Essays on Social Media Fundraising and E-Commerce

Xue Tan

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Yong Tan, Chair

Ming Fan

Foad Iravani

Program Authorized to Offer Degree:

Business Administration

University of Washington

Abstract

Essays on Social Media Fundraising and E-Commerce

Xue Tan

Chair of the Supervisory Committee:
Professor Yong Tan
Information Systems and Operations Management at Foster School of Business

This dissertation has two components: social media fundraising and e-commerce. The first component of social media fundraising discusses social media users' charitable content generation in essay 1 and charitable giving in essay 2. In essay 1, we examine how reciprocity of followees affects social influence on users' charitable content generation. We find that within a specific charitable topic, reciprocal relationship has a positive impact on social influence, and non-reciprocal relationship has a negative impact on social influence that drives content generation following the topic. In essay 2, we investigate how individual donation decisions are influenced by reputation incentive design, peer effects, and popularity effects. We find that despite the platform designer's desire to improve fundraising performance, higher visibility of donors' contributions may have negative impact on fundraising. Peer effects are found to be

positive and, hence, provide a potential solution to the free-rider problem. It is also observed that while most users crowd to popular projects, a group of users who exhibit leadership features crowd out from popular projects. E-commerce is the second component, and it also includes two essays. In essay 3, we use a game-theoretical model to understand how membership free shipping (MFS) differs from contingent free shipping (CFS). We find that in a duopoly competition, sellers' choice between CFS and MFS resembles a prisoner's dilemma. While they will both earn a higher profit when they both choose MFS, one seller may have an incentive to deviate from cooperation. In essay 4, we examined the role of live chatting tools in online marketplaces. We find that live chat can work as a substitution for reputation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Yong Tan, for his long-standing support and guidance. He is the best role model and mentor. His enthusiasm for research has greatly influenced me, and his positive attitude toward work and life has enlightened me. Working with him during my doctoral study is a life-changing experience.

My thanks also go to the members of my dissertation committee. Their feedback and comments provide me with great directions to improve this dissertation.

I would like to extend my gratitude to all fellow doctoral students at Foster School of Business. Their participation at my mock job talk and feedback in my other talks have greatly improved my presentation of this dissertation.

I am also grateful to my husband Songpei, my son Mingchen, my parents, and parents-in-law for their support and patience. Family has always been the place for me to relax, and the reason for me to work hard and efficiently.

Finally, I would like to thank my grandmother, Zhengshu, for her support and love in my childhood. The many stories she told me have inspired me to develop my research interest in fundraising.

Chapter 1. INTRODUCTION

With the growing popularity of social media, social networking sites have become an important channel for online donor engagement and charitable fundraising. The first component of my dissertation, social media fundraising, discusses social media users' charitable content generation in essay 1 and charitable giving in essay 2. In social broadcasting sites, users' charitable content generation is affected by both their followees (to whose content they subscribe) and followers (people who subscribe to these users' broadcasts). In the first essay of charitable content generation, we examine how reciprocity of followees affects social influence on users' charitable content generation. We employ a latent instrumental variable approach which effectively controls for the endogeneity issue arising from followees' charitable engagement. We find that within a specific charitable topic, reciprocal relationship has a positive impact on social influence, and non-reciprocal relationship has a negative impact on social influence that drives content generation following the topic. Our findings suggest an effect of information overload caused by excessive messages generated by non-reciprocal nodes such as news media and nonprofit organizations. We also reveal that users usually generate charitable content out of their altruistic motivations. However, those users who include pictures in their previous charitable content are more likely to be driven by reputation-based motivations in generating charitable content.

In the second essay of social media charitable giving, we investigate how individual donation decisions are influenced by reputation incentive design, peer effects, and popularity effects. We find that despite the platform designer's desire to improve fundraising performance, higher visibility of donors' contributions may have negative impact on fundraising. Peer effects are found to be positive and, hence, provide a potential solution to the free-rider problem. It is also observed that while most users crowd to popular projects, a group of users who exhibit leadership features crowd out from popular projects. Investing more fundraising effort on this crowding-out group may alleviate the rich-get-richer problem.

E-commerce is the second component of my dissertation, and it also includes two essays. In the third essay, we use a game-theoretical model to understand how membership free shipping (MFS) differs from contigent free shipping (CFS). Free shipping policy is widely used by e-tailers to improve customers' online shopping experience. In MFS, consumers pay a membership fee upfront and enjoy unlimited, free expedited shipping for a certain period, whereas in CFS, consumers are eligible for free standard shipping if the total transaction amount of an order exceeds a pre-determined threshold. We develop an economic model in which two competing e-tailers choose their shipping strategies in a sequential order to maximize their profits. We find that when both e-tailers use the same shipping policy, MFS leads to a higher profit than CFS if faster shipping is employed. While e-tailers have an incentive to adopt MFS, under some conditions they find themselves in the prisoner's dilemma — both e-tailers end up choosing the less profitable CFS. Nevertheless, such a prisoner's dilemma no longer exists when consumers' valuation of expedited shipping offered by MFS is sufficiently higher than that of standard shipping attached to CFS.

In the fourth essay, we examined the role of live chatting tools in online marketplaces. Live chatting tools have emerged as a new channel for sellers and buyers to communicate with each other. Live chats facilitate human-based interactions to reduce information asymmetry and build trust between sellers and buyers, exhibiting great potential to complement the existing feedback system used in online marketplaces. This essay analyzes the context of Alibaba where live chatting tools are offered by the platform to every seller. We use clickstream data of 3000 customers' page viewing, live chatting, and purchasing activities to investigate the role played by live chats in online marketplaces. We find that 1) live chats positively impact repeat visits for sellers with low reputation; 2) live chats improve conversion rates; 3) live chats improve conversion more effectively for sellers with low reputation; and 4) people are more likely to initiate live chats when they face sellers with low reputation or products with few reviews. These findings indicate that live chats can work as substitutes for reputation.

Chapter 2. SOCIAL MEDIA FUNDRAISING – CHARITABLE GIVING

2.1 INTRODUCTION

Philanthropic crowdfunding has emerged as a new channel for nonprofit organizations and individuals to raise funds. For nonprofit organizations, crowd-based fundraising differs from traditional fundraising because it collects small contributions from a larger number of donors. For individuals, the unprecedented opportunity allows victims or their agents to describe their difficulties to a larger audience and to communicate with donors to make them more aware of the impact they make. One successful platform, GoFundMe.com, has raised over $2 billion in 2015. In countries where government-owned nonprofit organizations face a crisis of trust, individuals are even more in favor of philanthropic crowdfunding because of its transparency (Tan 2015). A Chinese crowdfunding marketplace, Weibo Philanthropy, embedded in the microblogging platform Sina Weibo, raised over $36 million from over 20 million individuals in the three years since its inception.

Given the large number of potential donors, social networks serve as an effective instrument in helping crowdfunding projects succeed (Lu et al. 2014). Crowd-based fundraisers encourage donors to share their donations on social networking sites. By doing so, users not only take the role of a donor, but also take a role as a solicitor. As a matter of fact, GoFundMe suggests in their website that one share of a charitable cause on Facebook will bring in an average donation of $30. Formally, we refer to promotions of charities over social networking sites as social media fundraising. The great power of social media fundraising is manifested by some keynote campaigns. The Ice Bucket Challenge requires participants to tag five friends, who are then challenged either to upload a video clip showing themselves dumping cold water over their heads or to donate to Amyotrophic Lateral Sclerosis Association (ALSA). This social media campaign went viral, and raised over $114 million for ALSA from July to September 2014, compared to $5 million for the same period in 2013 (Wikipedia 2014).

Further, when social networking sites take the lead in fundraising, the social features can be fully exploited to promote giving. Right after the 2015 Nepal earthquake, Facebook launched a donation button at the top of the timeline to raise funds to assist in the relief of the earthquake (Berenson 2015). Donors were asked to share their donations on their timelines. By doing so, they encourage their followers to donate. This Facebook campaign raised more than $15 million from 750,000 donors, amounting to nearly half the donations raised by American Red Cross overall. Similarly, Sina Weibo raised over $15 million in 72 hours after the 2013 Ya'an earthquake (Sina 2015), through 37 crowdfunding projects in its affiliated crowdfunding platform, Weibo Philanthropy (gongyi.weibo.com).

Social media fundraising presents two features that distinguish it from traditional fundraising. On the one hand, the social media platforms provide explicit reputation incentive for giving. On social media platforms, donors' contributions are more visible than ever before. Donors can declare their acts of giving through user-generated posts, automated acknowledgement from crowdfunding platforms, or aggregate-level display of contributions in their social profiles (Smith 2010). Donors' social reputation gain is contingent on their expectation of how others perceive them. When their contributions are more visible to others, they receive higher reputation gain as altruists (Satow 1975, Wiepking 2008). It is therefore commonly believed that an increase in visibility of charitable behavior will translate into better fundraising performance (Andreoni and Petrie 2004). On the other hand, social media provide a rich set of information about other donors to prospective donors. One set of information is the donations made by users' followees. The broadcasting system allows users to receive notifications of their followees' acts of giving. Users are pressured to give following their followees' pro-social behavior because their actions are visible to other users on the platform (Alpizar et al. 2008). We denote this impact as the peer effect. The other set of information is the donations made by the crowd. Online platforms provide aggregate-level popularity indicators like the donation count and amount for each project (Cai et al. 2007, Dewan et al. 2016, Tucker and Zhang 2011). For example, in the Facebook campaign for the Nepal earthquake, the pledging goal and the amount remaining to reach the goal were listed on a specific page. The popularity indicators are collective evaluations of projects' quality, such as efficacy and the extent to which the cause can inspire empathy. These popularity indicators can also alter people's donation choices. We denote this impact as the popularity effect.

Understanding the impacts of reputation incentive, peer effects, and popularity effects on charitable giving over social media environment is of critical value to researchers and practitioners because these factors are the bases of resolving the free-rider problem and the rich-get-richer problem . The free-rider problem is potentially serious, and leads to a pervasive under-provision of public goods (Meer 2011). The rich-get-richer problem is less a concern in traditional fundraising because funds collected by nonprofit organizations are usually allocated among people in need. However, when victims become the direct recipients of funds in crowdfunding, the rich-get-richer problem can be much more serious. As an evidence, projects that better engage the crowd in the early phases of fundraising campaigns are more likely to be successful (Agrawal et al. 2015). Despite their great value, motivations of giving in the social media environment have received scant attention.

Our study is designed to jointly measure the impacts of reputation incentive, peer effects, and popularity effects on individual donation choices to fill the research gap. Specifically, we seek to answer the following questions: 1) How reputation, peer effects, and popularity effects impact individual donation decisions differently on social media platforms, and what the interplay of these factors is; 2) How to leverage these mechanisms to incentivize individual donations with better incentive designs, and improve the efficiency of funds allocation.

To address these two research questions, we employ a dataset from a leading donation-based crowdfunding platform in China, Weibo Philanthropy (gongyi.weibo.com). This platform was developed by Sina Weibo, the largest microblogging website in China. On Sina Weibo, users can post microblogs and follow other users as in Twitter. Sina Weibo developed Weibo Philanthropy as an integrated component of its major microblogging platform to bridge users formally with nonprofit organizations. Nonprofits register charitable causes on Weibo Philanthropy and solicit donations from microblogging users. When users make donations, the system publishes a microblog to the donor's timeline in the hope of raising awareness and engaging the donor's followers. Donation transactions are documented and published with donors' ID, which allows researchers to link donor information with social network topology information in the microblogging platform. We are therefore able to track the flow of influence in donations.

