to adopt such apps and to join and exercise with a group are to improve self-motivation, exercise frequency and physical ability, and the overall health condition. Correspondingly, having a better running record or a competitive participation rate can be considered to be a primary objective for most fitness app users. A higher running record or participation rate will be the most relevant performance dimension for all users, and this common use purpose can be at least one of the similarity for users. Naturally, with these two determinants, we can expect either upwards or downwards social comparisons existing among fitness app users. Further, as the consequence of either upwards or downwards comparison (Collin et al. 1996, Manis et al. 1997), there will occur an upwards or downwards assimilation process if there exists a higher level of perceived similarity between a person and his/her comparison reference and an upwards or downwards contrast process if there exist a higher level of perceived differences. Social comparison research works (Collins et al. 1990, Manis et al. 1997, Wheeler and Suls 2007) have reached to the consensus that for the assimilation process, an upward reference can induce more positive behavioral responses than a downward reference, while for the construct process, a downwards reference can induce more positive behavioral responses than an upward reference. Moreover, Groothof et al. 2007 demonstrated that perceived similarity could form a group membership identification process, which can further encourage assimilation responses such as enhancing performance and proactive coping behaviors. Intuitively, when a group member thinks he/she is similar to the top runner or a reference to which he/she aspires to become, he/she will work out harder or follow the references' routine and record. As a consequence, the person will exercise more often and harder.

In our research context, in order to encourage users to aim higher and exercise more, the app developers enable users to select and join a group, exercise and attend offline social activities with group members. Moreover, the records of top 3 performance members as public to other group members on the leaderboard, yet lower rank records are hidden from observation. This design generates an upward comparison only scenario, which allows us to examine whether the assimilation or the contrast process overweighs the other. The group-joining is a user decision. Thus the perceived similarity between a user and his/her joined group is logically higher than that between him/her and the un-joined group. After joining and exercising with the group, the perceived similarity will be augmented through online and offline interaction with group members and gradually amplify shared distinctiveness and foster the group identification process (Mussweiler and Strack 2000, Groothof et al. 2007). The similarity is supposed to be much larger than the difference among group members. Therefore, we expect that the within-group social comparison will form an assimilation process. That is, as being exposed top records, group members perceive them and tend to evaluate themselves to be similar to upward targets in the group and will act positively toward the targets they aspire to be (Collins 1996 and Wood 1989). Subsequently, the top members' records, serve to be the upward comparison references and can logically evoke group members' proactive behaviors and positive coping responses such as running longer and more often.

Mainly through survey and lab experiments, findings of organization and sports studies support the positive consequences of upward social comparison. Brown, Ferris Heller, and Keeping 2007, found that given the definiteness of members' core self-evaluations, upward social comparisons in the workplace can result in workers' higher affective commitment and job satisfaction. Athlete role model studies, such as Lockwood and Kunda 1997 found that the performance and records of athlete role models can encourage their peers and rouse their self-enhancement, inspiration, and improved performance when the success cases and top records seem achievable.

Combining the logical reasoning based on the social comparison theory, related empirical and experimental studies, as well as our research context, we consider the top-performance records as desired benchmarks for other group members. Top-performance members can better define a group's behavioral norm and group membership distinctiveness by promoting upward assimilation, which can drive peers to generate positive coping responses. More specifically, the top performance members, similar to the athlete role models, can inspire and encourage their group peers to participate more in physical activities. Formally, we propose

H1a: The average record of the top 3 runners in a group can significantly improve the group members' exercise participation rate.

While top-performance members can serve to define the group behavioral norm functionally, group organizers, on the other hand, can also be considered as influencers among group members and contribute to socially defining a group's distinctiveness and promoting the group identification process through the assimilation process. In online hobby groups, the function of a group organizer is to serve, operate, and manage the virtual group, such as promoting online campaigns, maintaining the group page and online content, and holding offline meetup social activities or team building events. During daily group physical and social activities, group organizers have more opportunities to interact and bond with group members and most of the group organizer's running record is above the average of that of all the other group members. The social bonding and closeness (Tesser 1998, Miller et al. 1988) between the organizer and group members, can thus promote the upwards assimilation process and group members' positive coping responses. Additionally, organizers' behaviors are more accountable and have a higher chance and frequency to be exposed to and observed and followed by other members (Breukele et al. 2012; Singer, 1981). Therefore, theoretically, group organizers' records can also be an influential comparison reference and encourage their members to exercise more.

Mainly through surveys, organization and leadership studies such as Heckman and Morris 1975, Zaccaro et al. 2001, and Tagger and Ellis 2007 found that team leadership can contribute to define and clarify collective objectives by reducing cognitive conflict among team members. In a fitness app context, group members might have different motives to join a group. However, group organizers and their behaviors can remind group members to achieve the common objective, strengthen the group distinctiveness and identification, promote members' upward assimilation, and eventually encourage them to exercise more. Formally, we propose

H1b: Group organizer's activity records can significantly improve the group members' exercise participation rate.

3.2 Between-Group Social Comparison

A group's performance (based on the average running record of all its members) published on the group rank list is observable to its members and all the other groups. Given the assimilation within each group, group membership (Turner 1982), as a kind of social identification process, is inclined to influence group members' cohesion, loyalty, and the pride of belonging to the group. Given the group membership and shared distinctiveness, the individual value would convergence (Ellemers and Rink 2005) over time with the emergence of the group identity. Individual members then pay more attention to group task accomplishment and performance (Meeussen et al. 2013). Knippenberg et al. (2003) conceptualized this convergence as a process that group members' initial values merge to a collective value, which longitudinally predicts the group task performance. Subsequently, the performance record of each group can also be considered as comparison references for different group members. During the group identification process, while group members increasingly intensify their group uniqueness, they also progressively notice the difference between the performance of their belonged group and that of other groups. Besides to improve and maintain their self-esteem and performance from upward assimilation within a group, on the collective level, members of a group also lean to amplifying the between-group difference to maintain and improve their group esteem and positive distinctiveness (Mussweiler 2001, Groothof et al. 2007). The process of evaluating and perceiving oneself as different from others is conceptualized as the contrast process (Manis et al. 1997, Tesser 1988).

In our research context, the between-group comparison can be upward or downwards or both, and there we will explain by assuming three scenarios, the top performance group, the bottom performance group, and groups in the middle. For the top performance group, all the other groups are downward targets, and the top group members logically perceive their group distinctiveness and group esteem also as the top one. During the contrasting process, downwards comparison references can induce more positive responses than upward references, and upward references can prompt more negative responses than the downward references (Buunk 1990 et al. and Manis et al. 1997). More specifically, in our research context, in order to maintain their head esteem, positive distinctiveness, positive emotion, and collective honor, members will respond proactively and positively to enlarge the performance difference between theirs and the performance of other groups. For the bottom performance group, unfortunately, all the other groups are upwards targets and these group members might face either upward contrasting or upward assimilation process. Recalling conclusion of Buunk et al 1990 and Manis et al 1997, the upward reference for the contrasting process can prompt negative responses. For the bottom group members, their collective emotion and esteem can be deteriorated and passive, and their group distinctiveness can be detrimental. Maintaining such a negative distinctiveness can change their relevance of performance dimension (Tesser and Smith 1980). In other words, the bottom members may lose their interest in exercise and competition. In a more intuitive sense, members in a low-rank group might "throw the handle after the blade." If a bottom member no longer maintains their group membership and will approximate to better performance groups, he/she will leave the bottom group and form an assimilation process towards the group he/she aspires to join. For the groups in the middle, their members are in a reference- rich and dynamic social environment, which can foster group-level competition (Greenberg and LaPrelle 1985, Taylor and Lovel 1989). Aspinwall and Taylor 1993 emphasized the salience of the contrasting process in a competitive climate and demonstrated that the group's rank information could strengthen the group members' awareness and sensitivity of the competition level among different groups. Tajfel 1972, Turner 1986, Turner 1979 used the term group indemnity to conceptualize group identity value when there are multiple references. They further pointed out that group members evaluate and mentally account comparable attributes of their groups with these of other groups to calculate their group indemnity, and then members strive to maintain their current positive social identity based on favorable comparison or leave the current group if the social indemnity is not satisfactory. Theoretically, for groups in the middle rank, the downward contrasting process also overweighs the assimilation.

