The Dissertation Committee for Zhuping Liu
certifies that this is the approved version of the following dissertation:

Push and Pull: Targeting and Couponing in Mobile Marketing

Committee:

Vijay Mahajan, Supervisor

Raghunath Rao

Jun Duan

James Scott

Cotter Cunningham

Push and Pull: Targeting and Couponing in Mobile Marketing

by

Zhuping Liu

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2017

Dedicated to the memory of my mentor Frenkel.

Acknowledgments

I am deeply indebted to my advisor,Vijay Mahajan, for his guidance and support during my time at UT Austin. I learned from him not only how to be a scholar, but also how to be a decent human being. I would like to thank Frenkel ter Hofestede, who was co-chairing this dissertation before his sudden death. I am grateful to Frenkel for his enormous help in getting me started in academic research. I thank my committee members, Raghunath Rao, Jun Duan, James Scott and Cotter Cunningham, for their constructive feedback and insightful comments, which have helped improve this dissertation.

I thank my friends, Lan Liang, Jacob Suher, Sandeep Arora, Xinying Hao, Jerry Han, Jiyoung Lee, Niket Jindal, Shunyang Lee, Prasart Jongjaroenkamol, Jaehee Choi and many others, for their friendship which have made this journey so much more enjoyable. Finally I thank my parents,my wife and my son for their unconditional love. I especially want to thank my wife for her sacrifice.

Push and Pull: Targeting and Couponing in Mobile Marketing

Zhuping Liu, Ph.D.
The University of Texas at Austin, 2017

Supervisor: Vijay Mahajan

The prevalence of mobile marketing practices has profoundly changed the way consumers shop. Consumers are increasingly shifting to mobile coupons to enhance their shopping experiences. This shift to mobile has created unique opportunities for marketers to engage and target consumers who both actively pull coupons from the mobile app and passively receive targeted push messages about coupons. The literature in mobile marketing is new and numerous issues have not yet been studied. My dissertation examines two such issues in mobile marketing to advance our understanding of the role of mobile promotions in consumers shopping journeys and to explore effective personalization strategies in mobile marketing.

My first essay examines the effect of mobile promotions on foot traffic by capturing the dynamic interactions among shopper-initiated and publisher-initiated activities. Shoppers might receive targeted push messages based on either their individual historical behavior ("behavior-based push") or their

current location ("location-based push"). I develop a novel multinomial multivariate point process model, which predicts the dynamic interactions between activities. To overcome computational issues in estimation, I develop a new methodology that allows the model to zoom in to days that include activities and to zoom out of inactive days. My simulation of a 15-day period reveals the following insights. First, a behavior-based push leads to an increase in mobile engagement outside malls of more than 25% and an increase in shopping traffic to online stores of about 24%. Second, a behavior-based push would result in an increase in foot traffic to regional malls of about 5% but to strip malls of only about 0.5%. Third, a behavior-based push leads to an increase in mobile engagement inside malls of more than 19% and in coupon redemptions of about 18%, while a location-based push increases mobile engagement inside malls by about 40% and coupon redemptions by about 25%. Therefore, behavior-based push and location-based push play different roles in influencing shopper-initiated activities. I conclude with implications for publishers, mall owners, and retailers on how to leverage mobile marketing to increase mobile engagement, online traffic, foot traffic, and coupon redemptions.

My second essay studies the ranking and personalization of organic and sponsored mobile advertising (or coupons) that mix together when delivered to consumers. The publisher faces a tradeoff between placing sponsored ads from retailers to receive revenue from advertised retailers and selecting the right organic ads to keep consumers engaged. I propose a consumer mobile search model that can account for the unique factors in our empirical context and answer my research questions. I present model-free evidence for the influence of screen size, whether consumers are in a shopping mall and ad type. I also show how consumer sliding and clicking influence their exit decisions. The proposed counterfactual simulations explore different ways of personalization, including (i) selecting personalized ad contents from the vast amount of available ads by consumers affinity score, by methods like collaborative filtering, or by whether the advertised retailer has an store in the shopping mall a consumer is in; and (ii) ranking selected ads by consumers affinity score, by ad discount quality score, or by whether an ad is for online shopping or for in-store shopping. These simulations will provide marketers insights into whether and how much each type of personalization improves consumer responses to both organic and sponsored ads, thereby offering guidance to the publisher in optimizing their current practice.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

As smartphones have become more affordable and mobile networks have advanced, the global population of smartphone users has been expanding rapidly. By the end of 2017, about 2.5 billion people worldwide will have adopted smartphones (eMarketer, 2014b). The high penetration of smartphones offers marketers opportunities to engage and target a large number of consumers through mobile applications. Consumers are increasingly relying on their mobile devices to satisfy their digital needs. In 2016, an adult spends on average 3.1 hours on mobile per day in the United States, overtaking the time spent on computers and other devices combined (KPCB, 2017).

Consumers' shift to mobile has created tremendous mobile targeting and advertising opportunities. Marketers can leverage mobile apps to offer consumers personalized content, to segment consumers using their browsing history, and to push targeted marketing messages. In fact, marketers are increasingly interested in mobile targeting and personalization, which are the top priorities for digital marketing (eMarketer, 2015). In 2014, 75 percent of digital media and marketing professionals targeted their ads to specific consumer segments on smartphones (eMarketer, 2015). Recent surveys find that personalization is what many consumers expect and want, and what advertisers are using to increase response and engagement rates (eMarketer, 2016). Chapter 2 and 3 examine the role of mobile targeting and personalization in mobile marketing to enhance the understanding of these newly-emerged marketing levers.

Chapter 2 studies the effect of mobile promotions on foot traffic by capturing the dynamic interactions among shopper-initiated and publisher-initiated activities. Shoppers[1] might receive targeted push messages based on either their individual historical behavior ("behavior-based push") or their current location ("location-based push"). I develop a novel multinomial multivariate point process model, in which coupon pulls, online store visits, visits to each of the shopping malls, and coupon redemptions in these shopping malls can mutually influence each other; meanwhile, the publisher-initiated mobile pushes directly influence coupon pulls and indirectly influence mall visits. To overcome computational issues in estimation, I develop a new methodology that allows the model to zoom in to days that include activities and to zoom out of inactive days. To take shoppers' unobserved heterogeneity into account, I cast the model in a Bayesian hierarchical modeling framework. The proposed model offers mall owners and retailers a novel method to evaluate the effect of mobile promotions on foot traffic to shopping malls and on coupon redemptions. I estimate the proposed model through large-scale parallel supercomputing, using the high performance computing facility in the Texas Advanced

[1] I use "shoppers" and "consumers" interchangeably

Computing Center. Based on my estimation results, I simulate each shopper's events for 15 days, which reveals interesting insights into mobile marketing.

Chapter 3 investigates the ranking and personalization of organic and sponsored mobile advertising that mix together when delivered to consumers. I develop a consumer information search model where a forward-looking consumer maximize her utility by choosing whether to click an ad or slide the screen to move down the viewable ads on the screen. The proposed model can account for the unique factors in our empirical context and answer my research questions. I present model-free evidence for the impact of screen size, whether consumers are in a shopping mall and ad type. I also show how sliding and clicking influence consumers' exit decisions. The proposed counterfactual simulations explore different ways of personalization, including (i) selecting personalized ad contents from the vast amount of available ads by consumers' affinity score, by methods like collaborative filtering, or by whether the advertised retailer has an store in the shopping mall a consumer is in; and (ii) ranking selected ads by consumers affinity score, by ad discount quality score, or by whether an ad is for online shopping or for in-store shopping. These simulations will provide marketers insights into whether and how much each type of personalization improves consumer responses to both organic and sponsored ads, thereby offering guidance to the publisher in optimizing their current practice.

In Chapter 4, I conclude my dissertation and discuss some of the issues for future research.

Chapter 2

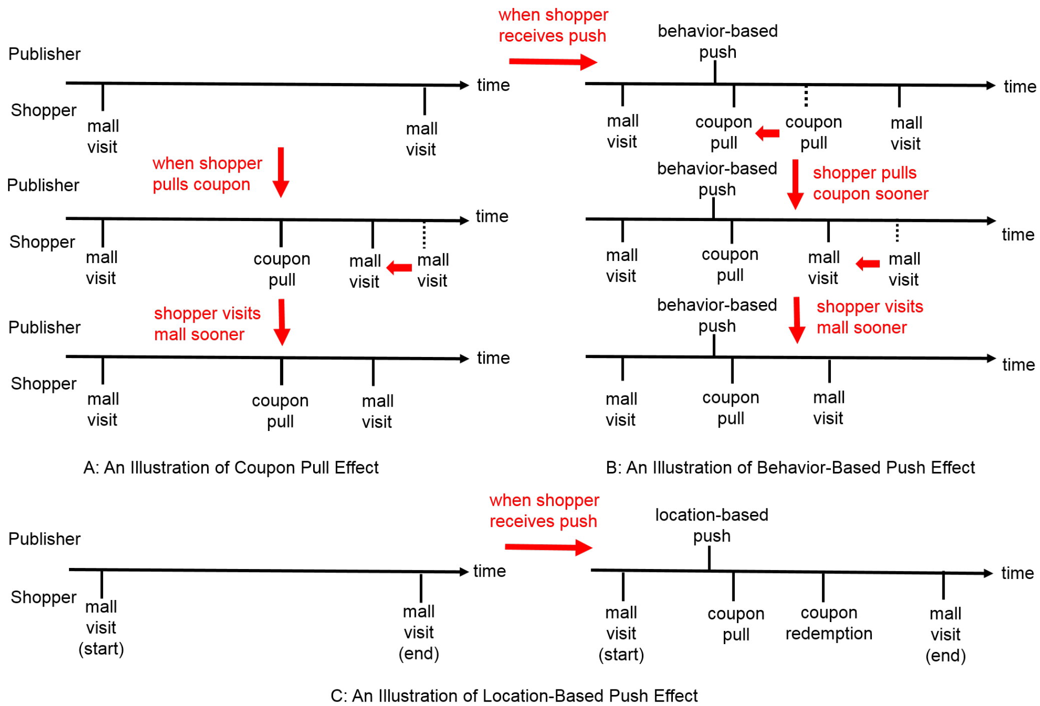
Engaging and Targeting Consumers on Mobile

2.1 Background

The prevalence of mobile marketing vehicle use, such as mobile coupons, has profoundly changed how consumers shop. Today, coupons are often offered through mobile apps. Consumers are increasingly shifting to mobile coupons to enhance their shopping experiences. The number of shoppers who redeem coupons via a mobile device worldwide is expected to pass 1 billion by 2019 (Juniper Research, 2014). This shift to mobile coupons has created unique opportunities for marketers to engage and target consumers who not only can actively pull coupons from the mobile app but also can passively receive targeted push messages based on their past behavior ("behavior-based push") or their current location ("location-based push"). These push and pull activities through mobile apps can influence shoppers' decisions on when to shop and where to shop, therefore influencing foot traffic to shopping malls.

Retail foot traffic, a key metric for shopping malls to charge rent, has declined significantly in the past few years because of the rise of e-commerce (Wall Street Journal, 2014; Forbes, 2016). Appreciating the new engaging and targeting opportunities that mobile marketing brings, shopping malls are ac-

Figure 2.1: Illustrative Examples of Coupon Pull and Mobile Push Effects



tively developing innovative mobile apps designed to engage and retain shoppers, fighting against the threat from online retailing. For example, Simon Property Group, Inc., the biggest mall owner in the United States, offers shoppers its own app to display promotions from stores in its malls (Wall Street Journal, 2011). The mobile app of the second largest mall owner — General Growth Properties, Inc. (GGP) — allows mall retailers to push out deals and coupons to drive in-store traffic (Mobile Marketer, 2011). Such mobile apps are deemed vital by mall owners and mall retailers at a time when shoppers increasingly consult their mobile devices to plan their shopping trips.

Although marketers are trying out different mobile marketing tactics to influence foot traffic, evaluating the results of such marketing activities remains a challenge for marketing researchers because they lack the appropriate data collection technology and a rigorous methodology for analysis. In fact, almost all previous research on mobile marketing collects data through “one-shot” experiments (Andrews et al., 2015; Luo et al., 2014; Danaher et al., 2015; Fong et al., 2015), which cannot generate an understanding of consumer dynamics in mobile marketing. In Figure 2.1A, we illustrate the interaction between a shopper-initiated coupon pull and a mall visit. When shoppers pull coupons from the mobile app, they might visit a mall sooner. In addition, they might be more likely to visit online stores after pulling coupons. Figure 2.1B demonstrates the interaction between a targeted behavior-based push and a coupon pull and the subsequent interaction between a coupon pull and a mall visit. A targeted behavior-based push message might trigger a shopper to pull coupons, visit online stores, and subsequently accelerate the next shopping trip. During a mall visit, a shopper might receive a targeted location-based push message that reminds her to pull and redeem coupons, as illustrated in Figure 2.1C. To the best of my knowledge, no prior research has studied how behavior-based push and location-based push influence mobile engagement and both online and offline shopping activities. This study aims to fill this gap by studying the following research questions:

1) How do behavior-based push and location-based push influence shoppers’ engagement with the mobile app?

2) How do behavior-based pushes influence shopping frequencies? Can they increase online or offline shopping frequencies?

3) How do behavior-based push and location-based push influence consumers' propensity to redeem coupons? Can they increase coupon redemptions in offline shopping trips?

To address these research questions, I leverage a unique data set from a major publisher that displays mobile coupons for online and offline shopping from retailers and that targets shoppers through behavior-based and location-based mobile pushes to improve mobile engagement. This data set records shoppers' coupon pulls, the mobile pushes received, their online store visits, their shopping mall visits, and their coupon redemptions in shopping malls, using state-of-the-art mobile technologies. In addition, I develop a new modeling approach that can properly characterize the dynamics of consumers' shopping behavior and that can account for the following unique properties and patterns of mobile promotion and offline shopping data.

First, shopper-initiated coupon pulls, online store visits, shopping mall visits, and coupon redemptions in shopping malls are stochastic activities over time that might interact with other occurrences of the same activity, as well as the other activities. To illustrate, pulling coupons from the mobile app might lead to subsequent coupon pulls, can also affect or interact with online store visits, and can lead to either more or fewer subsequent shopping mall visits and coupon redemptions. Visiting a shopping mall might increase a shopper's propensity to pull coupons to find out what deals are available. However, she may be less likely to use the mobile app if she has planned carefully before the shopping trip. Meanwhile, patronizing one mall might satisfy a shopper's needs and therefore reduce the probability of visiting another one in the next few days. However, when another mall is close by, the shopper might be more likely to visit it on the same trip. Therefore, properly modeling the dynamic interactions among both the same and different types of shopper-initiated activities (e.g., coupon pulls and mall visits) and allowing both positive (i.e., mutually exciting) and negative (i.e., mutually inhibitory) interactions are essential in understanding the effect of mobile promotions on foot traffic.

Second, in addition to pulling coupons, visiting online stores, visiting shopping malls, and redeeming coupons, shoppers are subject to mobile targeting (i.e., to receiving behavior-based and location-based targeted push messages). The influence of behavior-based push on shopping propensity is indirect because behavior-based push almost never indicates in the message whether it is for online or offline shopping, and the only way to find out about the coupon information associated with the push message is to pull the coupon from the mobile app. The effect of location-based push on coupon redemptions inside a shopping mall is also indirect because redeeming coupons also requires that shoppers pull the coupons from the mobile app. Different from shopper-initiated coupon pulls, online store visits, shopping mall visits, and coupon redemptions, these targeted behavior-based push and location-based push strategies are controlled by the publisher. Explicitly incorporating the "targeting rule" used by the publisher and allowing mobile pushes to influence shoppers' coupon pulls and mall visits form indispensable elements of the model.

Third, the effect of a mobile push on a coupon pull and the effect of a current coupon pull or mall visit on later ones should change over time. For example, a mobile push might have the highest effect on coupon pulls on the day when a shopper receives the push, leading to a monotonically decreasing shape over time. A time decay function is needed to capture these time-varying effects.