However, identifying the role of reputation on individual donations poses great challenges. The main reason is that reputation incentive is generally intrinsic and cannot be directly observed. Thus, most previous literature uses experimental or survey-based methods to identify the impact of reputation incentive on donation decisions (Andreoni and Petrie 2004, Clark 2002, Jones and Linardi 2014). In this research, we employ a unique website change on Weibo Philanthropy to address this challenge. At the end of 2013, the microblogging platform added a donation history widget to users' social profiles to display their past donations made through Weibo Philanthropy. The introduction of this widget imposes exogenous shock to users' reputation status because their donation counts are now publicly consumable to all users. This design change provides us with an identification source to understand the role of reputation, which is difficult to examine otherwise.

We further employ the Finite Mixture Model to capture unobserved individual heterogeneity and to segment users. A Finite Mixture Model employs users' observed decisions together with a set of covariates to infer their unobserved donor types. We identify three types of donors that exhibit different donation patterns. Segmenting prospective donors not only significantly improves our model's predictive power by 14.57%, but also allows us to infer the interplay of different motivating factors.

Despite the platform owners' desire to improve fundraising performance, not all users respond positively to the introduction of the widget. The first segment becomes more likely to donate, and the other two segments become less likely to donate after the introduction of this widget. The effect of the widget is moderated by donors' past contributions and their numbers of followers. With high counts of past donations, the first segment will be less incentivized, and the second segment will be more demotivated after the introduction of the widget. The demotivating effect from the widget is mitigated by the number of followers for the second segment.

We also find that while peer effects are positive for all segments, popularity effects are positive only for some segments. Specifically, popularity effects are positive for the first and third segments, and negative for the second segment. This second segment has more followers, and usually donates at earlier phases of fundraising. Individuals in this segment seem to exhibit "leadership" attributes.

Our paper makes several contributions to the literature of crowdfunding, philanthropy, and economics. To begin with, to the best of our knowledge, we are the first to simultaneously quantify and compare the influences of reputation, peer effects, and popularity effects on donation decisions in the social media environment at an individual level. Second, we capture a unique exogenous shock that allows us to isolate the impact of reputation on donation decision. It also allows us to evaluate the impact of displaying users' aggregate-level contributions on their profile pages. We find that this incentive design intended to improve donations through reputation may actually demotivate donors who used to contribute more. Third, the social network topology within the microblogging platform provides us a unique opportunity to track the flow of influence among donors, and separate peer effects from popularity effects on donation decisions. Last but not least, we employ the Finite Mixture Model to sort users based on their donation patterns. This method provides managerial insights to alleviating the free-rider and the rich-get-richer problems.

The remaining part of this essay is organized as follows: We first review related literature in section 2, then describe our research context in section 3. We formulate our model in section 4, and summarize our data in section 5. We present our results, segment characteristics analysis, and robustness checks in section 6. In section 7, we discuss our findings, limitations, and managerial implications.

2.2 LITERATURE REVIEW

2.2.1 *Crowdfunding*

Crowdfunding is an emerging phenomenon that has received much attention from academia in recent years. Kuppuswamy and Bayus (2014) divide crowdfunding projects into four categories: reward-based, equity-based, lending-based, and donation-based crowdfunding platforms (Kuppuswamy and Bayus 2014). Studies in donation-based crowdfunding platforms mainly focus on project-level determinants of successful projects (Meer 2014) and the impact of prior donations on subsequent donations (Bøg et al. 2012, Burtch et al. 2013, Koning and Model 2013). They find that prior donations, especially the initial donations, are vital to the success of crowdfunding projects because they signal the quality of the project. This is consistent with theoretical analysis that discusses the announcement of early donations (Andreoni 2006, Vesterlund 2003). In the field of online journalism, however, evidence of crowding out was found, which is consistent with traditional economic theory (Andreoni 1989, Andreoni 1990, Burtch et al. 2013). Saxton and Wang (2013) find that in the social media environment, the number of followers each nonprofit has will positively impact fundraising performance because of the network effects (Saxton and Wang 2013).

Research on individual-level decisions on crowdfunding platforms is relatively limited because individual contribution data is not commonly collectable on such platforms. Burtch et al. (2015) conducted a randomized field experiment and showed that delaying donation information reduces the contribution amount, but increases the donors' propensity to give (Burtch et al.

2015). We believe that this paper is the first work that investigates the impact of reputation, peer effects, and popularity effects on individual donation decisions jointly on charity-based crowdfunding platforms.

2.2.2 *Philanthropy*

Our work is related to a large body of literature that investigates the motivations of charitable giving, particularly reputation, peer effects, and popularity effects. Reputation has been shown to be a strong and dominant motivator of donation decisions (Bekkers and Wiepking 2010). Silverman et al. (1984) found that the announcement of donors' names can improve donors' reputation status, thus drive more donations (Silverman et al. 1984). Andreoni and Petrie (2004) demonstrated that donors prefer to reveal their identities for their donations to enhance their reputation; contributions in higher amounts can be collected in such settings. Closest related to our work is a stream of literature that studies the role of social comparison (Frey and Meier 2004, Harper et al. 2010, Schultz et al. 2007, Shang and Croson 2006). Jones and Linardi (2014) conducted an experiment to show that people are averse to both positive and negative reputation for being altruistic, and prefer to behave like an "average altruistic" person.

Popularity effects (or "social information" in some literature) have been recently studied as ways to encourage donations and resolve the free-rider problem in charitable fundraising. Popularity effects measure how information about prior donations may affect the donation decisions of subsequent agents. Some researchers find positive popularity effects where individual contribution increases with previous contributions from other donors (Gu et al. 2009, Shang and Croson 2009, Sugden 1984, Xia et al. 2012). This could be because individuals derive negative utility if they do not conform to social norms, according to which giving is valued (Shang and Croson 2009). An alternative explanation to this enforcement phenomenon is that previous donations can be perceived as a signal of the charity's quality (Vesterlund 2003). Negative popularity effects on individual donations are also documented in the previous literature (Adar and Huberman 2000, Burtch et al. 2013). This is mainly due to crowding out when the causes are taken care of by others.

While it is commonly believed that peer effects have a significant impact on donation decisions, research on the impact of peer effects on donation decisions is very limited, with a few exceptions (Carman 2003, Meer 2011, Smith et al. 2015). Meer (2011) examined the impact of solicitation from a person the prospective donor knows on his donation decisions. Carman (2003) provided an evidence of positive peer influence for charitable donations in the context of workplace teams. Castillo et al. (2014) found that public peer-to-peer solicitation is more effective than private peer-to-peer solicitation (Castillo et al. 2014). The reason for the scarcity of studies on peer effects is the difficulty in documenting network structure among donors in second-hand data research. As a result, income and other socio-demographic variables are used to define generic reference groups (Andreoni and Scholz 1998, Feldstein and Clotfelter 1976). The most closely related work to our study in terms of peer effects is Smith et al. (2015) which examined the impact of peer influence on donation amount in the context of fundraising for the London Marathon. However, they mainly examined the impact of previous donations on follow-up donations rather than measuring the peer effects based on social network structure.

Characteristics of the recipient, nonprofit organizations, and the solicitors also influence donation decisions. First, evidence was found that the urgency of need is positively correlated to the possibility of donations (Schwartz 1974). Second, nonprofits with high efficacy and trust from the society perform better in fundraising (Bowman 2004). These nonprofits are usually more connected to mass media, and can benefit from preferential attachment effect (Barabási and

Albert 1999). What is more, the solicitor plays an important role in fundraising. When solicited by family and friends, potential donors are more likely to give (Bekkers 2004, Schervish and Havens 1997). Further, donors are more likely to donate when they have a larger peer group (Einolf et al. 2013). Last, donors have heterogeneous preferences in charity types (Bekkers and Wiepking 2010).

2.3 Research Context

2.3.1 *Background*

Our research context is Weibo Philanthropy (http://gongyi.weibo.com/). Weibo Philanthropy was founded in 2012, and is the largest social-media crowdfunding marketplace in China. By 2015, more than $60 million in donations towards various charitable causes was raised through Weibo Philanthropy from 20 million individuals. This platform has become an increasingly important fundraising channel in China, particularly after a series of trust crises for government-owned nonprofit organizations in 2012 (Tan 2015). Similar to other crowdfunding marketplaces like GoFundMe, nonprofit organizations or qualified individuals register projects by setting up a webpage on Weibo Philanthropy (Figure 2.1 (A)). Nonprofits list descriptions of the charitable cause, the fundraising duration, and the location of victims on the webpage. They update the status of the recipient and the transfer of funds. Users visit the cause webpages and make donations electronically. The transaction-level donation history for each charitable cause is publicly accessible . Each charitable campaign has a pledging goal, and the money remaining to reach the goal is displayed next to it. Weibo Philanthropy applies the “Keep-It-All” (KIA) model, such that the amount raised will be delivered to the beneficiaries regardless of whether the goal is reached or not, or surpassed.

Weibo Philanthropy is an integrated component of the largest microblogging website in China, Sina Weibo (NASDAQ: WB). Users are required to log in to their Sina Weibo accounts before they can donate to Weibo Philanthropy. This allows us to uniquely identify donors with their microblogging IDs, and access their social network information on the microblogging platform (Figure 2.1 (C)). One key feature of the system is the auto-generated acknowledgement of donations. After a user makes a donation on Weibo Philanthropy, the system will automatically post (with donors' permission) a microblog on the donor's timeline to announce her act of giving. For instance, when a donor made a donation for the charity in Figure 2.1 (A), a microblog was generated on her timeline to show her donation to this charity cause (Figure 2.1 (B)). As a result, her followers on Weibo received this information when they browsed through their news feeds. Such a system enables propagation of charitable causes over donors' peer networks, and consequently provides researchers with opportunities to measure the impact of peer influence.

