ABSTRACT

Title of Document: **DESIGNING INFORMATION STRATEGIES FOR DIGITAL PLATFORMS: FINDINGS FROM LARGE-SCALE RANDOMIZED FIELD EXPERIMENTS.**

Lanfei Shi, Doctor of Philosophy, and 2019

Directed By: **Professor Siva Viswanathan**, Information Systems

The rise of digital platforms has transformed our economy and reshaped consumer behaviors and experiences. While practitioners and researchers have a growing interest in understanding digital platforms, there is still a dearth of research on how platforms can design effective information strategies to mitigate fundamental issues such as information asymmetry and search frictions by leveraging granular data. My dissertation seeks to fill this gap. Specifically, by focusing on significant real-world problems on digital platforms, I aim to examine IT-enabled and analytics-driven information strategies and study the impact of these strategies on the users as well as on the platforms themselves. In collaboration with two different online platforms, I design and conduct three randomized field experiments to investigate the impact of informational interventions and provide actionable suggestions. In Essay 1, I examine incentive strategies for motivating effective mobile app adoptions, by comparing monetary incentives against informational incentives. I find that the usage after app adoption depends on how customers are motivated, and only information induced adoption leads to long-term increase in purchases. In Essay 2, I investigate the role of "verification" when it is made optional, and find that it serves as a very effective signaling device, especially in markets that lack other mechanisms such as reputation systems. I also find that users on the two sides of online platform use the same signal very differently, and that this is attributable to the difference in the credibility of their primary signaling-attribute of each side, viz. income in males and beauty in females.

In Essay 3, I examine the effectiveness of three different recommendation systems in two-sided matching platforms with a focus on how the provisioning of potential candidates' preference information impacts focal user's decision-making and matching outcomes. I find that compared to "people you might prefer", users act strategically towards "people who might prefer you" and "people who you might prefer and who might prefer you" by actively reaching out to less desirable candidates, which leads to improved outcomes. In short, the three studies present new empirical evidence of how platforms can leverage information as a tool to design effective incentives, signaling mechanisms and recommender systems to facilitate users' decision-making, transactions and matching.

DESIGNING INFORMATION STRATEGIES FOR DIGITAL PLATFORMS: FINDINGS FROM LARGE-SCALE RANDOMIZED FIELD EXPERIMENTS.

By

Lanfei Shi.

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2019

Advisory Committee:
Professor Siva Viswanathan, Chair
Professor Peng Huang
Professor Kunpeng Zhang
Professor Bo Zhou
Professor Erkut Ozbay, Dean's Representative

Acknowledgements

I am in huge debt to people who have guided and helped me along this journey. I would not have gone so far without their support.

First and foremost, I would like to express my deepest gratitude to my advisor Professor Siva Viswanathan for the continuous guidance and support of my Ph.D. study, research and job search. Siva has guided me through the development of my own research taste and agenda, always providing insightful advice and constructive feedback on research or significant career decisions. The memory of some heated discussions is still so vivid, and I clearly see my growth toward a more critical, original and independent researcher. I also appreciate the flexibility and freedom Siva gives me in pursuit of the topics I am passionate about.

I would also like to thank Professor Peng Huang for strong support and valuable advice for my research and career development. The collaboration with Peng has been delightful and productive, and I learned a lot from him. My sincere thanks also go to Professor Kunpeng Zhang for being very supportive as my collaborator and friend. I enjoyed our exciting discussion about research, career, and life as we shared a common career path. I also would like to thank Professor Bo Zhou and Professor Erkut Ozbay for serving on my committee and providing helpful feedback and interesting new angles for my dissertation.

Besides my thesis committee, I am also thankful to Professor Ritu Agarwal and Professor Guodong Gao for helping me explore research ideas and supporting me along my Ph.D. journey. I would also like to thank Professor P.K. Kannan and Professor Kislaya Prasad for the great feedback on my dissertation work.

Finally, I am grateful to the help and support from other faculty at DOIT department and Smith School. I also appreciate all the inspiring conversations I had with industry people with various backgrounds and expertise. Last but not least, I am sincerely grateful to my Ph.D. colleagues, my family, and friends for the endless support and encouragement, which brings me so much joy and meaning.

Table of Contents

List of Tables

List of Figures

Chapter 1: Overview

Online markets and platforms have been gaining momentum and reshaping how people interact and trade. Most of today's biggest and most disruptive firms across industries, for instance, Google, Amazon, Ebay, Airbnb, and Uber to name a few operate as multi-sided markets or platforms that connect buyers to sellers, hosts to travelers, riders to drivers, etc. While there is a growing interest in understanding digital platforms, there is still a dearth of research that provides theoretical and empirical insights on information strategies that utilize analytics and technology to improve product adoption, user engagement and profitability. My dissertation seeks to fill this gap. It focuses on designing effective information strategies and artifacts to help mitigate fundamental problems of digital markets and to facilitate transactions and matches between two sides of the platforms. Specifically, I collaborate with one transactional platform and one matching platform to design and conduct three randomized field experiments that allow me to draw causal inferences about the effectiveness of information strategies and hence provide meaningful business implications.

Online platforms have a pressing need to promote their mobile channel to cater to the increasing mobile usage among consumers. In Essay 1, I examine effective strategies for mobile app adoptions. I focus on the *induced* adoption that platforms proactively influence through interventions as opposed to those organic adoptions where users convert on their own. I compare two most commonly adopted motivating strategies, information provisioning and monetary incentive, by designing and conducting a large-scale randomized field experiment on a transactional platform. I investigate the differential impacts of these two inducements on customers' mobile app adoptions and subsequent purchase behaviors. I find that i) providing monetary incentives as well as providing information can both lead to a significant increase in mobile app adoptions; ii) the causal effect of induced mobile app adoptions varies greatly depending on how customers are motivated. Although providing monetary incentives leads to a larger increase in mobile app adoptions, such induced adoptions do not result in more purchases in the long run. In contrast, providing information leads to effective mobile adoptions that sustainably increase customers' purchases, and overall profits for the firm. The further examining customers' multichannel purchase behaviors reveals a complementary effect between the mobile app and the desktop channel for information-induced app adopters, but a substitution effect between the mobile app and the mobile web channel for incentive-induced app adopters. For information-induced app adopters, the mobile app serves as a discovery tool and helps them find a greater variety of deals. The exploration of the underlying drivers of such differential impacts suggests that information, as compared to incentives serves as a better sorting device and can attract customers who have a greater need for the app and use it more effectively. My findings provide actionable insights for firms designing interventions to motivate effective mobile adoptions.

In Essay 2, I focus on a fundamental challenge faced by online matching platforms – information asymmetry, especially for markets with few alternative information mechanisms available such as reputation systems and quality assurance. While mandatory verification is widely understood as a good assurance in such case, in this study I seek to examine a different role of verification – its ability to serve as a credible signal for a user, when such verification is made optional and visible to others. In collaboration with an online dating platform, I design and conduct a randomized field experiment to introduce *optional* phone verification and I examine both *ex ante opt-in* decisions as well as the *ex post impact* of verification on individual users and the platform. I identify very interesting differential opt-in decisions across the two sides of the platform that only males are consistent with the conventional prediction of signaling with H-type (i.e., more popular) males being more likely to opt-in to verification. As for females I find M-type females are the most likely to opt-in to verification. In examining the underlying mechanism, I find that such differential opt-in decisions are related to the difference in the credibility of the existing key attribute of each side, viz. income for males and beauty for females. I extract beauty of females by applying state-of-the-art deep learning techniques using images. Furthermore, I identify an interesting two-pronged effect of verification on verified users. On one hand, I find that verified users, especially H-type males and M-type females, receive more messages from higher type partners. More interestingly, upon verification these users become more proactive and reach out to more and better potential partners. Further, the introduction of this voluntary verification signal facilitates desirable matching outcomes and benefits the platform as a whole. My study is among the first to document these differential opt-in decisions and impacts of verification across two sides of a matching platform and provide novel insights for platforms on optional verification mechanisms and signaling in two-sided markets.

On decentralized two-sided matching platforms, information seeking plays a critical role in determining the efficiency and quality of user matches. Platforms have resorted to information technologies to facilitate search and matches, among which recommender systems are often deemed as one of the most effective approaches. In Essay 3, I focus on user recommendation in two-sided matching markets, fundamentally distinctive from product recommendation in transactional markets. I am interested in understanding how recommender systems impact users' choices and final matching outcomes on a two-sided matching platform where a match is a bilateral decision involving agreement on both sides. Thus it is useful to examine users' choices and matching outcomes when users are presented with recommendations based on i) their own preferences, ii) the preferences of their potential matches, or iii) mutual preferences. I collaborate with an online dating platform to introduce three recommendation systems providing the aforementioned three types of recommendations based on different preferences. Interestingly, I find that users are sensitive to the new information containing potential partner's preferences, and they proactively reach out to these candidates despite the relatively low desirability. Such strategic behavior leads to greater increase in matching when recommendations are based on potential matches' preferences compared to recommendations based on focal user's preference. The findings provide valuable insights to two-sided matching platforms on how to design user recommendation systems beyond the current common practice that considers solely focal user's preference.

To summarize, my dissertation contributes to the literature regarding digital platforms on understanding the role of information strategies and information artifacts in users' decision-making and matching outcomes. Rich user behavior data helps me uncover the mechanisms underlying these outcomes. These findings are not only theoretically important but also useful to provide insights on how to effectively design information strategies to facilitate users' adoption, search, and matches on two-sided platforms. Essay 1 provides actionable insights on the role of information and monetary incentives on long-term purchases. It sheds light on multi-channel purchase behavior that highlights the complementary role between mobile app channel and desktop channel. Essay 2 documents differential opt-in decisions and the impacts of verification for males and females in a matching market. It highlights the value of optional verification that serves as a credible signal and adds understanding to the literature of signaling regarding how users on the two sides of a market may use the same signal *differently*. Essay 3 extends the current literature regarding recommender systems to *user* recommendation and examines how recommender systems generating candidates based on potential matches' preferences would impact user's choices and matching outcomes. The findings provide insights on how to design better information-provisioning strategies in two-sided markets.

Chapter 2: Motivating Effective Mobile App Adoptions: Evidence from a Large-scale Randomized Field Experiment

2.1 Introduction

The adoption and usage of mobile channels has not only grown significantly, but has also altered users' experience and behaviors in a multi-screen world (Einav et al. 2014, Fang et al. 2015, Ghose et al. 2012, Luo et al. 2013, Shankar et al. 2016, Xu et al. 2016). According to Forbes (2017), mobile commerce in the Nov-Dec 2017 holiday season is predicted to cross $5 billion, accounting for 54% of retail business spending. Recognizing this disruptive effect, firms have been increasing their investments in mobile channels, with a strong emphasis on the promotion of their own mobile apps (RetailMeNot 2016, Siwicki 2014).

Prior literature has established a positive association between mobile app adoption and business outcomes using observational data. The adoption of mobile apps may lead customers to make more purchases (Xu et al. 2016), be more socially engaged (Jung et al. 2014) and consume more news (Xu et al. 2014). Given the value of mobile app adoptions, a natural question for the firms is: can such findings be put into action? In other words, can firms actively influence customers to adopt mobile apps, and whether such *induced adoptions* can lead to an increase in customers' purchases? Further, *how* should firms induce such adoptions? Two interventions are commonly used by firms to encourage technology adoptions: providing *information* about the benefits of the technology (Guiteras et al. 2015) and providing *incentives* for adoption (Cohen et al. 2015). When choosing interventions, firms have two goals in mind: 1) increase mobile app adoptions ("coverage"); 2) attract customers who need the app more and will use it more effectively ("sorting"). The ideal intervention is one that helps achieve high coverage (increased adoption) as well as appropriate sorting (effective adoption). However, it is possible that there are inherent tradeoffs between the two goals. Despite the huge stakes involved and an active debate about the best intervention strategy (Techcrunch 2015a), no empirical research has rigorously investigated this problem.

My empirical study seeks to fill this gap. Specifically, my study addresses the following questions:

Q1a. (Coverage) What is effect of incentive vs. information on customers' app adoptions?

Q1b. (Effectiveness) What is the causal effect of *induced app adoptions* on customers' behaviors in the long run? Does the effect vary for adoptions induced by *incentive* vs. *information*?

Q1c. (Profitability) What is the impact of incentive vs. information on firm's overall profitability?

Q2 (Cross-channel Effect) How do incentive- and information-induced app adoptions affect customers' purchase behaviors across mobile and desktop channels?

Q3. (Mechanisms) What is the underlying driver of the differences between the causal effects of app adoptions induced by incentive vs. information?

Answers to these questions are valuable as they provide direct managerial implications and actionable insights to firms interested in designing active interventions to motivate effective mobile app adoptions and improve business outcomes.

It is pertinent to note that the prior studies on mobile channel adoptions have not examined the causal effect of *induced* mobile adoption (e.g. Q1b)[1]. From a firm's perspective, there are two types of mobile app adopters: 1) those who would adopt due to factors beyond the firm's direct intervention, such as WOM, app quality, etc. (defined as 'organic adopter'), and 2) those who would adopt the app only in response to the firm's direct interventions (defined as 'induced adopters'). Previous studies

[1] Q1a, Q1c, Q2, Q3 have also not been addressed in previous studies, most of which use observational data. Answer to those questions require both randomized experiments and active design of interventions.

investigate the effect of *observed* mobile adoptions (i.e. a mix of ‘organic’ and ‘induced’ adopters, Xu et al. 2016), rather than *induced* adoptions. As demonstrated in various contexts including technology adoption (Dupas 2014), customer acquisition (Datta et al. 2015), and multichannel purchases (Neslin and Shankar 2009), induced adopters may differ fundamentally from organic adopters in observed characteristics such as demographics and historical behaviors. In addition, organic adopters of technology may use it in ways different from induced adopters, as such organic adoptions may be driven by unobserved needs or preferences. Such unobserved needs may drive them to use the app in specific ways that would lead to more (or less) purchases. The same may not hold for customers who are nudged or incentivized to adopt the app. Consequently, it is not clear whether induced mobile adoptions would lead to desirable outcomes and be beneficial to the firms. Since the firms *can only* actively influence induced adoptions (rather than organic adoptions) using external interventions, it is crucial to understand the causal effect of induced adoptions on customers’ purchase behaviors. More interestingly, the effect of induced app adoptions may depend on *how* customers are motivated to adopt the app. Customers have private information about their need for the app (how and how often they use the app) and such unobserved ‘customer types’ may determine their long run behaviors. Different interventions may encourage different types of customers to adopt the app (“sorting”), and may lead to different customer behaviors and business outcomes in the long run. Therefore, it is important for firms to understand whether, and how, the effect of app adoptions varies for customers induced by different interventions, as well as the underlying mechanisms.

Identifying the causal effect of induced mobile app adoption (Q2), though relevant and important, is empirical challenging, for two reasons. First, just using observational data on app adoptions one cannot differentiate between organic adoptions and induced adoptions. Moreover, the effect of induced adoptions may be confounded by time-varying factors (e.g. concurrent marketing campaigns on adoption and purchase). Thus, without exogenous variations in the level of induced adoptions, it is extremely difficult to identify its causal effect on customers' purchase behaviors.

To address these challenges, I collaborate with a leading daily deal platform in the US and conduct a large-scale randomized field experiment to examine whether and how the firms can effectively induce app adoptions. I randomly choose over 230,000 customers who have never downloaded the firm's mobile app and randomly assign eligible users into one of three experimental groups: Control group with no information or incentives, Treatment 1 with an email highlighting incentives for adoption (5 deal bucks), and Treatment 2 with an email highlighting information about the benefits of discovering deals using mobile apps. The exogenous variation created by the experiment not only allows us to monitor the cohort of induced adopters over time and identify the causal effect of app adoptions induced by the different interventions (Q1b,c), but also facilitates a straightforward comparison of the effect of interventions on app adoptions and firm profitability (Q1a and Q1c). Specifically, I address Q2 (causal effect of induced app adoptions) using the framework of Local Average Treatment Effect (LATE, Angrist et al. 1996, i.e. using a random assignment of the test group as an instrument for app adoptions).

My experiment generates four main findings. First, both monetary incentives (T1) as well as information (T2) lead to a significant increase in customers' mobile app adoptions, with a relative increase of 466% and 144%, respectively. Second, while providing monetary incentives leads to a larger increase in mobile app adoptions, such induced adoptions do not result in more purchases in the long run. In contrast, providing information leads to more effective mobile adoptions that sustainably increase customers' purchases, even 12 months after adopting the app. Third, in examining customers' multichannel purchase behaviors, I find that information-induced adoptions (T2) lead to a complementarity in customers' purchases across the mobile app and the desktop channel, whereas incentive-induced adoptions lead to a substitution effect between mobile web and mobile web channel. Finally, I find that the information intervention (T2) can significantly increase overall profits of the firm (by 2-3%). In contrast, providing incentives (T1) does not lead to significant increase in overall profits.

I then explore the underlying drivers of the differences between the effect of app adoption induced by different interventions (Q4(a)), using a new field experiment and a series of analysis. I uncover evidence that indicates that the observed long-run differences in customers' behaviors across the two treatment groups are attributable to 'sorting'. As expected, I find evidence of sorting in customers' observable characteristics and behaviors, as well as in moderating effects. However, interestingly, after controlling for customers' observable characteristics in LATE, the causal effect of induced adoptions for both treatments remains qualitatively the same, indicating that the effect of induced adoptions is largely attributable to sorting on characteristics unobservable by the firm. These findings are surprising and managerially important. They imply that customers possess private information about their "type" (i.e., their need for the mobile app) that firms are unable to observe. Therefore, firms can only rely on appropriate interventions to encourage sorting, i.e. attracting those customers who would use the mobile app effectively to adopt it.

In summary, my study is among the first to investigate how firms can induce effective app adoptions. The findings of the study provide guidelines for designing interventions to motivate effective app adoptions, and add to our understanding of the mechanisms underlying the effect of induced app adoption (i.e. sorting). The rest of the paper proceeds as follows. Section 2 and 3 briefly discuss related literature and theoretical grounding. Section 4, 5 and 6 introduces the experiment design, empirical strategy and data. I present the results in section 7 and discuss managerial implications in section 8.

2.2 Prior Literature

There is a growing literature on the role of mobile devices in influencing customers' engagement and purchase behaviors (Jung et al. 2014, Xu et al. 2014, Xu et al. 2016). My study is closely related to three streams of research that spans information systems and marketing, among others.

The first and most relevant stream of literature is one that focuses on the causal effects of mobile app adoptions on customers' engagement and purchase behaviors. Using propensity score matching, difference-in-difference and other methods on observational data, previous studies have found that customers' adoption of mobile apps can lead to more purchases (Xu et al. 2016), more social engagement (Jung et al. 2014), more consumption of news (Xu et al. 2014) and higher demand for digital service (Liu et al. 2016). However, from the firms' perspective, a natural question is whether they can *actively* influence customers to adopt mobile apps, and whether such *induced* adoptions can increase customers' purchases and firms' profits in the long run. My study contributes to this research stream in two ways. First, previous studies focus on observed adoptions (Xu et al. 2016, Jung et al. 2014). My study complements these by focusing on the causal effects of *induced* app adoptions on customers' purchase behaviors. As noted earlier, induced adopters may differ from organic adopters in fundamental ways in mobile app usage. Since marketers can only actively influence induced adoptions, it is crucial to understand the causal effects of induced adoptions on customer behaviors and business outcomes. My nuanced results also provide practical guidelines for firms designing interventions to motivate *effective* mobile adoptions[2]. Second, my study extends previous research by designing

[2] In China, a huge amount of venture capital funding has gone towards subsidies for mobile app adoptions. Such subsidy in general leads to poor returns, because of low customer engagement in the long run

a new identification strategy. Specifically, the usage of a randomized field experiment allows us to cleanly identify the effect of different interventions in driving mobile app adoption (Q1) and customer profitability (Q3). The combination of a randomized experiment with an instrumental variable approach (LATE) allows me to cleanly identify the causal effect of induced adoptions on customers' purchase behaviors (Q2, see details in Section 3).

The second important stream of research relates to factors that drive technology adoption (Hann et al. 2016, Dupas 2014), especially mobile app adoption (Bang et al. 2013, Ghose and Han 2014, Han et al. 2016, Zheng et al. 2016), as well as subsequent usage after adoption (Retana et al. 2016, Son et al. 2016, Kato-Lin et al. 2015). While previous studies investigate the two outcomes separately, I highlight a central tradeoff between adoption and effective usage in the context of mobile apps. My study is among the first to identify the causal effect of external interventions on the two outcomes at the same time. The results imply that firms should use information-related interventions to strike a balance between motivating more adoptions (coverage) and ensuring effectiveness of those adoptions (sorting).

Finally, my study also complements the stream of literature on the role of mobile apps in influencing customers' purchase behaviors across multiple channels. Recent studies have demonstrated strong interdependence between different channels, in the form of substitution or complementarity (Brynjolfsson et al. 2009, Forman 2009). The interdependence has also been confirmed in the context of mobile commerce (Xu et al. 2016) and digital banking (Liu et al. 2016). A recent study using clickstream data (De Haan et al. 2015) hypothesized that mobile and alternative channels may be used separately to fulfill different flows (e.g. information vs. transaction) in a customer's online journey. My study complements this stream of literature with new evidence on the substitution/complementarity effects between mobile apps and alternate channels. My analyses also provide new insights that such channel complementarity and substitution may be closely related to how customers are induced to adopt the app (i.e. sorting).

2.3 Theoretical Grounding

The effects of monetary incentives and information on customers' purchase behaviors in the short run and long run (Figure 1a) have been studied in marketing (Ailawadi and Neslin 1998, Deighton et al. 1994) and economics (Nelson 1974, Dupas 2014). In addition, a recent stream of literature has investigated how the use of monetary and non-monetary interventions for customer acquisition may affect customers' long-term value (Figure 1b, Lewis 2007). A differentiating feature of my study is that rather than directly influencing customers' purchase behaviors, incentive or information is used to induce customers to *sort (or self-select)* into app adoption, and only such induced adoptions influence customers' purchase behaviors in the long run (Figure 1c). To the best of my knowledge, no study has investigated the mechanism underlying how incentive or information may influence customers' purchases in the long run through induced adoptions.