Fourth, interactions among mobile pushes, coupon pulls, and mall visits often occur at vastly different frequencies. For example, in Figure 2.1B, a shopper might take only a few seconds to launch the app after receiving a targeted behavior-based push message, but she also might not receive a targeted push or pull coupons, or visit a mall at all, in a few weeks. Therefore, aggregating data by day or week is problematic because it is unlikely to capture the interactions that happen within seconds. However, calibrating the model at the second level also is problematic because it significantly increases the computational burden for the days without any such activity, making the estimation computationally infeasible. Meanwhile, empirically speaking, shoppers tend to plan shopping trips by day, with some randomness on the timing.

In this study, I develop a novel model, which I call a Multinomial Multivariate Point Process model with Adaptive Piecewise Constant Intensity (mMPP-APC), in which coupon pulls, online store visits, visits to each of the shopping malls, and coupon redemptions in these shopping malls are treated as "shopper-initiated events" that can mutually excite or mutually inhibit; meanwhile, the publisher-initiated mobile pushes directly influence coupon pulls and indirectly influence mall visits. The adaptive piecewise constant intensity feature allows me to "zoom in" to days with events and "zoom out" from days without events to extract information from the data efficiently. Furthermore, I use an exponential time decay function to capture the time-varying effects. To take shoppers' unobserved heterogeneity into account, I cast the model in a Bayesian hierarchical modeling framework. My model offers mall owners and retailers a novel method to evaluate the effect of mobile promotions on foot traffic to shopping malls and on coupon redemptions. They can use my model to determine which shoppers they can engage and target to increase foot traffic and coupon redemptions. In addition, my model provides publishers and app owners a better alternative than the commonly used "reaction rate" metric — the percentage of users who received a mobile push and then clicked on it (Accengage, 2016) — to quantify the effect of targeted mobile pushes on shoppers' mobile engagement. My model also can shed light on substitutability and complementarity among shopping malls, which allows mall managers to evaluate their competitive position.

I estimate my model through large-scale parallel supercomputing, using the high performance computing facility in the Texas Advanced Computing Center. I implement the exact Markov Chain Monte Carlo (MCMC) algorithm in parallel on one of the most powerful supercomputers in the world. Based on my estimation results, I simulate each shopper's events for 15 days, which reveals interesting insights into mobile marketing. First, behavior-based push can substantially increase shoppers' mobile engagement and shopping traffic to online stores. A behavior-based push leads to an increase in coupon pulls of almost 25% and an increase in shopping traffic to online stores of about 24%. Second, behavior-based push is effective in increasing foot traffic to regional shopping malls. I find that a single behavior-based push would result in an increase in foot traffic to regional malls of about 5% but only a 0.5% increase to strip malls. Third, a behavior-based push increases coupon pulls inside malls by more than 19% and coupon redemptions by about 18%. In contrast, a location-based push leads to an increase in coupon pulls inside malls of about 40% and an increase in coupon redemptions of about 25%.

This study contributes to the marketing literature both substantively and methodologically. From a substantive point of view, this study is the first, to the best of my knowledge, to consider both mobile push and coupon pull to connect mobile promotions to foot traffic. Previous research has studied when pushed mobile coupons are more effective (Andrews et al., 2015; Luo et al., 2014; Danaher et al., 2015; Fong et al., 2015) and when mobile coupons are more likely to be pulled (Molitor et al., 2016). However, none of the previous research has studied how behavior-based push and location-based push influence in-app coupon pulls, online store visits, or offline shopping activities. In addition, previous research on mobile push or coupon pull has been conducted through "static" field experiments, which does not allow for the analysis of consumer dynamics. My study is therefore the first to examine the dynamic interactions among in-app coupon pulls, online store visits, and offline shopping activities. It is also the first to include both behavior-based push and location-based push at the individual level in one study.

In addition to the substantive contributions, this study also makes several methodological contributions. My model naturally integrates shopper-initiated events (e.g., coupon pulls, online store visits, shopping mall visits, and coupon redemptions) with publisher-initiated pushes (e.g., behavior-based push, location-based push) using a novel multinomial multivariate point process model that can accommodate both positive and negative interactions. To the best of my knowledge, this study is the first in marketing to incorporate both shopper-initiated and publisher-initiated activities into a unifying point process model while explicitly correcting for the non-randomness in publisher-initiated activities. It is also the first in applied statistics and econometrics to successfully incorporate both mutually exciting and mutually inhibiting interactions into a point process model and to estimate it using an exact MCMC algorithm through large-scale parallel supercomputing. Moreover, the adaptive piecewise constant intensity feature provides a solution to the modeling framework of drastically different frequency data. My model captures each shopper's event propensity from moment to moment and allows for decay over time in the effects of shopper- and publisher-initiated activities.

2.2 Literature Review

This study is related to the literature on mobile marketing. The mMPP-APC model I develop is built on the existing literature on point processes, so I also review the literature on point processes.

My study is related to the emerging literature on mobile marketing. In the past few years, mobile marketing has drawn increasing attention in academia. Researchers have shown that mobile phone use behavior is unique (Ghose and Han, 2011); that targeting consumers temporally and geographically and offering a substantial discount can increase sales (Luo et al., 2014); that targeting competitive locations with discounts can generate incremental sales (Fong et al., 2015); that location and time of mobile-coupon delivery and expiry length affect coupon redemption (Danaher et al., 2015); that targeted mobile ads to commuters in crowded subway trains are more effective than such ads in noncrowded ones (Andrews et al., 2015); and that geographical distance between a user and a store and coupon positions in the mobile app influence mobile coupon click rates (Molitor et al., 2016). However, little research has studied how mobile promotions influence consumers' dynamic shopping behavior. One exception is Ghose et al. (2015), who in a field experiment test the effectiveness of a new mobile advertising strategy based on consumers' offline movement trajectories. My study contributes to the literature of this emerging area by looking into the consumers' dynamic interactions among in-app mobile activities, online store visits, and offline shopping activities. I also study consumers' dynamic shopping behavior in light of mobile targeting by tapping into the rich history of their past behavior and their current location. To the best of my knowledge, no prior research has studied the effects of both behavior-based and location-based mobile targeting on consumers' in-app mobile activities, online store visits, and offline shopping activities. With a unique data set obtained from a natural setting, I offer much richer insights into consumers' shopping behavior as influenced by mobile targeting.

I propose a new model to characterize the dynamic interactions among different types of events, built on the literature on point processes. Point process is a stochastic process composed of a time series of binary events (Daley and Vere-Jones, 2003) and is often used to describe data that are localized at a finite set of time points. Point processes have been applied to research in seismology (Musmecl and Vere-Jones, 1992; Ogata, 2004; Ogata et al., 2003; Ogata and Zhuang, 2006), finance (At-Sahalia et al., 2015), homeland security (Porter and White, 2012), sociology (Mohler et al., 2011), ecology (Johnson et al., 2013), neuroscience (Truccolo et al., 2005), and many other areas. These studies do not consider individual heterogeneity; they use maximum likelihood estimation, often with approximations to simplify the likelihood function calculation. In addition, none of these studies allow for both exciting and inhibitory interaction effects and events that are externally controlled. This study is thus the first to consider both exciting and inhibitory interactions and externally controlled events while incorporating individual heterogeneity in an exact Bayesian estimation algorithm. Little work in marketing has applied point processes to marketing research. One exception is Xu et al. (2015), in which a mutually exciting point process is applied to an online advertising context to evaluate advertising conversion rates. In their model, the advertising effects are restricted to nonnegative values to ensure that the intensity function is positive. These restrictions do not apply to my context. In my empirical setting, the dynamic interaction effects among different types of events can be either positive or negative. I do not know the directions of the effect a priori. In addition, shoppers plan shopping trips in a discrete manner rather than a continuous one (e.g., in days). Nights serve as a natural separator in shoppers' decision-making. Setting location in the same city eliminates the possibility that shoppers visit a mall during the night, which contrasts with online marketing contexts where visitors from different time zones can "visit" around the clock. To take into account the unique characteristics in my empirical setting, I develop a novel discrete-time point process model that allows for both positive and negative interactions that decay over time. My model also explicitly accounts for the targeting rule that the publisher providing the data has set up for behavior-based pushes and allows for the influence of the publisher-initiated pushes on shopper-initiated events.

2.3 Data

The data used in this study come from a major publisher that publishes mobile coupons from major retailers and targets shoppers through a mobile app. Shoppers' coupon pulls, online store visits, and shopping mall visits are recorded through state-of-the-art mobile technologies. The data made available to the author were fully anonymized by the publisher to protect the privacy of the shoppers. To address my research questions, I selected a sample of 4,404 shoppers who both live and work in San Antonio, TX, and who opted to receive targeted push messages from the publisher from February 15, 2015, to June 15, 2015. These targeted mobile pushes fall into two categories: behavior-based push — a message informing shoppers of coupons for online or offline shopping, and location-based push — a general message about the number of coupons available at the current shopping mall. I also have 28 days of app use history prior to February 15, 2015, to calibrate the publisher's "targeting rule" for behavior-based push. I focus on the largest 29 shopping malls, including 4 regional malls and 25 strip malls, of the more than 300 malls in San Antonio. These 29 malls are the major destinations for non-grocery shopping in the San Antonio area.

Table 2.1: Summary Statistics

	Min	Mean	Max	Standard deviation	Total
Coupon pull outside mall	0	11.550	132	14.791	50,858
Online store visit	0	3.415	92	5.549	15,038
Regional mall visit	0	6.124	87	8.088	26,969
Strip mall visit	0	24.700	121	18.844	108,800
Coupon pull inside mall	0	2.082	87	2.931	9,167
Coupon redemption	0	1.088	24	1.814	4,793
Location-based push	0	20.280	98	15.520	89,292
Behavior-based push	0	14.620	56	11.757	64,390

In Table 2.1, I present the summary statistics of the data. The data include around 27,000 visits to the 4 regional malls; around 110,000 visits to the 25 strip malls; about 51,000 coupon pulls outside the 29 malls (coupon pull outside mall); more than 9,000 coupon pulls inside the 29 malls (coupon pull inside mall); and almost 5,000 coupon redemption clicks inside the 29 malls (coupon redemption). A coupon redemption click is the last click required to reveal a coupon code for redemption, serving as a proxy for coupon redemption in malls. Shoppers received almost 90,000 location-based push messages about the focal 29 malls from the publisher. In addition, these shoppers received about 65,000 behavior-based push messages reminding shoppers of coupons for online and offline shopping. The push message itself is decided solely by the publisher, while the promoted coupons associated with the push messages are often decided by the publisher along with the retailers. A glance at the summary statistics reveals a substantial amount of heterogeneity among shoppers. For example, one shopper might pull coupons as many as 132 times when outside malls, while another shopper never pulls coupons at all when outside malls. I take into account the unobserved heterogeneity among individual shoppers in the proposed model in Section 2.4.

The timestamps for each coupon pull outside malls, each online store visit, each regional and strip mall visit, each coupon pull inside a mall, each coupon redemption, each location-based push, and each behavior-based push are provided in the data. In the data, I observe a vast amount of different activity sequences. Consider a strip mall visit as an example. Before a shopper's next visit but after the last visit to a strip mall, she might pull coupons outside malls (CO), visit a regional mall (VR) or a strip mall (VS), visit online stores

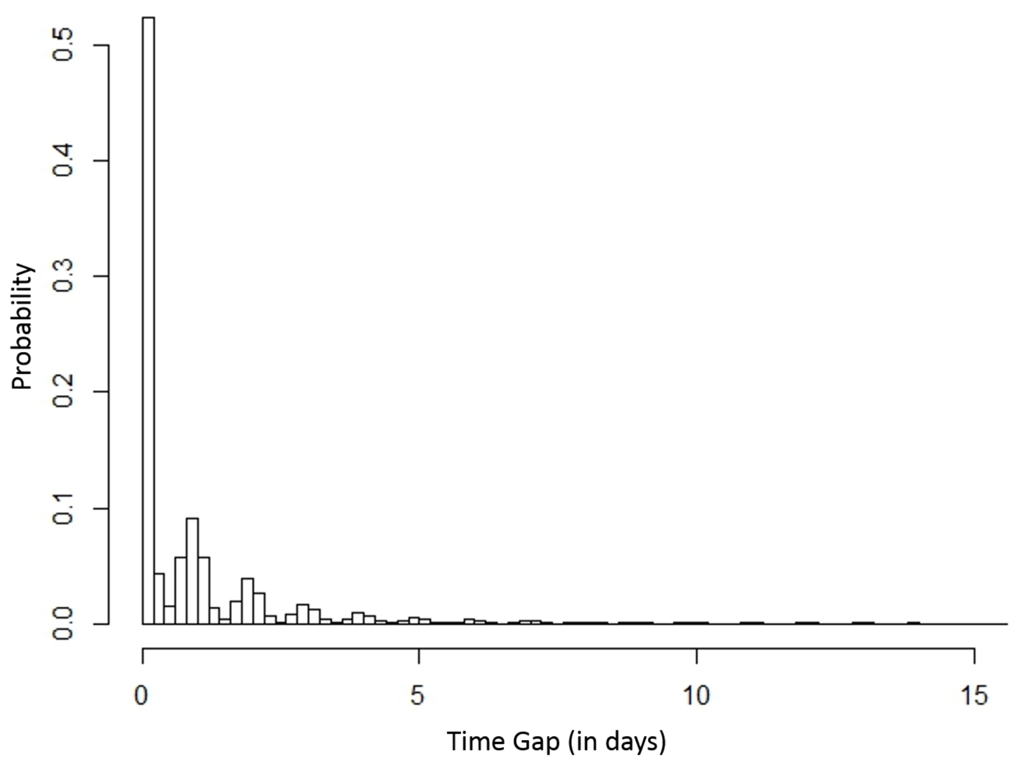
(VO), pull coupons inside a mall (CI), redeem coupons in a mall (RM), exit a regional mall (ER) or a strip mall (ES), receive a location-based push from the focal 29 malls (LF) or from other malls (LO), or receive a behavior-based push (BP). I observe numerous sequences of these events in my data and as an illustration present the most frequent sequences starting and ending with a strip mall visit in Table 2.2. A scrutiny of Table 2.2 reveals the common patterns of the variety of events in each of the sequences, indicating the necessity of capturing the interactions among different types of events. For example, the most common pattern shows that shoppers often visit a strip mall, receive a location-based push message, exit the strip mall, and then enter a different or the same strip mall later. Note that these activity sequences are extracted from the data without considering the time gap between the adjacent activities.

Table 2.2: Summary of Event Sequences for the Anonymous Shoppers Strip Mall Visits

Sequence	Count	Sequence	Count
VS-LF-ES-VS	19,522	VS-LF-ES-VR-LF-ER-VS	1,058
VS-ES-VS	14,598	VS-LF-LO-ES-VS	1,056
VS-LF-ES-LO-VS	6440	VS-LF-ES-CO-VS	814
VS-ES-LO-VS	2,857	VS-LF-ES-BP-LO-VS	752
VS-LF-ES-LO-LO-VS	2,515	VS-LF-CI-ES-VS	702
VS-LF-ES-BP-VS	2408	VS-LF-ES-LO-BP-VS	685
VS-ES-BP-VS	1,341	VS-ES-VR-ER-VS	655
VS-ES-LO-LO-VS	1,193	VS-ES-CO-VS	648
VS-LF-ES-LO-LO-LO-VS	1,071	...	...

In addition, the time gaps between any two adjacent data points vary substantially from a few seconds to a few weeks, which is illustrated in Figure

Figure 2.2: Histogram of Time Gaps Between Two Adajacent Data Points



2.2. For example, a shopper might take only a few seconds to launch the app after receiving a targeted mobile push message, but she might not receive a targeted push, pull coupons, or visit a mall at all in a few weeks. This data pattern poses challenges for extracting information from data while not significantly increasing the computational burden.

2.4 Model

In this section, I propose a novel mMPP-APC model to capture the unique characteristics and data patterns described in Sections 2.1 and 2.4. It accounts for the interactions among different types of events and naturally integrates publisher-initiated pushes and shopper-initiated coupon pulls, online store visits, shopping mall visits, and coupon redemptions. Different from previous research on point processes, it allows both mutually exciting and mutually inhibitory interactions. In addition, I cast the model in a Bayesian hierarchical framework to take into account heterogeneity among individual consumers. Built on a standard point process, mMPP-APC has several significant innovations. In this section, I first introduce the standard point process (Hawkes, 1971b,a), which is the basis of my modeling approach, and then describe the proposed model in detail.