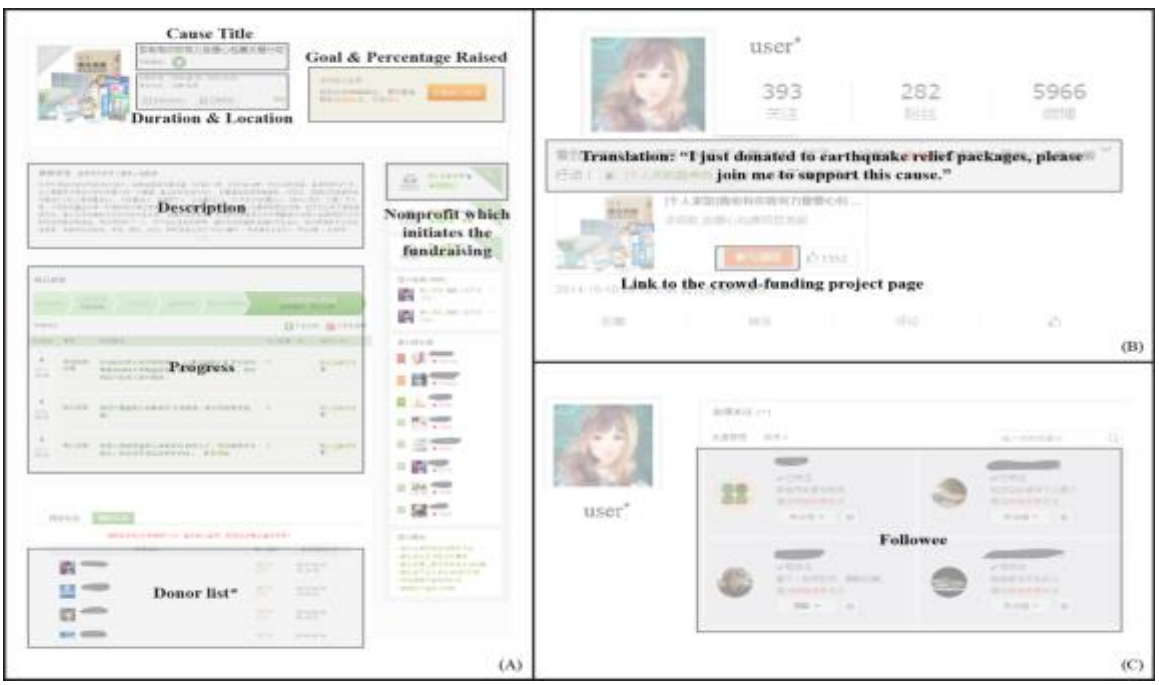


Figure 2.1. Donation-based Crowdfunded Market Weibo Philanthropy

Note: (A) donation-based crowdfunding page; (B) microblog acknowledging donation; (C) Followee List. * We anonymize donor's identities.

2.3.2 *Exogenous Shock*

At the end of 2013, the website introduced a charity history widget to each user's profile (Figure 2.2). This widget publicly displays the charitable projects supported by each individual. The total counts of past donations are highlighted at the top, and thumbnails of the projects are shown underneath. Since the widget is displayed on users' social profiles, it is publicly consumable for all users on the microblogging platform.

In contrast to the microblog-level visibility from the system-generated microblogs (Figure 2.1 (B)), this widget (Figure 2.2) greatly improves the visibility of users' contributions. First, this widget provides a single view that emphasizes the aggregate level of past donations. This makes it easy for visitors to view an accurate picture of the focal user's contribution level. Second, system-generated posts to acknowledge users' contributions are only broadcast to donors' followers. However, this widget makes this information available to any registered user of this platform. Third, system-generated microblogs may be superseded by new microblogs in news feeds quickly, while the widget will display all past donations all the time. To sum up, the introduction of this widget provides exogenous shock to people's reputation status while keeping other factors unchanged. It enables us to evaluate donors' responses to the higher visibility of their contributions.

Figure 2.2. Individual Donation History Widget

2.4 MODEL

In this section, we first discuss the theories behind the reputation incentive, peer effects, and popularity effects. We also describe the measures we use to examine these effects. We then introduce our baseline individual-level choice model, followed by a formal presentation of the Finite Mixture Model, which extends the baseline model to allow users to have heterogeneous donation patterns.

2.4.1 *Model Development*

2.4.1.1 Reputation Effect

The visibility of donors' contributions is a dominating factor with respect to giving behavior because it affects donors' reputation as altruists. Reputation as altruists is rooted in people's intrinsic need for social approval. When donors' pro-social behavior is more visible to others, they receive higher recognition (Bekkers and Wiepking 2010). These donors are regarded as having a higher level of generosity or income, which may result in reciprocity gain (Ariely et al. 2009). While a higher visibility of donors' contributions is commonly believed to impact people's donation behavior positively (Jones and Linardi 2014), the evidence of null or negative impact was also found (Reinstein and Riener 2012, Soetevent 2005). The explanation is that people are averse both to being too altruistic and too selfish (Jones and Linardi 2014). This theory is also referred to as social comparison (Frey and Meier 2004, Harper et al. 2010).

In our context, before the introduction of the donation history widget, people established their reputations as altruists only by the credible system-generated acknowledgement posted to their microblogging timelines. The introduction of the donation history widget on Weibo Philanthropy opened an additional channel to boost donors' reputation. Therefore, this widget may change users' donation patterns. Since the aggregate-level past contribution is highlighted by the widget, we expect that the impact of this new widget varies at different levels of past donation counts. We propose that the past donation count is likely a moderator of the treatment effect following the introduction of the widget. In a similar setting as ours, Frey and Meier (2004) have demonstrated the moderation effect of past contribution to social comparison. Donors' group size is also likely a moderator of the treatment effect from the introduction of the widget. This is because users' reputation is based on their expectation of how others perceive them. When they have more followers, this expectation is strengthened. When users have a larger group of followers, they have a more pronounced reputation gain if they increase donations following the treatment. On the other hand, they will suffer from a more pronounced reputation loss if they decrease donations after the treatment (Bekkers and Wiepking 2010, Einolf et al. 2013).

Our analysis is enabled by an exogenous shock to the visibility of users' contributions as we described in the last section. Since the exogenous shock is the introduction of a donation history widget, we use a treatment dummy (*Treatment*) that indicates whether the widget is implemented to identify the impact of this exogenous shock. In the heterogeneous treatment effect analysis, we interact Treatment with users' past donation counts (Past Donations) and number of followers (Followers) to examine their moderating effects.

2.4.1.2 Peer Effects

Peer effects refer to the social contagion of a node that influences its neighbors in a certain way (Aral et al. 2009). When a donor contributes to a charitable project, the system will post an automatic acknowledgement to announce her contribution and encourage her followers to give. While such peer effects are intuitive, empirical evidence was scant, due to the difficulty to document donors' social network topology (Carman 2003, Meer 2011, Smith et al. 2015).

Our work leverages the unique social network topology in Sina Weibo to examine peer influence in social media fundraising context. We denote the peer reference group as a focal user's followees whose microblogs are subscribed to by the focal user. The peer group's donations are instantly broadcast to the focal user's news stream to be read. In the diffusion of new products, marketing researchers often use the counts of past adoptions by the reference group members, termed "installed base", to examine peer effects (Manchanda et al. 2008). Following the same idea, we employ the logarithm of the number of followees who donate before the prospective donor in a certain charitable cause (*Donated Followees*) to examine the peer effect of giving.

We address the following challenges in the estimation of peer effects. It is notable that the availability of the exact sequence of donation transactions of each charitable cause eliminates the reflection problem (Manchanda et al. 2008, Manski 1993). To handle the homophily problem, we use an instrumental variable approach. We use the donation counts for the focal users' followees' followees as an instrumental variable. The idea is that followee's followees may affect fellowees' behavior, but not that of the focal user's (Narayanan and Nair 2013). However, in the Wald test of exogeneity following the instrumental variable estimation, we find that homophily does not play a significant role (Section 2.6.4.2). Therefore, we stick to a basic model in our analysis. Alternatively, assuming the existence of homophily, we used fixed effects estimation to control for the homophily problem. This is feasible because in our panel data we observe variations in the number of donated followees both at time level and project level for each individual. The results of our fixed effect estimation remain qualitatively the same as our current model. Last, other correlated unobservables like time effects can also contaminate the estimation of peer effects. We explicitly control for time effects in our robustness check (Section 2.6.4.5). We show that there are no significant time effects within the period of our study.

2.4.1.3 Popularity Effects

Popularity effects (or "social information" in some literature) refer to the impact of past contribution information, usually in the form of donation count and amount, to subsequent donors (Cai et al. 2007, Dewan et al. 2016, Tucker and Zhang 2011). Both positive and negative popularity effects were found in terms of donation choices. The negative effect arises from altruism (Andreoni 1989, Andreoni 1990). When the victims are supported by others, the altruistic donors will crowd out because they only care about the welfare of the victims. On the other hand, the positive effect was also found due to the following two reasons. First, a high level of aggregate donations signals higher collective evaluations of the charity, thus donors become more likely to contribute (Vesterlund 2003). This is usually termed "consumption externality" or "social spillover" effect. Second, individuals may follow the crowd to connect to the community. Since departure from the social norm may lead to disutility, users donate to projects that already received lots of donations (Bernheim 1994).

We measure popularity effects using the logarithm of total number of donors net of those who are directly followed by the focal user (*Donated Nonfollowees*). This measure is constructed for each individual, at each point in time, and for every project. We note that prior research has used a similar approach to identify the boundary between peers and the crowd (Lee et al. 2015).

2.4.1.4 Control Variables

There are project-level characteristics as well as individual-level characteristics that may influence people's donation behavior. Project-level characteristics include the remaining amount to reach the pledging goal (*Goal*), the charity category (*Category*), and the number of followers owned by the nonprofit organization or individual who initiated the project (*Nonprofit Followers*). Individual-level characteristics include users' number of followers (*Followers*), user type (*Type*), and past donation counts (*Past Donations*). On the microblogging site, users are categorized into four types: verified user, badge user, ordinary user, and corporate user. Verified users are people whose occupations are verified by their organizations. Badge users are those who actively engage in microblogging platform. We exclude corporate users because we observe too few of them.

Table 2.1. Variable Description

Variables	Descriptions
Independent Variables	
Donated Followees$_{ijt}$	The logarithm of the number of followees who donate to project *j* before user *i* at time *t*.
Donated Nonfollowees$_{ijt}$	The logarithm of the number of individuals, net of user *i*'s followees, who donate to project *j* before user *i* at time *t*.
Treatment	A binary variable indicating whether it is before the introduction of the widget or after. The value equals 1 after the introduction of the widget and vice versa.
Project-level Controls	
Goal$_{jt}$	The logarithm of the amount remaining to reach the pledging goal truncated at zero[1].

[1] There are two measures regarding the amount of project goal: one is the static pledging goal set by the nonprofit; the other is the remaining amount to reach the static goal. This is calculated as the pledging goal minus the raised

Category$_j$	There are four types of charitable projects: health projects, education projects, natural disaster projects, and human service.
Nonprofit Followers$_j$	The logarithm of the number of followers owned by the nonprofit/individual who initiates the charitable cause *j*.
Individual-level Controls	
Type$_i$	There are four types of users: verified, badge, ordinary, and corporate user.
Followers$_i$	The logarithm of the number of users following the user *i*.
Past Donations$_{it}$	The logarithm of the number of donations made by user *i* at time *t*.

Note: i refers to individual, j refers to charitable project, and t refers to time, which is weekly.

2.4.2 *Baseline Model*

Based on the discussions above, we construct a Logit model capturing individual donation choices. We use d_{ijt} to denote individual *i*'s donation decision on charitable project *j* at time *t*. $d_{ijt} = 1$ if the donation is made, and 0 otherwise. The donation probability can be written as: $\Pr(d_{ijt} = 1) = 1/\left(1 + \exp\left(-V_{ijt}\right)\right)$. Here, V_{ijt} can be further written as below:

$$\begin{aligned} V_{ijt} = {} & \beta_1 Donated\_Followees_{ijt} + \beta_2 Donated\_Nonfollowees_{ijt} \\ & + \beta_3 Treatment + \beta_4 Past\_Donations_{it} + \beta_5 Followers_i \\ & + \beta_6 Project\_Type_j + \beta_7 Nonprofit\_Followers_j + \beta_8 Goal_{jt}. \end{aligned} \tag{2.1}$$

To further account for project-level unobserved heterogeneity, we cluster our data at project level to allow correlations within each project.