To summarize the above three scenarios, if a group has a relatively high group rank (smaller rank order), its group members would have a positive group indemnity and consider the high rank as the prestige or a favorable status. Members would exercise more to maintain that encouraging status or, in other words, they would have more motives to keep their high rank. On the other hand, if a group's rank is relatively lower, its members have a negative group indemnity and would more likely to leave the group or to have "throw the handle after the blade." Formally, we propose

H2a: Group rank is significantly and negatively associated with the group members' exercise participation.

Given a positive group indemnity or a favorable discrepancy between a group with its upward targets, if there happens a shrink of the discrepancy, the shrink will theoretically motivate the group members to exercise more. Studies regarding competitive climate and leaderboards on the individual level support our expectations. For example, Garcia et al. 2013, Wu et al. 2015 and Sepehr and Head 2011 found that environmental factors like the change of competitive climate can significantly affect users' self-evaluation and coping behaviors. Bunnk et al. 1990 and Mussweiler 2000 demonstrated that the individuals could have active coping and adaptive responses to minimize the discrepancy between outcomes of self-evaluation and that of targets. Therefore, combining the theoretical reasoning based on social comparison theory, findings of competitive climate studies, as well as our research context, we deduce that while group users exercising within their belonged group, between-group contrasting process can fortify users' attention to group performance and their awareness of the discrepancy and the change of the discrepancy between the performance of their group and that of other groups. More specifically, if members observe their group rank drops, members will exercise more to regain their positive distinctiveness and prestige and to restore their group rank. Formally, we propose

H2b: The change of group rank is significantly and positively associated with the group members' exercise participation.

3. 3 Offline Social Activities

Besides working out together with other group members or holding local exercise competitions, group members' offline meetup events also include casual, leisure, or entertainment team building activities such as potluck, games, outing or movie night, etc. Logically, the casual or even entertainment team building activities can form a lite and positive collective mood for group members, and these causal offline activities might be able to strengthen the assimilation process between group members by increasing interpersonal psychologically closeness (Tesser 1988), boding with comparison references (Miller et al. 1988), and common personality attributes (Wills 1991). Following this logic, we might expect a buttressing moderation effect of offline team building activities on the group members' positive coping responses due to the strengthened assimilation effect. However, social comparison extensional research such as Tesser and Smith 1980, Cialdini et al. 1976, and Tesser 1988 provided another perspective to examine the underpinning determinants of the comparison and identification process. That is when an individual is closed and bonding with an upward target, there could be a self-evaluation enhancement on the comparison dimension through reflection or assimilation or both. The reflection is also described by "basking in reflected glory," a lift in both emotion and perceived self-evaluation due to the built social capital or having a social relationship with an upward reference. In a more intuitive way, a group member might feel happy and confident due to getting to know and being friends with the top runner. Collins et al. 2000 and Wheeler and Suls 2007 revealed that the degree of comparison dimension relevance is positively associated with the salience of the assimilation process and its consequences, whereas the interpersonal closeness is positively associated with the salience of the reference and its consequences. In our context, offline meetup events and team-building social activities are not limited to running or exercise. When group members participate in these activities, they can have a higher likelihood to improve psychological and emotional closeness and interpersonal relationships and to build bonding and friendship between each other. Meanwhile, as attending these social activities, group members' comparison dimension relevance, such as to compete for the running distance or to compete for the top rank, will be diluted. Accordingly, based on the findings of Collins et al. 2000 and Wheeler and Suls 2007, the assimilation process will be dominated by the reference process. More specifically, if a group member starts to build a friendship with the top-performance runner, or the organizer during team building activities, the member will not consider the influencer as a competitor. Instead, this member may keep exercise with the aim of enjoying running with a group of people or improving his/her ability rather than to achieve and compete with others' records. Based on the above reasoning, we naturally expect that offline activities can improve the reference process and its consequences within a group.

Moreover, several organizational and human resources studies show support of the reference process through other perspectives. For example, Barsade and Gibson 1998 and Zaccaro et al. 2001 conceptualized the group level collective mood generated during social activities as the team emotion. They stated that the positive team emotion among group members could nurture more cooperation, less conflict, and stronger cohesion as well as the creativity of task accomplishment (Carnevale & Isen 1986 and George 1996). Teambuilding research such as Carron and Spink 1993 and Carron 1998 concluded that casual team building activities could reduce social distance and boosting social support and bonding among group members, and thus can relieve the interpersonal competitiveness and tense within groups. In our context, except for working out with local in-group friends, most time, when members use the fitness application, they log the records online without having the chance to meet and socialize with other members. Frequently having the actual offline meeting with online friends can be considered as a valuable opportunity for group members to foster their social lives and to build or strengthen the bonding and social support among each other.

Combining the theoretical reasoning based on social comparison theories and findings of the organization studies, we formally propose

H3a: The offline team building activities can significantly suppress the effects of within-group social comparison on group members' exercise participation.