2.4.1 Standard Point Processes

A point process is a type of stochastic process that models the occurrence of events as a series of random events in time and/or in geographical space. In my empirical context, shopper-initiated coupon pulls, online store visits, shopping mall visits, and coupon redemptions can be modeled as random points in time. Denote such a sequence of N events, which occur at $\{t_1, \cdots, t_N\}$, as $\{E_1, \cdots, E_N\}$. Let $N(t)$ denote the number of occurrences of events up to, but not including, t, and $H_t = [E|t_E < t]$; then the conditional intensity function (or hazard) is

$$\lambda(t|H_t) = \lim_{\triangle t \to 0} \frac{E[N(t + \triangle t)|H_t]}{\triangle t} \tag{2.1}$$

which represents the expected instantaneous rate of future events at t.

Standard point processes are a class of mutually exciting point processes that include both the univariate (self-exciting) point process and the multivariate (mutual-exciting) point process. In the univariate case, I have the conditional intensity function, $\lambda(t|H_t) = \mu(t) + \int_{-\infty}^{t} g(t - \bar{t})dN(\bar{t})$, where $\mu(t)$ is the baseline intensity while $g(.)$ is a kernel function expressing the time-decay effect of previous events prior to time t. In the multivariate case, I have the conditional intensity function, $\lambda = [\lambda_1, \cdots, \lambda_D]$, where D is the number of coupling time series, and $\lambda_d(t|H_t) = \mu_d(t) + \sum_{d'=1}^{D} \int_{-\infty}^{t} g_{d'}(t - \bar{t})dN_{d'}(\bar{t})$, where $\mu_d(t)$ is the baseline intensity for event type d and $g_d(.)$ is a kernel function expressing the time-decay effect of all previous events on event type d at time t. One advantage of the point process framework is its capability of naturally accounting for the influence of past events on current events, which well suits my empirical context. If shoppers pulling coupons accelerates the next shopping trip, I want to see whether future shopping trips after the next will be decelerated, which requires considering multiple previous shopping activities. An important assumption in standard point processes is that events prior to time t always increase the conditional intensity at time t. This assumption guarantees that the conditional intensity is always nonnegative, but it severely limits the applicability of the model — especially in contexts where the interactions among events might be either inhibitory or exciting.

2.4.2 The Proposed mMPC-APC Model

In this section, I propose the new multinomial multivariate point process model, mMPP-APC, to study the effects of mobile promotions. I model shopper i's coupon pulls, online store visits, shopping mall visits, and coupon redemptions as shopper-initiated "events" in a point process. Specifically, I let $N^i_{O,t} = [N^i_{O,t,1}, \cdots, N^i_{O,t,K_O}]$ where N^i_{O,t,k_O} represents the total number of type k_O occurrences within the time interval $[0, t]$, and let $N^i_{I,t} = [N^i_{I,t,1}, \cdots, N^i_{I,t,K_I}]$, $k_O = 1, 2, \cdots, K_O$, when a shopper is outside malls; and let $N^i_{I,t} = [N^i_{I,t,1}, \cdots, N^i_{I,t,K_I}]$, where N^i_{I,t,k_I} represents the total number of type k_I occurrences within the time interval $[0, t]$, $k_I = 1, 2, \cdots, K_I$, when a shopper is inside a mall. In my empirical setting, each shopper has 31 types of shopper-initiated events when outside malls — coupon pull outside mall ($N_{O,1}$), visit to online stores ($N_{O,2}$), visit to each of the regional malls ($N_{O,3}, \cdots, N_{O,6}$), and visit to each of the strip malls ($N_{O,7}, \cdots, N_{O,31}$), and 4 types of shopper-initiated events when inside a mall — coupon pull inside mall ($N_{I,1}$), coupon redemption ($N_{I,2}$), regional mall exit ($N_{I,3}$), and strip mall exit ($N_{I,4}$). In addition, shoppers are subject to the influence of two types of publisher-initiated activities ($J = 2$): location-based push (T_1) and behavior-based push (T_2).

Although the standard point processes provide me with intuitions on modeling the interactions among different types of events, I have to extend the model in a number of ways to accommodate the unique characteristics in my empirical context. First, shoppers tend to plan shopping trips in a discrete-time manner rather than a continuous-time manner. The inactive state of shopping malls at night serves as a natural time separator in consumers' shopping decisions. However, the exact event time within the day is random. For example, a shopper might plan to shop tomorrow, but he likely is not certain about the exact shopping time. Therefore, I let my model look into each day and allow shopper-initiated events (if any) to occur randomly within the day.

I assume that at any small time period outside malls, a shopper can decide whether to pull coupons and visit online stores, and can choose which shopping mall to visit. I model shoppers' coupon pull decisions and online store visit decisions through Bernoulli distributions, and I model shopping mall visit decisions through multinomial distributions. Let $\triangle N^i_{O,t}$ define shopper i's choices in a time interval $\triangle^i_t$ from time t as a $K_O + 1$-dimensional binary vector — that is, $\triangle N^i_{O,t} = (\triangle N^i_{O,t,1}, \cdots, \triangle N^i_{O,t,K_O}, \triangle N^i_{O,t,K_O+1})$, where $\triangle N^i_{O,t,1:2}$ denotes shopper i's coupon pull decisions and online store visit decisions; $\triangle N^i_{O,t,3:(K_O+1)} = (\triangle N^i_{O,t,3}, \cdots, \triangle N^i_{O,t,K_O+1})$ denotes shopper i's shopping mall choices; and $\triangle N^i_{O,t,K_O+1}$ denotes shopper i's outside option of visiting a mall. Then, I have

$$\triangle N^i_{O,t,1} \sim Bernoulli\,(p^i_{O,t,1})$$

$$\triangle N^i_{O,t,2} \sim Bernoulli\,(p^i_{O,t,2})$$

$$\triangle N^i_{O,t,3:(K_O+1)} \sim multinomial\,(p^i_{O,t,3:(K_O+1)}), \tag{2.2}$$

where p^i_{O,t,k_O} is shopper i's probability of choosing event type k_O at time t when outside malls.

When inside a mall, a shopper chooses whether to pull coupons, whether to redeem coupons, and whether to exit either a regional mall or a strip mall. I model shoppers' decisions inside a mall through Bernoulli distributions. Let $\triangle N^i_{I,t}$ define shopper i's choices in a time interval $\triangle^i_t$ from time t as a K_I-dimensional binary vector ($K_I = 4$) — that is, $\triangle N^i_{I,t} = (\triangle N^i_{I,t,1}, \triangle N^i_{I,t,2}, \triangle N^i_{I,t,3}, \triangle N^i_{I,t,4})$ where $\triangle N^i_{I,t,1}$ denotes whether shopper i pulls coupons inside a mall; $\triangle N^i_{I,t,2}$ denotes whether shopper i redeems coupons, given that coupons have been pulled; $\triangle N^i_{I,t,3}$ denotes whether shopper i exits a regional mall; and $\triangle N^i_{I,t,4}$ denotes whether shopper i exits a strip mall. Then, I have

$$\triangle N^i_{I,t,1} \sim Bernoulli\,(p^i_{I,t,1})$$

$$\triangle N^i_{I,t,2} \sim Bernoulli\,(p^i_{I,t,2})$$

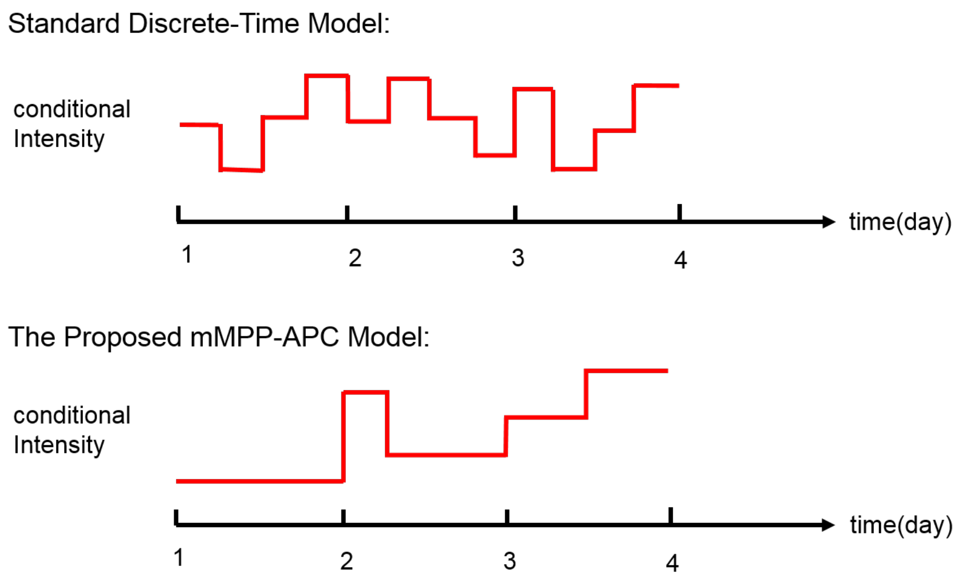
$$\triangle N^i_{I,t,3} \sim Bernoulli\,(p^i_{I,t,3})$$

$$\triangle N^i_{I,t,4} \sim Bernoulli\,(p^i_{I,t,4}) \tag{2.3}$$

Second, to accommodate the drastically different frequencies of event interactions, I let the conditional intensity be adaptively piecewise-constant. Standard discrete-time point process uses the same length of time interval as the discrete unit (Figure 2.3). In my empirical context, this time unit has to be at the second level to capture interactions that can happen within seconds for example, the ones between mobile pushes and coupon pulls. In this case, 120 days of data for model estimation would result in a total of 45,660,672,000 observations, which is computationally infeasible for a model that captures the dynamic interactions among mobile push, coupon pull, online visit, and shopping mall visit. Instead, I let the conditional intensity at time t change only when t is at the beginning of a day, right after a shopper-initiated event occurred, or right after a shopper received a publisher-initiated push. This adaptive piecewise-constant feature guarantees that I "zoom in" to look into every event and every day while "zooming out" to summarize the days without any events with one observation. For example, in Figure 2.3, when no event occurs in day 1, the conditional intensity stays constant over the entire day and only changes when day 2 begins. In day 2, only one event occurs, so the intensity changes only once during day 2 and jumps again when day 3 starts. This feature reduces the number of likelihood evaluations by more than 100,000 times, which makes the model feasible to estimate.

Because the conditional intensity is piecewise-constant, I can look into each time period where the intensity is constant and determine whether an event will occur. The probability that shopper i has event type k_O when

Figure 2.3: Illustration of Conditional Intensity in a Standard Model and in the Proposed Model



outside malls in the time period $\triangle_t^i$ from t can be denoted as

$$\begin{aligned}
p^i_{O,t,k_O} &= \frac{\lambda^i_{O,t,k_O}\triangle^i_t}{1+\lambda^i_{O,t,k_O}\triangle^i_t},\ k_O = 1,2;\\
p^i_{O,t,k_O} &= \frac{\lambda^i_{O,t,k_O}\triangle^i_t}{1+\sum_{k'_O}^{K_O}\lambda^i_{O,t,k'_O}\triangle^i_t}\ k_O = 3,\cdots,K_O;\\
p^i_{O,t,(K_O+1)} &= \frac{1}{1+\sum_{k'_O}^{K_O}\lambda^i_{O,t,k'_O}\triangle^i_t},
\end{aligned} \tag{2.4}$$

where λ^i_{O,t,k_O} is shopper i's conditional intensity of event type k_O when outside the malls and $\triangle_t^i$ is the length of the time interval from time t.

The probability that shopper i has event type k_I when inside a mall in the time period $\triangle_t^i$ from t can be denoted as

$$p^i_{I,t,k_I} = \frac{\lambda^i_{I,t,k_I}\triangle^i_t}{1+\lambda^i_{I,t,k_I}\triangle^i_t},\ k_I = 1,2,3,4 \tag{2.5}$$

where λ^i_{I,t,k_I} is shopper i's conditional intensity of event type k_I when inside a mall and $\triangle_t^i$ is the length of the time interval from time t. Because I specify an exponential function for the conditional intensity λ^i_{O,t,k_O} in Equation 2.6 and λ^i_{I,t,k_I} in Equation 2.7, Equation 2.4 and Equation 2.5 essentially become multinomial logit specifications.

Third, to allow both mutually exciting and mutually inhibitory interactions, and to account for the influence of past shopper-initiated events and publisher-initiated pushes, I model the conditional intensity function of shopper-initiated type k_O event of shopper i when outside malls as

$$\lambda^i_{O,t,k_O} = exp\{\delta^i_{O,k_O} + \Pi^i X^i_{O,t,k_O} + \sum_{l=1}^{K_O}\sum_{t>t'} \alpha^i_{O,lk_O} g(t-t'; \beta^i_{O,lk_O}) + \sum_{j=2}^{J}\sum_{t>t'} \gamma^i_{jk_O} Z^i_{jt'} g(t-t'; \phi^i_{jk_O})\} \tag{2.6}$$

for $k_O = 1, \cdots, K_O$. Here, δ^i_{O,k_O} is shopper i's baseline intensity of event type k_O; and X^i_{O,t,k_O} is shopper i's covariates, such as whether the time is on a weekend. The coefficient matrix Π^i measures the effects of these covariates on shopper i's intensity of event type k_O; $\alpha^i_{O_{lk_O}}$ measures the effects of event type l of shopper i prior to time t on her intensity of event type k_O at time t; and $\gamma^i_{jk_O}$ measures the effects of push type j that shopper i receives prior to time t on her intensity of event type k_O at time t. In my empirical context, $\gamma^i_{jk_O} = 0$ for $k \neq 1$, which I explain in Section 2.5. The time-decay function $g(.)$ captures the decay effects of previous events over time; $\beta^i_{O_{lk_O}}$ and $\phi^i_{jk_O}$ are the parameters of shopper i for the interaction from event type l and j to event type k_O. The dummy variable $Z^i_{jt'}$ indicates whether shopper i receives a type j push at time t'. To improve the estimation efficiency, I let all the shopper-initiated events related to regional malls share the same values of $\alpha^i_{O_{lk_O}}$, $\beta^i_{O_{lk_O}}$, $\gamma^i_{jk_O}$, and $\phi^i_{jk_O}$ while controlling for event-specific effects through the intercepts δ^i_{O,k_O}, for $k_O = 1, \cdots, K_O$. This simplification substantially reduces the number of interactions in my model without losing all the key substantive elements.

I model the conditional intensity function of a shopper-initiated type

k_I event for shopper i when he is inside a mall as

$$\lambda^i_{I,t,k_I} = exp\{\delta^i_{I,k_I} + \sum_{l=1}^{K_O} \sum_{t>t'} \alpha^i_{O_l I_{k_I}} g(t-t'; \beta^i_{O_l I_{k_I}}) + \sum_{t>t'} \gamma^i_{jk_I} Z^i_{jt'} g(t-t'; \phi^i_{jk_I})\} \tag{2.7}$$

for $k_I = 1, \cdots, K_I$. Similarly, δ^i_{I,k_I} is shopper i's baseline intensity of event type k_I; $\alpha^i_{O_l I_{k_I}}$ measures the effects that event type l outside malls for shopper i prior to time t has on her intensity of event type k_I inside a mall at time t; and $\gamma^i_{jk_I}$ measures the effects that push type j, received by shopper i prior to time t, has on her intensity of event type k_I at time t. In my empirical context, $\gamma^i_{jk_I} = 0$ for $k \neq 1$, which again I explain in Section 2.5.

Equations 2.6 and 2.7 guarantee that the conditional intensity functions are non-negative and allow all α^i and γ^i to be positive or negative. In other words, both mutually exciting and mutually inhibitory interactions are allowed in the model. Notice that if I take the logarithm on both sides of Equation 2.6 or Equation 2.7, then $log(\lambda^i_{k_O})$ or $log(\lambda^i_{k_I})$ becomes linear in the events and covariates, which would simplify the computation in the estimation. The conditional intensity λ^i_{t,k_O} or λ^i_{t,k_I} also can be nicely interpreted as shopper i's propensity of having event type k_O or k_I at time t. In addition, I adopt the exponential decay function as $g(.)$, and $\phi^i_{jk_O}$,$\phi^i_{jk_I}$ are the parameters of $g(.)$.