'Sorting' has its origins in information economics (Stiglitz 1975, Lazear 2000), and refers to the fact that individuals (job applicants, patients, customers) could choose certain arrangements (e.g. labor contract, health insurance, technology adoption) based on their observable attributes and unobservable preferences or information. In my context, customers could have private information on their potential benefits from using the app. Customers who are induced by information to adopt the app could be different from customers induced by monetary incentive in both their observed characteristics (e.g. demographics and past behaviors) as well as unobservable need (e.g. deal discovery). Such observed and unobserved differences would then explain the differential effect of incentive-induced and information-induced adoptions on customers' purchase behaviors.

Specifically, previous literature on technology adoption has demonstrated that the provision of incentives may act as a double-edge sword: on the one hand, providing incentives may encourage more trials (Cohen and Dupas 2010), and facilitate habit formation (Charness and Gneezy 2009) thus increasing the effect of technology adoption; on the other hand, the use of short-run incentives may encourage adverse selection -- attracting those who will not use or do not need the technology (Ashraf et al. 2010), thus countering the effects of adoption in the long run. Such trade-off is especially salient in the case of mobile app adoptions. On the one hand, a mobile app is fundamentally an experience good; providing incentives may help customers overcome the fixed cost of downloading app, setting up payments, and learning. On the other hand, continuous use, rather than one-time adoption, of the app is required to drive purchases and contribute to the firm's overall profitability; providing incentives may attract customers who only enjoy short-run benefits and not use the app in the long run (adverse selection).

In contrast, highlighting the benefits of the mobile app may attract the right type of customers (whose needs for information are aligned with the benefits from using the app), and thus may lead to effective adoptions: information-induced adopters are likely to use the app more and in a more effective way, thereby benefiting more from their adoption. Such differences exist from the time of adoption and would lead to different usage patterns and purchase behaviors in the long run.

In summary, there are two types of sorting: sorting on observables and sorting on unobservables. Both processes may be at work at the same time. For instance, providing monetary incentives (T1) may attract those customers who are more sensitive to incentives; and such adoption might only lead to short-run increase on customers' purchases in first few weeks. On the other hand, app adoptions induced by information (T2) could attract customers who download mobile app for its own value (such private value cannot be predicted by the firm using observables) – thus those induced adopters are more likely to use the mobile app more in the long run, discover more deals and make more purchases (and on a wider range of deal categories). This leads to several testable implications. The first two implications are related to observable characteristics while the latter three are related to unobserved needs for the app.

1) Adopters in the different treatment groups differ in their observable characteristics;

2) Adopters in T1 (incentive) are more likely to make purchases through mobile app in the short run but the effect disappears in the long run (e.g. after few weeks);

3) After controlling for all difference in observable characteristics, the causal effect of induced adoptions by both treatments still follows similar pattern, indicating sorting on unobservable.

4) Information-induced adopters (T2) are likely to use the app more and discover more deals, which may lead to a greater variety of deals purchased, as well as more purchases in regions with higher deal densities.

5) Information intervention (T2) does not directly affect customers' purchases in the long run. (It only works by attracting the right type of adopters with strong need for app, i.e. sorting)

In summary, I theorize that the effect of induced app adoptions may crucially depend on how customers are induced, because of sorting. Incentive and information may attract adopters with different observed and unobserved characteristics, which would then result in differential effects of induced adoptions. I test the implications and further discuss sorting[3] in Section 7.

[3] It is important to clarify that the "sorting mechanism" discussed in this paper is different from "self-selection" in previous studies. There are two key differences. First, in previous observational studies, self-selection, which indicates customers who are more likely to make purchases have a higher propensity to adopt the app, may bias the identification of the causal effect of app adoptions. However, in my case, the treatment effect of induced adoptions is perfectly identified using the exogenous interventions in the experiment. The "sorting effect" simply implies that the identified treatment effects may vary for adopters induced by different interventions. Second, while self-selection means that customers' adoption decisions may be

2.4 Experiment Design

In collaboration with a leading daily deal-sharing platform in US, I conduct a large-scale randomized field experiment to understand how to motivate effective mobile app adoptions. The platform offers a wide range of daily deals for local services and standard products at a high discount and has a large customer base. Users can use three channels (desktop, mobile web and mobile app) to browse and purchase deals on the platform. The platform offers a mobile app to its customers. However, at the time of my experiment, only a small portion of the customers had downloaded the mobile app, though a much larger portion of customers has accessed the platform's email using their mobile device. The platform observes the customers' mobile device type, app adoption status, and can target them with information or incentives through email.

My experiment focuses on customers who a) have already accessed emails of the platform using an iPhone but have never downloaded its mobile app, and b) have made at least one purchase before experiment[4]. In practice, such customers are the target audience of mobile app adoption campaigns (DigiDay 2016). I randomly select over 230,000 eligible customers from the platform's database, and randomly assign them into one of three experimental treatments: (i) Control group with no information or incentives (137,195 subjects); (ii) Treatment T1 with email highlighting an incentive ($5 deal bucks) for app adoption (48,027 subjects), and (iii) Treatment T2 with email highlighting information about the ease of discovering deals using the app,

---

related to their purchase behaviors (Y), "sorting" implies that customers may make adoption decisions based on the potential private benefits of app usage (ΔY).

[4] I focus on active users who have made at least one purchase, for the following reasons. First, firms often target active users in their mobile campaigns as they may purchase more upon adoption (DigiDay 2016). This is especially true when firms have a large number of active users (as in my case). Second, the response rate is usually very low for app download campaigns. To have enough power to identify the causal effect of induced adoptions, I focus on active users who usually have a higher response rate.

but does not contain any incentive for app adoption (48,070 subjects). The template is provided in the Figure 3. The sample size of both treatment groups is smaller than the control group because of the relatively large costs involved in sending out emails and incentives.

The emails to the 96,097 customers in T1 and T2 are sent out in a single day at the same time. The large scale campaign allows us to create a large and exogenous shock in the number of induced adoptions within a very short period (i.e. within few days, as discussed in section 6). Customers in the treatment groups receive the email only once during the test period, and can click a link to download the app. Customers in T1 are informed that they will automatically get $5 deal-bucks after they download and log in to the mobile app. The email for T1 also states that 1) the offer to get the deal bucks will expire in a week; 2) the deal bucks can be used in deal purchase and would expire in two weeks upon assignment. I also designed other aspects of the experiment very carefully. First, I strictly control spillovers across different test groups. Specifically, all the promotion/information is provided only through the email channel; thus customers cannot participate in the campaign through alternative channels. In addition, the $5 incentive is automatically tied to account ID of the customers in T1; thus other customers outside T1 are not eligible for the promotion. Second, once customers in different groups adopt the app (C, T1, T2), they will have exactly the same experience and receive the same information in their mobile app. This ensures that any difference in their future behaviors can be attributed to how they are induced to adopt the app in the first place. The interventions used in the field experiment are common industry practice and used by other ecommerce platforms in their email campaigns. For instance, major retailers such as Amazon, Groupon and LivingSocial have offered monetary incentive ($5 or $10 credit) for mobile app adoptions in historical campaigns, and have also promoted apps regularly using informational email (Dedman 2011, Techcrunch 2015b). Thus, the external validity of my interventions is assured and my findings may be generalizable to other settings.

2.5 Identification Strategy

With the field experiment, I seek to understand three types of causal effects: (a) the effect of the two treatments (T1: incentive and T2: information) on customers' app adoption decisions (Q1a); (b) the differential effect of incentive- and information-induced adoptions on customers' purchase behaviors (Local Average Treatment Effect, or LATE); (c) the effect of the two treatments on overall profitability (Q1c, intention-to-treat effect or ITT). The relationship between the three types of causal effects ('Adoption', 'LATE', 'ITT') is illustrated in Figure 2. They correspond to questions Q1a-Q1c highlighted earlier. Specifically, I answer Q1a and Q1c through direct comparisons between test groups, and address Q1b using the framework of (Imbens and Angrist 1994) (i.e. using the random assignment of test group as an instrument for the effect of induced adoptions). A main advantage of the intention-to-treat (ITT) analysis in my study is that the findings can be directly applied to similar contexts because the interventions used in the field experiment are standard industry practice. The identification of causal effects in Q1a and Q1c is straightforward, with specifications in Equation (2) and Equation (3) in Section 7; therefore I focus my discussions on the motivation and intuition behind the LATE approach in identifying the causal effect of induced app adoptions (Q1b).

An interesting aspect of motivating mobile app adoption is that only a small portion of users would ever be induced to adopt the mobile app[5]. However, these are exactly the users that firms can influence through external interventions. Thus, identifying the causal effect of mobile app adoptions on purchase behaviors for this population of induced adopters is important for firms. As discussed in the

[5] This is true for both campaigns across different ecommerce platforms (i.e. as revealed from low conversion rate of app adoption), and for technology adoption in general (Ashraf et al. 2010, Dupas 2014).

introduction, it is challenging to use observational data to identify the causal effect of induced app adoptions: first, using observational data, one cannot differentiate the induced adoptions from organic adoptions, and furthermore, one cannot differentiate different types of induced adoptions[6]; second, the identification strategy used in recent observational studies (e.g. matching, difference-in-difference) is based on the assumption that all endogeneity can be controlled by observables. However, this is unlikely in my case. As discussed later, I find evidence of sorting on unobservable, i.e. customers have private information about their future need of the app and such information cannot be explained by observables. To address both of these challenges, I conduct a field experiment and use interventions to exogenously create variations in induced adoption, and identify the causal effect of induced adoption using LATE. The randomized experiment, with external interventions randomly assigned over control and treatment groups, creates two unique features that are missing in the observational data: 1) a control group where only organic adoptions happen and 2) an exogenous and large shock to generate variation in the induced adoptions. The control group serves as a counterfactual (with only organic adoption) and helps us isolate the additional induced adoptions in treatment groups. The exogenous variation in induced adoptions ensures that its effect on customers' purchases is causal.

LATE: I explain the intuition of Local Average Treatment Effect (LATE) using Figure 2. In all three test groups (C, T1, T2), a portion of the customers would

[6] I want to highlight that beyond incentive- and information-induced adoptions, there are potentially other types of induced adoption, such as firm-created word-of-mouth. However, the existence of WOM or other non-experimental induced adoptions would be taken care of by the random assignment in my experiment. The blue square in figure 2 may include both adopters from uncontrolled factors (e.g. WOM), but they are equally distributed across test groups. My experiment is not designed to identify the exact source of each adoption at the individual customer level. Rather, the randomized experiment assures that any organic or non-experiment induced adoptions are equally distributed across groups at an aggregate level. Therefore I can directly attribute the differences in customers' app adoptions and the purchase behaviors to the different treatments. Such comparison allows us to cleanly identify the differential effect of treatments tested in the experiment (incentive- vs. information- intervention), without the need to overcome attribution challenges and understand the source of each adoption (e.g., WOM, delayed effects or other uncontrolled factors).

adopt the app organically. They are denoted as “always-takers” in the LATE framework (Athey and Imbens 2017) and are represented by the solid squares. In addition, some customers in the two treatment groups may be induced to adopt the mobile app after exposure to the interventions. The adoption decision is contingent on the specific intervention used by the firm and customers’ own information about their need for the app. Such induced adopters are called “compliers” in LATE framework, and are represented by dotted square (for incentive-induced adopters) and dashed square (for information-induced adopters). LATE can causally identify the treatment effect on (different types of) compliers, based on the following logic. First, the experiment incorporates a control group with no intervention, and thus provides a perfect counterfactual. I can observe what *would happen* if users had not received any interventions (control in Figure 2). Therefore, I can isolate the compliers at an aggregate level by comparing the adoption decision of customers in the control group with those in each treatment (dotted square for compliers in T1 and dashed area for compliers in T2). Second, for compliers in each treatment, I can separately identify the effect of induced app adoptions on their future purchase behaviors. Technically, I run a two-stage least square regression using test group assignment as the instrumental variable (see Equation 1, Wooldridge 2010, Imbens and Angrist 1994). Since the intervention is randomly assigned, the identified effect is not correlated with any confounding factors (e.g. endogenous targeting) and thus causality is assured (Adomavicius et al. 2013, Chen et al. 2017, Cui et al. 2016, Qiu and Kumar 2017, Zhang and Zhu 2011).

$$Download_i = \sum \alpha_k * T_{ik} + \varepsilon_i$$

$$Purchase_i = \beta * Download_i + \sigma_i \qquad (1)$$

I want to highlight that the characteristics of compliers might be different from those of the average users on the platform. However, it is important to remember that 1) the identification of the treatment effect on those users is perfect because of the control group and exogenous interventions[7]; and 2) those users are exactly the *population of interest*, as firms can only actively induce / influence those users for app adoption using interventions. In contrast, the average users are hardly the (influence-able) adopters. Therefore, the causal effect identified by LATE for this population of compliers (dotted area or dashed area) are exactly what firms wish to know. The LATE approach offers two benefits: the causal effect identified by LATE is 1) specific to the compliers (the population of interest); 2) contingent on the instrument/intervention the firm uses (Athey and Imbens 2017). Thus firms can compare the effect of adoptions induced by different types of interventions.

The identification of local average treatment effect (LATE) is based on two key assumptions. The first is monotonicity (Imbens and Angrist 1994), which requires the probability of app adoption is increasing when a user is treated (Adoption(treated) > Adoption(control)). This assumption is satisfied in my context as my email campaign provides more information or incentive for app adoption. The second assumption is the exclusion restriction, which requires there is no direct effect of the treatment (receiving email with information or incentive about mobile app) on the outcome (e.g. purchase behaviors), without being mediated by the mobile app adoption. In other words, all changes in future purchase behaviors should be driven by the difference in mobile app adoptions. In my context, the information treatment (T2) is one-time only and only mentions the benefits of the app (rather than encouraging purchases). Thus it satisfies the exclusion restriction. The incentive treatment (T1) provides a monetary incentive for app adoption (i.e. $5 deal bucks) but the incentives expire within two weeks of assignment. Thus I would expect a short-run increase in customers' purchases in T1 due to the effect of the incentives within

[7] Again, I want to emphasize that the identification is perfect in my case and there is no confounding problem as in most observational studies. Please see Imbens and Angrist 1994 and Athey and Imbens 2017 for more details. The treatment effect may vary based on the interventions used to induce adoption. I term this difference as "sorting".

the first 3 weeks (i.e. 1 week to claim the offer and 2 weeks to use deal bucks before expiration). However, firms are interested in the long-term effects of induced app adoptions beyond this short window, for both T1 and T2. Therefore, I exclude all the purchases within the first 3 weeks after the experiment for all test groups when estimating LATE. In this way, I can ensure my interventions do not directly affect the dependent variable in the second stage (purchase behaviors). Thus, the exclusion restriction is satisfied. Any effect on customers' purchase behaviors in the long run (6 or 12 months) can be attributed back to the differences in induced app adoptions.

2.6 Research Context and Data

I collaborate with a leading daily deal platform in the US and conduct a large scale randomized field experiment to examine how the firms can actively influence customers' adoption of mobile apps, and whether such induced adoptions can lead to an increase in customers' purchases. The collaborating platform offer a wide range of deals on local services and standard products, including restaurants, entertainment, outdoor activities, home service, retailing products, fitness activities, travel, beauty and health service. The vast majority of deals on the platform are local deals. Depending on the location (e.g. big vs. small cities), the inventory of deals per city may vary between a few and a few hundred deals. Customers need to incur a search cost when looking through daily deals, especially in those cities with a large inventory.

The randomized field experiment was run on the platform for one day and I am able to collect information for the entire sample of more than 230,000 unique customers over a long period after the experiment. For every customer, I record information including the unique hashed identifier of the customers, the assigned test group, the mobile app adoption status (and adoption time), and all purchases before and after the experiment. For each purchase, I record detailed information including the purchased deal, the revenue/discount from the purchase, as well as the purchase channel (mobile web, desktop). I further augment the purchase dataset with rich deal characteristics (price, category, location, merchant). The resulting dataset enables us to analyze the effect of different interventions at an aggregate level as well as at a more granular level.

2.7 Results and Discussion

I first check the validity of my randomization. In Table 1 I provide the breakdown of major covariates in the three groups. As shown in the results, there are no significant differences across the groups on all the covariates (number of past purchases in total and across channels, number of units per purchase, number of purchased categories, average price of purchased deals, total revenue, customer tenure). The well-balanced sample indicates that my randomization is at work.

2.7.1 Q1a: The effects on app adoptions

I examine the effect of incentive (T1) or information (T2) in motivating mobile app adoption by estimating a Linear Probability Model (LPM) on the full sample, as shown in equation (2). A similar strategy is widely used in field experiment studies, as illustrated in Duflo et al. (2008). The results are robust under alternative models such as Logit and Probit (Table A2).

$$Download_i = \sum \alpha_k * T_{ik} + \varepsilon_i \qquad (2)$$

Download$_i$ is a dummy variable indicating whether the customer i has downloaded the mobile app within a certain time frame. Since customers can respond any time after receiving the email, I examine the results using different time frames to understand how the effect changes over time (e.g. 1 day, 3 days, 1 week and 2 weeks after the experiment). T_{ik} is the dummy variable of test group k that the customer i is randomly assigned to.

The results are presented in Table 2. Both incentive (T1) and information (T2) lead to a significant increase in customers' mobile app adoptions, and such effects are consistent across different time frames (*1day, 3day, 1week, 2week*)[8]. The magnitude of increase is economically significant: providing incentives can lead to an 466% increase in app downloads over that in control group; while providing information leads to an 144% increase over control group (based on downloads within 3 days). The stronger effect of monetary incentive (T1) on technology adoption is aligned with previous findings (Dupas 2014). In addition, consistent with the temporary nature of email communication, there is a large increase in mobile app adoption in both T1 and T2 on the first day, and the differences in app adoption (T1-C and T2-C) becomes stable within a week after the intervention (Table 2).

In summary, my results show that firms can effectively motivate customers to adopt mobile apps using external interventions, with monetary incentive leading to significantly more adoptions than pure information. The key question then is whether such app adoptions induced by external interventions can lead to a significant increase in customers' purchase behaviors, and whether incentive-induced adoption is more effective than information-induced adoption.

2.7.2 Q1b: The causal effect of induced app adoptions on long-term purchases

I am particularly interested in the causal effect of induced mobile adoptions, rather than organic adoptions, on customers' purchase behaviors, because firms can active influence the level of induced adoption by providing incentives or information. As discussed in Section 5, I adopt the Local Average Treatment Effect (LATE) framework (Angrist et al. 1996, Imbens and Angrist 1994) to identify the causal effect of induced adoptions. As described in the Empirical Strategy section, the causal effect of induced adoptions is identified by LATE wherein the exogenous treatment assignment serves as the instrument variable to isolate the induced adoptions from

[8] The effect of incentive and information on app adoption is also consistent and stable in the long run, for instance using alternative time windows (1, 3, 6, 9, 12months).

organic adoptions. As discussed above, the effect of both incentives and information on app downloads become stable after one week. Thus, I use the download within the first week after the experiment as my outcome variable in Stage 1 and the instrument in Stage 2. I choose the time frame to include as many induced adoptions as possible and also to exclude organic adoptions to maintain the power in second stage estimation[9]. To examine the long-term effects of induced app adoptions on customers' purchase behaviors, I examine their purchases in two time frames after the experiment -- within 6 months and within 12 months. Following discussions in Empirical Strategy section and consistent with my focus on behavior changes in the long run, I exclude all purchases within the first 3 weeks when constructing the purchase outcome (the results for purchases within first 3 weeks are separately presented in Table 6). The results are robust when I exclude purchases in alternative time windows: first 3 months and first 6 months (Table A1).

I present the results from LATE in Table 3. Interestingly, I find the only app adoptions induced by information (T2) lead to a significant increase in customers' purchases in the long run. In contrast, app adoptions induced by incentive (T1) have no causal impact on customers' purchase behaviors. The results suggest that the effect of mobile app adoptions heavily depends on how customers are induced to adopt the app in the first place. Though monetary incentives (T1) are effective in driving people to adopt the app, such recruitment approach does not lead to more purchases from customers after they download the app. In contrast, providing information leads to a smaller increase in app adoptions but such adoptions lead a sustainable increase in customers' purchases in the long term. These findings show the nuanced tradeoffs between motivating mobile app adoption and appropriating value from such adoption, and provide guidelines to firms on how to encourage effective app adoptions.

[9] I also estimated LATE using alternative instruments, i.e. download behavior of 3 days, 2 weeks and 3 weeks. The results are qualitatively and quantitatively similar to the results using download behavior of 1 week.

2.7.3 Q1c: The impact of incentive vs. information on overall profitability

I run an OLS model on the full sample to examine the effect of two interventions on customer profitability (i.e. number of total purchases as well as net profits) using equation 3.

$$Purchase_i = \sum \gamma_k * T_{ik} + \epsilon_i \qquad (3)$$

$Purchase_i$ is the number of purchases within a time frame for customer i. I use customers' purchases within 6 months and 12 months after the interventions (including the first three weeks) as the outcome measure to investigate long-term effect of my interventions on customer profitability. The results are displayed in Table 4. Consistent with the findings on the causal effect of induced adoption (LATE), I find providing information (T2) has a positive and significant effect on customers' purchases (Table 4) and net profits (Table A3) across different time frames (but does not increase product returns as shown in Narang and Shankar 2016, see Table A13). Such increase is about 2-3% of net profits from all customers in the treatment group (T2) and amounts to hundreds of thousands of dollars if the information intervention is scaled up to target all active users on the platform. In contrast, providing incentives does not lead to any significant increase in customer profitability, as measured by both total purchases (Table 4) and net profits (Table A3). After taking into account the cost of incentives for adopters, the firm may lose a good amount of investment on average on each customer in the treatment group. Overall, my results suggest that providing information may increase customer profitability while monetary incentives may not, though the latter may lead to more app adoptions. Given the lower cost of providing information compared to monetary incentives, my findings indicate that managers should use information provisioning as the main mechanism to encourage mobile app adopters with aligned need.