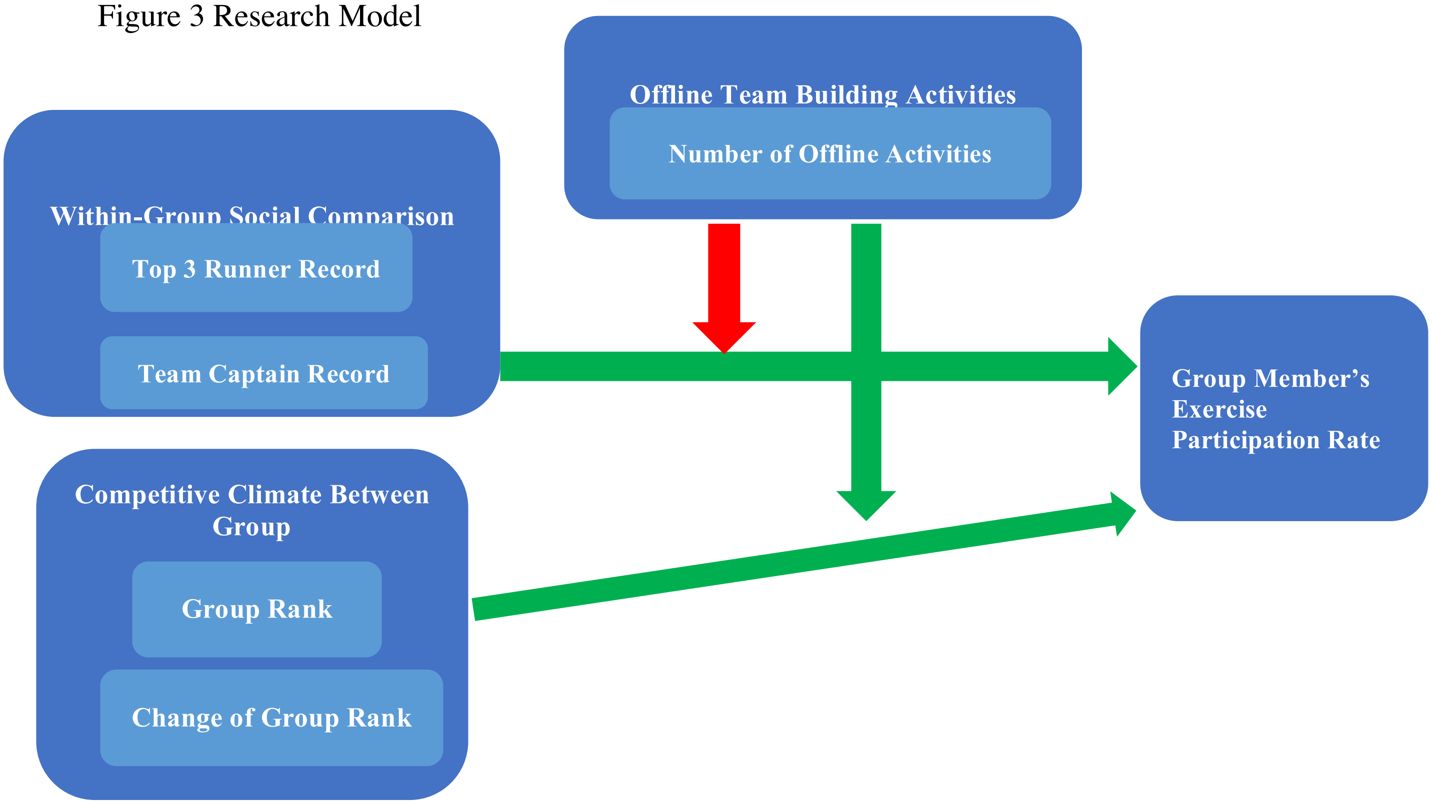
Regarding the influence of offline activities on the between-group comparison, according to the social comparison theories, as offline activities improving the psychological and emotional closeness, sense of belongings, and collective honor between group members, members will act positively to maintain their positive group identity and discrepancy with other groups. Offline meetup activities may suppress the in-group assimilation process, but the activities can promote group members to care more about their group performance, honor, and prestige. Thereby, following this logic, we can expect that offline activities can strengthen the contrast process and its consequences.

The findings of some organizational and social identity literature can support the above expectation. Tuckman 1965 proposed that team building activities can effectively enhance team member's sense of "togetherness" during the norming and performing stages of groups that are focusing on tasks. Accordingly, team members pay more attention to collective level objectives and performance. Widmeyer and Williams 1991 and Rideout and Richardson 1989 have come to similar conclusions. They revealed that teambuilding activities, especially for casual events, can improve group members' consensus to the high-performance reference and sense of responsibility for group outcomes, and sense of collective honor. Their findings show that offline activities could encourage group members to value more about their group performance and promote members to exercise more to keep or improve their group rank. Social identity literature uch as Turner 1982 and 1984 found that contextual factors can improve social identification, promote the construction of group membership, and result in improving group outcomes and performance. Ashforth and Mael 1989 further described the group membership construction as the process that while individual members gradually identity with the group, the collective value, norms, missions of the group they belong can be perceived as more positive, original, and distinctive compared with other groups. These positive collective perceptions and group identifications can consequently promote group members' commitment to maintain unique group culture, realize group objectives, and improve group performance. Hogg 2006 further specified the above generative identification process as a cognitive structure-- the consequences of "the meaning and value of social interaction and interdependence among group members." In other words, the contextual factors that promote interaction and interdependence among group members, can reinforce the group membership identification and improve the consequences of social identity. Under our research setting, offline team building activities allows social interaction among group members and can also nurture the interdependence among them and then help members to construct their unique group identity. This group identity cognition structure can eventually encourage members' commitment to group objectives and performance. Base on the above two perspectives from two streams of literature, we expect that team building activities can improve group members' sense of responsibility for group outcomes and the sense of collective honor. In order to keep the group honor, members will increase their exercise participation. Formally, we propose

H3b: The offline team building activities can significantly strengthen the effects of between-group social comparison on group members' exercise participation.

After summarizing the above hypotheses on the variable level, we re-structured our conceptual framework in Figure 3.

Figure 3 Research Model



4. Empirical Analysis

4. 1 Data, Measurement and Descriptive Statistics

Data: Since the app's API only provides timely records, to build longitudinal panel data, we manually observed and recorded group-level user's social and physical activity data from the running app every week, and the data collection time is for about a year. More specifically, the data sample spans 38 weeks from August 1, 2017, to April 29, 2018. We exclude groups whose size is less than ten members, and those were formed during our research time span. The sample includes 152 groups in Dalian, China.

Besides the high penetration level and the popularity of this fitness app, another reason we select this app is that running is one of the most affordable and beginner-friendly work-out. Runners can include professional athletes, amateurs, and beginners, and thus the user sample is naturally more representative. We select the city Dalian because it is listed as one of the most livable and populous four-season-city in China, and there is almost no extreme wealth in a year. The Dalian government holds various sport events every year, such as the coastline marathon and walking festival every year and the international marathon festival every May. Selecting group runners in the city Dalian as our observation, we can void potential confounding effects to some degree. We exclude groups whose size is less than ten members because the app does not rank these extremely small or newly created groups.

We summarize all variables of interests in Table 1. The dependent variable, $GPR_{i,t}$, Group members' exercise participation rate, is captured by the average number of unique group members who run at week t. We use the participation ration to describe group level activity rather than an absolute metric such as total distance or total participation mainly because groups are heterogeneous in running ability and habit, and group records like aggregated running distance can be confounded by the heterogeneity. Recalling the reasoning in the hypothesis development section, given the overwhelming group membership identification and similarity between group members, the upward assimilation will be the dominant cognitive process and members will more likely to response positively and proactively to their group star runner and other upward behavioral norms. The with-in group social comparison effect is the collective representation of desirable group conducts and behavioral norms, and aspired upward references. Here, social comparison are depicted by two metrics- the average record of the top-3 runners ($TOP3_{i,t-1}$) and the record of the group organizer ($CAP_{i,t-1}$). The group organizer is the one who form, manage and operate the group. He/she is usually not the top performance runner having the longest distance. The between-group social comparison effect is represented by the collective individual outcomes from a group to be contingent on the comparisons of their performance against that of the other group. The effects are thereby depicted by $RANK_{i,t-1}$ and ΔRANK. We measure the offline social activities by the number of offline social activities hold at t-1. Recalling the literature review section, in order to void potential confounding issue due to the decreasing scaling effects of social and temporal comparison and the local dominance effect, we will incorporate group age, group size, weather index and shortest distance as the control variables. The sport index is classified to 5 degrees and generally the lower the index, it is the safer and better to run outdoor. The descriptive statistics is summarized in Table 2.