Recall that the summary statistics in Section 3.3 show a large degree of heterogeneity among individual shoppers. To account for the heterogeneity, I let all the parameters in Equations 2.6 and 2.7 be individual-specific, and I let each set of parameters follow a multivariate normal or log-normal distribution.

Fourth, publisher-initiated behavior-based pushes are not randomly assigned. My discussions with the managers from the publisher reveal that behavior-based pushes are based only on shoppers' past coupon pulls through the mobile app. The timing of a behavior-based push is determined not by the publisher but by retailers and a third-party push delivery provider. I therefore treat the delivery timing as given and model the probability that a shopper receives a behavior-based push as a logistic function of the number of her past coupon pulls:

$$Prob(Z^i_{3:4,t'} = 1) = \frac{exp(X^i_{ht}\omega^i)}{1 + exp(X^i_{ht}\omega^i)}, \tag{2.8}$$

where X^i_{ht} is the number of coupon pulls of shopper i. Based on my discussion with the managers of the publisher, the number of coupon pulls in the past 28 days is a strong indicator of how active a shopper is and therefore is used as the most important factor in assigning a behavior-based push. The vector ω^i measures the effects of the past coupon pulls X^i_{ht}. This vector is individual-specific because the publisher values different types of coupon pulls differently. For example, if a shopper pulls coupons from Macy's ten times while another shopper pulls coupons from Sears, the two have a different likelihood of receiving pushed coupons from Macy's. Ignoring the fact that behavior-based push is not randomly assigned would result in biased estimates. For example, if the publisher sends more behavior-based pushes to relatively more active shoppers, then the estimated effects of behavior-based push are inflated. Therefore, incorporating the "targeting rule" into the model becomes necessary.

Given Equations 2.2 to 2.8, I can write the likelihood function for any realization of shopper i's point process as

$$
\begin{aligned}
&L_i(\alpha^i, \beta^i, \kappa^i, \gamma^i, \varphi^i, \phi^i, \delta^i, \Pi^i, \lambda_p^i, \zeta^i, \omega^i | Data) = \prod_{t=1}^{T} \prod_{k=1}^{K} \prod_{j=1}^{J} Prob(\triangle N_{t,k}^i | Z_{jt'}^i) Prob(Z_{jt'}^i) \\
&= exp\Bigg\{ \sum_{d=1}^{D} \sum_{m=1}^{M_O} \Bigg[\Bigg(-log(1 + \sum_{k_O=3}^{K_O} \lambda_{O,d,k_O,m}^i \triangle_{d,m}^i) + \sum_{k_O=3}^{K_O} \big(log(\lambda_{O,d,k_O,m}^i) + log(\triangle_{d,m}^i) \big) \triangle N_{d,m,k_O}^i \Bigg) \\
&\quad + \sum_{k_O=1}^{2} \big(-log(1 + \lambda_{O,d,k_O,m}^i \triangle_{d,m}^i) + \big(log(\lambda_{O,d,k_O,m}^i) + log(\triangle_{d,m}^i) \big) \triangle N_{d,m,k_O}^i \big) \Bigg] \\
&\quad + \sum_{d=1}^{D} \sum_{m=1}^{M_I} \sum_{k_I=1}^{K_I} \big(-log(1 + \lambda_{I,d,k_I,m}^i \triangle_{d,m}^i) + \big(log(\lambda_{I,d,k_I,m}^i) + log(\triangle_{d,m}^i) \big) \triangle N_{d,m,k_I}^i \big) \\
&\qquad + \sum_{f=1}^{F} \Bigg[\sum_{j=3}^{4} Z_{jf}^i X_h^i \omega^i - log(1 + exp(X_h^i \omega^i)) \Bigg] \Bigg\} \qquad (2.9)
\end{aligned}
$$

The detailed derivation of the likelihood function is presented in Appendix A. Note that my model treats the outcomes in each time period as random events that are influenced by the shopper's history; hence, the probability densities of all the outcomes directly enter the likelihood function. My model also treats behavior-based push as data generated from controlled processes, rather than data that are exogenously given.

To summarize, I construct a discrete-time mMPP-APC model that can accommodate both exciting and inhibitory interactions; I integrate naturally the publisher-initiated pushes with the shopper-initiated events; and I account for the decision process of assigning behavior-based pushes. Given the hierarchical nature of the model, I cast the model in the Bayesian hierarchical framework, which I summarize in Appendix B.

2.5 Model Estimation

To estimate the parameters in the model, I use the MCMC method for Bayesian inference. I apply the Metropolis-Hastings algorithm to sample the individual-level parameters and the Gibbs algorithm to sample the distribution parameters across individuals. The details of the MCMC algorithm are presented in Appendix C. Recall that in my model I consider the interactions among events and distinguish the sequence of events in the interactions for each anonymous shopper. The model also accounts for the time decay of all the interaction effects and looks into each day to determine whether any event will occur. I have about 4,400 shoppers in the data sample, and every individual-level parameter depends on the likelihood function because of the history-dependent feature of the model. These specifications can provide unique insights into how consumers shop with mobile promotions; however, they also pose significant challenges to the model estimation. In fact, without advanced computing techniques, estimating the model with the widely used R programming language within a reasonable amount of time becomes infeasible.

2.5.1 Large-Scale Parallel Supercomputing with R (Rcpp/RcppArmadillo)

I expedited the estimation process by taking the following steps. First, I computed the likelihood function in the more efficient C++ programming language and integrated the C++ code into R through the Rcpp and RccpArmadillo packages. The Rcpp package provides R users with an interface that allows C/C++ code within the R programming environment. Built on the

Rcpp package, the RccpArmadillo package creates an interface between R and the widely used Armadillo library in C++, making matrix operation in C++ straightforward and efficient in R. The application of these two packages led to significant time savings in the likelihood computation in each of the individual parameter estimations. However, these time savings in the estimations still were not significant enough to allow me to obtain the estimates within a reasonable amount of time.

To further expedite the estimation process, I leveraged a high performance computing facility the Stampede Supercomputer at the Texas Advanced Computing Center (TACC) and implemented the MCMC algorithms in an efficient way for parallel computing. I avoided using approximate algorithms, such as embarrassingly parallel computing, which ignores the dependency among the individual-level parameters and leads to finite-sample biases. Instead, I implemented the exact MCMC algorithm efficiently in parallel, which updates the population parameters and feeds back to each of the individual-level parameter samplers in every iteration. As a result, I eliminated the requirement of assigning a large number of individuals to each core, and my algorithm could be scaled to a large number of cores without finite-sample biases. The exact MCMC algorithms are especially attractive for estimating models with long time-series data.

To load the likelihood function in C++ efficiently to each of the cores, I wrote an R package that complies with the C++ function ahead of time and then loaded the package to each of the computing cores. This procedure eliminates the need for a C++ compiler in each core and also saves compilation time. I also used load-balancing techniques to reduce the task imbalance among the cores running in parallel so that I would achieve the maximum efficiency. In the end, I were able to use a single core as the master and to assign all the shoppers to the other 255 cores, thus achieving a speed that was at least 100 times faster than the single-core R environment while maintaining the same level of accuracy. For each model, I ran the sampling chain for 100,000 iterations, discarding the first 90,000 iterations to ensure convergence with visual inspections of the sampling chains.

2.5.2 Model Estimation Results

I estimated the mMPP-APC model using the data from the first 120 days. The posterior means and posterior credible intervals of the interactions among shopper-initiated events are shown in Table 2.3, and the corresponding time-decay parameters are shown in Table 2.4.

Recall from Equations 2.6 and 2.7 that the parameters α_{O,lk_O} and $\alpha_{O_l I_{k_I}}$ model the effects of historical shopper-initiated events on current events outside and inside malls, respectively. According to my discussions with managers from the publisher, online store visits from the mobile app are possible only after a shopper pulls coupons when outside malls, whereas coupon redemptions are possible only after a shopper pulls coupons when inside a mall. I capture these dependencies using the parameters $\alpha_{O_{12}}$ and $\alpha_{I_{12}}$. After pulling coupons outside malls and/or visiting online stores, shoppers might visit a regional mall or a strip mall. Thus, a coupon pull outside mall and an online store visit might influence shoppers' propensity to visit a regional mall or a strip mall; these effects are expressed in the parameters $\alpha_{O_{13}}$, $\alpha_{O_{23}}$ and $\alpha_{O_{17}}$,$\alpha_{O_{27}}$. After visiting a regional mall or a strip mall, shoppers might become more or less likely to visit another mall, which I model with $\alpha_{O_{33}}$, $\alpha_{O_{73}}$ and $\alpha_{O_{37}}$, $\alpha_{O_{77}}$. In addition, when shoppers pull coupons outside malls, they are more likely to pull coupons after arriving at a mall and more likely to pull coupons outside malls for future shopping trips. These effects are represented by $\alpha_{O_{11}}$ and $\alpha_{O_1 I_1}$ in the model. Note that in the data, almost all coupon pulls inside malls are for the current shopping trip. Because such pulls have little to do with future coupon pulls outside malls, I do not allow such effects in the model.

Table 2.3: Interaction Effects of Shopper-Initiated Events

Θ_a	Coupon pulls outside malls	Online store visits	Regional mall visits	Strip mall visits	Coupon pulls inside malls	Coupon redemptions
Coupon pull outside mall	0.435($\Theta_{a_{o_{12}}}$) (0.413, 0.466)	4.462($\Theta_{a_{o_{12}}}$) (4.335,0.4.585)	-0.012($\Theta_{a_{o_{13}}}$) (-0.042, 0.017)	-0.057($\Theta_{a_{o_{17}}}$) (-0.080, -0.036)	0.287($\Theta_{a_{o_1 I_1}}$) (0.249,0.331)	-
Online store visit	-	-	-0.133($\Theta_{a_{o_{23}}}$) (-0.171, -0.088)	-0.215($\Theta_{a_{o_{27}}}$) (-0.253, -0.168)	-	-
Regional mall visit	-	-	-0.835($\Theta_{a_{o_{33}}}$) (-0.877 -0.789)	0.253($\Theta_{a_{o_{37}}}$) (0.224, 0.279)	-	-
Strip mall visit	-	-	0.040($\Theta_{a_{o_{73}}}$) (0.010, 0.071)	-0.336($\Theta_{a_{o_{77}}}$) (-0.357, -0.315)	-	-
Coupon pull inside mall	-	-	-	-	-	7.699($\Theta_{a_{I_{12}}}$) (7.495,7.938)

Note: Posterior means and posterior credible intervals (in parentheses) are reported.

The estimation results in Table 2.3 are the posterior means of the individual parameters. Although interpret these parameter values as the "average effects" across individuals is tempting, the true "average effects" can only be derived through simulations because of consumer heterogeneity and the dynamic nature of shopper-initiated events. For example, for a group of shoppers, the effect from coupon pull outside mall on regional mall visit might be positive and large relative to their baselines. In contrast, for another group, the effect might be negative and relatively small. For the former group, the relatively large positive effects might translate into a large number of regional mall visits after these shoppers pull coupons outside malls if coupon pulls outside mall are strongly self-reinforcing. Meanwhile, the latter group might hardly see a reduction of regional mall visits after pulling coupons outside malls if the coupon pulls are weakly self-reinforcing. Overall, the true "average effects" might be positive even if the "average effects" from Table 2.3 are negative. Therefore, I refrain from interpreting these parameters here and measure the true "average effects" through simulations in Section 2.6. Note that the positive and negative estimates in Table 2.3 highlight the necessity of accounting for both inhibitory and exciting effects in studying the effect of mobile promotions on foot traffic. The significant effects between the same types of events demonstrate the importance of incorporating the dynamic interactions among both the same types and different types of shopper-initiated events.

The time decay parameters in Table 2.4 reveal how the interaction effects from Table 2.3 vary over time. Again, all the parameter values are the posterior means of the individual time-decay parameters, and I restrain refrain from interpreting them here.

Table 2.4: Time-Decay Parameters of Shopper-Initiated Events

Θ_a	Coupon pulls outside malls	Online store visits	Regional mall visits	Strip mall visits	Coupon pulls inside malls	Coupon redemptions
Coupon pull outside mall	0.546 (0.518, 0.586)	0.011 (0.010,0.013)	1.713 (1.289, 2.159)	5.845 (4.558, 7.267)	0.337 (0.311,0.382)	-
Online store visit	-	-	2.192 (1.910, 2.500)	5.327 (4.419, 6.437)	-	-
Regional mall visit	-	-	83.445 (70.165, 98.939)	0.966 (0.888, 1.065)	-	-
Strip mall visit	-	-	3.957 (3.352, 4.826)	142.935 (119.882, 170.750)	-	-
Coupon pull inside mall	-	-	-	-	-	0.720 (0.592,0.833)

Note: Posterior means and posterior credible intervals (in parentheses) are reported.

Behavior-based push and location-based push are targeted advertising, initiated by the publisher and delivered to shoppers' mobile phones through the mobile app. According to the publisher, shoppers who use the mobile app for discounts and promotions are very unlikely to visit a mall after receiving a targeted behavior-based push message if they have not pulled coupons from the app. The reason is that these savvy shoppers actively look for discounts or information before shopping. In addition, the push message itself almost never informs shoppers whether the offer is for online shopping or for offline shopping. The only way shoppers can find out is to see the details of the coupon by launching the mobile app. I therefore consider only the direct effects of the two types of behavior-based push on coupon pulls (i.e.,$\gamma^i_{jk_O} = 0$ for $k_O \neq 1, 2$). I account for the direct effects of location-based push on coupon pulls inside malls because when shoppers are inside a mall, location-based push might help to remind them to use the mobile app and so indirectly increase coupon redemptions. Meanwhile, I examine the indirect effects from behavior-based push on regional mall and strip mall visits through coupon pulls, as well as the indirect effects from location-based push to coupon redemptions, in Section 2.6.

Table 2.5: Interaction Effects of Publisher-Initiated Activities

Θ_γ	Coupon pull outside mall	Coupon pull inside mall
Behavior-based push	1.404 (1.351, 1.449)	-
Location-based push	-	0.386 (0.355, 0.418)

Note: Posterior means and posterior credible intervals (in parentheses) are reported.

Table 2.6: Time-Decay Parameters of Publisher-Initiated Activities

Θ_γ	Coupon pull outside mall	Coupon pull inside mall
Behavior-based push	1.991 (0.824, 3.103)	-
Location-based push	-	0.307 (0.158, 0.608)

Note: Posterior means and posterior credible intervals (in parentheses) are reported.

I present the interaction effects from behavior-based push and location-based push to coupon pulls in Table 2.5 and their corresponding time-decay parameters in Table 2.6. The estimates in these two tables are all positive and significant, but I do not interpret each of the parameter values here because they are the posterior means of the individual parameters across individuals. I quantify the effect of behavior-based push and location-based push on shopper-initiated events in Section 2.6.

In addition to the influence of publisher-initiated behavior-based push and location-based push, shoppers' propensity to visit a mall is also influenced by whether it is a weekend. My exploratory analysis shows that shoppers tend to shop more on weekends than on weekdays, I therefore include a weekend dummy variable as a covariate. I present the estimation results of the covariate parameter in Table 2.7.

Table 2.7: Effects of Covariates

Covariates	Coupon pull	Regional mall visit	Strip mall visit
Weekend	0.161	0.378	0.228
	(0.127, 0.188)	(0.339, 0.418)	(0.202, 0.250)

Note: Posterior means and posterior credible intervals (in parentheses) are reported.

The results in Table 2.7 seem to suggest strong weekend effects. Because these values are posterior means of the individual parameters, I defer the interpretations of the parameters to Section 2.6 and explore the influence of the covariate through simulations.

As discussed in Section 2.4, I account for the targeting rules for behavior-based push to address endogeneity concerns. I find that the probability of receiving a behavior-based push without any coupon pulls is low, and that shoppers who are more actively pulling coupons are more likely to receive behavior-based push messages, which is consistent with the rules the publisher has set up. I omit the estimates of the baseline parameters, which are available from the author upon request.