2.7.4 Q2: Multi-channel purchase behaviors

Although the previous analyses consistently support a positive effect of information-induced adoptions and reveal the underlying driver of the difference between the effects of information- vs. incentive-induced adoption, they do not shed light on the process underlying the treatment effect. In the rest of this section, I delve into the process and explore how app adoption affects customers' online shopping behavior across channels.

I first investigate the channel interdependencies for different types of induced adoptions, by decomposing customers' purchases into different channels. There are three channels that customers can use to browse products and make purchases – desktop (PC), mobile web and mobile app. The mobile web channel provides a smaller and customized view of the desktop website to fit the mobile screen. Mobile app offers the same set of products as the desktop and the mobile web but presents them in a way that is more convenient for mobile browsing and search.

My objective here is to understand the causal effect of induced mobile app adoptions on customers' purchases through these different channels. I follow the same empirical approach discussed in Q2 by changing my dependent variable to customer's purchases within each channel, i.e. *Desktop_Purch* for desktop purchases, *MobileApp_Purch* for mobile app purchases and *MobileWeb_Purch* for mobile web purchases. I use the LATE approach and leverage exogenous treatment assignment as the instrument variable for my identification. The results are presented in Table 8. Recalling my results on overall purchases, induced adoptions by information (T2) lead to a positive and significant effect on customers' purchases while those by incentive (T1) do not. In examining the results of the analysis of data broken down by channels, I find that adoption induced by monetary incentives (T1) has a significant negative impact on *mobile web purchases* while it has a positive impact on *mobile app purchases* (not statistically significant). The two channels substitute each other, resulting in a non-significant net effect. In contrast, app adoptions induced by information (T2) have a positive and significant impact on purchases through the *desktop channel*. This indicates that the desktop and mobile app channels are complementary to each other for information-induced adopters (T2). Such complementarity is aligned with findings in previous literature (Xu et al. 2016, De Haan et al. 2015). The above results indicate that the incentive-induced adopters merely shift their purchases from the mobile web to the mobile app channel (as the latter offers a better shopping experience on the same device), but do not significantly change their total purchases; whereas the information-induced adopters still use the desktop intensively, but add mobile app as a complementary channel for deal discovery in their online shopping process. This is consistent with recent observations in the industry: For instance, citing various reports, Forbe (2017) suggested "…retailers need to be aware that the customer journey is not simple. Many will view items on their mobile devices but only purchase on their desktop or laptop computers."

2.7.5 Q3: Underlying Mechanisms

As discussed in the Theoretical Grounding section (section 5), the differential impact of incentive-and information-induced adoption on customers' purchase behaviors is potentially driven by sorting, i.e. those information-induced adopters are different from those incentive-induced adopters in their observable characteristics and unobserved need (Figure 3). I now empirically test the derived implications in section 3 related to sorting on observable and unobservable:

First, I indeed find adopters are different in their observable characteristics (Table 5): the information-induced adopters, who benefit more from app adoption, make *fewer* purchases before the experiment, but on average, purchase more expensive deals. In addition, I also find that the incentive treatment (T1) increases customers' purchases in the short run (e.g. first 3 weeks) in the mobile app channel

(Table 6), but the effect disappears after the first 3 weeks. The evidence is aligned with the explanation that monetary incentive may attract users who do not value the mobile app as much, but are more interested in the monetary incentive.

Given the findings on the differences in observables for adopters across different groups I ask: could firms leverage customers' observable characteristic to predict who would need the mobile app most (or will increase their purchases most from mobile adoption)? If yes (i.e. sorting on observable), then the firm can actively target those customers and encourage mobile app adoptions using monetary incentives (which would lead to a higher adoption rate). This would enable the firm to achieve both coverage and effectiveness in motivating app adoptions. If the answer is no (i.e. sorting on unobservable), then the firm would need to design an appropriate intervention (i.e. information provisioning) to encourage sorting, and may need to make a tradeoff between coverage and effectiveness in app adoption. I formally test whether sorting is purely based on observables by adding observable covariates in both stages (Imbens and Angrist 1994, Angrist et al. 1996). Interestingly, I find that after accounting for various observable characteristics, there is still a strong positive relationship between indicator of informational treatment (T2) and the treatment effect (Table 7: the magnitude of the causal effect is almost the same and highly significant). The results indicate a strong form of sorting -- sorting on unobservables – driving induced adoptions. In other words, customers possess private information about their needs or potential usage of the mobile app that is unobservable by the firm. Thus firms can only use certain intervention to attract the customers with a higher need to sort into adoption, therefore increases customers' purchases in the long run.

I further test implication 4 related to sorting on unobservable. If sorting is at work, providing information (T2) would attract customers who download the mobile app due to an unobserved need, and therefore use the app more often and discover more deals. Consistent with increased app browsing as suggested by sorting, I find that those information-induced app adoptions are likely to lead to a larger variety of purchases in general (Table 9), Moreover, such induced adoptions are more likely to lead to a larger increase in customers' purchases in cities with higher deal densities (Table 10), where customers may benefit more from app usage.

Finally, my intervention (i.e. an information email) may affect the customers' purchase behaviors in the long run through two mechanisms: sorting and treatment (or influence). The treatment mechanism indicates that information-induced adopters would be similar to organic adopters (i.e. no sorting). The differences in the outcomes are only driven by the fact that they have received the information (email), which may directly influence their purchase behavior in the long run. It is notoriously hard to rule out the treatment mechanism using observational data (Wooldridge 2010), as the same intervention (i.e. information email) may induce sorting and a treatment effect at the same time. I therefore design and implement an additional randomized experiment to rule out the treatment/influence explanation. i.e. whether the information treatment would have a direct impact on customers' type and directly influence their behavior. To this end, I carefully choose over 2,700 users who have already adopted the mobile app (existing adopters), and randomly assign them into control and treatment group. All users in the treatment group would receive an information email about the app. Using existing app adopters as the experimental subjects allows me to turn off sorting and cleanly identify the effect of the information email (T2). I find that the effect of information treatment is not significant in driving customers' purchase (Table 11). These findings confirm implication 5 and provide further support of sorting mechanism as the driver behind the differential effect of induced app adoptions on customer behaviors.

2.8 Conclusion and Future Directions

In summary, my study is among the first to investigate how firms can actively influence customers' adoption of mobile apps and increase customers' purchases through induced app adoptions. My study confirms that firms may motivate effective app adoption and increase net profits, but only when using appropriate intervention. Contrary to the conventional wisdom and common practice, I find providing incentive may induce negative sorting in mobile app adoption and does not lead to long-run increase in customers' purchase and firm's profitability. In contrast, information-based intervention may attract the right group of customers who have strong need of the app and would use it effectively (i.e. positive sorting). By leveraging a carefully-designed randomized field experiment, my study is among the first to show that the causal effect of induced app adoptions may critically depend on *how* customers are motivated. I further look into the underlying driver of such difference in the effect of induced adoption (i.e. sorting on unobservable), and examine how adopters induced by different interventions behave differently in their purchase behaviors across multiple channels. The nuanced findings of the study not only provide guidelines for designing interventions to motivate effective mobile app adoptions, but also add to our understanding of the role of mobile apps in changing customers' online shopping behaviors.

I believe there are a few interesting directions for future research: First, my study demonstrates a fundamental tradeoff firms are facing when designing intervention to motivate app adoption: how to balance coverage (more adoption) and effectiveness (better adoption)? I show that offering incentives may boost coverage, but providing information, rather than incentives is what leads to effective adoptions. Ideally firms want to predict a subset of customers that would positively benefit from induced app adoption and target them with monetary incentive to enhance adoption. However, I find that the difference in treatment effects of induced adoptions cannot be fully predicted by observable characteristics. Thus firms may only rely on the appropriate intervention to encourage sorting. However, to a certain extent, firms may still be able to identify certain customer segments that would benefit more from induced app adoption. For instance, I find that information-induced adoptions (T2) are more effective for customers who have only used the desktop channel in their past purchases (Table 12). Future research can extend my study by combining prediction with active intervention in targeting application (Li et al. 2015) to achieve better coverage and higher effectiveness of mobile app adoption.

Second, in the current study, I keep in-app experience the same for adopters in three test groups and do not vary in-app intervention, as my goal is to identify the causal effect of induced app adoption (rather than in-app intervention) on customers' purchases. My study demonstrates that it is important for firms to 'get the right adoption' at the beginning, as such adoption may lead to long-run increase in customer profitability. However, upon adoption, firms may also use in-app interventions to further engage customers (Son et al. 2016, Kato-Lin et al. 2015). Understanding the effect of in-app intervention, contingent on various types of adoption, is thus crucial. Future research may extend my study by examining different types of mobile interventions (Chen et al. 2017, Wang et al. 2016, Li et al. 2015, Zhang et al. 2016), *after* the technology adoption (Retana et al. 2017). Researchers may also investigate whether and how mobile interventions can be customized for different types of induced adopters (Ma et al. 2007, Ghose 2017).

Third, my paper has focused on identifying the effect of app adoptions induced by two most commonly used interventions in mobile app campaigns: incentive and information (Techcrunch 2015a, Financial Times 2016). Future research may further investigate the effect of app adoptions induced by other types of interventions (e.g. firm-created word-of-mouth).

Chapter 3: Beauty and Signaling in Online Matching Markets: Evidence from a Randomized Field Experiment

3.1 Introduction

Online matching platforms have proliferated over the past decade, and a central challenge faced by these platforms is how to mitigate information asymmetry and facilitate matching among strangers. In response, online matching platforms have made significant investments in a variety of information mechanisms that seek to mitigate information asymmetry and increase overall credibility (Luo, 2002; Pavlou & Gefen, 2004; Urban et al., 2009). Among these mechanisms, phone verification of registrants is one of the most frequently used mechanisms to assure users are real and to avoid them changing identity via multiple accounts[10]. While mandatory verification is widely implemented in practice and well understood as an effective assurance for overall credibility and security, it is nevertheless, costly in terms of time and effort for participants, and thus requiring it for all participants can significantly impede the growth of the platform. More importantly, such mandatory verification may suppress useful information that is conveyed when users voluntarily choose to verify themselves. Taking these into account, in this study, I therefore seek to understand a different role of verification that has not been studied previously – its ability to serve as a credible signal for individual users, when such verification is made optional and visible to others.

My study focuses on non-transactional markets and online platforms such as online dating markets that lack alternative information mechanisms, such as reputation and transaction-assurance mechanisms (e.g., online reviews, escrow services, and money-back guarantees). In addition, users on these platforms also have to rely on self-disclosed information by potential partners. In such contexts, providing

---

[10] Phone number verification is more secured than commonly used email verification. It becomes the preferred and equivalent way to verify user identity.

verified and credible information of one's authenticity through the simple mechanism of phone verification takes on additional significance.

Unlike previous studies of other information mechanisms that largely focus on their *ex post* effectiveness, my study also examines the *ex ante* choices of opting in to verification. My study is also unique as it focuses on a two-sided matching market wherein both sides have the choice of opting in to the same signal. Therefore, I am interested in the differential implications for the two sides of the platform as the value of verification could differ across user types as well as across the two sides of the platform. Moreover, I focus on the role of paid verification wherein the platform charges a fee for verifying users, given the potential benefits to verified users (Ba et al. 2003; Goes and Lin 2012)[11]. I compare it with free verification wherein there are no monetary costs for verification. More specifically, I ask:

Q1. Verification Decisions:

Q1a: Given the option to verify, who would choose to verify? How do the opt-in decisions differ across the two sides of the platform?

Q1b: What is the underlying mechanism driving such decisions?

Q2. Impact of Verification:

Q2a: How does verification influence the potential partners of verified users?

Q2b: How does verification impact verified users themselves?

Q3. Verification Impact on the Platform:

Q3a: How does verification affect matching between the two sides of the platform?

Q3b: How does paid verification differ from free verification?

To examine these questions, I collaborate with one of the leading online dating platforms in the US. I choose an online dating market in particular for four main

[11] Phone verification is a credible signal since it is costly for spammers and scammers to pass this verification. Many of them create fake accounts in batches and use bots for email verification and communication (Salge and Karahanna, 2018). Research also suggests that verified information can inspire trust (Vosough et al., 2018).

reasons. Firstly, online dating markets have gone mainstream and have gotten increasing attention from researchers and marketers. One in five new relationships and one in six new marriages begin online[12]. As a business, it is worth $2.5 billion annually in the US alone[13]. More importantly, online dating markets, different from most online transactional matching markets that have well-established reputation systems, have no alternative signals to mitigate information asymmetry. Further, almost all user information is self-disclosed. As noted earlier, optional verification could play a more strategic role for users under such circumstances. Thus, the online dating platform I collaborate with provides an ideal environment to study optional verification and its differential impacts on the two sides. Last but not least, while there have been studies in other types of online matching markets, empirical studies and particularly randomized field experiments in these markets have been very limited due to a number of concerns and limitations (Coles et al. 2010; Hitsch et al. 2010).

I design and conduct a randomized controlled field experiment to draw causal inference. Before I introduce optional phone verification to users in the online dating platform, no verifiable information has been asked for, or has been disclosed on this platform. I randomly assign existing users to one of the three groups: one control group, and two treatment groups. The treatment I (T1) is a one-time invitation for verifying the user's account through their mobile phone with a charge of 50 virtual coins (equiv. $2) while treatment II (T2) is the same invitation but at no monetary cost. The invitation message reveals the visibility of verification to other users. After verification, these users get a badge prominently displayed on their profiles indicating

[12] Caitlin Stewart, The Dating Services Industry in 2016 and Beyond. 2016.

[13] Kapital Wire Team, Of Love and Money: the rise of the online dating industry. 2016.

the phone number is verified, which is visible to all users on the platform, including those in the control group.

I follow the signaling and related theories to guide my empirical investigation on the *ex ante* opt-in decisions as well as the *ex post* verification impacts of optional verification. In examining the opt-in decisions (Q1a), I find that only 8% of the users choose to pay a fee to verify in T1 while 10% of the users choose to verify for free in T2. Sharing personal information is costly, and additional charges further increase the cost and result in lower opt-in rates. I am particularly interested in understanding how the two sides on this platform – males and females – make verification choices and how it further affects their outcomes. I segment females and males into three tiers respectively– Low (L), Medium (M) and High (H) types - based on their popularity on the platform. Very interestingly, I find only the male side aligns with the conventional prediction of signaling. Specifically, I find that males have a monotonic pattern of verification, with higher type males being more likely to opt-in to verification. However surprisingly, I see a non-monotonic pattern of verification for females, wherein M-type females are more likely to opt into verification relative to L- and H-type females. Such a non-monotonic pattern has also been documented in an emerging literature in signaling and voluntary quality disclosure in other contexts. For instance, top business schools are less likely to disclose their ranking than middle tier schools as they confidently rely on existing favorable information/signal (Luca and Smith, 2015).

To understand this differential opt-in pattern among males and females (Q1b) within the same platform, I consider the working conditions for signaling - H-types would be more likely to adopt the new signal if it adds values to the existing signal (Riley, 2001; Spence, 1978; Spence, 2002). A new signal tends to be more valuable when the existing signal is less credible or less valuable (Boulding and Kirmani, 1993; Wells, et al., 2011). Following these theories, I examine the existing dominant signal for each side in my context- income of males and beauty of females, as borne out by prior research (Abramitzky et al., 2011; Hitsch et al. 2010b). I apply state-of-the-art deep learning framework to measure beauty using the photos of females. Consistent with the aforementioned theories, I find supporting evidence suggesting that in my context the decision whether to opt-in to verification depends on how credible the existing signal is. Beauty among females, assessed through images, is a credible and easily verifiable attribute. Consequently, H-type females with high beauty scores have a lower incentive to verify, while M-type females are more likely to verify to differentiate themselves from L-type females. However, income among males, as a self-disclosed value, is noisy and not easily verifiable. In keeping with this I find that males, especially H-type males, find verification to be a valuable signal and are more likely to verify as compared to M-type and L-type males.

In examining the outcomes of such verification on the potential partners of verified individuals (Q2a), I examine both the changes in the *quantity* of the messages received as well as the changes in *quality* (i.e., popularity) of the senders who initiate the messages. Consistent with rational *ex ante* opt-in decisions, verified users do benefit as they get more contacts from better users, and the tiers that are the most likely to verify (i.e., H-type males and M-type females) benefit the most. Moreover, interestingly, I find that upon verification the verified users become more proactive and are more likely to initiate messages to higher-quality potential partners (Q2b).

Going beyond individual users, I examine the impact of verification on the platform as a whole and find that it facilitates matching between the two sides (Q3a). Matching platforms often seek to balance out the attention among superstars and average users and to promote high quality users when the market is crowded and competitive (Roth, 2015). With the introduction of this voluntary verification signal, M-type females are able to move closer to H-type females while H-type males manage to further stand out from the crowd. This, in turn, allows females to easily identify these H-type males for desirable matching outcomes. Such improvement on matching quantity and quality result not only from the signaling effect of verification but also from increased pro-activeness among verified users. Further, I examine free verification and find that it has similar outcomes as paid verification, although the effects are weaker (Q3b). This highlights a trade-off between higher coverage of verification (free verification) and stronger signaling properties (paid verification).

My study makes a number of useful theoretical and empirical contributions to previous literature. First and foremost, although prior studies have examined the role of verification and its impact on market outcomes, verification in these studies is within the context of a reputation mechanism (Anderson and Simester 2014; Forman et al. 2008; Mayzlin et al. 2014). On the other hand, I focus on a context where there are no alternative reputation or third-party signals that can mitigate information asymmetry. I examine whether and how optional verification can serve an effective signal for further differentiation under such circumstances. More importantly, it is not clear how people strategically choose to adopt this signal. While prior research has examined the effectiveness of information mechanisms such as WOM and third-party certificates (Antony et al., 2006; Buttner and Goritz, 2008; Pavlou & Gefen, 2004; Ubran, et al., 2009) in other contexts, there are hardly any studies that examine users' adoption decisions of such mechanism. My study contributes to this literature on information mechanisms by examining both the user's adoption decision as well as its impacts on user outcomes.

Further, prior studies examining the role of information mechanisms have largely focused on one of the two sides in a platform (for e.g., the seller side; see, Ubran, et al., 2009; Anderson and Simester 2014), similar in the broader context of examining a signaling mechanism. However in a peer-to-peer platform where two sides can both adopt the same signal, such as online dating markets, the *ex ante* decisions to opt-in and the *ex post* benefits of verification may be different for the two sides. My study adds to the large body of work in signaling by examining the *differential* choices and subsequently *differential* impacts of verification across the two sides of an online platform (Luca and Smith, 2015; Riley, 2001). Moreover, I contribute to emerging research on applying deep-learning techniques to business problems, by using deep learning to predict human beauty and examine how beauty impacts females' adoption of verification (Malik et al., 2017).

Finally, my study generates interesting implications for the design of verification mechanisms for online matching platforms. It provides insights on how optional verification can serve as an effective signaling device for the different users on a matching platform and benefit the overall platform by helping M-type females and H-type males stand out. It also provides guidelines for platform designers on the tradeoff between paid verification and free verification. These actionable insights can be particularly valuable for those platforms that have no other complementary reputation systems or are in the cold-start stage with no other established reputation mechanisms.

3.2 Theoretical Motivation

My work builds on and contributes to three streams of literature: The first is the information systems and economics literature on matching markets and related empirical studies on specific marriage markets across economics, information systems, sociology, social psychology and anthropology. The second is the information systems and economics literature on verification and information mechanisms, and the seminal economics literature on signaling. The third is the emerging literature on applications of deep learning to business problems.

3.2.1 Matching and Marriage Markets

Research on matching markets, particularly those relating to market design, focus on designing efficient/ stable matching mechanisms (Coles et al. 2010; Hitsch et al. 2010; Roth 2015). Some examples include kidney transfer, matching of medical interns to hospitals, matching of gastroenterologists to hospitals etc. However, online dating markets have not been studied with equal vigor despite their popularity and importance to society. Online dating markets, in particular, suffer from significant information asymmetries due to self-disclosed information with low credibility as well as no alternative information signals. This is among the first study to examine the role of optional verification in an online matching market. Moreover, I design and conduct a randomized field experiment, which is very limited in other matching markets due to regulation and other constraints.

In particular, my study also draws upon prior research on specific matching markets for marriage and online dating. Most previous work on marriage markets has examined matching and sorting patterns (Bruch and Newman, 2018; Hitsch et al. 2010a; Hitsch et al. 2010b; Abramitzky et al., 2011) and identified gender differences in sorting preferences. For instance, prior research (Histch et al., 2010a; Histch et al., 2010b; Fisman et al., 2006) has shown that males place a higher value on physical appearance of females while females place a higher value on income of males. My study adds to existing research by showing how these differences are related to differential opt-in decisions for males and females. Finally, my findings also contribute to the growing research in Information systems on online dating markets. For instance, Bapna et al. conduct a controlled randomized field experiments to investigate how viewing a potential partner's profile anonymously works as a weak signal of preference (2016). My study in comparison, examines another important signal in online dating markets – signals of willingness to share personal information, which further leads to the transmission of preference signals (e.g., profile views and messages).

3.2.2 Signaling and Verification

My study builds on earlier research relating to verification as a basic mechanism to reduce information asymmetry. Researchers have studied the effect of verification on a number of different outcomes (Anderson and Simester 2014; Forman et al. 2008; Mayzlin et al. 2014). For instance, verified reviews are found to have a greater impact on subsequent review behaviors and product sales (Forman et al. 2008). The impact of verification is mostly examined within the context of reputation system. However, I focus on a strategic role that optional verification may play in a market where there are no alternative information signals, e.g. WOM. Besides examining the effectiveness of verification, I am also interested in how users across the two sides strategically adopt this signal when it is made optional and visible to others. While previous studies (Berger and Rand 2008; Forman et al. 2008; Ma and Agarwal 2007) consider verification as a self-enforcing norm or conforming to a community, I find optional verification can serve as a strategic differentiator in matching markets lacking alternate signals.