2.4.3 *Finite Mixture Model*

In equation 2.1, we assume a simple scenario where the impacts of covariates are the same across individuals. In this section, we relax this assumption and allow donors to be heterogeneous in their preferences. We employ the Finite Mixture Model where each donor is categorized to a

funds in each period. Since some donors choose not to reveal the amount of contribution, this measure is biased positively. However, since this measurement error does not seem to be correlated with the error term of individual choice, we believe that it will not lead to biased estimation of our choice model. In addition, our model results stay robust even after we take out this control variable.

certain segment with some probability. Individuals within the same segment are assumed to have a homogeneous taste, while the ones in different segments have different preferences (Dempster et al. 1977, Haaijer et al. 2000). By employing the Finite Mixture Model, we can better control for unobserved heterogeneity among donors, and significantly improve the predictive power (Bapna et al. 2011). In addition, this method offers deeper managerial insights to better guide decision making of the platform.

In the Finite Mixture Model, we assume that there are S segments of donors. Accordingly, there are S distinct sets of preference parameters, $\boldsymbol{\beta} = (\boldsymbol{\beta}^1, \boldsymbol{\beta}^2, \cdots, \boldsymbol{\beta}^S)$. We use $\boldsymbol{\beta}^s$ to represent the coefficient for segment $s \in \{1, 2, \ldots S\}$. The unconditional probability for an individual to belong to segment s is π_s. Thus the probability of observing a sequence of choices for individual i in segment s is:

$$L_{i|s} = \prod_{t=1}^{T} \prod_{j=1}^{J} \left(\frac{\exp(\boldsymbol{\beta}^s \mathbf{x})}{1 + \exp(\boldsymbol{\beta}^s \mathbf{x})} \right)^{d_{ijt}} \left(\frac{1}{1 + \exp(\boldsymbol{\beta}^s \mathbf{x})} \right)^{1 - d_{ijt}} . \tag{2.2}$$

As researchers, we don't observe which segment individual i actually belongs to. Thus, we can only infer the individual's probability of belonging to a segment based on the observed attributes and the individual's donation decisions. Therefore, we specify the unconditional likelihood of donors' sequence of choices with the weighted average of equation 2.2 over all segments:

$$L = \prod_{i=1}^{n} \sum_{s=1}^{S} \eta_{si} L_{i|s} , \tag{2.3}$$

where η_{si} is the conditional probability that individual i belongs to segment s.

Directly maximizing the log of equation 2.3 is difficult because of numerical issues, thus we maximize the following equivalent expression (Dempster et al. 1977):

$$\ln L = \sum_{i=1}^{n} \sum_{s=1}^{S} \eta_{si} \ln L_{i|s} . \tag{2.4}$$

It was proven that using an Expectation-Maximization approach, the maximization of equation 2.4 will always converge because the likelihood is non-decreasing (Dempster et al. 1977).

2.5 DATA

We used donation transactions four weeks before and four weeks after the introduction of the widget to conduct our study from December, 2013 to January, 2014. In our two months of data, we had 10,637 donations made to 549 charitable projects by 5,112 individuals. We aggregated donations at the weekly level and ended up with eight periods. In each period, all ongoing projects were included in each individual's choice set unless the individual had donated to the same project before. We randomly selected 75% of donors in our data as the estimation sample, and the remaining 25% of donors in our data as the holdout sample to validate the performance of the model calibrated by the estimation sample.

At an individual level, we included users who donated to at least two projects in their whole donation history from 2/6/2012 to 7/21/2014 to avoid purely impulse-driven donors. We also excluded users who have fewer than 10 followers or are following fewer than 10 users to avoid zombie microbloggers. We removed those who participated in a charity lottery project with a monetary reward because they had a potential monetary motivation which is not the focus of this paper. Further, we excluded donors whose total donation count from 2/6/2012 to 7/21/2014 exceeded 100 because these donors seem to be managed by multiple users. Note that 100 is the 99.7% percentile of the donation count distribution. In addition, we excluded 82 donors who donated to more than 10 projects during our focal period of two months because of their unusual donation patterns. We believe 10 is a good threshold because it is more than the average donation count for the two months (2.05) plus three standard deviations (2.42). Finally, donors who made at least one donation during the two-month period of our study were included as active potential donors. At a project level, we excluded extremely popular projects which received more than 1000 donations. For these projects, other factors like media exposure or celebrity endorsement may have come into play.

We list the summary statistics with the original scales in Table 2.2. Because all measurements are skewed, we took the log of these variables when we estimated the model. In terms of user type, we found that our sample constitutes 92% ordinary users, 6% badge users, and 2% verified users.

Among the 549 charitable projects, 414 of them were in the health category, 89 were related to education, 32 were in human service category, and 14 projects supported natural disaster causes. The goal of these projects ranged from ¥500 to ¥120,000. Since the fundraising is not capped, the amount remaining to reach the goal can be negative when the raised amount exceeds the goal. Because we use the remaining goal to measure urgency of need, we truncate the remaining goal at zero to better serve its purpose.

Table 2.2. Summary Statistics

	Covariates	Mean	S.D.	Min	Max
Individual Level (n=5112)	*Followers*	3546.09	36902.41	10.00	1250929.00
	Past Donations	9.00	13.70	1.00	99.00
Choice Level (n=7655721)	*Donated Followees*	0.04	0.39	0.00	88.00
	Donated Nonfollowees	60.97	72.76	0.00	397.00
	Goal (Remaining)	34920.32	24559.09	-1721.00	119900.00
	Treatment	0.51	0.50	0.00	1.00
Project Level (n=549)	*Nonprofit Follower*	37422.04	151952.7	0	2120734
	Goal (Pledging)	35585.55	24669.37	500	120000

To understand the overall funding status of Weibo Philanthropy better, we plotted the pledging goal of each project together with its final pledged amount in Figure 2.3. In this plot, we first sorted projects based on the pledging goal. We then sorted projects based on the pledged amount at each level of the pledging goal. For example, all projects with a pledging goal of ￥50000 are in Range I. The projects in Range II did not meet their goals, and the projects in Range III successfully met their goals. This plot showed strong evidence of under-contribution in giving. Many projects, especially those with high goals, didn't raise enough funds to meet the need, as shown in the dark grey area of Figure 2.3. Furthermore, the pledged amount at each goal level seems to follow a power law distribution, as shown at the edges of the light grey area at each pledging goal level. This indicates that fundraising over social media is a preferential attachment process where the rich get richer. It shows a suboptimal allocation of donors' funds because the projects that draw less attention may be better funded if there were no social interaction or social information on the platform. This plot presents the free-rider problem and the rich-get-richer problem that we seek to address in this work.

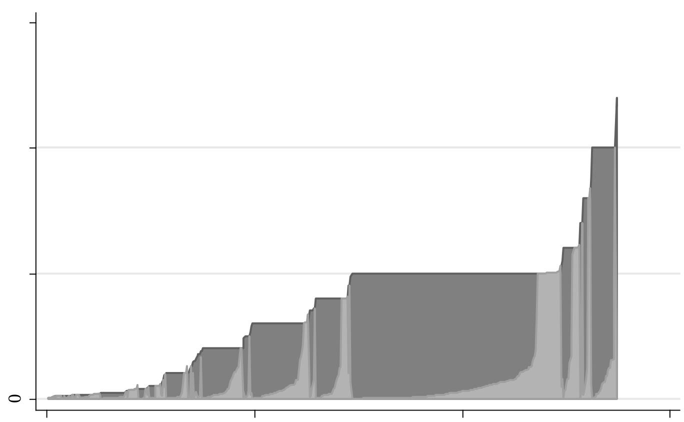


Figure 2.3. Project Funding Status

2.6 RESULTS

2.6.1 *Segmentations*

A critical question for the Finite Mixture Model is to decide the optimal number of segments. In this research, we use likelihood-based information criteria including AIC and BIC both for our estimation sample and holdout sample to determine the optimal number of segments for the model. Other than the information criteria, we also take into consideration interpretability and parsimony.

We report the fit statistics for both the estimation sample and the holdout sample in Table 2.3. Overall, the model with 3 segments outperforms the other models. In the holdout sample, the 2-segment model is the most preferred according to BIC, which is likely to result in a small model. The AIC for the 3-segment model is slightly higher than that for the 4-segment model, but the increment is relatively marginal considering the interpretability and parsimony (Kamakura and Russell 1989). In the estimation sample, the 3-segment model strictly outperforms the other models in terms of BIC, while the 4-segment model is the winner in terms of AIC. Considering all factors, we choose the 3-segment model as our main model. It is worth mentioning that even in the 4-segment model, our major results stay stable.

Table 2.3. Fitness Statistics

	Estimation Sample (75%)	Holdout Sample (25%)
Number of observations	5684962	1970759
1-Segment model		
LL	-52340.566	-18204.996
AIC	104711.13	36453.992
BIC	104914.43	36728.859
Pseudo R2	1.83%	1.72%
2-Segment Model		
LL	-50650.81	-17399.609
AIC	101429.62	34927.218
BIC	102297.03	**35726.829**
Pseudo R2	5%	6.07%
3-Segment Model		
LL	-49646.876	-16935.495
AIC	99581.753	34158.99
BIC	**101533.43**	35958.115
Pseudo R2	6.88%	8.58%
4-Segment Model		
LL	-49088.791	-16684.401
AIC	**98689.582**	**33880.802**

BIC	102159.24	37079.248
Pseudo R2	7.93%	9.93%

To infer the explanation power of different models, we calculated McFadden's Pseudo R2. The overall R2 sees a significant increment from the 1-segment to the 2-segment model, and to the 3-segment model. After the model reaches four segments, the Pseudo R2 doesn't see a noticeable improvement. This corroborates our choice of the 3-segment model.

Finally, we calculated the prediction accuracy for the holdout sample. The 3-segment Finite Mixture Model has improved the prediction accuracy by around 14.57% from the 1-segment model. Specifically, the classification accuracy of the 1-segment model is 56.87%, while the classification accuracy of the 3-segment model is 65.16%, with the accuracy for each segment being 59.06%, 75.26%, and 67.56%.

2.6.2 *Segment Characteristics*

In this section, we investigate donors' individual-level and choice-level characteristics and compare these characteristics across different segments. We took a snapshot of past donation counts right before the introduction of the widget (Past Donationst4) to avoid potential confoundedness. We also introduced Donation Order to reflect the degree of earliness in giving. We calculated donation order for each donor. It varies from 0 to 1, and a smaller value indicates that the donor makes donations earlier than others. It is calculated using relative donation sequence at each project. For example, if a project has 5 donations in total, the relative sequence for the first donor is 1/5, and that for the last donor is 5/5. We calculated the donation order for each individual i by averaging relative sequences across all projects she contributed. The summary statistics at both individual level and choice level are presented in Table 2.4.