2.6 Managerial Implications

In this section, I discuss the simulations I conduct and the managerial implications I derive. In my proposed mMPP-APC model, earlier events always influence later ones. To correctly evaluate the effects of any particular event, I have to account for both direct effects and indirect effects. For example, as shown in Section 2.5, a coupon pull outside mall directly influences a shopper's propensity to visit a regional mall. In addition to the direct effects, a coupon pull directly influences a shopper's propensity to visit a strip mall and then through the strip mall visit indirectly influences his propensity to visit a regional mall. Or a coupon pull outside mall directly influences a shopper's propensity to visit online stores and then indirectly influences his propensity to visit a regional mall. I are also interested in certain indirect effects. For example, I are interested in how behavior-based push influences regional and strip mall visits through coupon pulls outside malls. I also want to see how location-based push influences coupon redemptions through coupon pulls inside malls. In this section, I simulate the sequence of events over time for each shopper according to my model in Section 2.4 to account for both direct effects and indirect effects of an event or activity. In the following, I first describe how I quantify the true "average effects" across individuals, including both the direct effects and indirect effects through simulations, and then I present the effects of a coupon pull outside mall, a coupon pull inside mall, a behavior-based push, a location-based push, an online store visit, a regional mall visit, and a strip mall visit. I also quantify the effects of a weekend.

2.6.1 Quantifying the "Average Effects"

Given the parameter posterior from the model estimation results, I simulate the sequence of events in 15 days for 1,000 times per shopper in two different cases: base case and treatment case. In the base case, I use only the baselines and covariates to construct the conditional intensity function at the initial time point t_0 and then simulate the sequence of events from t_0 using the parameter posterior. Because no events occur prior to time t_0, I can eliminate the influence of any previous events on the sequence of events in the next 15 days from t_0. In other words, the sequence is only driven by the baselines and covariates. In the treatment case, similar to the base case, I also use only the baseline and covariate parameter estimates to construct the conditional intensity at the initial time point t_0 and then simulate the sequence of events from t_0. In addition, each shopper receives a "treatment" at the initial time point t_0. This "treatment" might be a coupon pull outside mall, a behavior-based push, a location-based push, and so on. Note that the base case when shoppers are outside malls at t_0 should be different from the case when shoppers are inside a mall at t_0. Therefore, I construct two separate base cases: one conditional on shoppers' being outside malls at t_0, the other conditional on shoppers' being inside malls at t_0. The former case is used to quantify the effects of a coupon pull outside mall, an online store visit, and a behavior-based push. The latter is used to evaluate the effects of a coupon inside a mall, a location-based push, and a regional mall or strip mall visit.

The simulation algorithm takes into account both the direct and indirect effects of events. By comparing the number of events of each type in the treatment case with the number of events in the corresponding base case, I can then find out the true "average effects" that the "treatment" has on each type of shopper-initiated event during the 15-day period. I choose 15 days because shoppers on average visit a regional mall about 0.38 times per week, and I want all the direct and indirect effects to play out in the simulation while not letting the baselines play a dominant role.

2.6.2 Effects of Coupon Pulls

In this section, I present the "average effects" of a coupon pull outside mall and a coupon pull inside mall, quantified through a series of simulations, as shown in Table 2.8. By comparing the "treatment" of a coupon pull outside mall with the base case when shoppers are outside malls at t_0, I find the following results. First, coupon pulls outside malls have strong self-reinforcing effects. A coupon pull outside mall increases future coupon pulls by more than 239% in the next 15 days, which indicates that marketing actions from the publisher that can drive immediate coupon pulls are likely to have a much greater total effect on coupon pulls. Second, because online store visits depend on coupon pulls outside malls, not surprisingly, I find that a coupon pull outside malls greatly affects online store visits. A single coupon pull outside mall can increase shopping traffic to online stores by more than 234%. Third, I find that following a coupon pull outside mall, the average foot traffic to regional malls in the next 15 days is about 55% higher than without a coupon pull outside malls. In contrast, the average foot traffic to strip malls is only less than 1% higher. These results suggest that a coupon pull outside malls is a stronger indicator of regional mall shopping, but much less so of strip mall shopping. Fourth, when shoppers pull coupons outside malls, their coupon pulls inside malls increase by about 200%, and their coupon redemptions increase by more than 130%, which demonstrate that coupon pulls outside malls are also a strong indicator of coupon pulls inside malls and coupon redemptions. Note that these large effects of coupon pulls outside malls mainly result from the fact that shoppers tend to be in the shopping mode when pulling coupons and therefore are more likely to go shopping either online or offline after using the mobile app.

Table 2.8: Effects of Coupon Pulls

Treatment	Coupon pulls outside malls	Online store visits	Regional mall visits	Strip mall visits	Coupon pulls inside malls	Coupon redemptions
A coupon pull outside mall	45.447 (239.144%)	41.816 (234.605%)	0.338 (54.870%)	0.077 (4.428%)	3.502 (199.090%)	1.089 (131.522%)
A coupon pull inside mall	0 (1.353%)	2.082 (1.323%)	87 (0.687%)	2.931 (0.667%)	9,167 (0.596%)	(128.319%)
Baseline (outside mall)	19.004	17.824	0.616	1.739	1.759	0.828
Baseline (inside mall)	17.737	16.632	0.582	1.648	1.677	0.791

Next, I turn to coupon pulls inside malls and compare the "treatment" of a coupon pull inside a mall with the base case when shoppers are inside a mall at t_0. I find the following effects. First, a coupon pull inside a mall has negligible effects on coupon pulls outside malls, online store visits, regional mall visits, and strip mall visits, which is consistent with the fact that coupon pulls inside malls are often related to the current shopping trip and have little to do with future activities outside malls. Second, when shoppers pull coupons inside malls, their coupon redemptions are more than 128% higher than without pulling coupons, which shows that that shoppers pull coupons mainly because they seek money savings through coupon redemptions.

2.6.3 Effects of Behavior-Based Push and Location-Based Push

I quantify the effects of behavior-based push and location-based push in Table 2.9 by comparing them with the two base cases, respectively. I find that a behavior-based push leads to an increase in coupon pulls outside malls of almost 25%, about 24% more shopping traffic to online stores, an increase in foot traffic to regional malls of about 5%, and about 0.5% more foot traffic to strip malls. In addition, it increases coupon pulls inside malls by more than 19% and improves coupon redemptions by about 18%. These results demonstrate that behavior-based push not only substantially helps to engage shoppers through the mobile app but also significantly improves shopping traffic to online s-tores and offline malls, as well as coupon redemptions offline. Interestingly, behavior-based push has a much larger effect on foot traffic to regional malls than to strip malls, which highlights the importance of separating regional malls from strip malls in the analysis. This result might be driven partly by the fact that regional malls tend to have many more retail stores and more promotional content on the mobile app and that shoppers are more likely to find what they need in regional malls than in strip malls.

Table 2.9: Effects of Behavior-Based Push and Location-Based Push

Treatment	Coupon pulls outside malls	Online store visits	Regional mall visits	Strip mall visits	Coupon pulls inside malls	Coupon redemptions
Behavior-based push	4.723 (24.853%)	4.188 (23.496%)	0.030 (4.870%)	0.009 (0.522%)	0.342 (19.443%)	0.151 (18.237%)
Location-based push	0.125 (0.705%)	0.110 (0.661%)	0.004 (0.835%)	0.008 (0.485%)	0.672 (40.072%)	0.201 (25.411%)
Baseline (outside mall)	19.004	17.824	0.616	1.739	1.759	0.828
Baseline (inside mall)	17.737	16.632	0.582	1.648	1.677	0.791

In contrast, location-based push has little effect on coupon pulls outside malls, on shopping traffic to online stores, and on foot traffic to regional malls and strip malls. However, when shoppers receive a location-based push, they pull coupons inside malls about 40% more and redeem coupons about 25% more than when they don't receive a location-based push. The significant effects of location-based push inside malls suggest that location-based pushes might serve as a reminder for shoppers to look for coupons and redeem them during shopping.

Table 2.9 shows that behavior-based push and location-based push play different roles in influencing shopper-initiated activities. Behavior-based push starts influencing shopping behavior when shoppers are outside malls, bringing them to online stores or shopping malls, and it continues its influence even when shoppers are shopping inside malls. In contrast, the influence of location-based push, which is delivered to shoppers only after they enter a regional or strip mall in my empirical context, is limited to shopping behavior inside malls.

2.6.4 Effects of Online and Offline Visits

Similar to the effects of coupon pulls, I quantify the effects of an online visit and an offline visit to a regional mall or a strip mall in Table 2.10 First, I find that after visiting online stores, shoppers visit a regional mall 25% more often and visit a strip mall about 3% more often, which demonstrates the synergy between online and offline shopping activities. Second, after shoppers visit a regional mall, they visit strip malls 19% more often. After shoppers visit a strip mall, I see a 14% increase in visits to a regional mall . These results indicate strong complimentarity between regional malls and strip malls, which might result from the fact that a regional mall is almost always surrounded by a few strip malls. Third, I find that a regional mall visit is followed by an increase in coupon redemptions of about 3% while a strip mall visit is followed by less than 0.4% increase, suggesting that shoppers are more likely to redeem coupons when they are shopping in regional malls. Note that shoppers tend to pull fewer coupons outside malls and to visit online stores less after visiting a regional mall or a strip mall, but the size of the negative influence is relatively smallfrom about 0.8% to 2.8%.

Table 2.10: Effects of Online and Offline Visits

Treatment	Coupon pulls outside malls	Online store visits	Regional mall visits	Strip mall visits	Coupon pulls inside malls	Coupon redemptions
Online store visit	-0.054	-0.021	0.154	0.059	-0.009	-0.005
	(-0.284%)	(-0.118%)	(25.000%)	(3.393%)	(-0.512%)	(-0.604%)
Regional mall visit	-0.497	-0.465	-0.009	0.319	0.009	0.023
	(-2.802%)	(-2.796%)	(-1.546%)	(19.357%)	(0.537%)	(2.908%)
Baseline (outside mall)	19.004	17.824	0.616	1.739	1.759	0.828
Baseline (inside mall)	17.737	16.632	0.582	1.648	1.677	0.791

2.6.5 Effects of a Weekend

Recall that the estimates in Table **??** seem to suggest positive weekend effects. I quantify the weekend effects by comparing the case that includes one fewer weekend with the base case of a 15-day period, which is presented in Table 2.11.

Table 2.11: Effects of One Weekend

Treatment	Coupon pulls outside malls	Online store visits	Regional mall visits	Strip mall visits	Coupon pulls inside malls	Coupon redemptions
Weekend	0.349 (1.836%)	0.340 (1.908%)	0.015 (2.435%)	0.028 (1.610%)	0.033 (1.876%)	0.016 (1.932%)
Baseline (outside mall)	19.004	17.824	0.616	1.739	1.759	0.828

Table 2.11 shows that a weekend has positive effects on shopper-initiated events, ranging from a 1.610% increase to a 2.435% increase, compared to the baseline. This finding is consistent with my expectation that shoppers tend to shop more on weekends and therefore pull and redeem coupons more often on weekends.

Chapter 3

Ranking and Personalizing Organic and Sponsored Mobile Advertising

3.1 Background

In the age of Internet, advertisings, such as search advertisings from Google Adwords, are often presented to consumers as listings. An ad's position of the listing is critical for its performance. For example, in Google search advertisings, the clickthrough rate for ads in position one on average is more than twice as high as position two and more than three times as higher as position three (Insights, 2013). Previous research on keyword search advertising has also documented that ads at higher positions attract more clicks from consumers (Ghose and Yang, 2009). Such ranking effects are even higher on mobile phones where links that appear at the top of the screen are much more likely to be clicked (Ghose et al., 2013). Another important factor on mobile phones is screen size. A smaller screen displays fewer ads at a time, so consumers have to scroll down more often to go through all the ads available, leading to lower clickthrough rate for all the ads. On the other hand, these consumers may be more focused on what is being displayed on the screen than those who own phones with big screens, therefore driving up the clickthrough rates.

To improve consumer response rate to advertising, publishers are increasingly designing sponsored contents in a similar style and format to the organic contents — “native advertising” — on their digital platforms. For example, sponsored search advertisings on Google mimic the style and format in the organic search results; Amazon presents sponsored products in the same way as those non-sponsored products that sell on Amazon.com. In these cases, however, publishers distinguish sponsored contents with “Ad” or “Sponsored” to disclose to consumers that the sponsored contents are paid by a third-party. The sponsored contents are often placed separately from the organic contents: sponsored ads on Google appear on the top of the screen followed by organic search results; sponsored products on Amazon.com are located at the bottom of the screen after the non-sponsored products. The separation of sponsored contents from organic ones prevents marketers from examining how consumers behave when the two types of contents are mixed together. It is therefore intriguing to study how to rank organic and sponsored contents when they are allowed to be mixed together.

In addition, consumers are more interested in receiving personalized communications, and marketers agree that personalized contents are more effective than uncustomized ones in driving response rate (eMarketer, 2014a). While it is difficult to personalize sponsored search advertising and organic search results due to the vast variety of consumers' search queries, marketers are very capable of delivering personalized contents like mobile ads to consumers, often through a mobile app. Due to the high consistency in contents on a mobile app over time, marketers are able to infer consumers' preferences from past history and therefore pair consumers with contents they are very likely to engage with. Therefore, personalization provides marketers another lever for ranking of sponsored and organic mobile ads.

In this study, I study ranking and personalization of mobile advertising using an unique data set from a publisher that present mixed sponsored and organic mobile ads to consumers. Sponsored ads pay a placement fee to the publisher while organic ads were selected by the publisher for the consumers' benefits. The publisher ranks sponsored and organic ads by balancing placement revenue from the sponsored ads against consumer engagement from the organic ads. In addition, the ranking of these ads is personalized: each consumer may receive the ads in a different order on the same day. The unique context provides me an opportunity to study the following research questions:

1) Given the sponsored ads, how should the publisher personalize the organic ads?

2) Given the sponsored and organic ads, how should the ads be ranked?

To answer these research questions, I develop a consumer information search model where a forward-looking consumer maximize her utility by choosing whether to click an ad or slide the screen to move down the viewable ads on the screen. My model, built upon previous research on consumer information search, incorporates unique factors such as mobile phone screen size and consumers' screen sliding actions, which pose unique methodological challenges.

Since the model is based on the utilitymaximizing framework, I can conduct counterfactual simulations to investigate policies that the publisher may use to rank the ads and to personalize the organic ads. To the best of my knowledge, this study is the first in marketing to examine consumer information search behavior under mixed organic and sponsored contents. It is also the first to study how mobile phone screen sizes influence consumer search. My research aims to compare different ranking and personalization policies publishers may use given organic and sponsored contents are mixed. My results will provide unique insights into this newly-emerged marketing practice.

3.2 Literature Review

My study is related to the previous literature on the relationship between organic listings and sponsored search advertising. Yang and Ghose (2010) analyze this relationship and find that clickthroughs on organic listings are positively interdependent of sponsored listings and vice versa. Jerath et al. (2014) find that consumer clicks after a search is low and mainly on organic listings but searches of keywords with low search volumes lead to more clicks and a large portion of sponsored clicks. Bentley et al. (2015) study the impact of organic search results on the performance of sponsored search advertising by allowing consumers to learn from organic results. They find that consumers, advertisers and the search engine are all significantly better off when the search engine provides "free" organic results to let consumers learn about the products and services associated with the each query. Different from previous work in this area, I study how organic and sponsored listings influence consumer search when they are mixed together without distinction.

This study is closely related to previous research on consumer information search. Weitzman (1979) studies the sequential search problem where each alternative yields an uncertain reward and show that the optimal stopping rule is to terminate search whenever the maximum reward from the searched alternatives is larger than the reservation price of unsearched alternatives. Kim et al. (2010) model an individual-level optimal sequential search process where each consumer maximizes her expected utility when making choices. Aggregating across consumers over individual choices are corresponding to the observed market-level product search data. They find that consumer search cost is significant and lowered for those products that appear on Amazon.com more frequently. Jeziorski and Segal (2015) show that consumers may click ads in a nonsequential order and the identity of competing ads matters to ad clickthrough rates. They find that more clicks would occur if no competing ads existed and that the optimal matching of ads to positions can significantly raise welfare. Chen and Yao (2016) proposed a consumer sequential search model that can integrate consumers decisions of search and refinement. They find that the refinement tools encourage more searches and increase the utility of purchased products, and that informing consumers of the default ranking rule reduces consumer search but improves consumer welfare. Complement to the previous research on consumer search, my study study consumer information search behavior on mobile devices where consumers are given personalized ranking and contents, and see different number of alternatives on a screen determined by different screen sizes and choose whether to slide the screen to see more alternatives.