Table 1. Metrics and Measurement (For group i at week t)

	Variable	Denotation	Description
Group members' exercise participation rate		$GPR_{i,t}$	$\mathrm{GPR}_{i,t} = \dfrac{\text{Number of unique members who run at week t}}{\text{Group size at t}}$
Within-Group Social comparison	Average of top 3 runners' records	$TOP3_{i,t-1}$	Average kilometers the top three members run at t-1

	Organizer's record	$CAP_{i,t-1}$	The kilometers the group organizer run at t-1
Between-Group Competitive Climate	The change of group rank	$\Delta RANK_{i,t-1}$	$\Delta RANK_{i,t-1} = RANK_{i,t-1} - RANK_{i,t-2}$
	Group rank	$RANK_{i,t-1}$	The rank of the group at t-1
Offline Team Building	Weekly number of offline social activities	$OfflineACT_{i,t-1}$	Number of offline social activities hold at t-1
Control variables	Weather	$Weather\_INDEX_t$	An index variable to show whether it is unsafe and not good for out-door exercises, e.g. very high/low temperature, wind force and extreme weather
	Group age	$GroupAge_{i,t}$	Number of weeks since the group was formed till t

Table 2. Data Description

Variable	Min	Max	Mean	s.e.
$GPR_{i,t}$	0	70.14%	5.59%	6.81%
$TOP3_{i,t-1}$	0	225.517	28.84	30.47
$CAP_{i,t-1}$	0	160.36	8.06	20.28
$RANK_{i,t-1}$	1	186	80	47.22
$\Delta RANK_{i,t-1}$	-150	155	0.82	27.10
$OfflineACT_{i,t-1}$	0	3	0.01	0.14
$Weather\_INDEX_t$	6.714	16.857	11.84	3.31
$GroupAge_{i,t}$	2.86 (weeks)	207.71	111.82	54.18
$GroupSize_i$	5	1316	72	131
$ShortestDistance_i$	0.05 (km)	9311	76.64	752.16

4. 2 Identification Strategy and Results

4. 2. 1 Baseline Model- The Fixed Effects

According to the result of the Hausman test, the Chi^2 is 80.61(p=0.000<0.05), there is sufficient evidence to reject the null that there is random effect. Thus, we use the fixed effect model as our base model, which estimates the impacts of the within-group social and between-group comparison effect on the group members' exercise participation rate.

The baseline model specification to examine the direct impacts is:

$GPR_{it} = \alpha_0 + \alpha_1 TOP3_{i,t-1} + \alpha_2 CAP_{i,t-1} + \alpha_3 RANK_{i,t-1} + \alpha_4 \Delta RANK_{i,t-1} + \varphi_B Control\ Variables + \gamma_i + \omega_{i,t}$, where α_0 is the constant, γ_i is the unobserved time-invariant individual effect and $\omega_{i,t}$ is the error term.

We summarize the results of the fixed effect analysis in Table 3. One unit change of top 3 runners ‘records, organizer’s record, and change of group rank can improve 3.5%, 1%, and 0.9% of group members’ exercise participation rate, respectively. For within-group social comparison effects, both the top-3 runners' record and the organizer's running records improve the group members’ exercise participation rate. Thus, both hypotheses 1a and 1b are supported. For the between-group social comparison effects, the coefficient of group rank is significantly negative and indicates that the higher rank groups are more inclined to improve their exercise participate than the lower rank groups. Meanwhile, the change of group rank significantly improves the group members’ exercise participation rate. The weather index, the group size and the group age have a negative impact on group members’ exercise participation rate. The negative consequence of weather intuitively indicates that higher sport- index or less safe and comfortable weather is associated with less activity participation, and younger groups tend to be more exercise active. This also verifies the local dominance effect (Zell and Alicke 2010). Meanwhile, the negative influences of group age and group size validates decreasing scaling effects of social and temporal compassions.

To sum up, the baseline analysis results well support H1a, H1b, H2a, and H2b, and they are consistent with most of previous organizational or fitness app IS research. As the consequences of upwards comparison and assimilation, the top runner can be considered as an athlete role model by other group members and thus, the performance record of the top runners can encourage and simulate other members to exercise more. As a result, group level exercise participation will be eventually improved. Organizers serve the group, participate in more group events, and their name and exercise performance will be more often exposed to other group members. Their behaviors are also saliently influential and subsequently have significantly impacts on group level exercise participation. A noticeable fact is that the average running records of organizers is 8.057 kilometers per week. Compared with the that of top performance members, the 8.057 distance is a relatively more accessible and achievable goal and will theoretically have stronger improvement in the upward assimilation given the higher comparison relevance. However, the coefficient of group organizer is 71.43% lower than the coefficient of top-3 runners. Recalling from the research context and literature section, given the social role and function of organizers, group members have more chance to socially interact with organizers through online and offline communication. These social interactions might make members' comparison dimension relevance blur and slighter while bonding and interdependence higher. The reflect process might overwhelm the assimilation process. We can have a clearer view of the reflection process in the moderation analysis. Yet, eventually, the consequences of mixed reflection-assimilation process dominant the consequences of contrast process. It is evident to claim that organizer's records have a significant boost in the overall group exercise participation.

Table 3. Fixed effect analysis results … *** p<0.001, ** p<0.01, * p<0.05

Variables		Coef	s.e.	p-value
Within-Group Social Comparison	$TOP3_{i,t-1}$	0.035***	0.003	0.000
	$CAP_{i,t-1}$	0.010***	0.003	0.007
Between-Group Social Comparison	$RANK_{i,t-1}$	-0.018***	0.002	0.000
	$\Delta RANK_{i,t-1}$	0.009***	0.001	0.000
Control Variables	$Weather\_INDEX_t$	-0.248***	0.012	0.000
	$GroupAge_{i,t}$	-0.011***	0.004	0.007
	$GroupSize_{i,t}$	-0.009*	0.004	0.005
	$ShortestDistance_{i,t}$	-0.001	0.002	0.617

4. 2. 2 Moderation Analysis

To examine the moderating effect of group's team building on the impacts of within-group and between-group social comparison on the members' exercise participation rate, we incorporate the moderating relationships in the fixed effect model and the reform the specification to:

$$GPR_{it} = \beta_0 + \beta_1 TOP3_{i,t-1} + \beta_2 CAP_{i,t-1} + \beta_3 RANK_{i,t-1} + \beta_4 \Delta RANK_{i,t-1} + \beta_5 OfflineACT_{i,t-1} + \beta_6 OfflineACT_{i,t-1} * TOP3_{i,t-1} + \beta_7 OfflineACT_{i,t-1} * CAP_{i,t-1} + \beta_8 OfflineACT_{i,t-1} * RANK_{i,t-1} + \beta_9 OfflineACT_{i,t-1} * \Delta RANK_{i,t-1} + \varphi_B Control\, Variables + \mu_i + \varepsilon_{i,t}$$

where β_0 is the constant, μ_i is the unobserved time-invariant individual effect and $\varepsilon_{i,t}$ is the error term.