Table 2.4. Descriptive Statistics of Three-Segment Finite Mixture Model

		Low-contribution		Medium-contribution		High-contribution	
Individual Level	*Donation Order*	0.501	(0.146)	0.455	(0.158)	0.498	(0.130)
	Followers	5.881	(1.583)	6.061	(1.588)	5.850	(1.440)
	Past Donations^{t4}	1.012	(1.198)	1.164	(1.226)	1.449	(1.281)
	User.Normal User	0.927	(0.261)	0.931	(0.254)	0.928	(0.258)
	User.Expert	0.053	(0.224)	0.056	(0.230)	0.053	(0.224)
	User.VIP	0.020	(0.141)	0.013	(0.114)	0.019	(0.137)
Choice Level	*Donated Followees*	0.014	(0.108)	0.014	(0.104)	0.030	(0.184)
	Donated Followees' Followees	0.008	(0.128)	0.010	(0.135)	0.017	(0.178)
	Donate	0.001	(0.033)	0.001	(0.036)	0.001	(0.037)

Standard deviations in parentheses, log scales are used in this statistical summary.

Before starting our discussion, we label each segment based on Past Donationst4 for easier interpretation. The graphical illustration is in the right panel of Figure 2.4. The first segment is labelled as "low-contribution" because donors in this segment have the lowest past donation count (6.75 in original scale) before the introduction of the widget. This segment constitutes 39.80% of the total sample. The second segment is labelled as "medium-contribution" because donors in this segment contributed a medium level of past donations before the treatment (8.54 on original scale). This segment constitutes 20.38% of all donors in the sample. The third segment is labelled as "high-contribution" because donors in this segment have the highest number of past donation counts before the widget is implemented (11.86 in original scale). This segment takes 39.82% of all donors. It is also notable that the differences among segments are significant in terms of past donations. Specifically, the low-contribution segment is significantly lower than the medium-contribution segment in past donation counts (p-value=0.0038), and the medium-contribution segment is significantly lower than the high-contribution segment in this measure (p-value<0.0001).

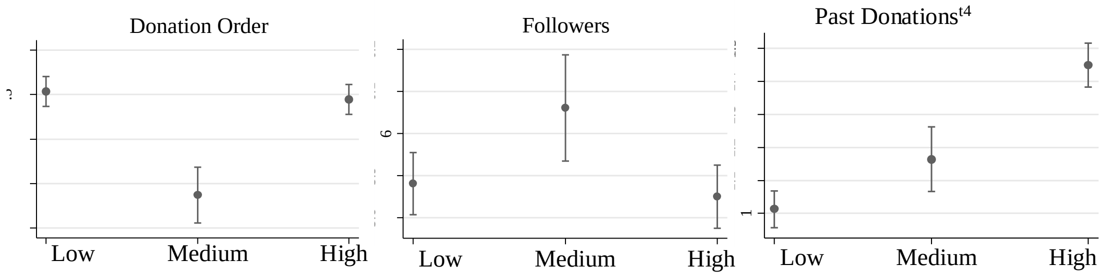


Figure 2.4. Segments Characteristics

We have interesting findings about the medium-contribution segment at individual level characteristics. In terms of Followers, we find that the medium-contribution segment has the highest number of followers (p-value=0.008 for greater than the low-contribution, and p-value=0.0017 for greater than the high-contribution segment). On the other hand, the low-contribution segment is not significantly different from the high-contribution segment in this measure (see the middle panel of Figure 2.4). In addition, the medium-contribution segment shows a significantly lower Donation Order than the other two segments (p-value<0.0001). However, the donation order of the low-contribution segment is not significantly different from that of the high-contribution segment (see the left panel of Figure 2.4). In other words, the medium-contribution segment usually donates at earlier phases of fundraising than the other two segments.

Besides individual-level characteristics, choice-level characteristics also reveal different features among segments (Table 2.4). Specifically, the average level of Donated Followees is equal across the low-contribution and the medium-contribution segments, and is much higher in the high-contribution segment. The high-contribution segment also has the highest Donated Followees' Followees' count. These findings imply positive peer effects.

2.6.3 *Parameter Estimations*

We present the parameter estimates (βs) of the 3-segment Finite Mixture Model in Table 2.5. The index s denotes the low-contribution segment when it takes the value of 1, the medium-contribution segment when it takes the value of 2, and the high-contribution segment while taking the value of 3. Note that this well-delineated segmentation is calibrated using all covariates with no constraints in how different factors affect the segmentation.

Table 2.5. Parameter Estimation of FMM Model

	Low-contribution		Medium-contribution		High-contribution	
Donated Followees	1.366***	(0.134)	2.689***	(0.122)	0.184+	(0.095)
Donated Nonfollowees	0.314***	(0.042)	-0.441***	(0.055)	0.404***	(0.047)
Treatment	**2.271*****	**(0.112)**	**-0.973*****	**(0.112)**	**-1.162*****	**(0.112)**
Past Donation	0.0160	(0.037)	0.0596*	(0.028)	0.226***	(0.030)
Goal	0.268***	(0.051)	0.191***	(0.039)	0.510***	(0.071)
Followers	0.00790	(0.014)	0.0257	(0.017)	0.00785	(0.011)
User.Expert	0.0732	(0.073)	-0.283*	(0.123)	-0.187**	(0.069)
User.VIP	-0.261*	(0.130)	-0.461*	(0.198)	-0.357**	(0.132)
Nonprofit Followers	-0.0648+	(0.039)	0.384***	(0.046)	-0.158***	(0.039)
Project.Health	-0.185	(0.220)	-0.818***	(0.167)	0.902***	(0.261)
Project.Natural Disaster	0.511+	(0.292)	0.688*	(0.329)	0.199	(0.498)
Project.Human Service	0.912*	(0.382)	0.446	(0.355)	0.366	(0.374)
Constant	-11.83***	(0.614)	-10.27***	(0.368)	-12.87***	(0.717)
Segment Share	39.80%		20.38%		39.82%	
R^2	7.42%		7.68%		5.95%	
McFadden's Pseudo R^2	6.88%					

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

First of all, we find that the introduction of the widget had significantly different impacts for donors in different segments. The coefficient of Treatment is positive and significant for the low-contribution segment (β_3^1=2.271, with p-value<0.001). The low-contribution segment's odds of donation increase by a multiplicative factor of 9.69 after the introduction of the widget. The coefficients of Treatment in the medium-contribution and high-contribution segments are negatively significant (β_3^2=-0.973 and β_3^3=-1.162, with both p-value<0.001). Following the treatment, the medium-contribution segment's odds of donation reduce by a multiplicative factor of 2.65, and the high-contribution segment's odds of donation reduce by a multiplicative factor of 3.2. It can be seen that the medium-contribution segment was less influenced by the introduction of the widget than the high-contribution segment. As we can see from the segment share, these three segments takes a percentage of 39.80%, 20.38% and 39.82%, respectively. Despite the platform owner's desire to improve fundraising, less than half of the donors responded positively to this new feature, while the majority of the donors had negative responses.

Combining this result with users' past donation counts before the introduction of the widget (Past Donationst4), we obtain an interesting interpretation about users' different responses to the introduction of the widget. We find that those who contributed less are incentivized by the widget, and those who contributed more are demotivated. This explicitly showed that people avoid being too altruistic or too egoistic (Jones and Linardi 2014). As stated in the theory of social comparison, the widget enables users to compare their past donation with that of others. Those who contributed less than the group mean are incentivized to contribute more, and those who contributed more to charities end up contributing less. Both groups of donors want to behave like an average-altruistic person. It is surprising that such a widget can stimulate social comparisons. This is because usually social comparisons are considered effective only when the incentive design explicitly displays users' percentiles or rankings of contributions.

Next, we examine the popularity effects. To our surprise, the medium-contribution segment has a significantly negative coefficient for Donated Nonfollowees (β_2^2=0.441, with p-value less than 0.001). This indicates that they will crowd out when projects are very popular. This crowding-out behavior signals their higher motivation from altruism or "behaving altruistic"

(Andreoni 1989, Andreoni 1990). When users care more about the welfare of victims, they have less incentive to support projects well taken care of by others. Instead, they will contribute to the less popular projects because the victims in these projects are more in need. The existence of this segment provides potential opportunities to resolve the rich-get-richer problem. The coefficients of Donated Nonfollowees for the low-contribution and the high-contribution segments are both positive and significant (β_2^1=0.314 and β_2^3=0.404, with p-value both less than 0.001). This suggests that donors in these two segments tend to donate to charitable causes that are well-liked by the public (Frey and Meier 2004). One possible reason is that the donation count is a signal of quality, and the low-contribution and high-contribution segments rely more heavily on such a signal because they have less private information about the true quality. The medium-contribution segment, whose members are considered "leaders", can better evaluate the quality of the projects. It is also possible that individuals in these two segments simply have a higher tendency to conform to the social norm of the community.

Finally, we interpret the results of peer effects. The coefficients of Donated Followees are all positive for the three segments (β_1^1=1.366, β_1^2=2.689, with p-value both less than 0.001, β_1^3=0.184, p-value=0.051). In other words, peer effects are generally positive in the social media fundraising environment. It is worth mentioning that donors in the high-contribution segment are much less influenced by the donation choices of their followees than donors in the other segments. For donors in the high-contribution segment, one followee's donation has even less influence than one nonfollowee's influence. However, for donors in the low-contribution segment, one followee's donation is equivalent to over 4 nonfollowee's donations in terms of donation influence. Overall, peer effects have demonstrated an effectiveness in resolving the free-rider problem (Meer 2011).

2.6.3.1 Heterogeneous Treatment Effects

To further examine heterogeneity in the treatment effect, we investigate the moderating roles of Past Donation and Followers, with respect to Treatment. According to Frazier et al. (2004), we let both covariates enter the model at the same time to control for inflated Type I error (Frazier et al. 2004). We report the results in Table 2.6.

Table 2.6. Parameter Heterogeneous Treatment Effects

	Low-contribution		Medium-contribution		High-contribution	
Donated Followees	1.359***	(0.134)	2.694***	(0.122)	0.185+	(0.095)
Donated Nonfollowees	0.314***	(0.042)	-0.441***	(0.055)	0.404***	(0.047)
Treatment	2.923***	(0.233)	-1.258***	(0.261)	-1.430***	(0.208)
Past Donation	0.304***	(0.040)	0.111**	(0.034)	0.220***	(0.034)
Treatment×Past Donation	**-0.321*** **	**(0.052)**	**-0.187** **	**(0.061)**	**0.0266**	**(0.057)**
Followers	0.0266	(0.027)	-0.00632	(0.021)	-0.00162	(0.013)
Treatment×Followers	-0.0205	(0.030)	0.0983**	(0.032)	0.0353	(0.023)
Segment Share	39.80%		20.38%		39.82%	

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

First, the coefficients of the interaction Treatment×Past Donation are negatively significant for both low-contribution and medium-contribution segments. In other words, the low-contribution segment will be less incentivized, and the medium-contribution segment will be more demotivated by the treatment when donors have higher counts of past donations. The high-contribution segment sees an insignificant effect from the interaction. This is possibly because the past donation count within this segment does not have much variation.