This study is also related to previous research on consumer behavior on mobile devices. Ghose et al. (2013) explore the differences in Internet browsing behavior between mobile phones and personal computers and find that links that appear at the top of the screens and stores located close to a users home are more likely to be clicked on mobile phones than personal computers. Daurer et al. (2016) construct a typology for mobile search based on search intensity and location to study when, where and how of consumer search on mobile devices. They find that consumers search at all times and locations, not just when they are in or close to a store. Liu et al. (2017) study the effect of mobile targeting on foot traffic by capturing activities on mobile devices and offline shopping mall activities and find that both behavior-based and location-based targeting can increase consumer responses. My research is the first in marketing to study how different screen sizes influence consumer search behavior on mobile devices

3.3 Data

The data used in this study come from a publisher that lists ads from major non-grocery retailers on its mobile app. These ads typically communicate discount information to consumers and include ads for online shopping and for in-store shopping. For example, “20% off one item in store at Bed

& Bath Beyond" is an ad for in-store shopping while "Free 30 day Prime at Amazon" is an ad for online shopping. The publisher earns revenue by selling some of the ad slots to retailers and engages consumers by adding non-paid organic ads to the rest of the slots. On each day, all consumers receive the same set of ads from the publisher but the ranking of the ads is personalized.

To study the effect of different screen sizes while minimizing the impact of other factors related to mobile phones, I focus on iPhone devices and randomly select 500 consumers from each of the following types: iPhone 4 and iPhone 4s (screen size: 3.5 inches, referred to as iPhone 4 thereafter); iPhone 5 and iPhone 5s (screen size: 4 inches, referred to as iPhone 5 thereafter); iPhone 6, iPhone 6s and iPhone 7 (screen size: 4.7 inches, referred to as iPhone 6); and iPhone 6 Plus, iPhone 6s Plus and iPhone 7 Plus (screen size: 5.5 inches, referred to as iPhone 6 Plus). The 2000 consumers in the sample have used the mobile app for at least 30 days before the observation period. Therefore, they are familiar with the features of the mobile app.

Given a consumer opened the mobile app, she has three different options: she can click any of the viewable ads on the screen, slide the screen up by one slot which throws out the ad on the top position and adds one new ad from the bottom, or exit the mobile app. I observe the ads delivered to each consumer on each day and consumers' screen sliding and clicking behavior on each app usage or session. I also observe the timestamps for each app session, including session start time and session end time. The summary statistics is presented in Table 3.1. Each type of devices have about 10,000 sessions. Consumers with iPhone 4 on average slide the most and the average app session duration is the longest while those with iPhone 6 Plus slide and click the least.

Table 3.1: Summary Statistics

Device	No. of consumers	No. of sessions	Slides per session	Clicks per session	Session duration (in seconds)
iPhone 4	500	9,738	6.247	0.274	2.201
iPhone 5	500	10,869	5.868	0.282	1.405
iPhone 6	500	10,877	5.183	0.272	1.328
iPhone 6 Plus	500	10,273	4.882	0.245	1.339

In addition, I observe whether consumers launched the mobile app inside a shopping mall through the geofence mobile technologies. When a consumer uses the app inside a shopping mall, he or she is probably looking for specific discounts for the current shopping trip. Therefore, consumers' app activities during shopping in malls may well be different from those when they are not. To explore these differences, I separate the data for app activities during shopping and for those not during shopping and present the data summary in Table 3.2. It shows that consumers slide and click more when they are not shopping in malls but stay longer in the app during shopping. The number of slides per session decreases over screen size almost monotonically while the number of clicks per session changes non-monotonically.

Next, I turn to consumers' last action before exiting from the mobile app. Table 3.3 shows the following: among over 230,000 slides, consumers choose to exit after sliding the screen about 17% of the times, whereas they exit over 43% of the times after clicking a viewable ad on the screen. In other

Table 3.2: Summary Statistics of Shopping and Non-Shopping

Device	Shopping	No. of sessions	Slides per session	Clicks per session	Session duration (in seconds)
iPhone 4	No	7,451	6.687	0.298	1.970
iPhone 4	Yes	2,287	4.812	0.195	2.953
iPhone 5	No	6,961	6.689	0.340	1.027
iPhone 5	Yes	3,908	4.407	0.177	2.078
iPhone 6	No	6,584	6.081	0.346	0.981
iPhone 6	Yes	4,293	3.805	0.157	1.860
iPhone 6 Plus	No	6,139	5.802	0.299	0.968
iPhone 6 Plus	Yes	4,134	3.517	0.165	1.890

words, consumers are more likely to exit after clicking an ad than after sliding the screen. Clicking an ad may satisfy or disappoint consumers, leading to the termination of the current session. The termination resulted from the ads on the top positions may reduce consumers' responses to the ones on the lower positions. Therefore, the ads on the screen may not independent.

Table 3.3: Last Action Before Exit

Action	No. of actions before exit	Total No. of actions	Percentage
Slide	39,184	231,137	16.953%
Click	4,854	11,202	43.332%

I also have information about whether the discount displayed on an ad is for online shopping or for in-store shopping. Table 3.4 shows the exit rates after consumers click an ad for online shopping and for in-store shopping. It demonstrates that consumers are more likely to exit after clicking an ad for online shopping than clicking one for in-store shopping. It may be driven by the design of the mobile app. Ads for online shopping are often diverting consumers to retailers' websites while ads for in-store shopping usually reveal the details about the discount within the mobile app. Therefore, taking into account the ad type could significantly influence consume responses to the ads, and potentially improve publisher's revenue.

Table 3.4: Ad Type and Clik to Exit

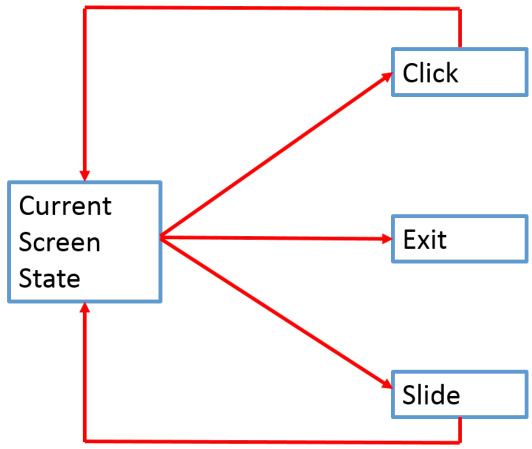
Ad type	Click to exit	Percentage
Ad for in-store shopping	1,742	35.888%
Ad for online shopping	3,112	64.112%

To measure the discount depth of an ad, I obtain the proprietary ad discount quality score data from the publisher. In addition, the publisher has also constructed a consumer affinity score for each consumer to each category based on each individual's historical app activities.

3.4 Model

In this section, I propose a consumer mobile search model that can account for the unique factors in my empirical context and answer the research questions in Section 3.1. I focus on the consumer information search process given a consumer has opened the mobile app. Future research may want to consider the dynamic decision process of mobile app openings. Recall that in any given period after opening the mobile app, a consumer can click any of the viewable ads on the screen, slide the screen to reveal one more ad for the bottom, or exit the app. After a consumer clicking an ad, the viewable ads and their positions on the screen does not change. In contrast, sliding the screen moves the current viewable ads up by one position, throwing the ad on

Figure 3.1: Consumer Information Search Process



the top position out of the screen and adding a new add to the bottom of the screen. Exiting the app terminating the search process. This dynamic search process is shown in Figure 3.1.

Consider consumer i who receives J_s number of ads on a mobile phone with a screen size that can display s number of ads from the top to the bottom. Consumer i's utility of clicking ad $j = 1, \cdots, J_s$ is

$$u_{ij} = v_{ij} + \epsilon_{ij} \tag{3.1}$$

with

$v_{ij} = X_{ij}\gamma_i, \epsilon_{ij} = N(0, \sigma_{ij}^2)$

where X_{ij} is a row vector of ad characteristics to consumer i and γ_i is a row vector that captures invididual-specific valuation of the ad characteristics, including discount quality score of an ad, consumer i's affinity to the advertised category, whether consumer i is inside a shopping mall or not, etc.ϵ_{ij} is the idiosyncratic shock to consumer i's valuation about ad j. I assume that the consumer knows her mean valuation v_{ij} but does not know ϵ_{ij}, and the goal of clicking an ad is to resolve ϵ_{ij}.

Consumers search information at a cost. I assume that at period t, the search cost of locating an ad after kth slide c^s_{ijk} is a function of its slot position on the current screen P^k_j since consumers tend to view the screen from the top down on mobile phones (Yao and Mela, 2011). Then, I have

$$-c^s_{ijk} = \theta_i P^k_j + e_{ijk} \tag{3.2}$$

where $k = 0, 1, \cdots, K, \theta_i$ is individual-specific cost parameters on the slot position and e_{ijk} is invididual-specific cost shock, independently distributed across individuals and ads.

In addition to the search cost, consumer i also incur a sliding cost c^d_{ij} when she slides up the ads by one slot.

Therefore, when consumer i click a subset $C = C_0 \cup C_1 \cup \cdots \cup C_K$ of the total ad slots, she obtatins a gross utility of

$$U_i(C) = ((\sum_{j \in C_0} u_{ij}^{1+H_i})^{\frac{1}{1+H_i}} - \sum_{j \in C_0} c^s_{ijk}) + \sum_{k=1}^{K} d_{ik}((\sum_{j \in C_k} u_{ij}^{1+H_i})^{\frac{1}{1+H_i}} - \sum_{j \in C_k} c^s_{ijk} - c^d_{ij}) \tag{3.3}$$

where the first bracket contains the utility of clicking on the first viewable screen C_0 while the second bracket shows the ulitility of clicking on the kth viewable screen C_k. Note that consumers can slide up the das by only one slot a time, therefore the ads after a slide overlap with the ones before the slide but are displayed in different positions on the screen. $d_{ik} = 1$ indicates consumer i slides up the ads by one slot. K is the number of viewable screens consumer i exposes to. H_i is a parameter that captures the substitutability of different ads to consumer i (Jeziorski and Segal, 2015). When $H_i = 0$, the utility of clicking becomes additive, meaning that consumer i's clicking decisions on different ads are independent. When $H_i > 0$, the ad clicks are substitues. When $H_i < 0$, they are complements.

Given the gross utility function, I now look at he timing of consumers' decisions. Consumer i make search decisions in the following order:

(i) Consumer i receives ads $j = 1, \cdots, J_s$

(ii) At time t, Consumer i reads brief descriptions of the s viewable ads on the screen to learn their mean valuations v_{ij}.

(iii) Consumer i decides whether to click on ad j in a position P_j^k on the current screen

(iv) Consumer i observes ϵ_{ij} of a clicked ad j on the current screen

(v) Consumer i decides whether to slide up the ads by one slot or exit if none of the ads on the current screen were clicked on.

(vi) $t = t + 1$, go to (ii)

I assume that consumers are forward-looking and maximizing their expected utility. The consumer search process can be modeled as a dynamic programming problem. The value function of consumer i in state (C, k) must satisfy the following Bellman equation:

$$V_i(C,k) = max[U_i(C),\ max_{c\in C_k, c\in \bar{C_k}}\ EV_i(C\cup c,k),\ max\ EV_i(C,k+1)] \quad (3.4)$$

where $k = 0, 1, \cdots, KK$. KK is the largest possible number of slides, which equals to the number of delivered ads minus one in my empirical setting. $\bar{C_k}$ denotes the subset of viewable ads after the kth slide that have not been clicked. $U_i(C)$ is the utility of exiting after clicking a subset of ads C. $max_{c\in C_k, c\in \bar{C_k}} EV_i(C\cup c,k)$ is the utility of clicking an additional ad displayed on the current screen while $max\ EV_i(C,k+1)$ is the utility of sliding up the ads by one slot.

3.5 Model Estimation and Counterfactual Simulations

I will estimate the model using Bayesian estimation which allows for rich unobserved heterogeneity. I first draw consumer-specific parameters and then compute each consumer's optimal policy by solving Equation 3.4 using backward induction. Given the parameters estimated from the proposed consumer information search model, I will conduct counterfactual simulations to study the impact of personalization in ad ranking and ad contents. These simulations will provide the publisher guidance on how to personalize ad ranking and contents to encourage consumer search.

To provide a benchmark of comparison, I first simulate consumers' sliding and clicking behavior according to the scenario where no personalization is applied to either the ad ranking or the ad contents. In other words, both the organic ads and sponsored ads are the same across individuals and are as given in the data while the ranking of those ads are ranked randomly or by the discount quality score to the advertised category. This model can be considered as the "NULL" model and show us how consumers respond to delivered ads which cannot be observed from the data.

Next, I will examine the impact of different personalization strategies. First, the publisher may personalize the ad ranking using each consumer's affinity score, but select the same set of ads for all consumers. Both the organic and sponsored ads are as given in the data. This scenario will show us the impact of ranking personalization by consumers' affinity score. Second, the publisher may explore the influence of personalized ad contents. Since the publisher receives revenue from the sponsored ads and often want to maximize the number of consumers who are exposed to those ads, he may have to keep the sponsored ads, as given in the data, in every consumer's app. The publisher, however, is able to personalize the organic ads delivered to consumers. He may select the same number of organic ads as given in the data, using consumers' affinity score or using methods like collaborative filtering. By simulating consumers' search behavior under personalized contents and compare with scenarios where every consumer receives the same contents, I will find out whether and how much content personalization benefit the publisher. In addition, on top of personalized organic ad contents, consumers may also receive the selected ads ranked by consumers' affinity score or the ad discount quality score. Third, the publisher may also personalize the organic ads by looking at whether the advertised retailer is close by and rank the delivered ads by ad type, i.e., whether an ad is for online shopping or for in-store shopping. Recall that data description in Section 3.3 shows that consumers behave very differently when they are in a shopping mall from when they are not. Therefore, presenting consumers ads from retailers that have stores in the same shopping mall may increase consumers' response to the organic ads, which may also lead to increased response to the organic ads. Section 3.3 also shows that consumers are more likely to exit after clicking an ad for online shopping than for one for in-store shopping. The publisher may rank the ones for in-store shopping higher on the screen to reduce consumer exits, given all the other factors are the same. Fourth, the publisher may vary the number of organic ads by device screen size and individual behavior. As is shown in Section 3.3, screen size has significant impact on consumers' sliding and clicking behavior by allowing different number of ads to appear on the screen. Meanwhile, individuals using the same type of device may be very heterogenous. Taking into account both the device types and individual behavior may help improve consumer responses to both organic and sponsored ads.

The counterfactual simulations I propose can answer many of the relevant questions the publisher wants to ask but could not test empirically. They will provide the publisher rich insights into how to change their current practice to improve the consumer response to both organic and sponsored ads.

Chapter 4

Conclusions

This dissertation examines the marketing implications of mobile targeting and personalization in mobile marketing to enhance the understanding of these newly-emerged marketing levers.It quantifies the impact of mobile promotions on consumers' shopping behavior, and provides a framework to marketers to rank and personalize mobile advertising.

Chapter 2 studies the dynamic interactions among shopper-initiated coupon pulls, online store visits, shopping mall visits, coupon redemptions, and publisher-initiated location-based and behavior-based pushes to connect mobile promotions to foot traffic to shopping malls. I contribute to the literature both methodologically and substantively.