The opt-in decisions of verification as a signal stem from the seminal theoretical and empirical studies of signaling. Traditional signaling models suggest that higher types (high in productivity, wealth, or some other valued attribute) are more likely to send a costly signal to separate from lower types (Jovanovic, 1982; Riley, 2001; Spence, 1978; Spence, 2002). Following the intuition of signaling, however, there is emerging literature that suggests H-type users sometimes are less likely to signal comparing with M-type users (Feltovich et al., 2012; Gambetta and Székely, 2014; Luca and Smith, 2015). For example, in Luca and Smith's study (2015), top business school are much less likely to disclose their ranking compared to middle-tier schools as middle-tier schools have more incentive to signal their type and differentiate from low-tier schools. This phenomenon could occur when H-type confidently rely on a favorable existing signal whereas the existing signal of M-type may not be a good enough differentiator.

In general, the decision of adopting a new signal depends on the additional value of the new signal comparing to the exiting signal. Users resort to the new signal when it is relatively more credible or adds new differentiation value as compared to the existing signal (Boulding and Kirmani, 1993; Wells, et al., 2011). H-type users may not adopt the new signal if the existing signal is sufficient. The emerging literature on the non-monotonic pattern largely focuses on theoretical models or lab experiments while my study provides novel evidence of such interesting pattern in an empirical setting. More importantly, my study provides novel evidence of such interesting patterns in an empirical setting. More importantly, while prior studies have provided evidence of either a monotonic pattern or a non-monotonic pattern in different contexts, my study is among the first that documents a heterogeneous pattern across two sides of the same market – wherein females and males strategically use the same signal differently. Moreover, existing studies focus only on the voluntary decisions of adopting the new signal (Luca and Smith, 2015). My study is among the first that also examines and validates the *ex post* benefit of adopting the signal through a well-designed controlled field experiment.

Finally, there is an increasing attention to compare voluntary and mandatory mechanisms in the literature of quality disclosure. Recent studies have identified the role of voluntary mechanisms as a signaling device in other contexts. For instance, the firms that choose to audit, when auditing is no longer legally required, attract upgrades to their credit ratings as such voluntary auditing sends a positive signal of low risk to the public (Lennox and Pittman, 2011). My study contributes to this thread of literature by examining the role of optional verification in its ability to serve as a credible signal for individual users in a matching market, particularly when there no alternative assurance mechanisms and users have to rely on self-disclosed information.

3.2.3 Deep Learning in Business

I adopt state-of-the-art deep learning techniques to construct beauty of females using their photos. Deep neural networks have been applied to various image-related tasks, such as image classification and face recognition, due to the superior ability to learn discriminative features (Krzhevsky et al., 2012; Taigman et al., 2014). There is emerging research that uses deep learning to predict beauty. Most papers in computer science seek to develop better algorithms (Liang et al., 2018; Xu, et al., 2018). Meanwhile, researchers in business and social science have also begun to value this technique's high scalability and consistency to mine values from pictures (Malik et al., 2017; Liu, et al., 2016). For instance, one study used it to examine bias in career development (Malik et al., 2017). I contribute to this emerging line of work by applying a state-of-the-art deep learning architecture to extract beauty of female users in my samples, and examine the implications of beauty for user's decisions to opt-in to verification.

3.3 Experiment

3.3.1 Research Context

To examine the research questions highlighted earlier, I collaborate with one of the leading online dating platforms in the U.S, which has over 1 million registered users in the U.S. Similar to most online dating websites, it offers the following features to its users: Users create their online profiles where they introduce themselves as well as reveal their preferences on seeking partners. User profiles typically also include photos. The platform offers a decentralized system wherein users perform a targeted search for potential partners that filters the profiles by age, location, and other demographic variables. Users are also able to browse others' profiles without limitations and at no cost.

3.3.2 Experimental Design

I randomly select over 20,000 users for my experiment and randomly assign each user into one of three groups. The treatments are verification invitations sent just once to each user who is currently online. It is to maximize the probability that the subject gets the message and reads it as compared to offline messages or emails. Acknowledging gender differences as documented in literature, I block on gender to make sure males and females are distributed proportionally across groups (Histch et al., 2010b; Fisman et al., 2006). The invitation message describes that the verification uses one's phone number, and the purpose of the verification is to reduce 'scams' and to make sure that others know that 'you' are real. The message for Treatment I (T1) says it will *charge* users 50 virtual coins (equiv. $2) if they pass the verification while the message for Treatment II (T2) validates the phone number for *free*. I also emphasize, in the message, that the platform will never disclose or share this information with anyone else. This is to alleviate people's concern about privacy that may hinder their willingness to verify. Each phone number can be used to verify only one user's profile. Specifically, if one chooses to verify, the user is directed to a cell phone verification page upon clicking the verification link in the message. Once the user successfully validates the verification code sent to the phone through SMS, he/she will get a verification badge that is displayed on his/her profile page as well as the search page[14]. The detailed design is illustrated in Figure 1 and the verification badge is as shown in Figure 2.

3.4 Measures

3.4.1 Variables

According to my research questions, I track three sets of outcomes. I track the verification decisions for all subjects in the treatment groups to answer Q1. I also track the messaging behavior among participants for 2 weeks for Q2 since it is the ultimate indicator, in an online platform, of strong interest with higher potential for offline date and matching. Specifically, as shown in Table 1, I track the initiation stages of the messages as well as the ultimate matching. I examine conversations originating from both sides; a conversation can be initiated by the focal users or by the potential partner. Similarly, a conversation match, defined as a three-round back-and-forth conversation following prior literature (Hitsch et al. 2010, Bapna et al.

[14] Please note that verification is not a factor in ranking. In other words, being verified or not doesn't affect where a user locates in others' search results.

2016), can be initiated by the focal user or by the potential partner. I calculate these outcome variables within a certain time window with the notation of *outcome_Xtime*. For instance, I mainly use *verify_1day* and *verify_1week* to study verification decisions within 1day after treatment and within 1 week after treatment respectively. Similarly, I use *msg_sent_1week*, *msg_received_1week* to measure the messages outward and inward within a week respectively. I take log transformation for all outcomes of *msg_sent, msg_received, msg_sent_match* and *msg_received_match* since they are count numbers.

To further study the heterogeneous treatment effects and associate them with different user types, I also link the experiment data with subjects' information disclosed to others, e.g., income, as well as statistics summarized by the platform. I list the detailed coding of each variable in Table 1.

3.4.2 Beauty using Deep Learning

Beauty is a valuable trait for users seeking a match. However, beauty is a more important attribute for females, as it ranks highest in terms of what males look for in potential partners (Hitsch et al. 2010b; Fisman et al., 2006). To obtain this missing measure, I use a deep-learning technique to predict the beauty of female users in my sample, based on their photos. There are a number of techniques to code beauty, including manual coding. However, such manual coding techniques do not scale well and also lead to privacy concerns. Increasingly, supervised machine learning approaches, rather than human raters, are growing in popularity due to their high scalability and consistency on rating criteria. I choose deep learning over other models due to its superior performance on image-related tasks (Malik et al., 2017; Krzhevsky et al., 2012; Taigman et al., 2014).

I follow a recent paper that tackles the problem of facial beauty prediction (Liang et al., 2018). The authors collected a dataset of face images and annotated each image with a beauty score between 1 and 5. They propose to use deep neural networks for the task, due to their extraordinary ability to learn powerful image representations. Thanks to the public availability of their dataset, I are able to train my own deep neural networks to predict beauty scores without having to risk unintended leakage of private user data. For my task, I only consider the female faces in the dataset that have the same racial composition as in my test examples. I modify the last layer of DensetNet-121 (Huang, et al., 2017), one of the state-of-the-art deep learning models, to directly output a scalar value as the predicted score. I also experimented with a few other models and found that DenseNet-121 performs the best due to its commensurateness with the size of my training set and hence its ability to avoid both under-fitting and over-fitting. The trained model predicts for each input image a score between 1 and 5. In Table 2, I show the results obtained from five-fold cross validation. My final model achieves a mean absolute error of 0.24 and a mean squared error of 0.10. Both errors are close to what the authors report in the paper.

To use my trained model for facial beauty prediction on a new image, I apply a preprocessing step to crop out the face(s) in the image using a state-of-the-art face detection algorithm. Images with no face or multiple faces detected are dropped out (roughly 2% of all). I extract beauty score from each female user's leading profile picture[15]. Another alternative is to average across multiple images for a user, and this generates consistent results.

3.5 Results and Discussion

Before analyzing the results, I first examine whether the randomization works well by comparing the key covariates across groups. I find no statistically significant differences among the covariates across groups. Summary statistics are presented in Table 3 where I demean variables across groups to respect the non-disclosure agreement (NDA).

[15] It is shown in the search page and is the landing picture on the profile page.

3.5.1 Question 1 – Verification Decisions

Q1a: Given the Option to Verify, Who Would Choose to Verify?
Since the treatments are assigned randomly to each user, I estimate an OLS to obtain the causal effect of the two treatments on the users' choices of verification. Since gender difference has been suggested in previous literature (Hitsch et al., 2010; Bapna et al., 2016), Irun all the regression analyses for males and females respectively.

$$Verify_i = \sum \alpha_k * T_{ik} + \varepsilon_i \quad (1)$$

As shown in Table 4, I observe that on average 8% of users choose to verify their phone numbers in T1 while an additional 2% of users choose to verify in T2. Such differences are directly attributable to the differences in the costs of verification. By evaluating the verification decision over different time windows, I find that most users make their decision on the same day of treatment and it stabilizes after week 1.

An interesting question related to this optional choice of verification is which types of users choose to verify. Signaling theories posit that given information asymmetry, the higher type users are more likely to verify to differentiate themselves from lower type users. In keeping with this, I segment males and females respectively into Low-, Medium- and High- types, based on their charm score on the platform[16]. As in Formula 1, I estimate an OLS regression to obtain the average rate of verification across quality-tiers, i.e., L-type, M-type and H-type. As shown in Table 5, I find that males display a pattern that matches the conventional prediction of signaling theories – a monotonic distribution of increasing opt-in rate in higher-type males. However, interestingly, I observe a different pattern for females. It is the M-type females that have the highest opt-in rate, rather than the H-type or L-type females. The verification rate among H-type females is actually statistically significantly lower than verification rate among M-type females. There is emerging empirical evidence of such non-monotonic pattern in recent literature of signaling and voluntary quality

[16] The charm score is created and used by the collaborating platform to track each user's popularity.

disclosure. However, my study is among the first to document differential adoption decisions across two sides of the same platform where users on the two sides strategically use the same signal differently.

Q1b: What is the Underlying Mechanism of the Differential Decisions?

I further examine the underlying mechanism of such differential opt-in decisions among males and females. As the literature in signaling points out H-type users are more likely to adopt a new signal when the existing signals are noisy or weak and when the new signal adds value to the existing signal or attributes (Jovanovic, 1982; Riley, 2001; Spence, 1978). However, the opposite is more likely when the existing signal is strong and credible (Feltovich et al, 2002; Gambetta and Székely, 2014; Luca et al., 2015). I therefore examine the existing signals in my setting, which in my setting are the attributes accessible to all users on the profile page, and associate them with the differential opt-in decisions. Following the prior literature of dating and marriage on gender differences (Histch et al., 2010b; Fisman et al., 2006), I specifically examine *beauty* in females and *income* in males as they play a dominant role in sorting preferences. As shown in Figure 3, income of H-type males and beauty of H-type females both rank highest within their gender, compared to M- and L- types. H-type females with their existing favorable dominant signal would be less likely to adopt the new signal compared to M-type (Luca et al., 2015). Therefore, in my context, I conjecture that H-type females rely on their high status of beauty so that verification signal has a much higher value to M-type females than to H-type females. To formally test this hypothesis, I examine how beauty plays a role in H-type females' opt-in decisions. As shown in Table 6 (a), I find that among H-type females, more beautiful users (above-median) are even less likely to verify than below-median users, consistent with the underlying rationale in this emerging thread of literature. As a robustness check, I find the opposite pattern within the M-type females, wherein the more beautiful M-type users (above-median) are more likely to verify, further indicating that M-type females use this verification as a signal to differentiate from lower types.

The next question asks, why do H-type males, on the other hand, behave differently from H-type females when they also have higher values of income than M-type males. To examine this, I follow the same format to investigate how income correlates with the males' opt-in decisions. Consistent with the monotonic pattern in the traditional prediction of signaling, I find that higher income males (above-median) are more likely to opt-in to verification than lower income males (below median) both within H-type males and within M-type males, as shown in Table 6 (b). In line with the theories of signaling, I find that the decision to signal or not in my context depends on whether the existing key signal is credible or not (Boulding and Kirmani, 1993; Feltovich et al, 2002; Spence, 1978). Specifically, income of males, which is self-disclosed, value, and although valuable to the female side, tends to be uncertain and noisy. Thus, H-type males have a stronger incentive to adopt this new signal of willingness to share personal information to complement their existing attributes and to further differentiate from lower-types. On the other hand, beauty inferred from photos, is much less noisy and more easily verifiable. H-type females relying on high values on beauty are hence, less likely to opt-in to verification while M-type females who rank lower in beauty are more likely to verify to enhance their profiles. In other words, the differential opt-in decisions between males and females in my context are driven by the difference in the credibility of the existing dominant signals for males and females.

I want to clarify how the new verification signal works differently than income or beauty. The vertical differentiation within males or females has been established given the existing signals, mainly contributing by income of males and beauty of females. Due to the difference in the credibility of these key attributes, male and female users have different opt-in decisions about the new signal – phone verification. The new signal has nothing to do with the differentiation along income or beauty, but helps to indicate a user's willingness to share personally identifiable information, which is also a favorable attribute to see in a potential partner, especially in online settings. As mentioned earlier, users consider whether or not to leverage this new signal to further differentiate along the "willingness to share" dimension in order to improve the overall popularity, depending on whether their existing signals are sufficient to serve the overarching goal of getting more demand. It suggests that platforms can easily implement such very simple phone verification mechanism to mitigate the noisiness of online information and to facilitate further differentiation among users.

I seek to collect more supporting evidence for understanding what verification signal represents by conducting a survey among platform users. I conjecture that the verification badge indicates willingness to share personal information and it tends to serve as a more credible and stronger signal than information already disclosed online. This is because it is the only verifiable information available and because the other information shared publicly does not tie to a person's identity so close as one's phone number. It is more privacy sensitive and more likely to be used for identifying a real-world person. In order to validate my conjecture, I design a survey to ask for people' LinkedIn information, similar to how I ask for phone verification. Between Facebook and LinkedIn, I choose the latter as it is even more sensitive and relates to professional life. If my hypothesis holds, I expect to observe that users who have verified their phone number are more likely to verify LinkedIn account compared to users who opt out phone verification. Among the 3450 users who respond to my survey, I do find a favoring pattern that users who are willing to verify phone number tend to be open to verify LinkedIn. The pattern is consistent in both female and male users.

3.5.2 Question 2 – Verification Impact

I then examine the causal impacts of verification on verified users on different sides in terms of both quantity and quality. I further decompose these impacts based on user types to understand the differential benefits and validate *ex ante* expectations. To formally investigate *ex post* effects, I briefly introduce my identification strategy. I randomize on group level that makes the recipient of a treatment exogenous. My focus is on the impact of the treatment on the treated users who opt-in for verification, rather than on the overall population who get assigned treatments T1 or T2. Such treatment effect on the treated (TOT) can be obtained in two ways. One is a standard instrumental-variable approach with randomized treatment as an instrument for verification, or a Local Average Treatment Effect (LATE) framework (Angrist et al., 1996; Angrist et al., 2008). This randomized treatment is a strong instrument since it is an exogenous assignment that only correlates with verification but is uncorrelated with subsequent behavior. The alternative is to adopt the intention-to-treat framework to get the intention-to-treat (ITT) overall effect on treatment group level. I then use the opt-in ratio to attribute the treatment effect back to verified users only, the portion that fully accounts for the incremental change. These two frameworks are the same in my setting where there are no 'always-takers' who can opt-in verification in the control group, as shown in Formula 2 (Angrist et al., 1996; Angrist et al., 2008). Therefore, I report using ITT to TOT framework that has the additional advantage to compare between T1 and T2 and across user types. Results using LATE as robustness check are consistent.

$$Verify_i = \sum \alpha_k * T_{ik} + \varepsilon_i$$

$$Msg\_X_i = \beta * Verify_i + \sigma_i \qquad (2)$$

The main outcomes I choose are the number of messages shared between potential partners since it is the ultimate indicator, in an online platform, of strong interest and the potential for offline date and matching (see Table 2). I focus on conversations originating from both sides - initiated by the focal users or by the potential partner. I take log transformation since these message measures are count numbers. All estimations follow the same framework as in Formula 2 but varying dependent variables, *msg_X*. I focus on the paid verification T1 in the first two sections and discuss its comparison with free verification T2 in the end.

Q2a: How does Verification Impact the Potential Partners of Verified Users?

Following signaling theories, I hypothesize that *ex ante* choices of focal users should align with *ex post* responses of potential partners. Specifically, verified users should receive increased attention and from higher quality users. I therefore examine both the changes in *quantity* of the messages received as well as the changes in *quality* (popularity) of the senders who initiate the messages. As shown in Table 7, I confirm that verification leads to an increase in messages received from others, and more importantly, it results in an increase in sender quality as well, measured by the average charm score of the message senders.

To understand the heterogeneity of the effects, I further decompose the impact across each type of users for males and females respectively. Theoretically, the benefit should align with *ex ante* differential opt-in decisions that H-type males should benefit the most among verified males while M-type females should benefit the most among verified females. I follow the same regression above but adding user types and interaction terms with treatment assignment, T1 or T2. As shown in Table 8, I confirm that H-type males benefit the most from verification among males as the coefficients of *H-type*T1* for both quantity and quality are significantly greater than coefficients of *M-type*T1* and of *L-type*T1*. On the other hand, M-type females benefit the most from verification among females both quantitatively and qualitatively since the coefficients of *M-type*T1* are significantly greater than the other two interaction terms. Moreover, such increased attention on M-type females does not harm H-type females, as there is no significant decrease in either messages received or sender quality among H-type females. I further show that with the help of verification, M-type females successfully attract more attention from males and such improvement in their popularity brings them closer to H-type females. However, H-type females still have more messages received than M-type females as shown in Table 9 – suggesting that beauty as reflected in the images serves as a more credible signal for females.

Q2b: How does Verification Impact the Verified Users Themselves?

Besides examining that verification is an effective signaling device, I am also interested in whether and how verification impacts the verified users themselves. Interestingly, I do observe a behavioral change of verified users as compared to their 'counterfactuals' in the control group. As shown in Table 10, verified users become more proactive and are more likely to initiate messages to more potential partners. More interestingly, they become more selective and contact higher type users, measured by the average charm score of the message receivers. Such increased pro-activeness on the female side is very meaningful to the platform since generally females are more passive and less likely to proactively engage in online dating (Bapna et al. 2016). I further decompose this behavioral change across user types as shown in Table 11. I find that H-type males and M-type females send out the most messages to better types compared to other types within the same gender, consistent with the outcomes of message receiving. All results reported here use data within 1 week, and using other time windows generates consistent results, e.g., 2 weeks.

3.5.3 Question 3 – Impacts of Verification for the Platform

Q3a: How does Verification Impact Matching between the Two Sides?

I first discuss the sending and receiving behaviors of verified users as these are the initial step that would lead to a match. Among males, H-type males who are most likely to verify, are now able to further differentiate themselves from M-type males, leading to an average increase in both quantity of messages received and the quality of the senders of those messages. On the other hand, verified M-type females are able to further differentiate themselves from L-type females and receive more messages from higher type males. These positive outcomes for H-type males and M-type females are further strengthened as these verified users become more proactive in sending out messages to potential partners.

Such improvements for M-type females and H-type males are desirable to online dating platform as platforms are eager to design effective mechanism to balance the attention between the most popular users and the less popular ones to enhance market coverage (Roth, 2015). At the same time, platforms also strive to facilitate high quality matches that require appropriate information mechanism to promote high quality users (Roth, 2015). Specifically, I find that optional verification can help enhance M-type females' profiles without hurting H-type females. It also helps H-type males to stand out so that females can more easily identify a potential good match.

Next I examine how optional verification impacts matching of the two sides on the platform. As shown in Table 12, consistent with the initial stage, there is a significant increase in matching quantity for both verified males and verified females. In particular, as shown in Table 13 it is the H-type verified males and M-type verified females who benefit the most from verification. Such beneficial effects do not solely come from an increase in matches that are initiated by others, but more importantly, are driven by the increased pro-activeness of the verified users. This optional verification has a two-pronged effect. It not only helps verified individuals to get more matches through better differentiation but also causes verified users to reach out to more and better potential partners. It also benefits the platform as a whole by balancing the attention among females, and increasing the overall coverage.

Q3b: How does Paid Verification Differ from Free Verification?

A practical question for a platform designer is whether to charge for verification or not given its benefit to users. To provide actionable implications, I design two treatments in my experiment to causally examine the difference between paid verification and free verification on users' opt-in decisions as well as subsequent impact on messaging behaviors. In examining the opt-in decisions, paid verification results in smaller verification rate due to its additional monetary cost (Table 4). In examining the impacts of verification, I find that both types of verification are effective, but free verification (T2) generates weaker results in both receiving and sending messages (see Table 7-8 and 10-11). It is pertinent to note that these differences among verified users between the two treatment groups do not affect the identification of the causal impact of verification. I use TOT to identify the causal impact of paid verification and free verification respectively. To summarize, free verification and paid verification are both effective with a tradeoff that free verification induces more coverage of verification while paid verification leads to more pronounced individual benefit. Platform designers may take their current goal into consideration, e.g., increase overall credibility, or promote more proactive individual behaviors, to decide which one to implement.