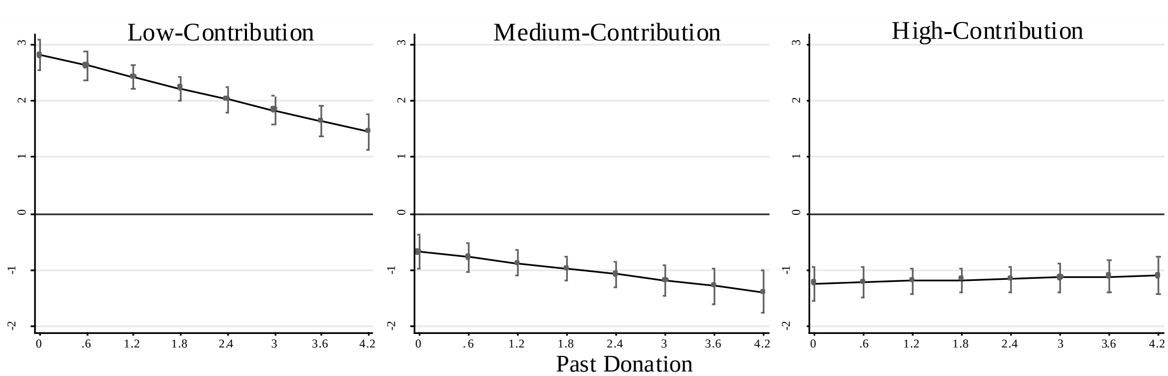


Figure 2.5. Average Marginal Effects of treatment=1 with 95% Cis

To illustrate the moderation role of Past Donation better, we plotted the average marginal treatment effect at different levels of past donations over the odds ratio of donating in Figure 2.5. As can be seen, the effect of the widget is positive for the low-contribution segment and negative for the other two segments. Further, both curves in the low-contribution and medium-contribution segments see a downward trend, indicating that donation probabilities generally decrease at a higher level of past donations, following the introduction of the widget. One explanation is that for individuals with a large number of past donations, their altruistic image has been well established by the introduction of the widget. Since their past contributions are displayed in their profiles in a stable manner, they have less incentive to keep contributing.

Second, we find that the coefficients of Treatment×Followers are significant and positive for the medium segment, while insignificant for the other two segments. In other words, the negative treatment effect is mitigated by Followers for the medium-contribution segment. The interpretation is that the medium-contribution segment takes into account the possible reputation damage of donating less frequently. Such reputation damage can be caused by their followers' perceptions that they are not altruists. The finding implies that donors in the medium-contribution segment are more concerned with followers. It is consistent with our early finding that the medium-contribution segment has the highest number of followers. To leverage such a result, platform owner or nonprofits should invest more to engage users with lots of followers because once they give, they are likely to keep giving to avoid reputation damage.

2.6.4 *Robustness Checks*

In this section, we conduct robustness checks to validate our assertions by ruling out alternative explanations.

2.6.4.1 Falsification Check

The treatment effects we discover can be spurious if users change their donation probability over time. They may reduce or increase donation probabilities at different phases of their donation lifecycles. In donors' lifecycles, they may put increasingly high efforts in the beginning phase, and drop out of the platform gradually as they approach the ending phases of their lifecycles. This trend is demonstrated in repeated public good experiments (Andreoni 1988). It can also explain why we find that donors with low past donation counts were incentivized, and others with high past donation counts were demotivated. To rule out this explanation, we conducted two falsification tests where we used false times as treatment. In the pre-treatment falsification test, we changed the treatment time to two weeks before the actual treatment (12/17/2013), and used only the first four weeks data when no website design change took place (12/3/2013-12/31/2013). In the post-treatment falsification test, we changed the treatment time to two weeks after the actual treatment (1/14/2014), and used the last four-weeks data (1/1/2014-1/28/2014), with the widget present at the beginning of this period. We illustrate the details of the falsification tests in Figure 2.6, and compare the results of these two falsification tests with the full model result in Table 2.7. Notice that we employ the 1-segment model rather than the 3-segment finite mixture model for a more straightforward comparison.

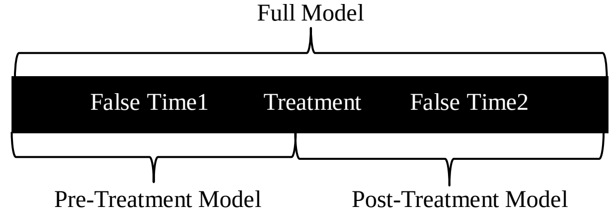


Figure 2.6. Illustration of Falsification Test

Table 2.7. Robustness Check – Falsification Test

	Full Model		Pre-Treatment		Post-Treatment	
Donated Followees	0.822***	(0.080)	0.671***	(0.099)	0.972***	(0.099)
Donated Nonfollowees	0.225***	(0.036)	0.194***	(0.055)	0.244***	(0.043)
Treatment	**0.562*** **	**(0.130)**	**-0.0974**	**(0.147)**	**0.302+**	**(0.160)**
Past Donations	0.276***	(0.029)	0.292***	(0.039)	-0.0161	(0.047)
Treatment×Past Donations	-0.313***	(0.045)	-0.00215	(0.050)	-0.135*	(0.056)

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

From the estimated parameters, we find that the treatment effect is positively significant for the full model, with a negatively significant interaction effect from Treatment×Past Donations. This indicates that for people with a low level of past donations, this widget encouraged them to contribute more, and for people with a high level of past donations, this widget discouraged them from contributing. The treatment effect of the pre-treatment model is insignificant. This indicates that the donation probability remains unchanged without the intervention of this widget. In the meantime, the treatment effect of the post-treatment model is significant at 0.1 level, with a smaller magnitude than that of the full model. This shows that the treatment effect lasts more than two weeks, and it decays in its effect over time. This is likely because users may take some time to discover the new feature. In sum, the results from these falsification tests verify that the driver of users' behavior change at the end of 2013 is not the natural evolvement in people's donation behavior, but the introduction of the widget .

2.6.4.2 Homophily

Empirical research always faces the problem of identifying peer effects on individuals' choices from homophily (Aral et al. 2009). In our context, individuals may follow their followees' choices either because they are influenced by their followees or because they are intrinsically similar to them. In this robustness check, we show that homophily actually does not play a significant role in the social contagion phenomenon in our context (Smith et al. 2015).

To test this assertion formally, we constructed another variable - donation counts for the focal users' followees' followees (Donated Followees' Followees) – as an instrumental variable for Donated Followees. The rationale is that followees' followees will influence followees' donations, but not the focal user's. This instrumental variable construction is widely used to account for the endogeneity problem in peer influence literature (Narayanan and Nair 2013). We used a standard IV estimation procedure for binary outcome variables, and checked the Wald test for exogeneity (Wooldridge 2010). It is shown that we cannot reject the null hypothesis of exogeneity (p-value = 0.7720), and a regular Logit model is a more proper specification. A possible explanation is that Sina Weibo is a general social media platform, and charity is only a small component of it.

By checking the correlation between Donated Followees' Followees and Donated Nonfollowees, we find that these two covariates are only weakly correlated (correlation= 0.0727). Thus, introducing Donated Dollowees' Dollowees does not cause multicollinearity problems.

2.6.4.3 Narrower Time Window

We also checked our model's robustness to the length of time frame by narrowing our time window to two weeks before and two weeks after the treatment. It turns out that the major effects are qualitatively the same as in our main model. The only difference is that the high-contribution segment no longer sees positive peer effects. Overall, we have confirmed that our results are robust by varying the time frame of the dataset.

Table 2.8. Robustness Check – Shorter Timeframe

	Low-contribution		Medium-contribution		High-contribution	
Donated Followees	1.385***	(0.206)	2.743***	(0.240)	-0.0157	(0.160)
Donated Nonfollowees	0.264***	(0.061)	-0.542***	(0.101)	0.387***	(0.069)
Treatment	3.094***	(0.168)	-1.358***	(0.161)	-1.445***	(0.140)
Segment Share	39.80%		20.38%		39.82%	

Standard errors in parentheses

2.6.4.4 Advertisement Effects of the Widget

The positive treatment effect in the low-contribution segment can be spurious because the reputation change following the treatment confounds with the advertisement effects. The website design change is implemented through a widget that lists users' past donation counts on their profile pages. At the same time, the widget lists thumbnails of charity projects donated to by the focal user (Figure 2.2). For example, there are the four projects to which the focal user had contributed listed in the front display of the widget in Figure 2.2. By clicking the Next Page arrow, visitors can browse through all projects donated to by the focal user. It is possible that displaying the projects leads to visitors clicking through and donating to these charities.

To test the robustness of our results in the presence of advertisement effects, we included the interaction between Donated Followees and Treatment. This interaction can control for advertisement effects because after the widget came online, users were able to browse the thumbnails of charitable projects donated to by their followees. If advertisement effects exist, users would be more likely to donate to those projects, since those projects have more potential exposure to them. Our results stay robust across all segments after controlling for the potential advertisement effects (Table 2.9). This result also confirms that there does exist positive advertisement effects for the medium-contribution and high-contribution segments. However, the low-contribution segment, which is incentivized by the treatment, doesn't see significant advertisement effects.

Table 2.9. Robustness Check –Advertisement Effect

	Low-contribution		Medium-contribution		High-contribution	
Donated Nonfollowees	0.343***	(0.043)	-0.397***	(0.056)	0.424***	(0.040)
Treatment	2.282***	(0.116)	-1.025***	(0.115)	-1.159***	(0.117)
Donated Followees	1.319***	(0.152)	2.520***	(0.139)	0.0189	(0.103)
Treatment×Donated Followees	-0.00220	(0.185)	0.407*	(0.185)	0.383*	(0.169)
Segment Share	39.80%		20.38%		39.82%	

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

2.6.4.5 Time Effects

To account for time effects that may affect the whole community at the same time, we include weekly time dummies in our model. Individual donation probabilities are modeled with $\Pr(d_{ijt}=1)=1/\left(1+\exp\left(-V_{ijt}\right)\right)$, where is updated as follows:

$$V_{ijt}=\boldsymbol{\beta}^{s}\mathbf{x}+\boldsymbol{\gamma}\mathbf{z}_{\mathbf{t}} \quad . \tag{2.5}$$

Here, $\mathbf{x}$ refers to a row vector of the covariates as in equation 2.5, and $\mathbf{z}_{\mathbf{t}}$ refers to a row vector of time dummies for each week. Note that the time effect $\boldsymbol{\gamma}$ is homogeneous across segments. This is in accordance with the literature on peer influence. The results are provided in Table 2.10.

From the parameter estimation results in Table 2.10, we can see that our major results stay robust. In the result, Week 4 is omitted because it is equivalent to the treatment dummy. Week 8 is also omitted because of collinearity. We note that there are no significant time effects because all time dummies are insignificant. This corroborates our main model specification which does not control for time effects.