We summarize the moderation analysis results in the Table 4. The moderation works oppositely on two direct impacts such that the interactions between offline activities and between-group social comparison are significant and works oppositely on within-group and between group social comparisons. More specifically, $OfflineACT_{i,t-1}$ has a significant positive impact (2.949***) on the group level exercise participation, and it has a significant positively moderation impact (-0.045** and 0.047**) on the relationship between the between-group social comparison and the group level exercise participation. H3b are well supported. The interactions between offline team social activities and within-group social comparison are not all significant. H3a is partially supported. These distinct interaction impacts of offline social activities are consistent with findings of previous organization and teambuilding research (Carron 1998, and Widmeyer and Williams 1991). According to the teambuilding framework proposed by Tuckman 1995, casual social activities and team building activities can improve the sense of cohesion, social ties, and bonding between group members. Under our research setting- Fitness apps, most of the offline teambuilding activities are not limited to exercise or fitness meetups. Teambuilding can also be social, casual and entertainment events. Through the meetup social parties and casual social interactions, the individual level workout competition can be diminished by bonding and social support, or even friendships between group members. Friends' running records may not necessarily be a target to be compared with and to be broken through, and group members will lose to the motive to compete with each other as the intimacy grows. Moreover, there is an interesting fact that the positive moderation on the impact of group organizers' records is much more salient than that of top runners' record. Recalling the main analysis result, the social comparison effect of group organizer's lower than that of top runners. The moderation distinction and the lower main effect of group organizer can all be the result of different level of reflection process. Again, since the primary role of group organizer is to manage and operate the group and thus have more likelihood to interact and bond with other group members. Organizers, compared with top runners, contribute to define the group's social norm more likely. The social nature of the duty of group organizers foster the emotional similarity between organizer and members and the basking in glory (reflection process), whereas the social nature of organizers' operations and administration can dilute the strength of the upward assimilation process. Therefore, the offline social activities, especially the causal events, can strengthen the consequences of the reflection process, which is more salient on the main effect of social comparison when members compare with organizers.

On the other hand, as the cohesion improves though offline team building activities, members focus more on group performance and have a stronger sense of having responsibility for group task outcomes (Widmeyer and Williams 1991). Due to the awareness of the collective honor, sense of belongings and positive group identification, there will be incentives to exercise and participate more to keep the group's high rank or to sustain a positive prestige. Thus, the interaction between offline activities and between-group social comparison are salient. Besides the interesting findings of moderation effects, the main effects of within-group and between-group social comparison and the influences of control variables on group activity participation are all consistent with the results of the baseline model.

Table 4. Offline social activities moderation with fixed effect analysis results *** p<0.001, ** p<0.01, * p<0.05				
Variable		**Coef**	**s.e.**	**p-value**
Within-Group Social-Comparison	$TOP3_{i,t-1}$	0.036***	0.003	0.000
	$CAP_{i,t-1}$	0.329***	0.047	0.000
Between-Group Social-Comparison	$RANK_{i,t-1}$	-0.015***	0.02	0.000
	$\Delta RANK_{i,t-1}$	0.007***	0.002	0.000
Offline Team Building	$OfflineACT_{i,t-1}$	3.298***	1.086	0.002
Moderators	$OfflineACT_{i,t-1} \times TOP3_{i,t-1}$	-0.008	0.008	0.310
	$OfflineACT_{i,t-1} \times CAP_{i,t-1}$	-0.385**	0.185	0.038
	$OfflineACT_{i,t-1} \times RANK_{i,t-1}$	-0.045**	0.019	0.021
	$OfflineACT_{i,t-1} \times \Delta\mathrm{RANK}$	0.047**	0.022	0.036
Control variables	$Weather\_INDEX_t$	-0.236***	0.0134	0.000
	$GroupAge_{i,t}$	-0.014***	0.004	0.001
	$GroupSize_{i,t}$	-0.011**	0.004	0.019
	$ShortestDistance_{i,t}$	-0.0005***	0.0001	0.000

4. 3 Validation and Robustness Check

We conduct several additional tests to validate the robustness of the results.

4. 3. 1Reversal Causality Check through the Panel VAR Granger Causality Test

If there is reversal causality, the group activity participation rate would dynamically impact on the within-group and between-group social comparisons. To avoid the threat of reversal causality to the internal validity of our empirical findings, we conduct the Granger causality test procedure in the Panel Vector Auto-Regression analysis. Table 5 shows the Granger test results for the baseline model. Based on accumulated impacts between covariates, the results illustrate that both within and between group social comparisons significantly Granger cause the group activity participation rate, and meanwhile, there is no significant reversal Granger causality. Through this test, we can exclude the threat of reversal causality to the internal validity of the results of our baseline model. Thereby, the findings and support for the first two hypotheses are validated. Similarly, we conduct the PVAR Granger test for the moderation analysis findings by entering the moderators in the PVAR estimation, and test results are summarized in Table 6. The interactions between offline team building activities and between and within-group social comparisons significantly Granger cause group activity participation, and there is no reversal causality either. This test also strengthens the validity of the findings of the moderation analysis.

4. 3. 2 Alternative Measure of Offline Team Building Activates

We collect the number of aggregated number of pictures of offline activities from another panel of the fitness app. Besides initiating and recording offline meetup events, the app also allows users to upload pictures later on. Only recent event's information (the weekly number of offline meetup events) is shown on the group board, yet, all past event pictures are listed on the board. More specifically, given our research setting, the number of offline activities listed on the group information board will be refreshed by the fitness app weekly and any historical offline activities earlier than a week ago are not accessible to group members. Thereby, we consider the number of offline activities as the short-term measure of offline social activities. Since all previous meetup event photos can be accessible and observable by all group members, we consider the total number of activities pictures as a longer-term measurement and representation of offline social activities than the number of offline activities.

Table 5. Panel VAR-Granger Causality Wald Test-Baseline Model Findings_ Chi2 (Significant Level, SC= Social Comparison)

Ho: Excluded variable does not Granger-cause Equation variable	
Ha: Excluded variable Granger-causes Equation variable	
Causality Validation	Reverse Effect Check

Response to	$GPR_{i,t}$		Response to	Within-Group SC		Between-G	
Within-Group SC	$TOP3_{i,t-1}$	5.124**					
	$CAP_{i,t-1}$	2.994*	$GPR_{i,t}$	$TOP3_{i,t-1}$	$CAP_{i,t-1}$	$RANK_{i,t-1}$	
Between-Group SC	$RANK_{i,t-1}$	24.28***		0.906	2.11	0.723	
	$\Delta RANK_{i,t-1}$	4.627**					

Table 6. Panel VAR-Granger Causality Wald Test-Moderation Model Findings (Chi2 (Significant Level))

Causality Validation			Reverse Effect Check				
Response to		$GPR_{i,t}$	Response to	Moderation on Within-Group SC		Moderation on Between-Group SC	
Moderation on Within-Group SC	$OfflineACT_{i,t-1}\times TOP3_{i,t-1}$	0.313 *					
	$OfflineACT_{i,t-1}\times CAP_{i,t-1}$	4.885 **	$GPR_{i,t}$	$OfflineACT_{i,t-1}\times TOP3_{i,t-1}$	$OfflineACT_{i,t-1}\times CAP_{i,t-1}$	$OfflineACT_{i,t-1}\times RANK_{i,t-1}$	$OfflineACT\times\Delta$RANK
Moderation on Between-Group SC	$OfflineACT_{i,t-1}\times RANK_{i,t-1}$	6.326 *		0.521	0.137	1.055	0.584
	$OfflineACT_{i,t-1}\times$ ΔRANK	0.963					

We expect that this longer-term offline social activities measure (in terms of the total number of activities pictures), denoted as $PIC_{i,t}$ for group i at time t, has moderating impacts, which would be consistent with that of $OfflineACT_{i,t-1}$ That is $PIC_{i,t}$ can significantly moderate the effects of within-group social comparison and the effects of between-group competitive climates on the group activity performance. We estimate the moderation model by using $OfflineACT_{i,t-1}$and summarize the results in Table 7.