3.6 Conclusion

While online platforms place information symmetry at the center, markets with few alternative information mechanisms may face an even bigger challenge. Unlike mandatory verification mechanism that helps enhance overall security and credibility of the platform, optional verification in my study enables individual users to strategically adopt a positive signal to differentiate themselves from other users on the platform. I examine the effectiveness of such optional verification in an online dating market where there is no reputation system available and all information is self-disclosed. More importantly, I investigate who would adopt this signal, and I see interesting and differential opt-in patterns between males and females that they use the same signal differently, depending on their existing signals. This study is among the first to design and conduct a large-scale randomized field experiment to examine these two important questions in an online matching platform.

My paper contributes to, and compliments, previous literature in a number of ways. In keeping with the literature on voluntary disclosure and signaling, theorists suggest "peaches" tend to signal to differentiate from "lemons". There is some emerging evidence that it is not always the case – sometimes M-types have a stronger incentive to signal, but the majority of the work is either theoretical or lab experiments. I provide empirical evidence based on a large-scale study to add to this emerging line of work. More importantly, my context is unique in that users on both sides of the platform can adopt the same signal. I am among the first to document *heterogeneous* opt-in patterns between two sides of a matching market whereas existing work identifies a homogenous pattern that H-type is either more likely or less likely to signal than lower types within the subjects of interest. I also add to the literature on the effects of verification. Previous literature has examined the effects of verification but overlooked its strategic role as a signal in a market with no alternative signals. I am among the first to design a randomized controlled experiment to draw causal inference, and I carefully choose online dating market to isolate the standalone effect of verification. I am add to the broader literature on the impacts of voluntary disclosure by identifying an interesting two-pronged effect of verification – increased attention from the others as well as increased pro-activeness of the verified users themselves. Finally, I contribute to the nascent body of work that applies deep learning to business research.

My findings provide useful insights for matching market designers and platform owners on how to effectively design and implement verification. I demonstrate the value of *optional* verification that maintains the minimal mandatory screening cost and effort but serves as a very effective signal to users on the platform by fostering desirable engagement and matching outcomes. Moreover, the comparison between free verification and paid verification suggests that even free verification can foster a similar, albeit weaker, impact - there is a tradeoff between higher coverage of free verification and lower pro-activeness. Platforms can choose the type of verification that fits the best of their current need.

Chapter 4: Your Preference or Mine? A Randomized Field Experiment on Recommender Systems in Two-sided Matching Markets

4.1 Introduction

Peer-to-peer two-sided matching markets have become major players across many industries, e.g., labor markets, crowd-funding, and online dating. With the rapid growth of these markets, the choices for users expand exponentially exacerbating search frictions. Consequently, platforms resort to personalized recommender systems as one of the most effective approaches to improve the efficiency of search and matching. While researchers and practitioners generally focus on *product* recommendations in transactional markets, there is a dearth of research that studies *user* recommendations in two-sided matching markets.

User recommendation in two-sided matching markets differs from *product* recommendation in transactional markets due to some fundamental characteristics that distinguish the two types of markets. First, a match on a two-sided matching platform is a bilateral decision, as opposed to a purchasing decision in E-commerce, that eventually depends on the preferences of both sides of the markets - focal users on one side vs. potential matches on the other side (e.g., employers vs. employees in a labor market, or men vs. women in a dating market). Given this two-sided nature, focal users may make different choices when the recommendations are generated based on the other side's preference. Another distinction relates to the bandwidth issue of recommendations, especially those in high demand. In transactional markets, a popular item (say, a best-seller) can be recommended to multiple users. However, it is not ideal for matching markets to recommend the same popular candidate to many potential partners since only a few are likely to get a response. This congestion may lead to fewer matches for the platform as the whole.

Clearly, when designing *user* recommendations on two-sided matching platforms, more attention needs to be paid on the candidate pool regarding what preference information is used and how it impacts the platform as a whole. Yet, most online matching platforms provide recommendations similar to that of transactional platforms – their recommendations are largely based on the preferences of the focal user (Horton, 2017). It may not be the optimal practice considering the differences between the two types of markets. Previous studies in Economics and Information Systems have provided some theoretical and empirical pointers to the potential benefits by including the other side's preferences. The findings suggest that providing the information about the other side's preferences can lead to strategic behaviors of the focal users and such provisioning is likely to improve matching outcomes (Avery & Levin, 2010; Cole et al., 2013). However, there have been no field studies that empirically examine its implications for the design of recommender systems in two-sided matching markets.

I seek to bridge this gap and start by investigating how the usage of potential matches' preferences in recommender systems impacts focal users' decision-making. Specifically, I seek to understand what preference information should be used in recommender systems and how it impacts user decision-making and matching outcomes. From the perspective of choices and preferences, the research question can be viewed as in two-sided matching markets, whether and how people make different choices when the choice set includes or precludes the other side's preferences. From the perspective of information provisioning, I can think of the question as to whether and how users' decisions are affected when the other side' preferences are made available. The choice sets and the preference information presented to the focal users here are generated by recommender systems.

Three recommender systems are developed to examine how focal users respond differently. The first system uses the focal users' preference ("Your

Preferences”). The second system is based on potential matches’ preferences (“Potential Matches’ Preferences”). The third takes mutual preferences into consideration (“Mutual Preferences”). All other elements of the three recommender systems are held the same, including the feature set and the prediction model, to make comparisons meaningful. In other words, the three recommender systems only differ in the input data for candidate generation, by using the preference information of the different sides.

To examine the research question based on real user behavior, I collaborate with one of the leading online dating platforms in the U.S. and design and conduct a randomized field experiment. I choose online dating as the representative context not only due to its prevalence but also due to its flexibility to conduct randomized field experiments compared to other matching markets (Coles et al. 2010; Hitsch et al. 2010). Besides the three treatment groups that respectively implement the three recommendation systems, there is a control group with a baseline model. As opposed to the common baseline choice that shows a random list of users, I modify it to randomly show the top popular users to serve as a “higher” baseline. To avoid carry-over effects, I adopt a between-subject design that makes sure every subject is assigned only to one group of the experiment. Furthermore, the treatment contains two inseparable elements[17]: 1) recommendations generated by the assigned recommender system, which, compared to the recommender systems in other treatment groups, only differs in the preference information used as input, and 2) the associated title of the assigned recommender system to inform users what preference information is used to generate the recommendations. This study seeks to provide design implications for two-sided matching platforms on what preference information should be used so the recommendation content needs to be truly based on different

17 It would be interesting to separate the two elements and examine only one of them, however, these are beyond the interest and goal of this paper and can be pursued as future directions.

preferences rather than merely a manipulation of framings without changing the content. This is consistent with the literature on preference signals (Avery & Levin, 2010; Cole et al., 2013), which notes that the information has to be 'transparent' to the focal users for it to be effective. Finally, users are also informed that the ordering of the candidates is based on the fitness of the designated preference. For instance, in "People who might prefer you" group, candidates on the top have a higher likelihood of preferring the focal users than the candidates on the bottom.

In examining the effects, I find that users are responsive to "the other side's preferences", which leads to both quantitative and qualitative impacts on the platform. Specifically, users in the "People who might prefer you" group and "Mutual Preferences" group are as proactive as users in the "Your Preferences" group; they all initiate more messages to the recommended candidates than in the control group. Interestingly, I further observe that focal users in the "Potential Matches' Preferences" group and "Mutual Preferences" groups tend to choose the candidates who are more likely to prefer them regardless of these candidates' desirability whereas people in "Your Preferences" group tend to seek highly desirable candidates. It is worth mentioning that in my study, users are not bounded by the limited recommendation choices; the platform provides a target search tool that ensures that every subject has the same opportunity to look for desirable partners. Interestingly, given the equal search access to all users on the platform, the qualitative difference of message receivers across groups only happens among those recommended candidates, not the candidates from search results. It further assures that those focal users in "Potential Matches' Preferences" and "Mutual Preferences" are not switching to lower desirable users in general but are responding to the recommended candidates who are more likely to "prefer" them. My results indicate that users value the other side's preference and act upon it when such information is available.

The positive effect on message initiation by using "the other side's preference" is further amplified in the examination of responses and matches as the message receivers in "Potential Matches' Preferences" and "Mutual Preferences" groups respond more to the message proposals. I find that while "Your Preferences" group receives more responses than the control group, "Potential Matches' Preferences" and "Mutual Preferences" groups even outperform "Your Preferences" group. Therefore, providing the other side's preferences ("Potential Matches' Preferences" and "Mutual Preferences") does lead to more matches than only considering the focal user's preferences ("Your Preferences"). Such an increase may result from the novelty and diversity of choices generated by "Potential Matches' Preferences" and "Mutual Preferences" recommender systems, motivating focal users to explore and finally convert to matches. It may also be likely that users react strategically to the newly added information of the other side's preference; they tend to maximize the replies and matches. Further, there are heterogeneous effects wherein users who search broadly benefit more from "People who might prefer you" and "Mutual Preferences" recommendations whereas "Your Preferences" benefits users who search narrowly.

This study contributes to the literature in several ways. First, to the best of my knowledge, it is among the first to examine the design and impact of *user* recommendation in a two-sided matching market that is fundamentally different from product recommendations. Second, it extends the literature on preference information disclosure and preference signaling to a new setting where the preference on the other side is provided by the platform's recommendation system. Further, I am among the first to design and conduct a randomized field experiment to investigate user recommendations in a two-sided matching market. It allows us to observe users' real-world choices and matching outcomes. These findings provide valuable implications to two-sided matching platforms and highlight the significance of including the other side's preference in the recommender systems. Such inclusion not only helps to improve user engagement and matching outcomes, but also potentially reduces the disproportionate focus on the most popular users due to the diversity of users that are recommended as potential candidates for matching.

The paper is organized as follows. First, prior work is reviewed to outline my contributions. The research context is then described to provide details on the online dating platform as a representative of two-sided matching market. It is followed by discussion of my experimental design as well as the details of the recommender systems I deploy. Variables and results are presented and I conclude with managerial implications.

4.2 Prior Literature

My paper closely relates to three streams of research. The first two streams of work examine recommender systems from different perspectives; one from the business perspective of recommender systems on how they impact users and platforms, and the other from the technical perspective of optimal design of recommender systems. The third stream of research draws upon studies on preference information provision and preference signaling to serve as the theoretical underpinning for how focal users may make choices differently when the recommendations are generated using different preference information.

There is emerging literature in the domains of information systems and economics that examines how recommender systems change users and online markets. Researchers have been focusing on either the quantitative or the qualitative side of the impact. On the one hand, researchers have found a positive effect of recommender systems on sales (Fleder and Hosanagar, 2009; De et al, 2010; Oestreicher-Singer and Sundararajan, 2012). On the other hand, some studies investigate how recommender systems shape consumers' choices – whether the introduction of recommender systems leads to more fragmented or unified choices collectively. However, mixed results are reported in different markets and contexts. For instance, Brynjolfsson et al. (2011) find that recommender systems lead to an increase in sales diversity while Hosanagar et al (2014) find that it leads to an increase in commonality in music choices. Moreover, several studies have shown the co-existence of an increase in diversity and an increase in commonality, albeit on different levels of analysis (Fleder and Hosanagar, 2009; Lee and Hosanagar, 2014). The existing studies along this line focus mostly on product recommendations in transactional markets. Given distinctive market characteristics of two-sided matching markets, my study seeks to be among the first to examine user recommendation in this setting. I complement this line of research by examining both the quantitative and qualitative impacts on user's decision-making in a two-sided matching market.

In contrast to studying the business impact, studies from computer science focus on the performance of recommender systems algorithms. Some recent papers have proposed recommendation algorithms for matching problems (Pizzato, et al, 2010; Xia, et al et al, 2015). As my focus is to investigate when using the same algorithm, how different sources of preferences would impact users' choices and matches differently, I adopt an existing algorithm in Pizzato et al's (2010) to obtain established recommendation performance. From the design aspect, I make additional effort to reduce biases of favoring popular users and confounding factors of inferring preference. From the evaluation aspect, existing studies evaluate new algorithms using secondary data while I design and conduct a randomized online field experiment to observe users' real choices and matches.

Finally, I draw from the emerging literature on preference disclosure and preference signaling in Economics and Information Systems to provide supporting evidence that incorporating the other side's preferences in the recommender systems may be beneficial to the users and the platform. There is empirical evidence that presenting a focal user with information regarding the preferences of another user tends to increase the chance of a match between the two (Avery & Levin, 2010). Such provision of the other side's preferences serves as a weak signal that prompts focal users to proactively connect with potential matches. Some theoretical work also suggests that the focal users would be more likely to accept one's proposal if it comes along with a credible signal of preferences (Cole et al., 2013). These theoretical and empirical evidences tend to support the fact that focal user's decision-making is affected by the awareness of the other side's preference in a matching market. However, in the existing studies, the preference signal sent to the focal users is directly from another user, e.g., a proposal or a profile visit. It is not clear how focal users react when the preference of the other side is based on predictions, and when the preference signal is sent by the platform. My study, therefore, extends this line of literature with a relaxed condition of predicted preference information.

4.3 Methods and Data

4.3.1 Research Context

I collaborate with one of the leading online dating platforms in the U.S, which has more than 1 million registered users. As with most online dating websites, it offers the following features to users. First, users can create their own online profiles to introduce themselves. User profiles typically also include photos. Moreover, the platform offers a decentralized search tool wherein users can filter profiles by age, location and other demographic attributes to find potential matches. These targeted search results can be sorted based on location distance or user tenure from registration. Users are able to browse others' profiles without limitations and at no cost. There is no personalization or recommendation on this platform before my experiment.

4.3.2 Experimental Design

Based on preference information from the two sides on the platform (focal users on one side vs. potential matches on the other side), I compare three recommendation algorithms that are based on 1) focal users' preference ("Your Preferences"), 2) the other side's preferences ("Potential Matches' Preference"), and 3) mutual preferences ("Mutual Preferences") respectively. Since I are interested in how information about the different preferences impact user's decision-making, I use the same features and the same recommendation model but only alter the input information by leveraging preferences from different sides of the market. The control group outputs the top popular users in a random order to create a better baseline group than the commonly used benchmark - "generating a random list of users" as it makes sure the provided options are of high quality. The recommendation system is newly added to the platform and the targeted search function remains in use without any changes. I also make careful design considerations to account for other factors contributing to users' choices. As suggested by the literature on decision-making, the size of the choice set plays a role. Therefore, I fix the number of recommendations for all the four recommendation algorithms. I also limit this number to be a reasonable size (i.e., 100 users) because too many choices may increase the complexity in decision-making due to bounded rationality.

To inform users how the recommendations are generated, a title is provided. "Your Preferences" group shows "People you might prefer" while "Potential Matches' Preferences" group uses "People who might prefer you". "Mutual Preferences" group displays "People who you might prefer and who might prefer you", and control group says "System Recommendation". It is important to disclose this information to make sure that users are aware of whether or not the choices are incorporated with the other side's preference. Otherwise, the strategic behavior documented in the previous literature would not be induced. In addition, the users are also informed regarding the sorting of the recommendations in each group, which is based on how compatible these candidates are with the designated preference. Specifically, the candidates shown at the top in "Your Preference" group have a higher chance of fitting focal users' own preferences than those candidates at the bottom. The recommendations displayed at the top in "Potential Matches' Preferences" are more likely to prefer the focal users than those ranked at the bottom.

Following a between-subject design, I randomly assign users to one of the four groups. I focus on the users who have interacted with others to be able to extract their revealed preference. Once a user is assigned to a group, I always generate recommendations using the assigned recommendation system to assure each subject is exposed to only one treatment. The recommender system refreshes every day so each user will get updated recommendations on a daily basis. New users with no historical data will get their recommendations once they start engaging on the platform.

4.3.3 Design of the Recommender Systems

My interest in this study is not to design new algorithms but to leverage existing algorithms and investigate how different preference information would impact users' choices and matches differently. Researchers in computer science have used two types of models for matching problems; one is profile-based similarity ranking (Pizzato, et al, 2010) and the other is collaborative filtering (Xia, et al et al, 2015)[18]. With careful consideration, I decide to not use the collaborative filtering-based model as researchers have found that it tends to favor superstars (Lee and Hosanagar, 2014), which may potentially exacerbate the congestion among superstars in two-sided matching markets. Following the profile-based similarity ranking approach, I implement the recommendation algorithm based on Pizzato et al (2010)'s model, which looks for "similar" candidates who are compatible with the preference information based on user attributes on profile pages. The algorithm treats users with

[18] Other supervised machine learning techniques may also be applied to recommender systems in two-sided matching markets but have not been studied in previous papers. At least in my context, there is data limitation that only positive cases are available (i.e., who likes whom) but no negative ones (i.e., who dislike whom) that makes supervised models unfeasible. Based on the data description, my conjecture that is also the case in the two cited paper in computer science.

same profile attributes equivalent despite the fact that these users may differ in demand and may be considered differently using collaborative filtering. For details of the algorithm, please refer to the original paper (Pizzato, et al, 2010).

Specifically, the recommender system consists of three parts – input preference data, feature set and model as shown in Figure 2. For the three recommender systems, I use the same feature set and recommendation model to make sure the only difference across "Your Preferences", "Potential Matches' Preferences" and "Mutual Preferences" recommender systems are the input preference data. Specifically, the input preference data in "Your Preferences" recommender system is extracted from those people who are visited or contacted by the focal users. The input preference data in "Potential Matches' Preferences" recommender system is extracted from those people who have initiated visits and messages to the focal users, and that in "Mutual Preferences" recommender system is extracted from the historical partners with back-and-forth visitation and conversations. I then extract features from these profiles to form preferences.

The specific features I extract to represent one's preference are primarily based on the profile information since in general people rely on these profile attributes to make decisions. The features include age difference between the focal user and potential candidates, location proximity, number of photos, income, education, length of self-introduction and immigration status. I also include tenure-length as one feature as users, especially long-time users, are very familiar with all the other old users on the platform and thus they may pay more attention to new users.

I create potential selection pool for each focal user using the active users over the last 2 weeks to make sure the potential candidates have been recently active, to maximize the response and engagement. For each user, I exclude the ones that they have visited over the last 3 months to create a customized selection pool for each user that aims to generate useful recommendations rather than redundant information. I calculate a compatibility score of each potential candidate within the selection pool and I output only the top one hundred compatible candidates for the focal user. The three recommender systems work in the same way and only differ in the source of the input preference data.

Although the focus of this paper is not to develop the best performing recommender systems for matching markets, I still seek to improve the existing Pizzato's. Besides using profile-based ranking to mitigate potential biases, the other improvements are listed as follows. First, while some prior work use stated preferences that are described by users in their profile, I mainly use the revealed preference based on historical behaviors of each user, which better reflect their true preferences. Along this line, I also carefully pre-process the historical information to pick only the initial visitation and messages between each pair as this indicates a strong preference compared to visiting back. Moreover, while existing studies evaluates new algorithms using secondary data, I design and conduct a randomized field experiment to observe the real choices of users. Finally, in order to ensure user engagement and observe how users use recommender systems, I update the recommendations on a daily basis.

4.3.4 Variables

As I focus on the impact on the focal users, I track their subsequent engagement behaviors upon receiving the experiment interventions. To obtain a comprehensive understanding on how different recommender systems may play a role in focal user's decision-making, I collect outcomes along the messaging funnel from message initiation to the other side's response and to the final match. I follow the previous literature in online dating to define matching as a three-round back-and-forth conversation since it indicates initial mutual interest of both sides (Hitsch et al. 2010, Bapna et al. 2016).

I am interested in not only the number of messages initiated by the focal users but also who the focal users send the messages to. The number of messages is a direct measure of user engagement to indicate the performance of the recommender systems while the qualitative aspect of these choices uncovers whether and how the focal users choose the candidates differently. These two dimensions working together provides us a better understanding of how the usage of different preference information in the recommender systems will impact the interaction and matching outcomes on the platform. I use *charm* to measure the overall desirability of each candidate, which is developed by the collaborating platform to track each user's popularity or demand.

I calculate these outcome variables within a certain time window denoted by *outcome_Xtime*. For instance, I focus on the outcomes within one week after treatment, so I calculate *msg_rec_1week*, *response_rec_1week* and *match_rec_1week* to examine the messages initiated by the focal user, the messages responded by the message receivers, and the final matches formed between the focal users and the message receivers. I take log transformation of these message counts. Further, as focal users can also use targeted search as an alternative way to identify potential candidate, I also look at the same series of outcomes initiated from the search. These outcomes in parallel serve as an additional check on how users are impacted by the introduction of recommender systems.

To further study the heterogeneous treatment effects, I link the experiment data to subjects' historical behaviors. I are particularly interested in categorizing the users based on their prior-experiment decision-making in searching. I segment the users based on search diversity – whether the users search broadly or narrowly. I speculate this may relate to their openness to explore "who are interested in me". For each user on the platform, I calculate the standard deviation of the charm scores of those who the user visits to within two weeks prior to the experiment. I adopt a data-driven approach and choose the median of this distribution as the cutoff for "*broadness*". The subjects in the experiment are labeled as low or high in breadth depending on whether the value is below or above the threshold. The detailed coding of each variable is listed in Table 1.

4.4 Results and Discussion

Since the randomization is done at the focal user level, I use post-experiment individual-level data to run OLS regressions across experiment groups. I focus on the subsequent behaviors within one week after the treatment. Since gender difference has been noted in previous literature (Hitsch et al., 2010; Ravi et al., 2016), I block on gender in the randomization and run all the regression analyses for males and females respectively.

I first examine the quantity change across groups to see if the introduction of recommender systems leads to more message initiation from the focal users. Presumably, if the recommender system provides personalized choices that better fit one's needs, it should outperform the baseline algorithm even though I choose a relatively high baseline using top popular users with customization. As shown in Table 2, users in "Your Preferences" group on average initiate more messages than those in the control group, which further assures that the model and features in use work well in practice. More importantly, "Mutual Preference" and "Your Preference" groups also outperform the control group. There is no statistically significant difference in message initiation across the three treatment groups with different recommender systems, indicating that providing recommendations using the other side's preference has an equivalent scale of positive effect on the engagement of focal users.