Methodologically, I develop a novel multinomial multivariate point process with adaptive piecewise-constant intensity, "zooming in" to the days on which events occur and "zooming out" from the days without events to extract information from the data efficiently. My model framework provides a solution to modeling multivariate data where different types of interactions occur with dramatically different frequencies. Considering the influence from past events reveals unique insights into the effects of mobile promotions on consumers' dynamic shopping behavior over time. Through a specification similar to the proportional hazard function in the literature, my model can accommodate both mutually exciting and mutually inhibitory interactions, which expands the applications of standard point processes. I estimate the model through large-scale parallel supercomputing on one of the most powerful supercomputers in the world. Using the state-of-the-art computing techniques and facility, I efficiently implement the exact MCMC algorithm in parallel and scale to a large number of computing cores. My estimation procedures can be directly applied to numerous big data marketing problems without change.

Substantively, I generate interesting insights into mobile marketing. A behavior-based push leads to an increase in coupon pulls of about 25% and about a 24% increase in shopping traffic to online stores. Behavior-based push is effective in increasing foot traffic to regional shopping malls. I find that a single behavior-based push would result in an increase in foot traffic to regional malls of almost 5% but only a 0.5% increase in foot traffic to strip malls. A behavior-based push increases coupon pulls inside malls by about 19% and coupon redemptions by about 18%. In contrast, a location-based push leads to about a 40% increase in coupon pulls inside malls and an increase of about 25% in coupon redemptions.

In addition, I find that after pulling coupons outside malls, shoppers are substantially more likely to visit online stores, visit a shopping mall, and redeem coupons. In contrast, shoppers are more likely to redeem coupons after pulling coupons inside a mall. Shoppers are more likely to shop in a regional mall or a strip mall after visiting online stores. Shoppers are more likely to visit a strip mall after a regional mall, and vice versa.

This study also has several limitations, which are offered as opportunities for future research. First, my data record only the mall visits of anonymous shoppers and a proxy of coupon redemptions inside malls, which prevents me from evaluating the effect of mobile promotions on offline transactions. To connect mobile promotions to offline transactions, the publisher has to sign a contract with almost every retailer in the United States to acquire transaction-level data with unique consumer identifiers that can be linked to mobile app users. In addition, access to the complete transaction data of each retailer is necessary for researchers to determine the effects of mobile promotions. When such data become available, future research might explore how much mobile promotions affect retailer's performance and how retailers should run their mobile promotions. Second, I do not consider the interactions between targeted mobile pushes and shoppers' locations because of a lack of real-time location data, but future research might look into such interactions to share more insights into location-based targeting. For example, shoppers might be more likely to respond to a promotion when it is delivered at work around lunch time but might be more likely to redeem a promotion the next day when it is delivered at home around 8pm. Third, I do not consider the "over-targeting" effects of publisher-initiated mobile pushes. That shoppers might be annoyed when they receive too many targeted push messages is certainly possible. Depending on how useful a mobile app is and how annoying the targeted pushes are, shoppers might choose to share or not to share their locations, to receive targeted pushes, and even to delete the mobile app, which has serious consequences for publishers. Future research might segment shoppers using historical data and explore the optimal cap for each segment through field experiments. Fourth, in my empirical context, shoppers are targeted through behavior-based push using their past browsing history in the mobile app. Future research might study whether a better strategy is to target shoppers using both past browsing history in the app and offline shopping mall visits through field experiments. For example, after spending a few hours shopping at a mall yesterday, a shopper might dislike receiving targeted pushes that inform him or her of promotions inside malls, thus resulting in annoying "over-targeting" effects. Fifth, I do not consider the role of content in targeted pushes. Different messages that link to the same promotion might have dramatically different effects on shopping behavior. For example, using language emphasizing the urgency of the promotion (e.g., "today only!" or "hurry!") and adding touchable pictures to the message might significantly influence the results of targeted mobile marketing. Future research might test these factors through large-scale field experiments and find an effective way to deliver targeted push messages. Sixth, the geofence in my context was designed to capture shoppers' mall visits rather than bringing potential shoppers nearby to a shopping mall. Future research might explore the possibilities of targeting shoppers at a geofence with promotions from nearby geofences or construct a larger geofence to include nearby locations in addition to the geofence that captures the precise locations of a shopping mall.

Overall, this study is one of the first in marketing to look beyond "one-shot" experiments and to examine dynamic shopping behavior under behavior-based and location-based mobile targeting. It reveals new insights into mobile marketing, which helps marketers to understand how mobile marketing works. It also provides a new model that can be used to study many of the online-to-offline and offline-to-online marketing problems. The large-scale parallel supercomputing techniques used in this study offer marketing researchers a solution to big data marketing inference.

Chapter 3 studies the ranking and personalization of mobile advertising that mix sponsored and organic mobile ads when delivered to consumers. The publisher faces a tradeoff between placing sponsored ads from retailer to receive revenue from advertised retailers and selecting the right organic ads to keep consumers engaged. The unique empirical context provides me an opportunity to study the impact of ad content personalization and ad ranking personalization. In this paper, I propose a consumer mobile search model that can account for the unique factors in my empirical context and answer my research questions. Through data description, I present model-free evidence for the influence of screen size, whether consumers are in a shopping mall and ad type. I also show how sliding and clicking influence consumers' exit decisions. These model-free evidence provides motivations for me to include these factors in my model. The counterfactual simulations I propose based on model estimates explore different ways of personalization, including

(i) selecting personalized ad contents from the vast amount of available ads by consumers' affinity score, by methods like collaborative filtering, or by whether the advertised retailer has an store in the shopping mall a consumer is in; and (ii) ranking selected ads by consumers' affinity score, by ad discount quality score, or by whether an ad is for online shopping or for in-store shopping. These simulations will provide marketers insights into whether and how much each type of personalization improves consumer responses to both organic and sponsored ads, thereby offering guidance to the publisher in optimizing their current practice. To the best of my knowledge, this study is the first in marketing to examine consumer information search behavior under mixed organic and sponsored contents. It is also the first to study how mobile phone screen sizes influence consumer search. Note that the estimation and counterfactual simulations have not been implemented yet. I will first use simulated data to test my algorithm in estimating the model parameters and to refine my proposed model. Then I will implement my estimation algorithm and conduct the proposed counterfactual simulations.

While Chapter 3 studies interesting research questions about ad personalization on mobile devices, it does have limitations which provide directions for future research. First, this study focuses on consumer information search process given that a consumer has opened the app, which basically assumes that app sessions are independent of each other. However, it is possible that a more engaged consumer from an earlier app session may return to the app sooner, therefore increasing the number of responses like clicks to the sponsored and organic ads in the long term. Second, I do not model consumers' app opening decisions, which are subject to the publisher's marketing activities. For example, the publisher may target consumers through mobile push tools with ads from their favorite retailers and therefore encourage consumers to open the app more often. It would also be interesting to see how to personalize targeted mobile push to maximize consumers' app openings. Third, I observe only whether a consumer is in a shopping mall or not when using the app, which limits my ability to study the effects of distance from consumers to the advertised retailer location. When such data become available, future research may rank ads by distance from consumers' current locations to the nearest retailer store if exits. Fourth, this study examines screen sliding and ad clicks only since I do not have access to purchase data from all the major retailers. In the future, researchers may study how the personalization strategies I propose influence each step in the conversion funnel, if richer data become available. Fifth, this study is based on observational data from the publisher and therefore cannot explore whether presenting the sponsored ads in the way as the organic ads or in a different way is more effective in increasing consumer responses. Future research may conduct field experiments which compares the two different formats to reveal insights into how consumers behave differently in the two conditions.

Appendices

Appendix A

Likelihood Function

The likelihood function for individual i can be written as follows:

$$
\begin{aligned}
& L_i(\alpha^i, \beta^i, \gamma^i, \varphi^i, \phi^i, \delta^i, \Pi^i, \omega^i | Data) \\
& = \prod_{t=1}^{T} \left\{ \left[\prod_{k_O=3}^{K_O} Prob(\triangle N^i_{O,t,k_O}) \prod_{k_O=1}^{2} Prob(\triangle N^i_{O,t,k_O}) \right] \left[\prod_{k_I=1}^{K_I} Prob(\triangle N^i_{I,t,k_I}) \right] \prod_{j=3}^{4} Prob(Z^i_{jt}) \right\} \\
& = \prod_{d=1}^{D} \left\{ \prod_{m=1}^{M_O} \left([1 - \sum_{k_O=3}^{K_O} p^i_{d,m,k_O}]^{1-\sum_{k=3}^{K} \triangle N^i_{d,m,k_O}} \prod_{k_O=3}^{K_O} (p^i_{d,m,k_O})^{\triangle N^i_{d,m,k_O}} \right) \right. \\
& \left([1 - \sum_{k_O=1}^{2} p^i_{d,m,k_O}]^{1-\sum_{k_O=1}^{2} \triangle N^i_{d,m,k_O}} \prod_{k_O=1}^{2} (p^i_{d,m,k_O})^{\triangle N^i_{d,m,k_O}} \right) \\
& \left. \prod_{m=1}^{M_I} \prod_{k_I=1}^{K_I} \left([1 - p^i_{d,m,k_I}]^{1-\triangle N^i_{d,m,k_I}} (p^i_{d,m,k_I})^{\triangle N^i_{d,m,k_I}} \right) \right\} \\
& \prod_{f=1}^{F} (\frac{exp(X^i_h \omega^i)}{1 + exp(X^i_h \omega^i)})^{\sum_{j=3}^{4} Z^i_{jf}} (\frac{1}{1 + exp(X^i_h \omega^i)})^{1-\sum_{j=3}^{4} Z^i_{jf}} \\
& = exp \left\{ \sum_{d=1}^{D} \sum_{m=1}^{M_O} \left[\left(log(1 - \sum_{k_O=3}^{K_O} p^i_{d,m,k_O}) - log(1 - \sum_{k_O=3}^{K_O} p^i_{d,m,k_O}) \sum_{k_O=3}^{K_O} \triangle N^i_{d,m,k_O} + \sum_{k_O=3}^{K_O} log(p^i_{d,m,k_O}) \triangle N^i_{d,m,k_O} \right) \right. \right. \\
& \left. + \left(log(1 - \sum_{k_O=1}^{2} p^i_{d,m,k_O}) - log(1 - \sum_{k_O=1}^{2} p^i_{d,m,k_O}) \sum_{k_O=1}^{2} \triangle N^i_{d,m,k_O} + \sum_{k_O=1}^{2} log(p^i_{d,m,k_O}) \triangle N^i_{d,m,k_O} \right) \right] \\
& + \sum_{d=1}^{D} \sum_{m=1}^{M_I} \sum_{k_I=1}^{K_I} \left(log(1 - p^i_{d,m,k_I}) - log(1 - p^i_{d,m,k_I}) \triangle N^i_{d,m,k_I} + log(p^i_{d,m,k_I}) \triangle N^i_{d,m,k_I} \right) \\
& \left. + \sum_{f=1}^{F} \left[\sum_{j=3}^{4} Z^i_{jf} X^i_h \omega^i - log(1 + exp(X^i_h \omega^i)) \right] \right\}
\end{aligned}
$$

$$
\begin{aligned}
&= exp\Bigg\{\sum_{d=1}^{D}\sum_{m=1}^{M_O}\Bigg[\Bigg(log(1-\sum_{k_O=3}^{K_O}p^i_{d,m,k_O}) + \sum_{k_O=3}^{K_O} log\frac{p^i_{d,m,k_O}}{1-\sum_{k_O=3}^{K_O}p^i_{d,m,k_O}}\triangle N^i_{d,m,k_O}\Bigg) \\
&\qquad + \Bigg(log(1-\sum_{k_O=1}^{2}p^i_{d,m,k_O}) + \sum_{k_O=1}^{2} log\frac{p^i_{d,m,k_O}}{1-\sum_{k_O=1}^{2}p^i_{d,m,k_O}}\triangle N^i_{d,m,k_O}\Bigg) \\
&\qquad + \sum_{d=1}^{D}\sum_{m=1}^{M_I}\sum_{k_I=1}^{K_I}\Bigg(log(1-p^i_{d,m,k_I}) + log\frac{p^i_{d,m,k_I}}{1-p^i_{d,m,k_I}}\triangle N^i_{d,m,k_I}\Bigg) \\
&\qquad + \sum_{f=1}^{F}\Bigg[\sum_{j=3}^{4} Z^i_{jf}X^i_h\omega^i - log(1+exp(X^i_h\omega^i))\Bigg]\Bigg]\Bigg\} \\
&= exp\Bigg\{\sum_{d=1}^{D}\sum_{m=1}^{M_O}\Bigg[\Bigg(-log(1+\sum_{k_O=3}^{K_O}\lambda^i_{O,d,k_O,m}\triangle^i_{d,m}) + \sum_{k_O=3}^{K_O}\Big(log(\lambda^i_{O,d,k_O,m}) + log(\triangle^i_{d,m})\Big)\triangle N^i_{d,m,k_O}\Bigg) \\
&\qquad + \Bigg(-log(1+\sum_{k_O=1}^{2}\lambda^i_{O,d,k_O,m}\triangle^i_{d,m}) + \sum_{k_O=1}^{2}\Big(log(\lambda^i_{O,d,k_O,m}) + log(\triangle^i_{d,m})\Big)\triangle N^i_{d,m,k_O}\Bigg) \\
&\qquad + \sum_{d=1}^{D}\sum_{m=1}^{M_I}\sum_{k_I=1}^{K_I}\Big(-log(1+\lambda^i_{I,d,k_I,m}\triangle^i_{d,m}) + \Big(log(\lambda^i_{I,d,k_I,m}) + log(\triangle^i_{d,m})\Big)\triangle N^i_{d,m,k_I}\Big) \\
&\qquad + \sum_{f=1}^{F}\Bigg[\sum_{j=3}^{4} Z^i_{jf}X^i_h\omega^i - log(1+exp(X^i_h\omega^i))\Bigg]\Bigg]\Bigg\}
\end{aligned}
$$

Appendix B

Bayesian Hierachical Model

Layer 1: Shopper Decisions

$$\triangle N^i_{s=O,t,0:2}|\delta^i, \Pi^i, Z^i_{t'}, \alpha^i, \beta^i, \gamma^i, \phi^i, \omega^i \sim multinomial\,(p^i_{s=O,t,0:2})$$

$$\triangle N^i_{s=O,t,3:(K_O+1)}|\delta^i, \Pi^i, Z^i_{t'}, \alpha^i, \beta^i, \gamma^i, \phi^i, \omega^i \sim multinomial\,(p^i_{s=O,t,3:(K_O+1)})$$

$$\triangle N^i_{s=I,t,k_I}|\delta^i, \Pi^i, Z^i_{t'}, \alpha^i, \beta^i, \gamma^i, \phi^i, \omega^i \sim Bernoulli\,(p^i_{s=I,t,k_I}),\ k_I = 1, 2, \cdots, K_I \tag{B.1}$$

Layer 2: Event Probabilities

$$p^i_{O,t,k_O} = \frac{\lambda^i_{O,t,k_O}\triangle^i_t}{1+\sum^2_{k'_O}\lambda^i_{O,t,k'_O}\triangle^i_t},\ k_O = 1, 2;\ p^i_{O,t,0} = \frac{1}{1+\sum^2_{k'_O}\lambda^i_{O,t,k'_O}\triangle^i_t};$$

$$p^i_{O,t,k_O} = \frac{\lambda^i_{O,t,k_O}\triangle^i_t}{1+\sum^{K_O}_{k'_O}\lambda^i_{O,t,k'_O}\triangle^i_t}, k_O = 3, \cdots, K_O; p^i_{O,t,(K_O+1)} = \frac{1}{1+\sum^{K_O}_{k'_O}\lambda^i_{O,t,k'_O}\triangle^i_t};$$

$$p^i_{I,t,k_I} = \frac{\lambda^i_{I,t,k_I}\triangle^i_t}{1+\lambda^i_{I,t,k_I}\triangle^i_t},\ k_I = 1, 2, 3, 4; \tag{B.2}$$