We find the long-term measure of offline team working can significantly moderate the impacts of between-group competitive climates on group exercise participation. However, the moderation effects on the relationship between within-group social comparison and the group exercise participation are non- significant. These results are consistent with the moderation effects of the short-term measure of offline team building and again support H3b. The main effects of within-group social comparison and between-group competitive climate are consistent with that of the results of using the number of offline activities. Thereby, hypothesis one and two are supported by the results alternative measure as well.

Table 7. Offline Team Building Moderation With Fixed Effect Analysis Results-Alternative Measure
*** p<0.001, ** p<0.01, * p<0.05 (Unit=Percentage)

Variable		Coef	s.e.	p-value
Within-Group Social-Comparison	$TOP3_{i,t-1}$	0.019***	0.003	0.000
	$CAP_{i,t-1}$	0.169***	0.046	0.000
Between-Group Social-Comparison	$RANK_{i,t-1}$	-0.005**	0.002	0.012
	$\Delta RANK_{i,t-1}$	0.005***	0.001	0.000
Offline Social Activities	$PIC_{i,t}$	0.364**	0.158	0.022
Moderators	$PIC_{i,t}$*$TOP3_{i,t-1}$	0.001	0.002	0.528
	$PIC_{i,t}$*$CAP_{i,t-1}$	-0.001	0.001	0.683
	$PIC_{i,t}$ * $RANK_{i,t-1}$	-0.004***	0.001	0.003
	$PIC_{i,t}$ * $\Delta RANK_{i,t-1}$	0.004**	0.001	0.019
Controls	$Weather\_INDEX_t$	-23.9***	0.013	0.000
	$GroupAge_{i,t}$	-0.2***	0.001	0.000
	$GroupSize_i$	-0.006	0.005	0.210
	$ShortestDistance_i$	-0.001	0.001	0.233

4. 4 Additional Analysis

4. 4. 1Dynamic Effects Analysis through Panel VAR IRF and FEVD

To study the dynamics association between within-group social comparison, between-group competitive climate and the group activity participation, we conduct a comprehensive Panel VAR analysis. The Panel VAR specification for the baseline model is:$Y_{it} = \sum_{k=1}^{p} A_k\, Y_{it-k} + BX_{it} + e_{it}$, where $Y_{it} = \begin{pmatrix} GPR_{i,t} \\ TOP3_{i,t} \\ CAP_{i,t} \\ RANK_{i,t} \\ \Delta RANK_{i,t} \end{pmatrix}$, $X_{it} = \begin{pmatrix} Weather\_INDEX_t \\ GroupAge_{i,t} \\ GroupSize_i \\ ShortestDistance_i \end{pmatrix}$, $i \in \{1,2, \dots N\}$, $t \in \{1,2, \dots T\}$.

N is the number of groups and T is the time span. $\boldsymbol{Y_{it}}$ is a vector of endogenous covariates. $\boldsymbol{X_{it}}$ is a vector of control variables. The matrix A and matrix B are parameters to be estimated through Panel VAR model. Here the idiosyncratic errors e_{it} are assumed to have stable trend s.t. $E[e_{it}] = 0$, $E[e_{it}' e_{it}] = \sum$, and $E[e_{it}' e_{is}] = 0,\ \forall\, t \neq s.$ Based on Love (2015), Hayakawa (2015), Canova & Ciccarelli (2013), and Holtz-Eakin et al. (1988), we take Helmert transformation for all covariates to avoid heteroscedasticity and to remove scale effect and following these PVAR analysis, we have the following procedure: 1). Pre-estimation unit root test for stability check. 2). Select optimal lag terms. 3) GMM estimation. 4). Eigenvalue test for IRF estimates stability check. 5). Granger causality test for causal and reverse effect. 6). Impulse response function (IRF). 7). Forecast error variance decomposition (FEVD). All covariates pass the ADF test and based on the minimized information criteria (MQIC*= -63.526 and J-Statistics*= 47.025) and the maximal determination coefficient (0.999), the optimal highest order is 2. After checking causality validity and reverse effects through the Granger test and given the parameters 'estimates (will be provided if requested) for all lag terms, we generate the impulse response functions (IRF). The impulse response function Φ_i can be captured by the re-shaping the reduced form Panel VAR model to infinite vector moving average form and the VMA parameters $\Theta_i = \begin{cases} I_k & , i = 0 \\ \sum_{j=1}^{i} \varphi_{t-j} A_j , & i = 1,2 \ldots \end{cases}$. The IRF with estimated coefficients can gauge the net effects of one unit unexpected change in between/ within group social comparisons metrics *i* on $GPR_{i,t}$ at time t. Standard errors are generated by Monte Carlo simulation with 500 repetition and coefficients' significance is tested by 0.95 confidence interval. We summarize the combined IRF graphs of key covariates in Figure 4. Based on innovation simulation, IRF illustrates how many unit y-axis metric changes given one unit change of x axis metric. The dotted lines form a cone, which represents the significance level. Here, the range includes two standard deviation. In Table 8, we summarize the immediate predictive elasticity and the accumulated value that combines all significant effects across the forecasting periods from each IRF.