The result becomes even more interesting when I couple it with the qualitative analysis of the message receivers in each group as shown in Table 3. By comparing the chose candidates with the unchosen candidates for each focal user across groups, I find that the average desirability or charm scores of the message receivers in "Your Preference" group is higher than the charm scores of the unchosen candidates. However, the charm scores of chosen candidates in "Potential Matches' Preference" group and "Mutual Preference" group are not always higher than the unchosen candidates. In other words, without the other side's information provided, focal users tend to pick the more desirable users from the list of recommended candidates. Yet when the matching side's preference is available, they value such information and are willing to choose those less desirable candidates who have a higher chance of preferring them. The increases in reaching out in "Potential Matches' Preferences" group and "Mutual Preferences" group are as significant as the increase in "Your Preference" group despite the fact that the increases are potentially driven by different mechanisms. Users in "Potential Matches' Preferences" and "Mutual Preferences" react on the access to candidates who are more likely to prefer themselves while the users in "Your Preferences" group become more proactive due to a good fit to their own preference. It is possible that users are curious about the novel and diverse choice sets generated by leveraging the other side's preferences, and browsing these profiles may lead to conversions to conversations. It is also possible that users tend to utilize the prediction of the other side's preference and act upon it to maximize the response rate.

Furthermore, users are not bounded by the limited recommendation choices at all as they have a search tool to locate users they prefer. They have the equal access to desirable partners with the same search cost using the generic target search. In other words, each user has full access to everyone on the platform using the search tool plus an additional subset of recommended candidates. The focal users contact candidates with lower charm scores in "Potential Matches' Preferences" is not because users have no access to other more desirable candidates but because they intend to choose these candidates who may be "less popular" but are more likely to be "interested in themselves”. As robustness checks, I further examine how the quantity and quality of message initiation using the search tool are affected at the meanwhile. As shown in Table 4, I find there is no statistically significant difference across treatment groups in both the numbers of initiated messages and the desirability of the message receivers. It means when the provided choices (e.g., from target search) do not contain the other side’s preference, the focal users in “Potential Matches’ Preferences” and “Mutual Preferences” groups make decisions similarly as their counterparts in “Your Preferences” group.

I further examine responses and matches along the messaging funnel. As shown in Table 5, I find that the positive effect of the recommender systems carries on, which leads to an increase in replies in all treatment groups. More importantly, “Potential Matches’ Preferences” and “Mutual Preferences” groups get even more responses than “Your Preferences” group. Similarly, I find the introduction of the recommender systems leads to an increase in final matches in all treatment groups, and “Potential Matches’ Preferences” and “Mutual Preferences” groups benefit from even more matches on average than “Your Preferences” group. Overall, providing the other side’s preferences (“Potential Matches’ Preferences” and “Mutual Preferences”) does lead to more matches than only using the focal user’s preferences (“Your Preferences”). This outcome gap in matching between “Your Preferences” and the other two groups is mainly driven by two aspects. Firstly, conditional on similar numbers of messages sent out, users in “Potential Matches’ Preferences” and “Mutual Preferences” groups are more likely to get a response, which plays an important role in the conversion of final matches. Secondly, the chosen candidates in “Your Preferences” group are more popular than those in “Potential Matches’ Preferences” and “Mutual Preferences” groups, and thus these candidates from “Your Preferences” group tend to have less bandwidth than those in the other two groups to deal with the extra incoming messages due to the introduction of recommender systems.

Finally, I investigate the heterogeneous effects of each recommender system on different user types to gain more insights on what user type would benefit the most from which recommender system. I do find that recommender systems have a differential impact on different users. I am particularly interested in segmenting users based on their search patterns. Specifically, here the user type is based on whether a user searches narrowly or broadly. Interestingly, as shown in Table 7, I find that users who search broadly have a significant increase in message initiation when offered with “Potential Matches’ Preferences” and “Mutual Preferences” whereas “Your Preference” leads to a significant increase in message sending among users who search narrowly. This is consistent with the trend where users who search broadly are more likely to be more open-minded to candidates who are not typically their own “type” while people who are particular about choices and have a narrow search may tend to stick to their own preferences. In addition, given documented gender differences in online dating, it is worth mentioning that female and male users have a consistent pattern in response to each recommender system.

4.5 Conclusion

User recommendation is often deemed as one of the keys to mitigate search friction and matching inefficiency in two-sided matching markets, but much less attention, both in industry and in academia, has been paid compared to product recommendations in transactional markets. With an emphasis on the fundamental characteristics of user recommendation in two-sided matching markets, my study seeks to fill this gap by starting at examining whether and how the provision of potential candidates’ preference can positively impact users’ decision-making and overall matching on the platform.

In collaboration with a leading online dating platform, I carefully design recommender systems with the same algorithm but only alter the preference information in use. I design and conduct a randomized field experiment to investigate how the recommender system using only the focal user's preference (i.e., "Users who you might prefer") plays a different role than the recommender systems using the other side's preference (i.e., "Users who might prefer you" and "Users who you might prefer and who might prefer you"). Very interestingly, I find that focal users are willing to initiate messages to less desirable users than their counterparts when they are aware that these recommended candidates are likely to be interested in them. These focal users end up sending no fewer messages to these candidates "who may be interested in them" compared to their counterparts sending to those "who they may be interested in". Moreover, when it comes to responses and matches, the advantage of incorporating the other side's preference is further consolidated; the focal users in "Your Preference" group get a smaller increase in responses and matches than users in the other two groups. Clearly, users are sensitive and responsive to "the other side's preference" and value candidates who are likely to prefer them regardless of these candidates' desirability. It leads to a higher volume of matches since the message receivers in "Potential Matches' Preferences" and "Mutual Preferences" groups have a higher probability of responding. Further, these recommender systems display differential impact on different users based on the diversity of their historical search. Users who search broadly are more responsive to "People who might prefer themselves" and "Mutual Preferences" while users who search narrowly are more interested in "People who I might prefer".

My work contributes to the literature in several ways. First, I are among the first to acknowledge the fundamental characteristics of user recommender systems in two-sided matching markets and study the design and impact of user recommendations. The study extends the existing literature on the impact of recommender systems. Second, there is emerging literature that studies how the provision of the preference information from the sender will affect the decisions of the receiver, but there is no study examining the implications on recommender systems. My findings therefore complement this line of work and add empirical evidence to a different setting where the preference information is prediction and the preference signal is not directly sent by the sender. Finally, in terms of identification strategies, I are among the first to design and conduct a randomized field experiment to examine the impact of user recommendations in a two-sided matching market.

My findings provide practical insights to the platform designers. The results suggest multiple benefits of incorporating the other side's preference into the recommender systems. Besides the greater volume of user engagement and final matches, more importantly, these recommender systems facilitate the discovery of "seemingly unusual" matches. Without any information on the other side, the focal user can only act on their own preferences and look for their preferred "types". However, with the other side's preference information available upfront, the focal users increase the efficiency of their search but also have access to a broader array of "types" they would not have reached out to otherwise. Driven either by curiosity or efficiency improvements, the focal users get a chance to learn about these candidates by browsing their profiles and talking to them, which in turn leads to more matching opportunities for the focal users. This also mitigates the overloading problem of superstars as more matches are discovered in this manner. Future work can examine how the different recommendation systems impact the user and effectiveness of targeted search mechanisms for different user segments.

Chapter 5: Conclusion

Digital platforms have thrived over the last decade. Data analytics and information technologies offer new opportunities to these platforms to mitigate the fundamental problems of online markets such as information asymmetry and search friction. The revolutionary advancements in mobile technologies and machine learning further provide competitive advantage to user acquisition and user retention. Every platform has to ask the same fundamental question of how to design and provide information to users via different mechanisms, including information incentives, differentiation signals, personalization, etc. Yet, it is still not clear in this emerging context with distinctive characteristics how these information strategies would impact platform users and platforms themselves.

My dissertation therefore seeks to examine the design of effective information provisioning strategies for digital platforms to mitigate some real-world significant challenges they face and to facilitate user decision-making and matching. I collaborate with one transactional platform and one matching platform to examine the design of optimal information-provisioning strategies. I conduct three large-scale randomized field experiments to causally identify the impact of the introduced interventions on customers' engagement behaviors as well as on matching outcomes.

The first essay examines whether platforms can effectively induce mobile app adoption through information provisioning, and compares with another widely adopted strategy - monetary incentive. I find that while both strategies are effective in motivating mobile app adoption, only information provisioning is effective in driving long-term increase in sales. I also identify different patterns of multi-channel usage induced by different motivating strategies. These findings provide platforms with guidelines on the design of optimal motivating strategies to induce effective mobile app adoption that leads to long-term increase in profitability.

The second essay investigates optional verification mechanisms to mitigate information asymmetry especially for non-transactional markets that lack common information mechanisms such as reputation systems and quality assurance. I focus on a different role of verification by making it voluntary and visible to other users and find that even simple phone verification plays a significant role in these non-transactional markets and serves as an effective signaling device. I find, however, that male and female users as the two sides of the matching market use the same signal in very different ways. Such differential patterns are related to the disparity in differentiation ability of each side's existing key attribute, i.e., income for males and beauty for females. I also observe that verified users become more proactive and reach out to more and better potential partners, which further improves desirable matching outcomes and benefits the platform as a whole. My study is among the first to document these differential opt-in decisions and the impacts of verification across two sides of a matching platform and to provide novel insights on optional verification and signaling in two-sided markets.

The third essay targets user recommendation to reduce search friction in decentralized two-sided matching markets. I seek to understand which side's preference should be considered for recommendation purposes as it relates to the fundamental characteristic of matching – a bilateral decision. I design and conduct a randomized field experiment to compare how users make choices differently when the recommendations suggest 1) who they may be interested in, 2) who may be interested in them, and 3) who they may be interested in each other. I implement three recommender systems that only differ in what preference information is leveraged. Notably, I find users act strategically when they are given candidates who are more likely to be interested in themselves, and they are willing to lower their selectivity and proactively reach out to those candidates. Recommender systems based on potential matches' preference or mutual preference lead to better matching outcomes than recommender systems solely relying on focal user's preference, in terms of not only quantity but also offloading superstars and promoting other users. The findings provide useful design suggestions for two-sided matching platforms as the current practice often neglect the preference of the other side when designing recommender systems.

In each of these studies, in addition to conducting randomized field experiments for a clean identification of the causal effects, I also integrate a rich set of user characteristics that further allows me to uncover the underlying mechanisms of the identified phenomena. To summarize, my dissertation provides both empirical contributions and managerial implications to information-provisioning strategies for digital platforms and two-sided matching markets.

Appendix I. Figures

Figure 1.1: the Contribution of My Study to the Literature on How Monetary Incentive and Information may Affect Customers' Purchases

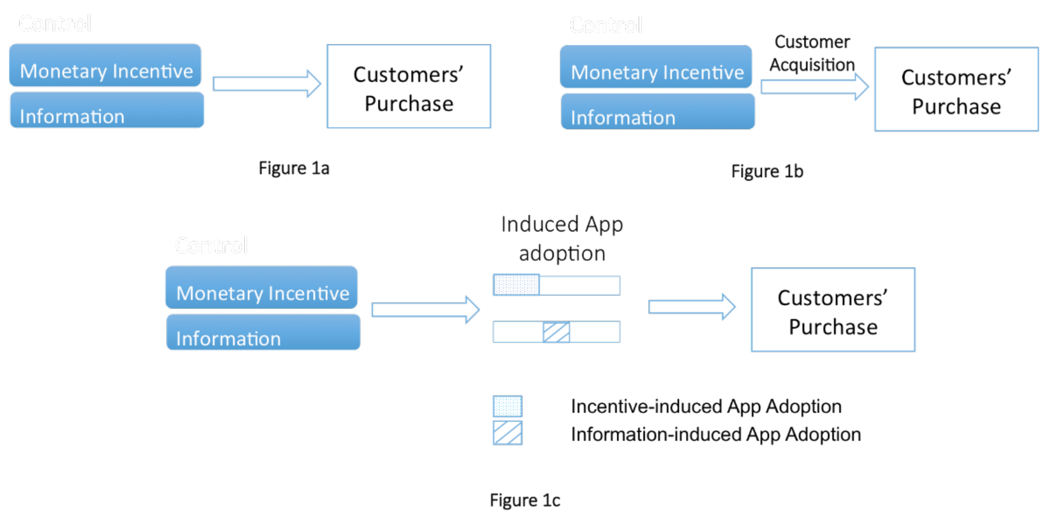
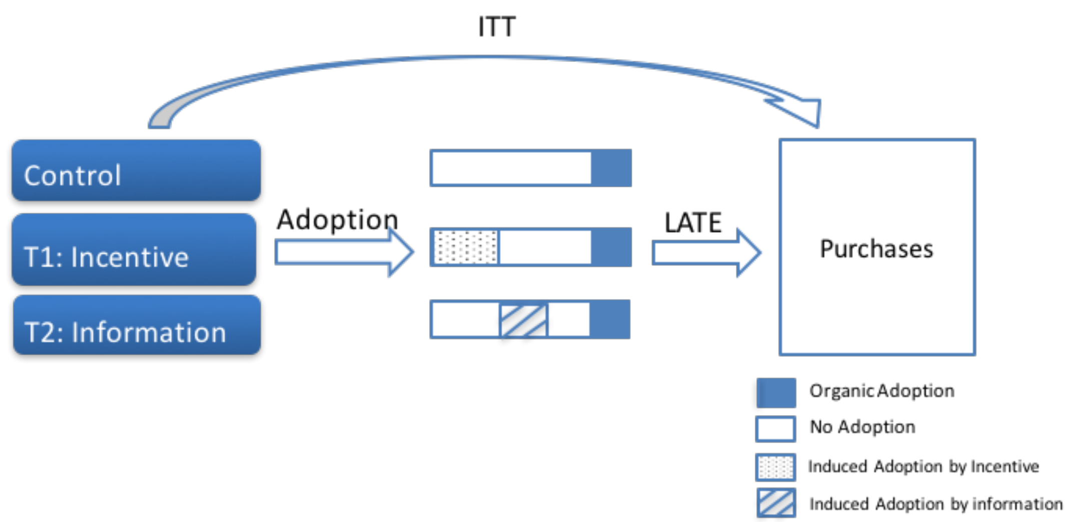


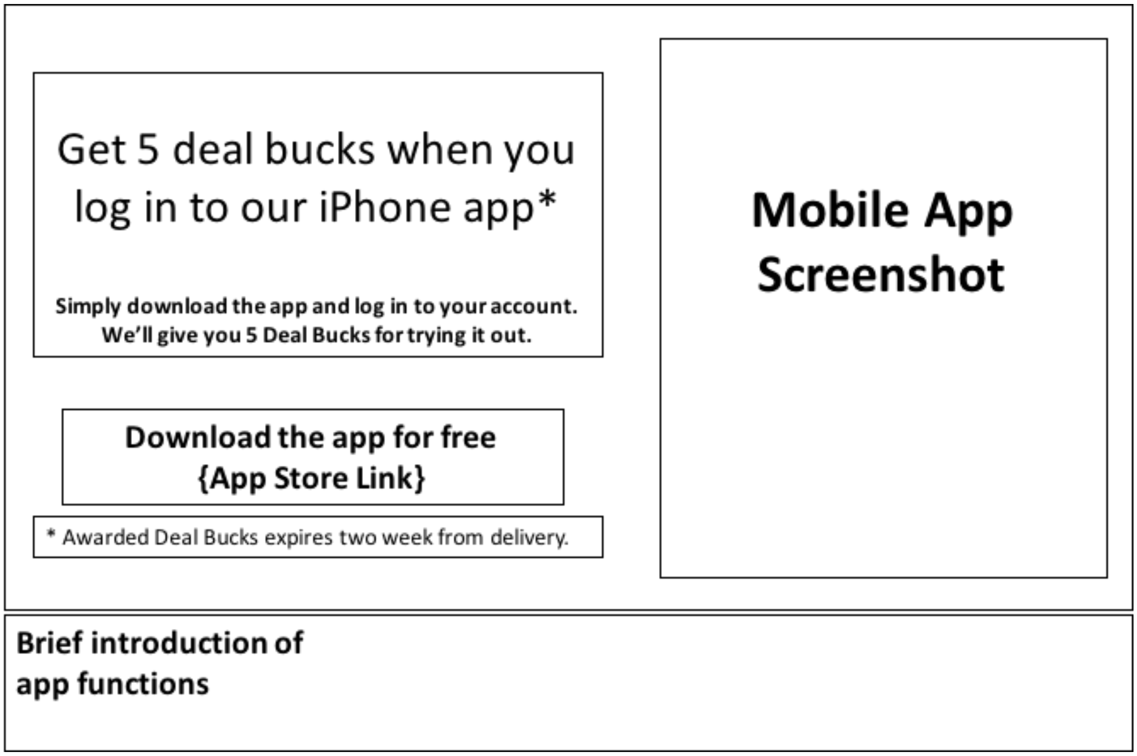
Figure 1.2: Relationship between Effect of Treatment on Adoption, Effect of Induced Adoptions on Purchases (LATE), and Effect of Treatment on Purchases (ITT)

The three effects correspond to the Question Q1a, b and c in the Introduction section.



* App adoption consists of organic adoptions (solid part) and induced adoptions (dotted part in T1 or dashed part in T2). The 'induced adoptions' are influenced by firm's specific interventions. I use the 'local average treatment effect' (LATE) approach to identify the causal effect of such 'induced adoptions' on customers' purchase behaviors (Q2), for both incentive treatment (T1) and information treatment (T2).

Figure 1.3: Relationship between Effect of Treatment on Adoption, Effect of Induced Adoptions on Purchases (LATE), and Effect of Treatment on Purchases (ITT) (T1) Email Template for Treatment 1: **Highlight incentive for app adoption**



(T2) Email Template for Treatment 2: Highlight information about the benefits of the app