Layer 3: Conditional Intensities of Shopper-Initiated Events

$$\lambda^i_{O,t,k_O} = exp\{\delta^i_{O,k_O}+\Pi^i X^i_{O,t,k_O}+\sum^{K_O}_{l=1}\sum_{t>t'}\alpha^i_{O,lk_O}g(t-t';\beta^i_{O,lk_O})+\sum^{J}_{j=2}\sum_{t>t'}\gamma^i_{jk_O}Z^i_{jt'}g(t-t';\phi^i_{jk_O})\}$$

$$\lambda^i_{I,t,k_I} = exp\{\delta^i_{I,k_I} + \sum^{K_O}_{l=1}\sum_{t>t'}\alpha^i_{O_lI_{k_I}}g(t-t';\beta^i_{O_lI_{k_I}}) + \sum_{t>t'}\gamma^i_{jk_I}Z^i_{jt'}g(t-t';\phi^i_{jk_I})\}$$

$$Prob(Z^i_{3:4,t'} = 1) = \frac{exp(X^i_{ht}\omega^i)}{1+exp(X^i_{ht}\omega^i)} \tag{B.3}$$

Layer 4: Prior Distributions

$$\delta^i \sim MVN_K(\Theta_\delta, \Sigma_\delta),\ \pi_m^i \sim MVN_K(\Theta_{\pi_m}, \Sigma_{\pi_m}),\ m = 1, 2, \cdots M$$
$$\alpha^i \sim MVN_{KK}(\Theta_\alpha, \Sigma_\alpha),\ \gamma^i \sim MVN_{JK}(\Theta_\gamma, \Sigma_\gamma),\ \omega^i \sim MVN(\Theta_\omega, \Sigma_\omega)$$
$$\beta^i \sim log - MVN_{KK}(\Theta_\beta, \Sigma_\beta), \phi^i \sim log - MVN_{JK}(\Theta_\phi, \Sigma_\phi)$$
$$\Sigma_q \sim IW(\bar{S}^{-1}, \bar{\nu}),\ \theta_q \sim MVN(\bar{\theta}_q, \bar{\Sigma}_q),\ q = \delta, \pi_m, \alpha, \beta, \gamma, \phi, \omega. \quad \text{(B.4)}$$

Appendix C

MCMC Samplers

The following steps show the details of the MCMC samplers used in this study:

Step 1: Sample α^i. I consider the prior distribution of α^i following $MVN_{K^2}(\theta_\alpha, \Sigma_\alpha^2)$. I use the Metropolis-Hastings algorithm to sample α^i. The accepting probability of the proposed α^{i*}, drawn from a MVN_{K^2} distribution, is given as

$$min\{\frac{L_i(\alpha^{i*}, \beta^i, \gamma^i, \phi^i, \delta^i, \pi^i, \omega^i|Data^i).MVN_{K^2}(\alpha^{i*}|\theta_\alpha, \Sigma_\alpha^2)}{L_i(\alpha^i, \beta^i, \gamma^i, \phi^i, \delta^i, \pi^i, \omega^i|Data^i).MVN_{K^2}(\alpha^i|\theta_\alpha, \Sigma_\alpha^2)}, 1\}.$$

Step 2: Sample β^i. I consider the prior distribution of β^i following $log-MVN_{K^2}(\theta_\beta, \Sigma_\beta^2)$. The accepting probability of the proposed β^{i*}, drawn from a log-normal distribution, is given as

$$min\{\frac{L_i(\alpha^i, \beta^{i*}, \gamma^i, \phi^i, \delta^i, \pi^i, \omega^i|Data^i).log-MVN_{K^2}(\beta^{i*}|\theta_\beta, \Sigma_\beta^2)\prod_{k,k}\beta_{kk}^{i*}}{L_i(\alpha^i, \beta^i, \gamma^i, \phi^i, \delta^i, \pi^i, \omega^i|Data^i).log-MVN_{K^2}(\beta^i|\theta_\beta, \Sigma_\beta^2)\prod_{k,k}\beta_{kk}^{i}}, 1\}.$$

Step 3: Sample γ^i. I consider the prior distribution of γ^i following $MVN_{JK}(\theta_\gamma, \Sigma_\gamma^2)$. I use the Metropolis-Hastings algorithm to sample γ^i. The accepting probability of the proposed γ^{i*}, drawn from a MVN_{JK} distribution, is given as

$$min\{\frac{L_i(\alpha^i, \beta^i, \gamma^{i*}, \phi^i, \delta^i, \pi^i, \omega^i|Data^i).MVN_{JK}(\gamma^{i*}|\theta_\gamma, \Sigma_\gamma^2)}{L_i(\alpha^i, \beta^i, \gamma^i, \phi^i, \delta^i, \pi^i, \omega^i|Data^i).MVN_{JK}(\gamma^i|\theta_\gamma, \Sigma_\gamma^2)}, 1\}.$$

Step 4: Sample ϕ^i. I consider the prior distribution of ϕ^i following $log-MVN_{JK}(\theta_\phi, \Sigma_\phi^2)$. The accepting probability of the proposed ϕ^{i*}, drawn from a $log - MVN_{JK}$ distribution, is given as

$$min\{\frac{L_i(\alpha^i,\beta^i,\gamma^i,\phi^{i*},\delta^i,\pi^i,\omega^i|Data^i).log-MVN_{JK}(\phi^i|\theta_\phi,\Sigma^2_\phi)\prod_{j,k}\phi^{i*}_{jk}}{L_i(\alpha^i,\beta^i,\gamma^i,\phi^i,\delta^i,\pi^i,\omega^i|Data^i).log-MVN_{JK}(\phi^i|\theta_\phi,\Sigma^2_\phi)\prod_{j,k}\phi_{jk}},1\}.$$

Step 5: Sample δ^i. I consider the prior distribution of δ^i following $MVN_K(\theta_\delta,\Sigma_\delta)$. The accepting probability of the proposed δ^{i*}, drawn from a MVN_K distribution, is given as

$$min\{\frac{L_i(\alpha^i,\beta^i,\gamma^i,\phi^i,\delta^{i*},\pi^i,\omega^i|Data^i).MVN_K(\delta^{i*}|\theta_\delta,\Sigma_\delta)}{L_i(\alpha^i,\beta^i,\gamma^i,\phi^i,\delta^i,\pi^i,\omega^i|Data^i).MVN_K(\delta^i|\theta_\delta,\Sigma_\delta)},1\}.$$

Step 6: Sample π^i. I consider the prior distribution of π^i following $MVN_K(\theta_\pi,\Sigma_\pi)$. The accepting probability of the proposed π^{i*}, drawn from a MVN_K distribution, is given as

$$min\{\frac{L_i(\alpha^i,\beta^i,\gamma^i,\phi^i,\delta^i,\pi^{i*},\omega^i|Data^i).MVN_K(\pi^{i*}|\theta_\pi,\Sigma_\pi)}{L_i(\alpha^i,\beta^i,\gamma^i,\phi^i,\delta^i,\pi^i,\omega^i|Data^i).MVN_K(\pi^i|\theta_\pi,\Sigma_\pi)},1\}.$$

Step 7: Sample ω^i. I consider the prior distribution of ω^i following $MVN_2(\bar{\theta}_\omega,\bar{\Sigma}_\omega)$. The accepting probability of the proposed ω^{i*} drawn from an MVN_2 distribution, is given as

$$min\{\frac{L(\alpha^i,\beta^i,\gamma^i,\phi^i,\delta^i,\pi^i,\omega^{i*}|Data).MVN_2(\omega^{i*}|\bar{\theta}_\omega,\bar{\Sigma}_\omega)}{L(\alpha^i,\beta^i,\gamma^i,\phi^i,\delta^i,\pi^i,\omega^i|Data).MVN_2(\omega^i|\bar{\theta}_\omega,\bar{\Sigma}_\omega)},1\}.$$

Step 8: Sample θ_n. I consider the prior distribution of θ_n following $MVN_Q(\bar{\theta}_{\theta_n},\bar{\Sigma}_{\theta_n})$, where $\bar{\theta}_{\theta_n}=0$ and $\bar{\Sigma}_{\theta_n}=10^6 I_Q$. θ^*_n is drawn from a multivariate distribution

$$\theta^*_n \sim MVN_Q(M,\ N),$$

where $M=N'((\sum_{i=1}^{I}n^i)'\Sigma_n^{-1}+\bar{\theta}'_{\theta_n}\bar{\Sigma}^{-1}_{\theta_n})'$, $N=(I\Sigma_n^{-1}+\bar{\Sigma}^{-1}_{\theta_n})^{-1}$, $n=\alpha$,$log(\beta)$,γ,$log(\phi)$,δ,π, and $Q=K*K$ for $n=\alpha$,$log(\beta)$, $Q=J*K$ for $n=\gamma$,$log(\phi)$, and $Q=K$ for

$n = \delta,\pi$.

Step 9: Sample Σ_n. I consider the prior distribution of Σ_n following $IW(\bar{S}^{-1},\bar{\nu})$, where $\bar{S} = I_Q$ and $\bar{\nu} = 1$. Σ_n^* is drawn from an inverse Wishart distribution

$$\Sigma_n^* \sim IW((\sum_{i=1}^{I}(n^i - \theta_n)(n^i - \theta_n)' + \bar{S})^{-1}, I + \bar{\nu}),$$

where $Q = K * K$ for $n = \alpha,log(\beta)$, $Q = J * K$ for $n = \gamma,log(\phi)$, $Q = K$ for $n = \delta,\pi$, $Q = 2$ for $n = \omega$.

Bibliography

Accengage. 2016. The push notification benchmark. Http://www.accengage.com/benchmarkyourpushnotificationoptinand-reactionratesoniosandandroidbyindustryandcountry/.

Andrews, Michelle, Xueming Luo, Zheng Fang, Anindya Ghose. 2015. Mobile ad effectiveness: Hyper-contextual targeting with crowdedness. *Marketing Science* **Articles in Advance** 1–17.

At-Sahalia, Yacine, Julio Cacho-Diaz, Roger J.A. Laeven. 2015. Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics* **117** 585–606.

Bentley, Taylor, Tat Y. Chan, Young-Hoon Park. 2015. How search advertising works: A model of consumer information search in organic and sponsored links. *Working paper* .

Chen, Yuxin, Song Yao. 2016. Sequential search with refinement: Model and application with click-stream data. *Management Science* Articles in Advance, 1–22.

Daley, D.J., D. Vere-Jones. 2003. *An Introduction to the Theory of Point Processes*. 2nd ed. Springer, New York.

Danaher, Peter J., Michael S. Smith, Kulan Ranasinghe, Tracey S. Danaher. 2015. Where, when and how long: Factors that influence the redemption of mobile phone coupons. *Journal of Marketing Research* **52**(5) 710–725.

Daurer, Stephan, Dominik Molitor, Martin Spann, Puneet Manchanda. 2016. Consumer search behavior on the mobile internet: An empirical analysis. *Ross School of Business Paper No. 1275* .

eMarketer. 2014a. Marketers respond to consumer demand for personalization. Https://www.emarketer.com/Article/Marketers-Respond-Consumer-Demand-Personalization/1010618.

eMarketer. 2014b. Smartphone users worldwide will total 1.75 billion in 2014. Http://www.emarketer.com/Article/SmartphoneUsersWorldwide-WillTotal175Billion2014/1010536.

eMarketer. 2015. Despite concerns, interest in mobile audience targeting rises. Http://www.emarketer.com/Article/DespiteConcernsInterestMobile-AudienceTargetingRises/1012471.

eMarketer. 2016. Internet users expect more personalized mobile apps, sites.

Fong, Nathan M., Zheng Fang, Xueming Luo. 2015. Geo-conquesting: Competitive locational targeting of mobile promotions. *Journal of Marketing Research* **52**(5) 726–735.

Forbes. 2016. Retail's dirty open secret: Store traffic declines are shaking up the shopping experience. Https://www.forbes.com/sites/barbarathau/2016/06/24/.

Ghose, Anindya, Avi Goldfarb, Sang Pil Han. 2013. How is the mobile internet different? search costs and local activities. *Information Systems Research* **24 (3)** 613 – 631.

Ghose, Anindya, Sang Pil Han. 2011. An empirical analysis of user content generation and usage behavior on the mobile internet. *Management Science* **57 (9)** 1671–1691.

Ghose, Anindya, Beibei Li, Siyuan Liu. 2015. Mobile advertising using customer movement patterns. *Working paper* .

Ghose, Anindya, Sha Yang. 2009. An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science* **55 (10)** 1605–1622.

Hawkes, Alan G. 1971a. Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society. Series B (Methodological)* **33**(3) 438–443.

Hawkes, Alan G. 1971b. Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58**(1) 83–90.

Insights, Smart. 2013. New ad ranking research shows importance of position. Http://www.smartinsights.com/paid-search-marketing-ppc/paid-search-analytics/clickthrough-rate-per-advert-position-in-adwords/.

Jerath, Kinshuk, Liye Ma, Young-Hoon Park. 2014. Consumer click behavior at a search engine: The role of keyword popularity. *Journal of Marketing Research* **51 (4)** 480–486.

Jeziorski, Przemyslaw, Ilya Segal. 2015. What make then click: Empirical analysis of consumer demand for search advertising. *American Economic Journal: Microeconomics 2015* **7(3)** 24–53.

Johnson, Devin S., Mevin B. Hooten, Carey E. Kuhn. 2013. Estimating animal resource selection from telemetry data using point process models. *Journal of Animal Ecology* **82** 1155–1164.

Juniper Research. 2014. Mobile coupon users to pass 1 billion by 2019, juniper research finds. Http://www.juniperresearch.com/press-release/coupons-pr1.

Kim, Jun B., Paulo Albuquerque, Bart J. Bronnenberg. 2010. Online demand under limited consumer search. *Marketing Science* **29 (6)** 1001–1023.

KPCB. 2017. Internet trends 2017. Http://www.kpcb.com/internet-trends.

Liu, Zhuping, Frenkel ter Hofstede, Vijay Mahajan. 2017. Engaging and targeting consumers on mobile: Connecting mobile promotions to foot traffic. *Working paper* .

Luo, Xueming, Michelle Andrews, Zheng Fang, Cheewei Phang. 2014. Mobile targeting. *Management Science* **60 (7)** 1738–1756.

Mobile Marketer. 2011. Mall owner ggp: The higher the offer the more successful the campaign. Http://www.mobilemarketer.com/cms/news/database-crm/11527.html.

Mohler, G. O., M. B. Short, P. J. Brantingham, F. P. Schoenberg, G. E. Tita. 2011. Self-exciting point process modeling of crime. *Journal of the American Statistical Association* **106** 100–108.

Molitor, Dominik, Philipp Reichhart, Martin Spann, Anindya Ghose. 2016. Measuring the effectiveness of location-based advertising: A randomized field experiment. *Working paper* .

Musmecl, F., D. Vere-Jones. 1992. A space-time clustering model for historical earthquakes. *Ann. Inst. Statist. Math.* **44**(1) 1–11.

Ogata, Yosihiko. 2004. Space-time model for regional seismicity and detection of crustal stress changes. *Journal of Geophysical Research* **109** 1–16.

Ogata, Yosihiko, Koichi Katsura, Masaharu Tanemura. 2003. Modelling heterogeneous space-time occurrences of earthquakes and its residual analysis. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **52**(4).

Ogata, Yosihiko, Jiancang Zhuang. 2006. Space-time etas models and an improved extension. *Tectonophysics* **4** 13–23.

Porter, Michael D., Gentry White. 2012. Self-exciting hurdle models for terrorist activity. *The Annals of Applied Statistics* **6**(1) 106–124.

Truccolo, Wilson, Uri T. Eden, Matthew R. Fellows, John P. Donoghue, Emery N. Brown. 2005. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *Journal of Neurophysiology* **93** 1074 –1089.

Wall Street Journal. 2011. Malls test apps to aid shoppers. Http://www.wsj.com/articles/SB10001424052748704336504576258740640080926.

Wall Street Journal. 2014. Stores confront new world of reduced shopper traffic. Http://www.wsj.com/articles/SB10001424052702304419104579325100372435802.

Weitzman, Martin L. 1979. Optimal search for the best alternative. *Econometrica* **47 (3)** 641–654.

Xu, Lizhen, Jason A. Duan, Andrew Whinston. 2015. Path to purchase: A mutually exciting point process model for online advertising and conversion. *Management Science* **60**(6) 1392–1412.

Yang, Sha, Anindya Ghose. 2010. Analyzing the relationship between organic and sponsored search advertising: Positive, negative, or zero interdependence? *Marketing Science* **29 (4)** 602–623.

Yao, Song, Carl F. Mela. 2011. A dynamic model of sponsored search advertising. *Marketing Science* **30 (3)** 447–468.