Figure 4. Impulse Responses for Panel VAR

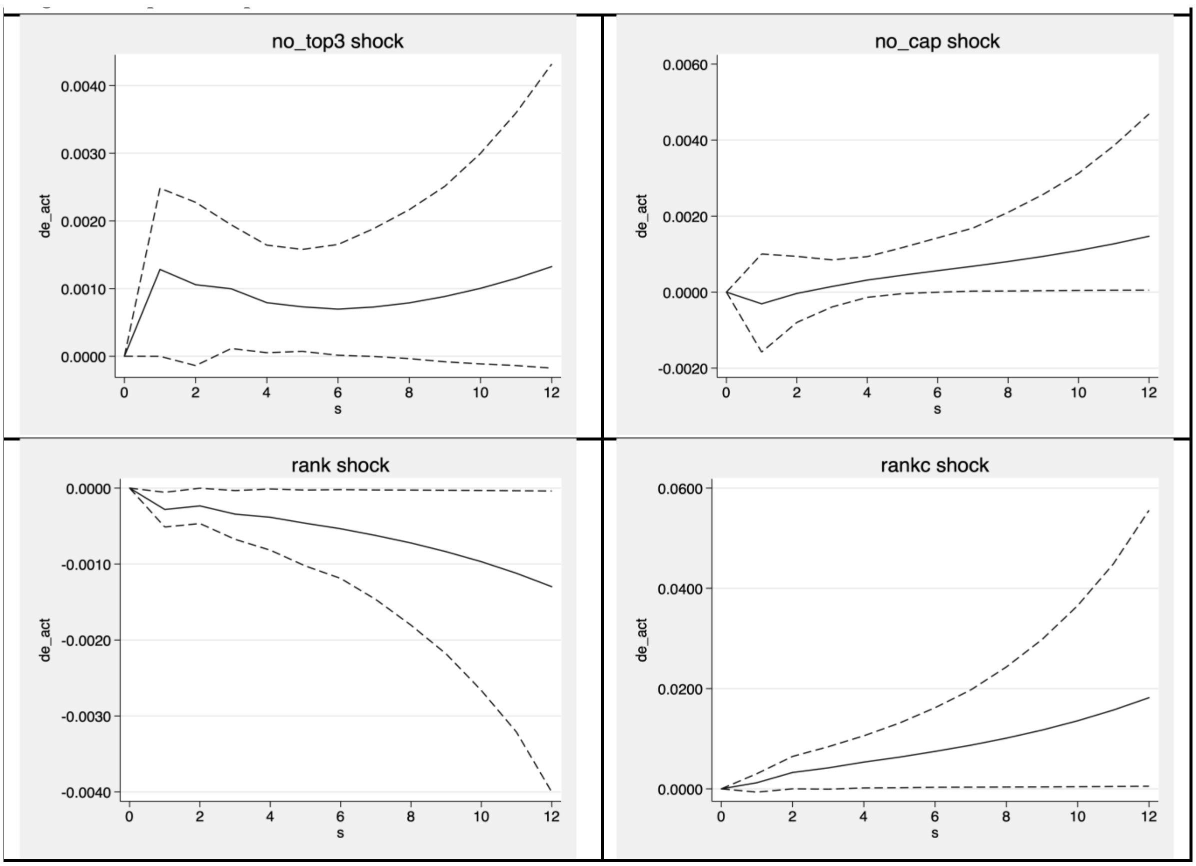


Table 8. Summary of IRF patterns (Unit=Percentage)			
	Group Activity Participation Rate: $GPR_{i,t}$		
		Immediate	Accumulated
Within-Group Social-Comparison	$TOP3_{i,t-1}$	0.13*	1.15**
	$CAP_{i,t-1}$	-0.03	0.74*
Between-Group Social-Comparison	$RANK_{i,t-1}$	-0.03*	-0.77**
	$\Delta RANK_{i,t-1}$	0.12	10.58***

The IRF results illustrate that top runners' performance records and the drop of group rank weakly improve the immediate value of group activity participation rate by 0.0013 and 0.003 units. While there are few significant and immediate impacts, all of within-group and between-group social comparison metrics have significant predictive relationships with group activity participation rate. More specifically, one unit $TOP3_{i,t}$, $CAP_{i,t}$, $RANK_{i,t}$ and ΔRANK (drop) can increase 0.0115, 0.0074, 0.0077 and 0.1058 units of group activity participation respectively. Based on the estimated PVAR parameters, we derive the forecast- error- variance-decomposition (FEVD) estimates to isolate the contribution of within and between group social comparison metrics. According to Abrigo and Love 2015, the h-step ahead forecast-error is described as $GPR_{i,t+h} - E\left[GPR_{i,t+h}\right] = \sum_{i=0}^{h-1} e_{i,(t+h-i)}\Theta_i$. After summating the normalized relative to the 10-period ahead FEVDs, the relative predictive values of $TOP3_{i,t}$, $CAP_{i,t}$, $RANK_{i,t}$ and ΔRANK (drop) are 0.273%, 0.012%, 2.137% and 0.019% respectively. Patterns in the results of PVAR analysis are consistent the findings of the fixed effect estimation. Besides, providing dynamic impacts of covariates, PVAR also further validate our empirical results in previous sections and further support our hypotheses about the main effects between group-level social comparisons and group participation.

4. 4. 2 Heterogeneous Effects of Top three members and Different Weathers

We enter the performance records of top 1, 2 and 3 members into the baseline fixed effect model and find that the impacts of first and second top performance members on group activity participation are not significant (0.067 and 0.079). Yet, the impact of the third top performance member is significant (0.157**). According to the descriptive statistics of top 3 members' performance, the average running distance of top 1 member is 44.16 kilometers per week, top 2 is 22.79 and top 3 is 17.46. Through pair comparison, we find the top 3 member's performance is significant lower than top 2's (p=0.012). Only the top 3 member's performance has significant impact on group participation rate may because that the athletic ability of the top-3 member is considered to be more related to that of average group members and thus top-3 members' records are considered to be relatively more comparable and reachable upwards role models. Additionally, by collecting more detailed weather data from the fitness app, we separate the $Wealther\_INDEX_t$ to four weather indicators such as high temperature, low temperature, wind force and extreme weather, and then we enter these weather indicators to the baseline model. We find that only high temperature weather and extreme weather significantly impact on group participation (-0.031**, -0.209***).

5. Implications and Contribution

This study is intended to examine the impacts of with-in group and between-group social comparison on group level users' exercise participation as well as how the offline social activities moderate the main effects. Our results show that both the within-group and between-group social comparisons significantly improve the group members' exercise participation. However, the offline social activities moderate on the main effect oppositely. Offline activities can significantly suppress the effects of within-group social comparison on group members' exercise participation but can significantly strengthen the effects of between-group social comparison on group members' exercise participation.

This study has several theoretical implications. First, we extend the social comparison theory to the group level and examines the role of "group" by differentiating the social comparison within group members and the comparison between different groups. Second, in previous literature, group identity is examined as an environmental factor and measured mainly through surveying people's perceived belongingness. Our study provides direct empirical evidence to quantify the distinct consequences of with-in and between group social comparisons. Third, there are few previous research studying the user's offline social interactions and their impacts on user's online retention and exercise participation. By combining the Teamwork theory, we explain distinctive moderating effects of offline team building activities on the effects of within-group and between-group social comparison. Fourth, we extend the generalization of the Teamwork theory from a working setting in the human resource field to an online-hobby-community setting in the IS field by verifying its central propositions through rich empirical evidence.

This study also has several managerial implications for app developers. First, our findings can help developers understand the role and importance of “group” and group-level social comparison. Based on our results, both within and between group comparison can have a salient boost in group user’s activity participation and retention. We suggest developers allocate more rewards for desired group outcomes and reinforce users’ awareness of the cohesion of the group when designing group level gamification mechanism. Second, our results demonstrate that offline team building activities only significantly moderate the effects of between-group competitive climates on group exercise participation. Motivated such we suggest managers initiate offline events, which can include more group competitions to trigger group members’ motive to exercise and participate more for the honor of groups. To sum up, the idea of strengthening the group concept and of luring users to form positive group identification can be significant to not only fitness but also all types of online community platforms’ and apps’ design of essential function, social function, and gamification. It can help to prevent users’ churning behaviors and improve users’ retention.

Conclusion

In conclusion, this study provides an initial step towards how the within-group and between-group social comparisons affect group level activity performance and how offline social activities moderate on these effects. Given the significance of the role of group and the importance of offline social activities, future research could further uncover the how offline social interaction and user behaviors can shape the growth of the online community and their online behaviors.