Table 2.10. Robustness Check – Time Effect

	Low-contribution		Medium-contribution		High-contribution	
Donated Followees	1.378***	(0.128)	2.716***	(0.120)	0.194*	(0.088)

Donated Nonfollowees	0.316***	(0.040)	-0.427***	(0.055)	0.406***	(0.046)
Treatment	2.337***	(0.181)	-0.916***	(0.181)	-1.094***	(0.173)
Week 1	0.228	(0.159)	0.228	(0.159)	0.228	(0.159)
Week 2	-0.0176	(0.166)	-0.0176	(0.166)	-0.0176	(0.166)
Week 3	-0.00684	(0.117)	-0.00684	(0.117)	-0.00684	(0.117)
Week 5	-0.0728	(0.177)	-0.0728	(0.177)	-0.0728	(0.177)
Week 6	-0.0840	(0.159)	-0.0840	(0.159)	-0.0840	(0.159)
Week 7	0.00515	(0.139)	0.00515	(0.139)	0.00515	(0.139)
Constant	-12.62***	(0.718)	-12.62***	(0.718)	-12.62***	(0.718)
Segment Share	39.80%		20.38%		39.82%	
McFadden's Pseudo R^2	6.9%					

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

2.6.4.6 Split-Sample Analysis for Social Comparison

In this robustness check, we split our sample based on the size of past donation counts, and confirmed that social comparison is a major driver for our results. Specifically, we calculated the 39.80% and 60.18% quantiles of past donation count, and used them as cutoff values to split the sample (they are 2 and 5 in original scale, respectively). These percentiles are determined using the segment share of our 3-segment Finite Mixture Model.

Table 2.11. Robustness Check – Split-Sample Analysis

Past Donations (Original Scale)	0-2		2-5		5+	
Treatment	0.824***	(0.116)	-0.460***	(0.117)	-0.307**	(0.105)
Segment Share	39.80%		20.38%		39.82%	

Standard errors in parentheses
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

In this naïve split-sample analysis in Table 2.11, we find that donors with low past donation counts responded positively to the introduction of the widget, and donors with medium and high past donation counts responded negatively to the widget. These findings are consistent with the major explanation of social comparison from our Finite Mixture Model.

One more thing to note is that although this split-sample method seems to reproduce results similar to our major finding in terms of donors' responses to the introduction of the widget, it does not provide interplays of reputation incentive, peer effect, and popularity effect. It could not identify the crowding-out behavior because it overlooks the heterogeneity of users' donation pattern as a comprehensive result affected by many factors. Thus, this naïve model is insufficient to substitute our main model.

2.7 Conclusions

In recent years, the development of social media networking sites has enabled nonprofit organizations to solicit donations from social media users in a revolutionary way. The reputation incentive, peer effects, and popularity effects inherent in the social media platforms provide natural sources to overcome the free-rider problem which is prevalent in online fundraising. Using a unique individual-level dataset, we examine these factors concurrently on a charity-based crowdfunding platform. Among our rich set of findings, two phenomena are of significant note. First, when platform designers increase the visibility of people's contributions, especially at aggregate level, they may face an undesirable outcome for a subgroup of users. Such a website change may discourage donors who have a high level of past contributions. Second, a certain group of donors are more likely to crowd out from well-supported charitable projects and contribute to the less popular projects. They can reach a larger audience (followers), and they usually donate at early phases of fundraising. To the best of our knowledge, our paper provides the first empirical evidence of such phenomena in social media fundraising environment.

2.7.1 *Implications*

Our work has meaningful implications both theoretically and practically. First, we examined the impact of increasing the visibility of users' aggregate-level contributions. Both positive and negative impacts were found in previous literature (Jones and Linardi 2014). Our work adds empirical evidence to the social comparison theory in the context of social media fundraising. More importantly, our work provides insights into the incentive design of online communities. Platform designers need to be cautious when increasing the visibility of users' aggregate-level contributions. Instead, a platform can display users' recent contributions to avoid the negative outcomes from social comparison. Further, platforms can customize incentive designs based on user segmentations, if possible. For example, by turning off the widget for the medium-contribution and high-contribution segments, the platform can improve its fundraising performance.

Second, we provide potential solutions to the rich-get-richer problem. With positive peer effects, attention-getting projects will attract funds from the less popular projects. This leads to the rich-get-richer problem (Meer 2011). Our finding implies that nonprofits or platform owners who seek to direct funds to the less-popular projects should invest their fundraising effort on people who exhibit more leadership attributes. For example, platform owners can make incentive designs to motivate these "leaders". A high-level engagement of these crowding-out donors will lead to a more efficient allocation of funds in the whole platform. Our work also provides an explanation for the disparate findings of popularity effects. From both theoretical and empirical analysis, popularity effects can be positive and negative (e.g., Bøg et al. 2012 and Burtch et al. 2013, respectively). Using the Finite Mixture Model, we are able to observe both positive and negative popularity effects. We find that donors who exhibit leadership attributes will crowd out from popular projects, while other donors will crowd into popular projects.

Last, given the general existence of positive peer effects, we address the free-rider problem. Nonprofit organizations and crowdfunding marketers should encourage donors to engage their friends by sharing their contributions over social media sites. This will alleviate the free-rider problem, and cultivate a social norm of giving for the community. Crowdfunding marketplaces that do not have strong social networking features within their own sites should connect to major social media platforms like Facebook and Twitter. For example, Justgiving.com allows users to log in with their Facebook accounts to see what projects are supported by their Facebook friends. Such functions will greatly improve fundraising performance.

2.7.2 *Limitations and Future Research*

Our work also has its limitations. First of all, the introduction of the donation history widget is only one possible form of reputation incentive designs, and it applies to all users in the platform. While we have controlled for time-level, project-level, and individual-level attributes to account for any possible confounders, we acknowledge our second-hand data limitation. We believe that a randomized field experiment may allow researchers to control for a broader spectrum of confounders and provide more comprehensive insights into different incentive design mechanisms. The incentive designs to induce users' contributions hold considerable value to practitioners. Second, our data is collected in China, and determinants of people's donation decision may vary across cultures. Given the availability of public data from Twitter and Facebook, we believe more analyses can be carried out in social media fundraising for different cultural backgrounds. Last, although we observe the strict donation order for each project, we didn't record the exact microblogs posted by the system or by individual users to engage followers. Since users can customize the microblogs posted to their timelines, we believe that leveraging the content of those microblogs will allow a deeper examination of peer influence.

2.7.3 *Concluding Remark*

Given the prospects of social media fundraising, it is important to understand how users' pro-social behavior is driven by reputation, peer effects, and popularity effects. Using appropriate incentive designs, nonprofit organizations, social media platforms, and crowdfunding platforms can improve fundraising performance and alleviate the free-rider problem. On the other hand, despite the virtues of social media fundraising, the positive peer effects and popularity effects may lead to inefficiency of funds allocation. The funds to charities are likely to concentrate on a handful of attention-getting projects, while leaving other projects under-contributed. Our study finds that there is a segment of donors who are more interested in supporting less popular projects. Their existence provides potential opportunities to alleviate the rich-get-richer problem. We hope that our work provides a starting point for future research in social media fundraising, which is of considerable value to social planners, crowdfunding marketers, and nonprofits.

Chapter 3. SOCIAL MEDIA FUNDRAISING – CHARITABLE CONTENT GENERATION

3.1 INTRODUCTION

Over the last few years, social media have emerged as new channels for charitable fundraising. Compared with a 3% growth rate for overall donations, the growth rate for social donations is at an astonishing 70% in 2014 (Stein 2015). Social donations are raised by individuals from their networks, family, and friends on behalf of charities. There are also remarkable campaigns demonstrating the power of social media fundraising. One example is the "Ice Bucket Challenge". The Ice Bucket Challenge is a social media campaign that successfully raised over

$15 million in three weeks (Townsend 2014). Participants film themselves dumping icy water over their heads, and upload the video to social media. In the post, they tag and challenge their friends to do the same, and/or donate to research on ALS (Amyotrophic Lateral Sclerosis). Altogether, 4,483,726 mentions were generated on the topic of the Ice Bucket Challenge. Practitioners attribute the success to its capability of stimulating user-generated charitable content (Wu 2015).

User-generated content is critical to the success of social media fundraising, as well as innovation adoption, social commerce, and information diffusion (Bapna et al. 2011). For one thing, user-generated content prompts online word of mouth, and increases the awareness of the information to be diffused (Duan et al. 2008, Stein 2015). In addition, user-generated content, especially content generated by noncommercial users, disseminates information in local cliques where tie strengths are usually strong, effectively increasing the adoption rate (Aral and Walker 2011, Bapna et al. 2011, Shriver et al. 2013). In terms of charitable giving, evidence has been found that soliciting donations from friends and family improves fundraising performance significantly (Bekkers 2004, Meer 2011). It was reported from studies on the Haiti earthquake that about half the donors tried to solicit donations from their family and friends, with a high conversion rate of over three quarters (Smith 2012).

Among the most important stimuli for users to generate charitable content are the broadcasts they receive from the social media broadcasting systems. The essence of social networking sites is a broadcasting system embedded in a social network (Shi et al. 2014). Users (*followers*) subscribe to the broadcasts of other users (*followees*), and receive all content generated by their followees. For example, users' followees post messages to announce their acts of giving, share their favorite nonprofit organizations, or just remind people how important it is to give back to the community. After consuming such information, users may choose to post similar messages and bolster the momentum of giving. While social networking sites like Facebook manifest an undirected graph where users are usually mutual friends, other sites like Twitter manifest a directed graph where node A can usually follow node B's broadcast without B's permission. The directed graph is especially efficient in the diffusion of information because a directed link often indicates a weak tie, which is powerful in bridging cliques in the networks (Granovetter 1973, Kwak et al. 2010, Wu et al. 2011).

Users receive charitable content from their reciprocal followees and non-reciprocal followees. Their decisions whether to follow up and generate similar charitable content are likely related to the *tie strength* between these focal users and their followees. Tie strengths are usually considered strong in a reciprocal relationship where both nodes follow each other, and considered weak in a non-reciprocal relationship where the following is asymmetric (Granovetter 1973). Strong tie strengths usually positively affect adoptions because of trust and social pressure (Bakshy et al. 2011, Lee et al. 2015). Weak tie strengths usually have less social influence because users face less social pressure in the absence of the subscription from their followees (Einolf et al. 2013). However, weak ties are powerful in information diffusion because novel information is more likely to be supplied by weak ties (Shi et al. 2014). In our context of charitable fundraising, the strong ties exist between users and their family, friends, and people that share similar interests. The weak ties usually exist between individual users and celebrities, organizations and news media. Other than the influence from followees, users' group size of followers also affects their content generation (Shriver et al. 2013, Toubia and Stephen 2013). This is because people are more likely to behave pro-socially when they are aware of the potential audience that can observe their acts (Bekkers and Wiepking 2010).

Our paper aims to answer the following questions: 1) How does reciprocity affect social influence in generating charitable content?[2] 2) Does the follower group size moderate the social influence? Understanding answers to these questions provides distinctive contributions to the literature. This is because charitable content generation is a unique provision of public goods that directly transforms word of mouth into benefits to the needy, and this has not been studied before. The answers to these questions are also fundamental in coordinating the fundraising efforts between organizations and individuals.

More importantly, answers to our research questions provide an assessment of the level of information overload.[3] One key difference between social media fundraising and traditional fundraising is the low cost of posting messages. Given the low cost of content generation, fundraisers and individuals post a myriad of tweets to engage their followers. The excessive posts likely lead to information overload or compassion fatigue because users usually follow hundreds and thousands of broadcasts (Hiltz and Turoff 1985). The general existence of information overload has been recognized by Twitter, and Twitter has started to show selected tweets to users in the hope of keeping them coming back (Isaac 2016). In addition, overexposure of charitable campaigns may lead to "crowding out", when users feel less need to engage because the charitable campaigns have attracted sufficient attention and potential donations (Andreoni 1989, Andreoni 1990).