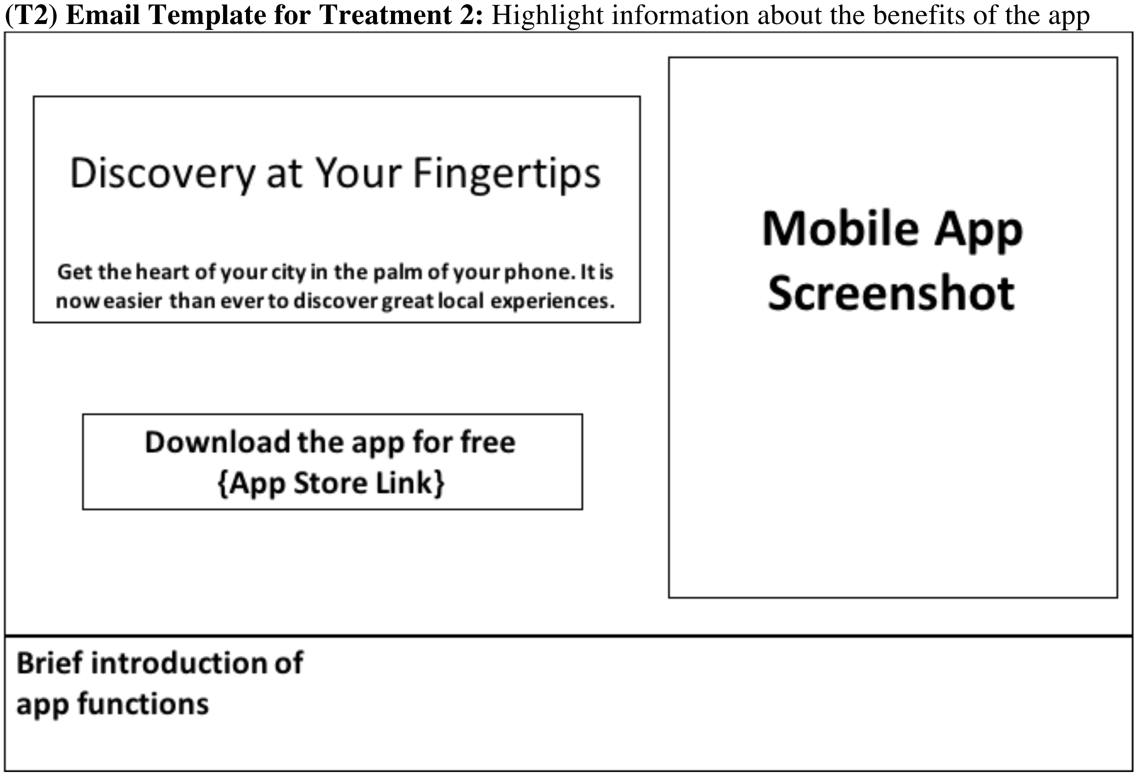


Figure 2.1. Experimental Design

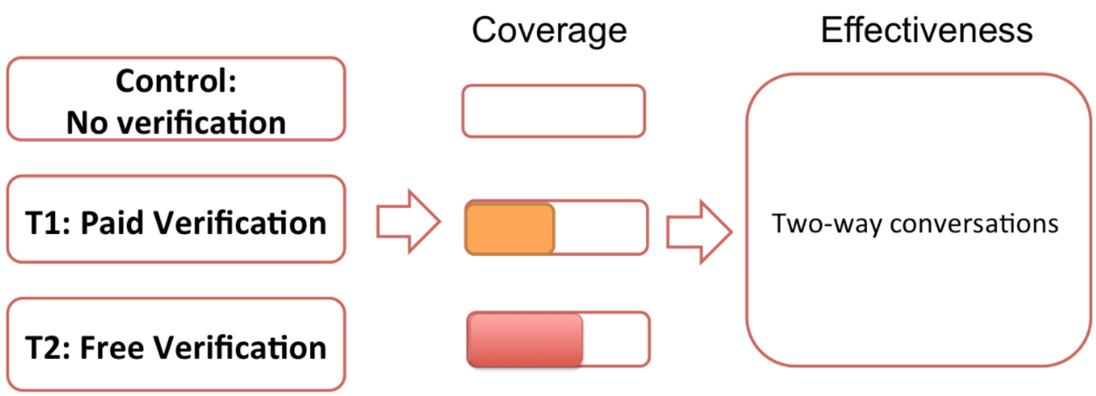


Figure 2.2. Verification Badge of Verified Users

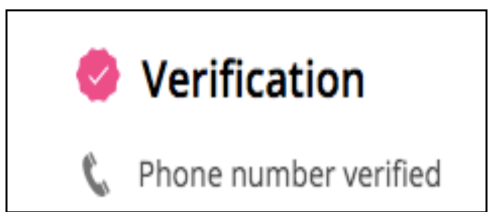


Figure 2.3. Income of Males and Beauty of Females across Types

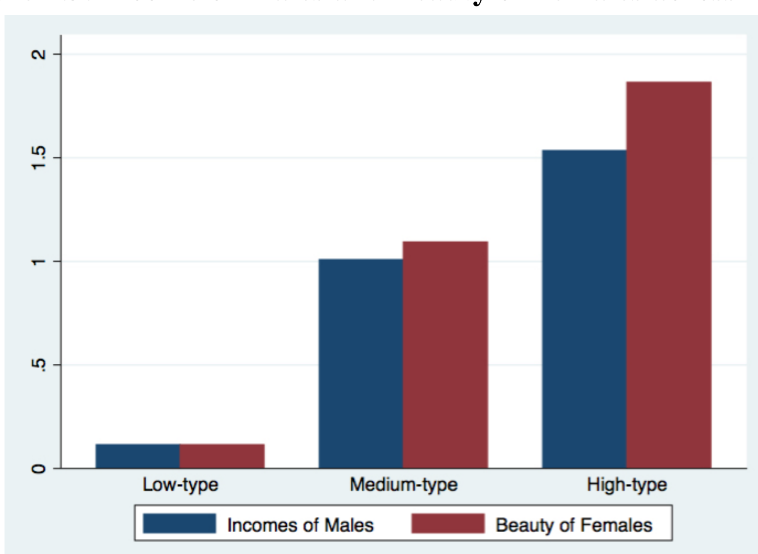


Figure 3.1 Experiment Design

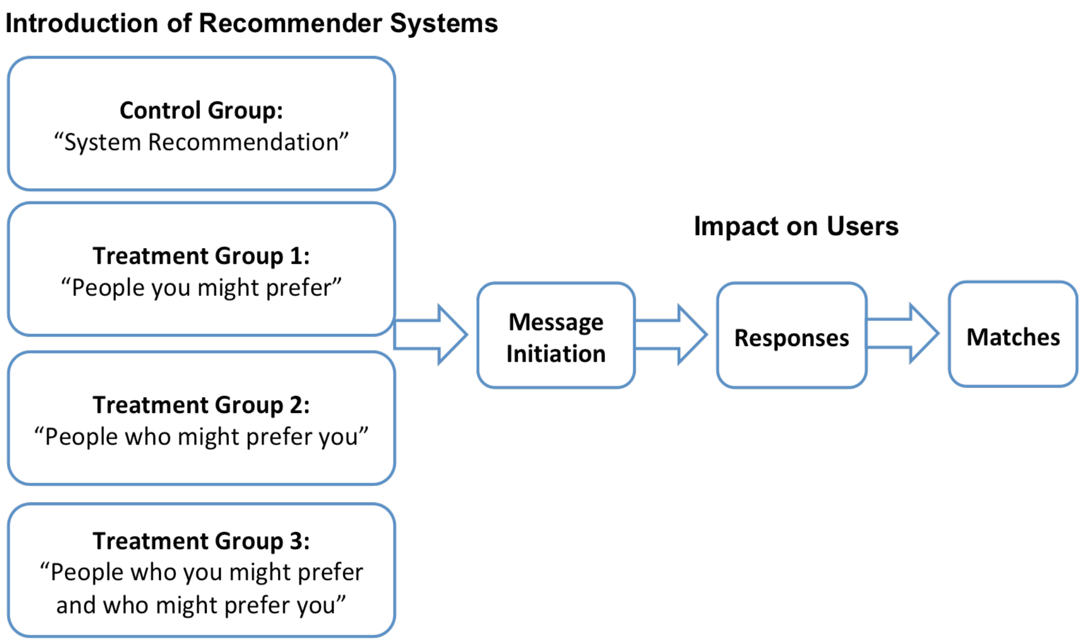
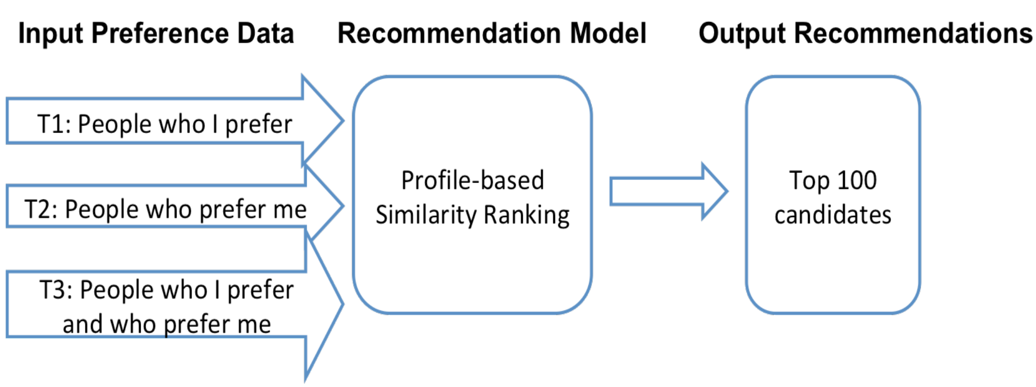


Figure 3.2 Components of Recommender Systems



Appendix II. Tables

Table 1.1: Randomization Check

		Control	**T1**	**T2**	**p-value (C=T1=T2)**
		n = 137,195	n = 48,027	n = 48,070	
Customer tenure (days)	Mean	0	-1.32	0.01	0.407
	Sd	193.4	192.9	193.97	
Total number of purchases	Mean	0	0.01	0.01	0.909
	Sd	3.82	3.69	3.74	
Total Desktop Purchases	Mean	0	0.02	0.01	0.726
	Sd	3.6	3.48	3.52	
Total Mobile Web Purchases	Mean	0	-0.01	0	0.384
	Sd	0.84	0.81	0.83	
Total revenue	Mean	0	-0.02	0.21	0.983
	Sd	237.37	228.19	236.89	
Avg. price of deal purchase	Mean	0	-0.39	-0.53	0.481
	Sd	89.88	82.10	80.83	
Total number of purchased categories	Mean	2.08	2.08	2.09	0.460
	Sd	1.44	1.46	1.45	

*The figures provided are demeaned values obtained by subtracting the mean value of treatment groups from that of control group. Demeaning preserves the difference in mean value between test groups as well as the t-test (i.e. randomization check).

Table 1.2: Effect of Treatments on Mobile App Adoptions (Q1a)

	download_1day	**download_3day**	**download_1week**	**download_2week**
T1	**0.00533*** **	**0.00931*** **	**0.00995*** **	**0.0101*** **
	(0.000347)	(0.000497)	(0.000543)	(0.000626)
T2	**0.00124*** **	**0.00287*** **	**0.00328*** **	**0.00322*** **
	(0.000189)	(0.000340)	(0.000398)	(0.000502)
Constant	0.000357***	0.00200***	0.00325***	0.00666***
	(5.10e-05)	(0.000121)	(0.000154)	(0.000220)
n	233,292	233,292	233,292	233,292
R-squared	0.003	0.003	0.003	0.002

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
* coefficients of T1 and T2 are significantly different

Table 1.3: Effect of Induced App Adoptions on Purchase Behaviors in the Long Run - LATE (Q1b)

	T1		**T2**	
	purch_6month	**purch_1year**	**purch_6month**	**purch_1year**
Induced Adoption	-0.274	-0.500	**5.081*** *	**10.25*** *
	(0.830)	(1.476)	(2.589)	(4.650)
Constant	0.877***	1.685***	0.859***	1.650***
	(0.00613)	(0.0109)	(0.0113)	(0.0201)
Observations	185,222	185,222	185,265	185,265

Robust Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
*I exclude customers' purchases within the first 3 weeks to maintain exclusion restriction and to focus on customers' purchase behavior in the long run. The results are robust when excluding purchases within first 3 months and 6 months (see Table A1).

Table 1.4: Effect of Treatment on Customer Profitability (Q1c), Measured by Average Number of Purchases

	purch_3month	**purch_6month**	**purch_1year**
T1	0.000407	0.000243	-0.00201
	(0.00557)	(0.00912)	(0.0154)
T2	**0.00939***	**0.0196****	**0.0365****
	(0.00567)	(0.00923)	(0.0157)
Constant	0.490***	0.990***	1.798***
	(0.00288)	(0.00472)	(0.00799)
Observations	233,292	233,292	233,292
R-squared	0.00001	0.00002	0.00003

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.5: Differences in observable characteristics between mobile app adopters in Control, T1 and T2

Pre-avg-price	**download_1day**	**download_3day**	**download_1week**	**download_2week**
T1	2.433	1.159	-0.611	-1.485
	(8.874)	(6.724)	(5.179)	(3.382)
T2	**29.93*****	**22.79*****	**15.67****	**9.644****
	(10.49)	(8.092)	(6.193)	(3.957)
p (T1 = T2)	0.0003	0.0025	0.0053	0.0061
Pre-total-purch	**download_1day**	**download_3day**	**download_1week**	**download_2week**
T1	0.598	**0.784****	**0.929*****	**0.840*****
	(0.78)	(0.376)	(0.313)	(0.244)
T2	-0.633	-0.255	-0.336	-0.153
	(0.918)	(0.452)	(0.373)	(0.285)
p (T1 = T2)	0.0584	0.0090	0.0003	0.0007

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.6: Customers' Purchases within 3 Weeks after App Adoption (short-run effect)

	Total	Decomposed by Channel		
	purch_3week	**Desktop_3week**	**Mobile_App_3week**	**Mobile_Web_3week**
T1	0.00297	0.00269	**0.00165*** **	-0.000842
	(0.00214)	(0.00184)	(0.000326)	(0.000845)
T2	0.00288	**0.00313***	**0.000631***	-0.00107
	(0.00214)	(0.00184)	(0.000326)	(0.000845)
Constant	0.114***	0.0871***	0.00243***	0.0198***
	(0.00109)	(0.000935)	(0.000166)	(0.000430)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.7: The Causal Effect of Induced Mobile App Adoptions on Customers' purchases (LATE), after Controlling for Observable Characteristics (Q1b)

LATE	T1		T2	
	purch_6month	**purch_1year**	**purch_6month**	**purch_1year**
Download_1week	-0.343	-0.629	**4.661****	**9.464****
	(0.753)	(1.314)	(2.321)	(4.087)
Pre_desktop	**0.180*****	**0.340*****	**0.181*****	**0.341*****
	(0.00105)	(0.00183)	(0.00111)	(0.00196)
Pre_mobile_web	**0.308*****	**0.584*****	**0.304*****	**0.572*****
	(0.00428)	(0.00747)	(0.00512)	(0.00902)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.8: The Causal Effect of Induced Mobile App Adoptions on Customers' Purchases (LATE), Decomposed by Channel (Q1b)

	Desktop	Mobile App	Mobile Web
T1 Money Induced App Adoption	0.611	0.176	**-0.622***
	(0.693)	(0.146)	(0.33)
T2 Info Induced App Adoption	**5.399****	**0.887****	-0.401
	(2.176)	(0.445)	(0.998)

Standard errors in parentheses,
*** p<0.01, ** p<0.05, * p<0.1
Here I report purchases within 6 months. The results are robust for the 1 year.

Table 1.9: The Causal Effect of Induced Mobile App Adoptions (LATE), on Customers' Purchase Diversity

	unique_categories_6month	unique_categories_1year
Induced Adoption by T1	0.261	0.549
	(0.554)	(0.755)
Induced Adoption by T2	**3.719****	**4.111***
	(1.718)	(2.319)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.10: The Causal Effect of Induced Mobile App Adoptions (LATE), Decomposed by Cities with High/Low Density Deals

	6 Month		1 year	
	Low (1~265)	**High (>265)**	**Low (1~265)**	**High (>265)**
Induced Adoption by T1	**-0.381***	0.148	-0.352	-0.117
	(0.198)	(0.767)	(0.333)	(1.333)
Induced Adoption by T2	**1.089***	**3.872***	**3.401*****	**6.697***
	(0.616)	(2.349)	(1.075)	(4.084)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.11: Effect of Information Intervention on Existing Adopter's Purchase Behaviors

VARIABLES	**purch_1month**	**purch_3month**
Providing Info.	-0.0305	-0.0901
	(0.0324)	(0.0707)
Constant	0.363***	1.058***
	(0.0177)	(0.0377)
R-squared	0.001	0.000

Robust Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.12: The Causal Effect of Induced Mobile App Adoptions on Customers' purchases (LATE) across Channels, Decomposed by Pre-Experiment Channel Usage

	Purchases within 6 Months			
Desktop Only Purchasers (n = 190,069)				
Treatment	**Total**	**Desktop**	**Mobile App**	**Mobile Web**
T1	0.128	1.031	0.118	**-0.724*** **
	(0.989)	(0.876)	(0.159)	(0.273)
T2	**5.884** **	**4.829* **	**0.882* **	0.833
	(2.915)	(2.577)	(0.476)	(0.798)
Mobile Purchasers (exclude desktop-only) (n = 43,134)				
Treatment	**Total**	**Desktop**	**Mobile App**	**Mobile Web**
T1	-1.491	-0.544	0.355	0.061
	(1.637)	(1.053)	(0.314)	(0.875)
T2	1.012	6.818*	0.943	-3.468
	(5.459)	(4.026)	(1.052)	(3.127)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.1. Individual-level Variable Description

Outcome Variables	Description
verify	whether a user verifies or not (dummy variable)
msg_sent	the total messages sent to others
msg_received	the total messages received from others
msg_sent_match	the total matches driven by the focal user contacting others
msg_received_match	the total matches driven by others contacting the focal user
User Characteristics	**Description**
charm	charm score based on popularity
self-description	how complete the user's self-description (0-100%)
education	education level (categorical variable, 0-5). 0 means non-disclosure
income	income level (categorical variable, 0-5). 0 means non-disclosure

Table 2.2. Five-fold Cross Validation Performance

	1	2	3	4	5	Average
mean absolute error	0.24	0.23	0.24	0.25	0.25	0.24
mean squared error	0.10	0.09	0.10	0.11	0.11	0.10

Table 2.3. Summary Stats and Randomization Check

	Control		**T1**		**T2**		**p-val**
covariates	mean	sd	mean	sd	mean	sd	(C=T1=T2)
charm	0	2.92	0.03	2.81	0.04	2.85	0.75
self-description	0	33.70	0.65	33.72	1.19	33.93	0.13
education	0	1.80	-0.01	1.79	0.01	1.79	0.67
income	0	1.48	-0.03	1.43	0.01	1.48	0.12

Table 2.4. Q1a – The Verification Decision among Groups

	Male		**Female**	
Var.	verify_1day	verify_1week	verify_1day	verify_1week
T1	**0.0665*** **	**0.0774*** **	**0.0650*** **	**0.0737*** **
	(0.00526)	(0.00560)	(0.00641)	(0.00677)
T2	**0.0983*** **	**0.112*** **	**0.0904*** **	**0.102*** **
	(0.00521)	(0.00555)	(0.00630)	(0.00665)
Observations	12,725	12,725	8,428	8,428

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.5. Q1a – Who Chooses to Verify

	Verified Male	Verified Female
Medium_type	**0.0960*** **	**0.0939*** **
	(0.00658)	(0.00780)
High-type	**0.162*** **	**0.0599*** **
	(0.0115)	(0.0138)
Constant	0.0247***	0.0279***
	(0.00528)	(0.00626)
p-val (M-type =H-type)	**0.000**	**0.002**

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.6. Q1b –Why Males and Females Have Differential Decisions

(a) How Beauty Correlates with the Non-monotonic Decisions among Females

	M-type females	H-type females
Var.	verify_1week	verify_1week
	0.0334* **	**-0.0500** **
	(0.0116)	(0.0252)
Constant	0.122***	0.117***
	(0.00723)	(0.0180)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
*beauty_above_median is a dummy variable that indicates whether or not a user is above median of beauty within H-type females and within M-type females respectively.

(b) How Income Correlates with the Monotonic Decisions among Males

	M-type males	H-type males
Var.	verify_1week	verify_1week
	0.0623* **	**0.0738** **
	(0.00928)	(0.0293)
Constant	0.105***	0.184***
	(0.00604)	(0.0193)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
*income_above_median is a dummy variable that indicates whether or not a user is above median of income within H-type males and within M-type males respectively.

Table 2.7. Q2a - The Causal Effect of Verification on the Messages Others Initiated to Verified Users and the Quality of the Senders

	Quantity		**Quality**	
	Male	**Female**	**Male**	**Female**
Var.	msg_received	msg_received	sender-quality	sender-quality
T1	**0.470*****	**1.651*****	**2.348*****	**4.942*****
	(0.119)	(0.253)	(0.704)	(1.039)
T2	**0.363*****	**0.940*****	**1.877***	**4.465*****
	(0.0810)	(0.187)	(1.036)	(1.406)
Constant	0.136***	0.376***	3.258***	2.929***
	(0.00688)	(0.0145)	(0.0598)	(0.0808)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.8. Q2a - The Causal Effect of Verification on the Messages Others Initiated to Verified Users and the Quality of the Senders – Decomposed by User Types

	Quantity		**Quality**	
	Male	**Female**	**Male**	**Female**
Var.	msg_received	msg_received	sender-quality	sender-quality
M-type	**0.152*****	**0.429*****	**4.332*****	**3.273*****
	(0.0141)	(0.0289)	(0.0952)	(0.154)
H-type	**0.480*****	**0.686*****	**6.957*****	**4.680*****
	(0.0247)	(0.0463)	(0.167)	(0.246)
L-type*T1	-0.0271	0.341	0.00718	0.0411
	(0.644)	(1.535)	(3.919)	(4.502)
M-type*T1	**0.426*****	**1.906*****	**3.090*****	**5.566*****
	(0.115)	(0.232)	(0.506)	(0.889)
H-type*T1	**0.915*****	0.339	**5.394*****	**0.158**
	(0.190)	(0.701)	(0.884)	(3.147)
L-type*T2	-0.0123	0.148	0.0538	0.153
	(0.578)	(0.847)	(4.364)	(8.158)
M-type*T2	**0.365*****	**1.056*****	**2.164*****	**4.549*****
	(0.0747)	(0.167)	(0.782)	(1.231)
H-type*T2	**0.728*****	0.433	**4.862*****	0.451
	(0.131)	(0.592)	(1.286)	(3.724)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.9. Q2a – Messages Received across user types among Females (in treatment groups)

Var.	Females msg_received
M-type	**0.594***** (0.0183)
H-type	**0.715***** (0.0322)
Constant	0.0756*** (0.0147)
p-val (M-type = H-type)	0.0001

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.10. Q2b - The Causal Effect of Verification on User' Initiated Messages to Others and the Quality of the Receivers

	Quantity		Quality	
	Male	**Female**	**Male**	**Female**
Var.	msg_sent	msg_sent	receiver_quality	receiver_quality
T1	**0.627*****	**0.777*****	**3.154*****	**7.342*****
	(0.141)	(0.160)	(0.986)	(1.361)
T2	**0.401*****	**0.395*****	**1.922*****	**4.564*****
	(0.0958)	(0.118)	(0.671)	(1.006)
Constant	0.0897***	0.108***	3.308***	3.414***
	(0.00814)	(0.00921)	(0.0570)	(0.0783)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.11. Q2b - The Causal Effect of Verification on User' Initiated Messages to Others and the Quality of the Receivers – Decomposed by User Types

	Quantity		Quality	
	Male	**Female**	**Male**	**Female**
Var.	msg_sent	msg_sent	receiver_quality	receiver_quality
M-type	**0.105*****	**0.137*****	**4.619*****	**4.694*****
	(0.0175)	(0.0196)	(0.0895)	(0.125)
H-type	**0.275*****	**0.159*****	**5.532*****	**6.142*****
	(0.0308)	(0.0313)	(0.157)	(0.200)
L-type*T1	0.128	0.251	1.223	2.693
	(0.802)	(1.038)	(4.099)	(6.647)
M-type*T1	**0.509*****	**0.889*****	**3.511*****	**7.294*****
	(0.144)	(0.157)	(0.734)	(1.003)
H-type*T1	**1.403*****	0.113	**7.063*****	0.309
	(0.236)	(0.474)	(1.208)	(3.035)
L-type*T2	0.136	0.116	1.209	1.295
	(0.720)	(0.573)	(3.681)	(3.669)
M-type*T2	**0.369*****	**0.431*****	**2.676*****	**4.476*****
	(0.0930)	(0.113)	(0.475)	(0.724)
H-type*T2	**0.810*****	0.0656	**4.431*****	0.543
	(0.163)	(0.400)	(0.831)	(2.564)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.12 Q3a - The Causal Effect of Verification on Final Matches Initiated by Others and by the Focal Verified User

	Male		**Female**	
Var.	match_ sent	match_ received	match_ sent	match_ received
T1	**0.348*****	**0.311*****	**0.281*****	**0.526*****
	(0.0871)	(0.0601)	(0.0916)	(0.135)
T2	**0.212*****	**0.224*****	**0.122***	**0.319*****
	(0.0592)	(0.0409)	(0.0677)	(0.1000)
Constant	0.0417***	0.0248***	0.0461***	0.0883***
	(0.00503)	(0.00347)	(0.00527)	(0.00778)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.13 Q3a - The Causal Effect of Verification on Final Matches Initiated by Others and by the Focal Verified User - Decomposed by User Types