Reference

[1] Bock, G. W., & Kim, Y. G. (2002). Breaking the myths of rewards: An exploratory study of attitudes about knowledge sharing. Information Resources Management Journal (IRMJ), 15(2), 14-21.
[2] Breukelen WV, Leeden RVD, Wesselius W, Hoes M 2012. Differential treatment within sports teams, leader–member (coach–player) exchange quality, team atmosphere, and team performance. J Organ Behav, 33(1): 43–63.
[3] Brown, D. J., D. L. Ferris, D. Heller and L. M. Keeping (2007). "Antecedents and consequences of the frequency of upward and downward social comparisons at work." Organizational Behavior and Human Decision Processes 102(1): 59-75.
[4] Buunk, B. P., Collins, R. L., Taylor, S. E., VanYperen, N. W., & Dakof, G. A. (1990). The affective consequences of social comparison: either direction has its ups and downs. Journal of personality and social psychology, 59(6), 1238.
[5] Carron, A.V. (1988). Group dynamics in sport. London, ON: Sports Dynamics.
[6] Carron, A.V., & Spink, K.S. (1992). Internal consistency of the Group Environment Questionnaire modified for an exercise setting. Perceptual and Motor Skills, 74, 304-306.
[7] Cavallo, D. N., Tate, D. F., Ries, A. V., Brown, J. D., DeVellis, R. F., & Ammerman, A. S. (2012). A social media-based physical activity intervention: A randomized controlled trial. American Journal of Preventive Medicine, 43(5), 527-532. doi: 10.1016/j.amepre.2012.07.019
[8] Festinger L (1954). "A theory of social comparison processes". Human Relations. 7 (2): 117–140. doi:10.1177/001872675400700202.
[9] Fisher, R. J., & Ackerman, D. (1998). The effects of recognition and group need on volunteerism: A social norm perspective. Journal of consumer research, 25(3), 262-275.
[10] Garcia, S. M., Tor, A., & Schiff, T. M. (2013). The psychology of competition: A social comparison perspective. Perspectives on Psychological Science, 8(6), 634-650
[11] Kelman, H. C. (1958). "Compliance, Identification, and Internalization: Three Processes of Attitude Change." Journal of Conflict Resolution 2(1): 51-60.
[12] Mussweiler, T., & Strack, F. (2000). The "relative self": Informational and judgmental consequences of comparative self-evaluation. Journal of Personality and Social Psychology, 79(1), 23-38.
[13] Mussweiler, T., Ruter, K., & Epstude, K. (2004). The man who wasn't there: Subliminal social comparison standards influence self-evaluation. Journal of Experimental SocialPsychology, 40, 689-696.
[14] Munson, S. A., & Consolvo, S. (2012, May). Exploring goal-setting, rewards, self-monitoring, and sharing to motivate physical activity. In 2012 6th international conference on pervasive computing technologies for healthcare (pervasive health) and workshops (pp. 25-32). IEEE.
[15] Richardson, C.R., Buis, L.R., Janney, A.W., Goodrich, D.E., Sen, A., Hess, M.L., Mehari, K.S., Fortlage, L.
[16] A., Resnick, P.J., Zikmund-Fisher, B.J., Strecher, V.J., and Piette, J.D. 2010. "An Online Community Improves Adherence in an Internet-Mediated Walking Program. Part 1: Results of a Randomized Controlled Trial," Journal of Medical Internet Research (12:4), pp. 138-153.
[17] Sepehr, S., and Head, M. 2011. "The Role of Competitiveness in the Cognitive Absorption of Video Games," The 10th Pre-ICIS Annual Workshop on HCI Research in MIS.

[18] Singer, E. (1981). Reference groups and social evaluations. In Rosenberg, M., Turner, R. H. (Eds.), Social psychology: Sociological perspectives (pp. 66–93). New York: Basic Books.
[19] Sonders and Ana L. (2016). These fitness app statistics show what's going right (and wrong) for Fitbit. SurveyMonkey Intelligence
[20] Tuckman, B. W. (1965). Developmental sequence in small groups. Psychological bulletin, 63(6), 384. Widmeyer, W. N., & Williams, J. M. (1991). Predicting Cohesion in a Coaching Sport. Small Group Research, 22(4), 548–570. https://doi.org/10.1177/1046496491224007
[21] Wu, Y., A. Kankanhalli and K. Huang (2015). "Gamification in Fitness Apps: How Do Leaderboards Influence Exercise?".
[22] Zhou, Y., A. Kankanhalli and K. W. Huang (2016). Effects of Fitness Applications with SNS: How Do They Influence Physical Activity. International Conference on Information Systems.

CHAPTER FIVE

Conclusion

This dissertation brings values to both the literature and industry for studying new UGC and online community technologies and their influences. The first essay investigates how the entry and penetration of LBAR mobile applications impacts on business performance of local restaurants. The second essay studies whether and how content consumption on the pre-record video platforms and the live streaming video platforms affect related product sales and customer's stickiness. The third essay examines how the between- and within-group social comparison influence group users' exercise participation and how offline social activities shape the direct impacts.

In the first essay, through a series of treatment efficiency and boundary conditions analyses, we validate the overall positive spillover effects of the entry of the LBAR app on all restaurants located in the proximity areas of in-app portals and reveal a rich-get-richer pattern of the effects internalization given different localized restaurants factors. This study reveals the underlying mechanism of the positive effects by decoding the AR technologies' bridging function to link the real physical world to virtual app portals as well as the LBS technologies' stimulating functions to encourage users to explore vicinity areas. This study also contributes to spillover literature with more refined findings in terms of the various influences of the spillover internalization. Further, we reveal the significance of locational dominance coming from spatial factors. This study also sheds light on the economic value of LBAR applications and recommends local business owners to proactively advertise and market their locational advantages as well as to utilize LBAR application as an assistive tool for proximity marketing.

The second essay contributes to IS, marketing as well as the communication literature in several ways. This study, as far as we know, is among the first to empirically validate the marketing efficiency of both the pre-recorded video platforms and the live streaming video platforms through a large, timely and unique sales data. This study is also the first to show the distinctiveness of two platforms and shows their idiosyncratic marketing efficiency. Moreover, this study extends the media richness theory and applies its key proposition to individual usage level and to the scenario that the user-generated videos, including intendedly or unintendedly product information, can also shape product sales and customer's stickiness. This study validates the social presences theory by providing direct empirical evidence and quantifying purchase and re-use intentions to sales and customer stickiness. For practitioners, this study provides several intuitive rules of thumb tips for practitioners, especially for hedonic products like video games. We recommend all managers to allocate certain marketing budgets on the online video platforms, to create and share vlog type of videos embedding with products and branding information and to sponsor and cooperate with influencers on such platforms for branding and inbound marketing.

The third essay aims to solve three research questions. How do within-group social comparisons affect group members' exercise participation? How do between-group competitive climates affect group members' exercise participation? How do offline team building activities moderate these effects? Theoretically, this paper extends the social comparison theory to the group level, examines the role of "group" by differentiating the social comparison into two angles, provides direct empirical evidence to quantify the distinct consequences of within and between group social comparisons, combines the findings from teamwork and offline online social activities with the reasoning based on the cognitive processes during social comparison to explain distinctive moderating effects of offline team building activities, extends the generalization of the teamwork literature from a working setting in the HR field to an online-hobby-community setting in the IS field. For managers and app developers, we suggest developers to allocate more rewards for desired group outcomes and reinforce users' awareness of the cohesion of the group when designing group level gamification mechanism. We also recommend managers to initiate offline events, which can include more group competitions to trigger group members' motives to exercise and participate more for the honor of groups.

To sum up, this dissertation demonstrates the importance and meanings of the newly emergent UGC and online community platforms and mobile technologies to IS literature. This dissertation also sheds light on the application and implication of such technologies for practitioners based on our findings. For the first paper, future research can measure the direct impacts of LBAR applications on local businesses if real-time traffic data is available. Further, future research can collect more source data to examine the role of LBAR technologies in the ecosystem of local business and social media. For the second paper, if there are viewers' click-through data, future research can specifically track their subsequent behaviors on both the UGC and the e-commerce platforms. Future research can also reveal more values and applications of live streaming platforms. For the third paper, if there the discussion board text and graphic data are available, future research can further define within-group interaction and examine how different offline activities are associated with users' online behaviors.