	Male		Female	
Var.	match_ sent	match_ received	match_ sent	match_ received
M-type	**0.0494*** **	**0.0267*** **	**0.0597*** **	**0.116*** **
	(0.0109)	(0.00751)	(0.0113)	(0.0165)
H-type	**0.138*** **	**0.102*** **	**0.0781*** **	**0.200*** **
	(0.0192)	(0.0132)	(0.0181)	(0.0264)
L-type*T1	0.0204	0.0204	0.131	0.359
	(0.499)	(0.344)	(0.599)	(0.875)
M-type*T1	**0.248*** **	**0.300*** **	**0.319*** **	**0.595*** **
	(0.0894)	(0.0617)	(0.0904)	(0.132)
H-type*T1	**0.939*** **	**0.507*** **	0.0169	0.0677
	(0.147)	(0.101)	(0.274)	(0.400)
L-type*T2	0.0347	0.0173	0.0164	0.0802
	(0.448)	(0.309)	(0.331)	(0.483)
M-type*T2	**0.188*** **	**0.219*** **	**0.136** **	**0.363*** **
	(0.0930)	(0.113)	(90.81)	(88.94)
H-type*T2	**0.810*** **	0.0656	**464.1*** **	86.85
	(0.163)	(0.400)	(158.7)	(314.9)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.1. Individual-level Variable Description

Outcome Variables	**Description**
msg_rec_1week	the total messages initiated to recommended candidates
response_rec_1week	the total responses of recommended candidates
match_rec_1week	the total matches of recommended candidates
msg_search_1week	the total messages initiated to candidates from search
response_search_1week	the total responses of candidates from search
match_search_1week	the total matches of candidates from search
User Characteristics	**Description**
charm	charm score based on popularity

Table 3.2 The Number of Messages That Focal Users Initiate to Recommended Candidates

VARIABLES	**Male** **msg_rec**	**Female** **msg_rec**
Your Preference	**0.0272***	**0.0238****
	(0.0155)	(0.0101)
Potential Matches' Preference	**0.0514*****	**0.0234****
	(0.0155)	(0.0101)
Mutual Preference	**0.0236***	**0.0336*****
	(0.0163)	(0.0106)
Constant	0.0725***	0.0231***
	(0.0109)	(0.00711)
Observations	5,559	5,196
p-value("Your"= "Potential")	0.119	0.968
p-value("Your"= "Mutual")	0.825	0.356
p-value("Potential"= "Mutual")	0.112	0.336

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.3 Desirability Comparison Between Chosen Candidates and Unchosen Candidates across Groups

VARIABLES	Male	Female
Your Preference	-0.0145**	-0.00628
	(0.00614)	(0.00416)
Potential Matches' Preference	0.00941	0.0117***
	(0.00610)	(0.00416)
Mutual Preference	0.00676	0.00313
	(0.00656)	(0.00442)
chosen	0.0220***	0.0201***
	(0.00763)	(0.00576)
Your Preference & chosen	**0.0542*** **	**0.0316*** **
	(0.0114)	(0.00853)
Potential Matches' Preference & chosen	**-0.0492*** **	**-0.0594*** **
	(0.0111)	(0.00840)
Mutual Preference & chosen	**-0.0444*** **	**-0.0375*** **
	(0.0137)	(0.00954)
Constant	8.558***	7.854***
	(0.00431)	(0.00299)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.4 The Number of Messages That Focal Users Initiate to and the Desirability of Receivers Using Search

VARIABLES	**Male msg_search**	**Female msg_search**	**Male charm**	**Female charm**
Your Preference	-0.00719	-0.000929	-0.0107	-0.0255
	(0.0208)	(0.0192)	(0.0176)	(0.0378)
Potential Matches' Preference	0.0167	0.00926	-0.00615	0.0188
	(0.0207)	(0.0192)	(0.0176)	(0.0377)
Mutual Preference	0.0251	0.00736	0.0114	0.0379
	(0.0218)	(0.0201)	(0.0185)	(0.0395)
Constant	0.261***	0.229***	8.026***	7.633***
	(0.0146)	(0.0135)	(0.0124)	(0.0266)
Observations	5,559	5,196	5,559	5,196
p-value ("Your"= "Potential")	0.250	0.597	0.795	0.243
p-value ("Your"= "Mutual")	0.140	0.681	0.234	0.110
p-value ("Potential"= "Mutual")	0.700	0.925	0.344	0.629

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.5 The Number of Responses and Matches That Focal Users Received from Recommendation

	Responses		Matches	
VARIABLES	**Male**	**Female**	**Male**	**Female**
Your Preference	**0.0120***	**0.00933****	0.00819	0.00566
	(0.00663)	(0.00396)	(0.00545)	(0.00349)
Potential Matches' Preference	**0.0365*****	**0.0219*****	**0.0192*****	**0.0136*****
	(0.00661)	(0.00395)	(0.00543)	(0.00349)
Mutual Preference	**0.0259*****	**0.0188*****	**0.0199*****	**0.0134*****
	(0.00696)	(0.00414)	(0.00572)	(0.00365)
msg_rec_1week	0.0644***	0.118***	0.0575***	0.0869***
	(0.000861)	(0.00119)	(0.000708)	(0.00105)
Constant	0.0138***	-0.00255	0.00631	-0.00302
	(0.00467)	(0.00278)	(0.00384)	(0.00246)
Observations	5,559	5,196	5,559	5,196
p-value("Your"= "Potential")	<0.001	0.001	0.043	0.024
p-value("Your"= "Mutual")	0.046	0.023	0.042	0.036
p-value("Potential"= "Mutual")	0.131	0.446	0.906	0.956

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.6 Messages That Focal Users Initiate to Candidates -Segmented by Search Broadness

	Male		Female	
	Narrow	**Broad**	**Narrow**	**Broad**
VARIABLES	msg_rec	msg_rec	msg_rec	msg_rec
Your Preference	**0.0468****	-0.0139	**0.0546****	0.0230
	(0.0224)	(0.0414)	(0.0230)	(0.0154)
Potential Matches' Preference	0.0319	**0.251*****	0.0317	**0.0291***
	(0.0224)	(0.0418)	(0.0225)	(0.0154)
Mutual Preference	0.0193	0.0330	**0.0402***	**0.0432*****
	(0.0230)	(0.0436)	(0.0237)	(0.0160)
Constant	0.103***	0.0535*	0.0199	0.0326***
	(0.0158)	(0.0291)	(0.0159)	(0.0108)
Observations	3,560	761	1,173	2,923
R-squared	0.001	0.064	0.005	0.003

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Bibliography

Adomavicius, G., Curley, S. P., Gupta, A., & Sanyal, P. (2013). Impact of information feedback in continuous combinatorial auctions: An experimental study of economic performance. MIS Quarterly, 37(1), 55-76.
Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. Journal of the American Statistical Association, 91(434), 444-455.
Angrist, J. D., & Pischke, J. S. (2008). Mostly harmless econometrics: An empiricist's companion. Princeton university press.
Agrawal, A. K., Lacetera, N., & Lyons, E. (2013). *Does information help or hinder job applicants from less developed countries in online markets?* (No. w18720). National Bureau of Economic Research.
Antony, S., Lin, Z., & Xu, B. (2006). Determinants of escrow service adoption in consumer-to-consumer online auction market: an experimental study. *Decision Support Systems*, *42*(3), 1889-1900.
Ashraf, N., Berry, J., & Shapiro, J. M. (2010). Can higher prices stimulate product use? Evidence from a field experiment in Zambia. *American Economic Review*, *100*(5), 2383-2413.
Athey, S., & Imbens, G. W. (2017). The econometrics of randomized experiments. Handbook of Economic Field Experiments, 1, 73-140.
Ba, S., Whinston, A. B., & Zhang, H. (2003). Building trust in online auction markets through an economic incentive mechanism. *Decision Support Systems*, *35*(3), 273-286.
Bang, Y., Han, K., Animesh, A., & Hwang, M. (2013). From Online to Mobile: Linking Consumers' Online Purchase Behaviors with Mobile Commerce Adoption. The 2013 Pacific Asia Conference on Information Systems (PACIS).
Bederson, B. B., Jin, G. Z., Leslie, P., Quinn, A. J., & Zou, B. (2018). Incomplete disclosure: Evidence of signaling and countersignaling. *American Economic Journal: Microeconomics*, *10*(1), 41-66.
Berger, J., & Rand, L. (2008). Shifting signals to help health: Using identity signaling to reduce risky health behaviors. *Journal of Consumer Research*, *35*(3), 509-518.
Brandimarte, L., & Acquisti, A. (2012). The economics of privacy. *The Oxford handbook of the digital economy*, *20*.
Brynjolfsson, E., Hu, Y., & Rahman, M. S. (2009). Battle of the retail channels: How product selection and geography drive cross-channel competition. *Management Science*, *55*(11), 1755-1765.
Brynjolfsson, E., Hu, Y., & Simester, D. (2011). Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science*, *57*(8), 1373-1386.
Charness, G., & Gneezy, U. (2009). Incentives to exercise. *Econometrica*, *77*(3), 909-931.
Chen, Y., Li, X., & Sun, M. (2017). Competitive mobile geo targeting. *Marketing Science*, *36*(5), 666-682.
Chen, P. Y., Hong, Y., & Liu, Y. (2017). The Value of Multi-Dimensional Rating Systems: Evidence from a Natural Experiment and Randomized Experiments. Management Science, forthcoming
Cheong, I., & Kim, J. Y. (2004). Costly information disclosure in oligopoly. *The Journal of Industrial Economics*, *52*(1), 121-132.
Coles, P., Cawley, J., Levine, P. B., Niederle, M., Roth, A. E., & Siegfried, J. J. (2010). The job market for new economists: A market design perspective. *Journal of Economic Perspectives*, *24*(4), 187-206.
Coles, P., Kushnir, A., & Niederle, M. (2013). Preference signaling in matching markets. *American Economic Journal: Microeconomics*, *5*(2), 99-134.
Cohen, J., & Dupas, P. (2010). Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment. The Quarterly Journal of Economics, 1-45.
Cui, R., Zhang, D. J., & Bassamboo, A. (2016). Learning from Inventory Availability Information: Field Evidence from Amazon. Working Paper
Datta, H., Foubert, B., & Van Heerde, H. J. (2015). The challenge of retaining customers acquired with free trials. Journal of Marketing Research, 52(2), 217-234.

DigiDay (2016) 3 ways retailers are improving their mobile apps. Accessed at http://digiday.com/marketing/3-ways-retailers-improving-shop-mobile-apps/
De Haan E., Kannan PK., Verhoef P., & Wiesel T., (2015) The Role of Mobile Devices in the Online Customer Journey, Marketing Science Institute Working Paper Series 2015 Report No.15-124
Dedman, C. (2011) Groupon Mobile Rewards Offer: Purchase from Mobile App Get $10 Credit, Accessed at http://blog.al.com/bargain-mom/2011/05/groupon_mobile_rewards_offer_p.html
De, P., Hu, Y., & Rahman, M. S. (2010). Technology usage and online sales: An empirical study. *Management Science*, *56*(11), 1930-1945.
Dupas, P. (2014). Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. Econometrica, 82(1), 197-228
Duflo, E., Glennerster, R., & Kremer, M. (2008). Using randomization in development economics research: A toolkit. Handbook of Development Economics, 4, 3895-3962.
Einav, L., Levin, J., Popov, I., & Sundaresan, N. (2014). Growth, adoption, and use of mobile e-commerce. American economic review, 104(5), 489-494.
Eric T. Anderson and Duncan I. Simester (2014) Reviews Without a Purchase: Low Ratings, Loyal Customers, and Deception. Journal of Marketing Research: June 2014, Vol. 51, No. 3, pp. 249-269.
Fang, Z., Gu, B., Luo, X., & Xu, Y. (2015). Contemporaneous and delayed sales impact of location-based mobile promotions. *Information Systems Research*, *26*(3), 552-564.
Feltovich, N., Harbaugh, R., & To, T. (2002). Too cool for school? Signalling and countersignalling. RAND Journal of Economics, 630-649.
Financial Times (2016), China tech: Renminbi to burn, Accessed at: https://www.ft.com/content/eac1ff4a-fc9d-11e5-b5f5-070dca6d0a0d
Fleder, D., & Hosanagar, K. (2009). Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management science*, *55*(5), 697-712.
Forbes (2017), Holiday Ecommerce To Hit Record $107B in 2017; Mobile Will Lead In Visits, Accessed at https://www.forbes.com/sites/johnkoetsier/2017/11/02/holiday-ecommerce-to-break-100b-for-first-time-says-1-trillion-visit-adobe-study/#3ab945393774
Forman, C., Ghose, A., & Goldfarb, A. (2009). Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management science*, *55*(1), 47-57.
Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information systems research*, *19*(3), 291-313.
Ge, R., Gu, B., & Feng, J. (2017, January). Borrower’s self-disclosure of social media information in P2P lending. In *Proceedings of the 50th Hawaii International Conference on System Sciences*.
Goes, P., & Lin, M. (2012). Does Information Really'Unravel'? Understanding Factors That Motivate Sellers to Seek Third-Party Certifications in an Online Labor Market. *Understanding Factors That Motivate Sellers to Seek Third-Party Certifications in an Online Labor Market (November 1, 2012) .NET Institute Working Paper*, (12-02).
Ghose, A., Goldfarb, A., & Han, S. P. (2012). How is the mobile Internet different? Search costs and local activities. Information Systems Research, 24(3), 613-631.
Ghose, A., & Han, S. P. (2014). Estimating demand for mobile applications in the new economy. Management Science, 60(6), 1470-1488.
Ghose, A. (2017). Tap: Unlocking the Mobile Economy. MIT Press.
Guiteras, R., Levinsohn, J., & Mobarak, A. M. (2015). Encouraging sanitation investment in the developing world: a cluster-randomized trial. Science, 348(6237), 903-906.
Hann, I. H., Hui, K. L., Lee, T. S., & Png, I. P. (2003). The value of online information privacy: An empirical investigation. *Industrial Organization, Econ-WPA*.
Han, S. P., Park, S., & Oh, W. (2016). Mobile app analytics: A multiple discrete-continuous choice framework. MIS Quarterly, 40(4), 2016, 983-1008
Hann, I. H., Koh, B., & Niculescu, M. F. (2016). The double-edged sword of backward compatibility: The adoption of multigenerational platforms in the presence of intergenerational services. Information Systems Research, 27(1), 112-130.

Hitsch, G. J., Hortaçsu, A., & Ariely, D. (2010). Matching and sorting in online dating. *American Economic Review*, *100*(1), 130-63.
Hosanagar, K., Fleder, D., Lee, D., & Buja, A. (2013). Will the global village fracture into tribes? Recommender systems and their effects on consumer fragmentation. *Management Science*, *60*(4), 805-823.
Imbens, G., & Angrist, J. (1994), Identification and Estimation of Local Average Treatment Effects, Econometrica, Vol. 61, No. 2, 467-476
Dranove, D., & Jin, G. Z. (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature*, *48*(4), 935-63.
Jovanovic, B. (1982), "Truthful Disclosure of Information," Bell Journal of Economics, 13:36-44
Jung, J., Umyarov, A., Bapna, R., & Ramaprasad, J. (2014). Mobile as a channel: Evidence from online dating.
Kato-Lin, Y.-C., Padman, R., Downs, J., & Abhishek, V. (2015). Evaluating Consumer m-Health Services for Promoting Healthy Eating: A Randomized Field Experiment. AMIA Annual Symposium Proceedings, 2015, 1947–1956.
Lee, D., & Hosanagar, K. (2014). Impact of recommender systems on sales volume and diversity.
Li, Y., Xie, Y., & Zheng, E. (2015). Competitive Analytics of Multi-channel Advertising and Consumer Inertia. The 2015 Americas Conference on Information Systems (AMCIS), Puerto Rico
Liu, J., Abhishek, V., & Li, B. (2016). The Impact of Mobile Adoption on Customer Omni-Channel Banking Behavior. The 2016 International Conference on Information Systems (ICIS), Dublin, Ireland
Luca, M., & Smith, J. (2015). Strategic disclosure: The case of business school rankings. Journal of Economic Behavior & Organization, 112, 17-25.
Luo, X., Andrews, M., Fang, Z., & Phang, C. W. (2013). Mobile targeting. Management Science, 60(7), 1738-1756.
Ma, Z., Pant, G., & Sheng, O. R. L. (2007). Interest-based personalized search. ACM Transactions on Information Systems (TOIS), 25(1), 5.
Ma, M., & Agarwal, R. (2007). Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities. *Information systems research*, *18*(1), 42-67.
Mayzlin, D., Dover, Y., & Chevalier, J. (2014). Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, *104*(8), 2421-55.
Miller (2014), "Manage Your Incentivized Mobile Budget for Maximum Volume" Accessed at: http://secretsaucehq.com/blog/mobile-marketing/manage-your-incentivized-mobile-budget-for-maximum-volume/
Milgrom, P. (1981), "Good News and Bad News: Representation Theorems and Applications," The Bell Journal of Economics, 12: 380-391.
Moreno, A., & Terwiesch, C. (2014). Doing business with strangers: Reputation in online service marketplaces. *Information Systems Research*, *25*(4), 865-886.
Narang, U., & Shankar, V. (2016). The Effects of Mobile Apps on Shopper Purchases and Product Returns
Neslin, S. A., & Shankar, V. (2009). Key issues in multichannel customer management: current knowledge and future directions. Journal of Interactive Marketing, 23(1), 70-81.
Pizzato, L., Rej, T., Chung, T., Koprinska, I., & Kay, J. (2010, September). RECON: a reciprocal recommender for online dating. In *Proceedings of the fourth ACM conference on Recommender systems* (pp. 207-214). ACM.
Oestreicher-Singer, G., & Sundararajan, A. (2010). Recommendation networks and the long tail of electronic commerce. *Available at SSRN 1324064*.
Qiu, L., & Kumar, S. (2017). Understanding Voluntary Knowledge Provision and Content Contribution Through a Social-Media-Based Prediction Market: A Field Experiment. Information Systems Research. Forthcoming
Bapna, R., Ramaprasad, J., Shmueli, G., & Umyarov, A. (2016). One-way mirrors in online dating: A randomized field experiment. *Management Science*, *62*(11), 3100-3122.
RetailMeNot (2016) "The Rise of Mobile Marketing Spend in Retail: 5 Trends to Watch in 2016"

Retana, G. F., Forman, C., & Wu, D. J. (2016). Proactive customer education, customer retention, and demand for technology support: Evidence from a field experiment. Manufacturing & Service Operations Management, 18(1), 34-50.
Retana, G. F., Forman, C., Narasimhan, S., Niculescu, M. F., & Wu, D. J. (2017). Technology support and post-adoption IT service use: Evidence from the cloud, Working Paper
Roth, A. E. (2015). *Who Gets What—and Why: The New Economics of Matchmaking and Market Design*. Houghton Mifflin Harcourt.
Shankar, V., Kleijnen, M., Ramanathan, S., Rizley, R., Holland, S., & Morrissey, S. (2016). Mobile shopper marketing: Key issues, current insights, and future research avenues. Journal of Interactive Marketing, 34, 37-48.
Siwicki B (2014) Mobile commerce will be nearly half of e-commerce by 2018. Internet Retailer, Accessed at www.internetretailer.com/2014/03/10/mobile-commerce-will-benearly-half-e-commerce-2018.
Son, Y., Oh, W., Han, S. P., & Park, S. (2016). The Adoption and Use of Mobile Application-Based Reward Systems: Implications for Offline Purchase and Mobile Commerce. The 2016 International Conference on Information Systems (ICIS), Dublin, Ireland
Stiglitz, J. E. (1975). The theory of" screening," education, and the distribution of income. The American economic review, 65(3), 283-300.
Stanton, Christopher, Catherine Thomas. (2012). Landing the first job: The value of intermediaries in online hiring. Available at SSRN 1862109
Spence, M. (2002). Signaling in retrospect and the informational structure of markets. *American Economic Review*, *92*(3), 434-459.
Taubman, S. L., Allen, H. L., Wright, B. J., Baicker, K., & Finkelstein, A. N. (2014). Medicaid increases emergency-department use: evidence from Oregon's Health Insurance Experiment. *Science*, *343*(6168), 263-268.
Techcrunch (2015a) What's Better? Incentivized or Non-Incentivized App-Install Campaigns. Accessed at: techcrunch.com/2015/09/10/whats-better-incentivized-or-non-incentivized-app-install-campaigns/
Techcrunch (2015b) Amazon rolls out a referral program to encourage more customers to shop on mobile, Accessed at https://techcrunch.com/2015/12/17/amazon-rolls-out-a-referral-program-to-encourage-more-customers-to-shop-on-mobile/
The Dating Services Industry in 2016 and Beyond. The http://blog.marketresearch.com/dating-services-industry-in-2016-and-beyond
The rise of the online dating industry/ http://wire.kapitall.com/investment-idea/of-love-and-money-the-rise-of-the-online-dating-industry/
Wang, Q., Li, B., Wang, P., & Yang, J. (2016). Using TB-Sized Data to Understand Multi-Device Advertising. The 2016 International Conference on Information Systems (ICIS), Dublin, Ireland
Xu, J., Forman, C., Kim, J.B. and Van Ittersum, K., (2014). News media channels: complements or substitutes? Evidence from mobile phone usage. Journal of Marketing, 78(4), pp.97-112.
Xu, K., Chan, J., Ghose, A. and Han, S., (2016). Battle of the Channels: The Impact of Tablets on Digital Commerce. Management Science, 2016, 63(5):1469-1492
Yu, Mo, et al. "Recommendation in reciprocal and bipartite social networks–a case study of online dating." International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction. Springer, Berlin, Heidelberg, 2013.
Yu, M., Zhao, K., Yen, J., & Kreager, D. (2013, April). Recommendation in reciprocal and bipartite social networks–a case study of online dating. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction* (pp. 231-239). Springer, Berlin, Heidelberg.
Zhang, X., & Zhu, F. (2011). Group size and incentives to contribute: A natural experiment at Chinese Wikipedia. The American economic review, 101(4), 1601-1615.
Zheng, J., Qi, Z., Dou, Y., & Tan, Y. (2016). How Mega Is the Mega? Measuring the Spillover Effects of WeChat by Machine Learning and Econometrics. Working Paper