To address the above research questions, we conducted analyses using a novel dataset generated by a new global day of giving, #GivingTuesday. The hashtag #GivingTuesday is an annual charitable campaign that takes place on the first Tuesday after Thanksgiving in the United

---

[2] In social media environment, information diffuses through user-generated content. We use these two terms interchangeably to refer to user content generation following charitable topics in this manuscript.
[3] We cannot distinguish the effect of information overload from compassion fatigue or crowding out in this work, and will mainly use information overload to denote the effects from all these factors.

States. On this day, nonprofit organizations solicit donations on social networking sites like Twitter, Facebook, and Instagram. People share their donations, and engage their friends to give back. The influence of #GivingTuesday has been recognized by President Obama, "#GivingTuesday is a grassroots movement that illustrates the impact we can have when we all pull together." By checking whether users included the hashtag #GivingTuesday in their tweets, we can verify their participation. We further extract users' network topological structures on Twitter to understand the roles of social influence and reciprocity.

Evaluating the effect of social influence in engaging users is complicated due to simultaneity and homophily which cause spurious correlation between followees' charitable content generation and followers' engagement (Aral et al. 2009, Manski 2000). Prior research has employed causal instrumental variable approaches (Shriver et al. 2013, Tucker 2008) and experimental approaches (Aral and Walker 2014, Bapna and Umyarov 2015) to handle the endogeneity issue. We, on the other hand, used a Latent Instrumental Variable (LIV) approach to address this problem. The LIV approach decomposes the endogenous regressor into two components: one component is the exogenous latent categorical variable, and the other component is an error term correlated with the error term that directly contributes to the outcome. The classic LIV approach has a continuous outcome (Ebbes et al. 2005), and we revised the model to derive an estimator for a binary outcome (Rutz et al. 2012). This paper is one of the first applications of LIV in the information systems literature.

Our research discovers findings that are both interesting and surprising. First, we found that social influence from reciprocal relationships is *positive*, and social influence from non-reciprocal relationships is *negative* in the engagement of charitable content generation. Second, we found that social influence is moderated by the group size (followers). When users have more followers, they are more subject to the influence from their followees' engagement, regardless of the reciprocity types.

Given the positive moderating role of users' group size, we conducted a field experiment to understand whether the moderating effect sources from users' altruism or from reputation-based motivations. Our findings indicate that users' motivations for charitable content generation is mainly a result of their altruistic motivations. We also have a side finding that users who included a picture in their previous charitable content are more likely to have reputation-based motivations for charitable content generation.

Our work makes substantial contributions to the literature of information systems, philanthropy, and media communication. To begin with, we are the first to empirically examine the impact of social influence in charitable content generation. The dedication to this topic is important because information diffusion on different topics exhibits very different patterns (Wu et al. 2011). Second, by employing the unique graph information on Twitter, we divide followees into reciprocal followees and non-reciprocal followees. This allows us to examine their different impacts. Third, the counter-intuitive result of the negative impact from non-reciprocal followees' engagement has salient theoretical value because it supports the argument of Granovetter (1973) by teasing out the novelty of content. It also has substantial practical value in the coordination between fundraising resources. Fourth, we provide an application of the LIV approach to account for endogeneity, which can be applied to many problems in information systems. Last but not least, we employ a randomized field experiment to examine the factor of group size, and reveal the underlying motivations for peoples' charitable content generation.

This manuscript is organized as follows. In Section 2 we introduce the research context. We describe the data in Section 3, and present our empirical analysis in Section 4. We illustrate our field experiment in Section 5, and report the robustness checks in Section 6. In Section 7, we conclude our work.

3.2 BACKGROUND

3.2.1 *Twitter*

Twitter is a microblogging site that allows users to generate and broadcast content. On Twitter, users create text messages ("tweets") that are up to 140 characters long. Users (followers) can follow other users (followees) to subscribe to their feeds of tweets. When users log on to Twitter, their landing page is a list of tweets posted by their followees. These tweets are sorted in reverse chronological order, and are contained in a web holder termed a "timeline". Users can see their number of followees and followers in their profile status bars. The number of followees indicates the out-degree, and the number of followers indicates the in-degree. Twitter manifests a directed social network, where the out-degree is different from the in-degree. Statistics show that only 22.1% of user pairs have a reciprocal relationship in Twitter, while this number is 95% in Facebook. Twitter provides a two-way communication system where the audience can "retweet" or "favorite" the tweets posted by others as feedbacks.

Users usually include hashtags in the content to mark their tweets topically so that others can find this tweet based on topic search. The use of the hashtag has become a convention to associate twitter content with events or context. Specifically, a keyword is prefixed by a # symbol and included in the text. It is the basic function for Twitter to organize information. Users also include pictures in their tweets to better express their ideas and emotions. One of the reasons is the 140-character limit.

3.2.2 *#GivingTuesday*

The hashtag #GivingTuesday was launched in 2012 by 92nd Street Y[4] and the United Nations Foundation. It is a movement that celebrates the generosity of giving on the Tuesday after the U.S. Thanksgiving in late November, following Black Friday and Cyber Monday. It harnesses social media technology, and takes place mainly on Twitter, Facebook, and Instagram. In 2014, #GivingTuesday was trended for eleven hours on Twitter, and was used 754,600 times. In 2015, 1.3 million social media mentions related to #GivingTuesday were generated. According to the biggest donation processor provider Blackbaud (MacLaughlin 2015), #GivingTuesday received donations of $10.1 million in 2012 (11/27/2012), $19.2 million in 2013 (12/3/2013), $26.1 million in 2014 (12/2/2014), and $39.6 million in 2015 (12/1/2015). Overall, this movement has engaged more than 10,000 organizations.

Individual users participate in this movement by posting related content, ususally including the hashtag #GivingTuesday. Typical engaged tweets are: "It's #GivingTuesday, which nonprofit organizations do you support?", or "I just donated to @americanredcross, please join me to support their work #GivingTuesday!" Users are also encouraged to include pictures in their posts to promote the event. The pictures are mostly selfies of donors in scenes of giving.

3.3 DATA

3.3.1 *Data Collection*

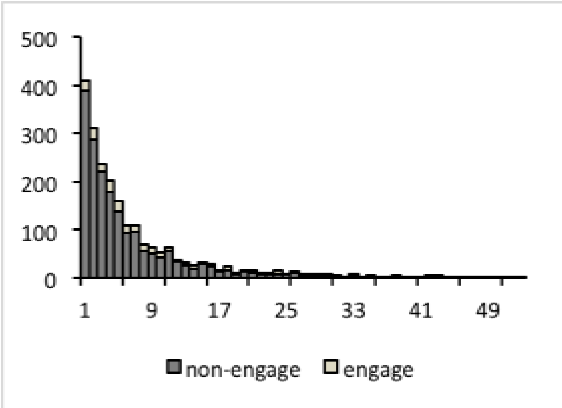
We collected data from Twitter through Twitter's application programming interface (API). Our goal is to select users who participated in #GivingTuesday 2014, and examine their engagement in #GivingTuesday 2015. We started by crawling 75,668 tweets that contain the hashtag

[4] 92nd Street Y is a cultural center in New York City that brings people together toward the value of giving back.

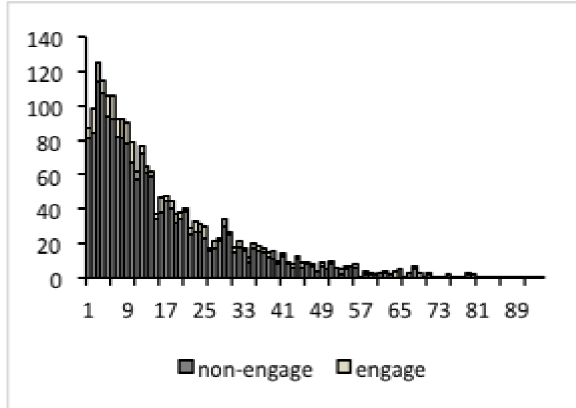
#GivingTuesday and were posted on 12/2/2014, when GivingTuesday 2014 took place. These tweets were posted by 31,593 users. They are a random sample of all the participating tweets on 2014. Twitter accounts can be classified into commercial accounts and noncommercial accounts; commercial accounts usually have monetary incentives in posting tweets (Toubia and Stephen 2013). In our study, we mainly focus on noncommercial accounts (or private citizens) because these users represent the grassroots power that is much harder to engage in traditional fundraising. We employed three research assistants to remove organization Twitter accounts or individuals who work in nonprofits. We did not specifically remove celebrities because it is trivial to define a celebrity. We only kept users who posted at least one tweet within one month before #GivingTuesday 2015 because other users were likely to be inactive, or had already left the platform. We excluded "private" accounts because we are unable to collect their social network relationships. This left us with 2,164 users to conduct our empirical analyses. We further collected these users' social network information, and found their 1,040,463 followees from 1,812,715 pairs of links. Lastly, for the 2,164 users and their followees, we tracked their charitable content generation on 12/1/2015 (this is when GivingTuesday 2015 took place) by crawling all their tweets posted regarding #GivingTuesday 2015. We ended up with 41,433 tweets containing the keyword "GivingTuesday" or "giving tuesday". These tweets were generated by 15,739 distinct users.

3.3.2 *Model-Free Evidence*

Out of the total sample of 2,164 users who contributed charitable content for #GivingTuesday 2014, 260 of them also engaged in #GivingTuesday 2015. We plot the distribution of participated followees with respect to focal users' decisions in Figure 3.1.



a. Number of Participated Reciprocal Followees



b. Number of Participated Non-reciprocal Followees

Figure 3.1. Engaged Followees Distribution by Engagement

The distribution of reciprocal followees in Figure 3.1 (a) shows the pattern that the proportion of engaged users increases while the number of reciprocal followees who participated in #GivingTuesday 2015 increases. On the contrary, the proportion of engaged users does not increase at a higher level of participated non-reciprocal followees, as shown in Figure 3.1 (b). It is notable that altogether, only 46 users did not receive any charitable engagement information from their followees; the rest of the users were all exposed to information about this movement.

In Figure 3.2, we plot the aggregate dyadic relationships with respect to users' engagement in #GivingTuesday 2015. It can be seen that the 260 engaged users have 2,681 undirected links and 5,424 directed links from #GivingTuesday participants. The ratio between directed and undirected links is around 2. For the non-engaged users, they have 30,320 directed links and 10,339 undirected links from #GivingTuesday participants. The ratio between directed and undirected links is about 3. The significant difference between followee compositions for engaged and non-engaged users indicates the salient role played by reciprocity. In our main analysis, we will formally examine reciprocity after controlling for other factors.

We can also calculate the total links for the engaged and non-engaged users by adding the undirected links and the directed links. Links for engaged users total 8,105; the average number of links for engaged users is approximately 31. Links for non-engaged users total 40,659, and the average number of links for non-engaged users is approximately 21. The difference in the number of links implies the positive role of social influence. However, this simple model-free evidence does not take into consideration the potential endogeneity issue, which will be controlled in our main analysis.

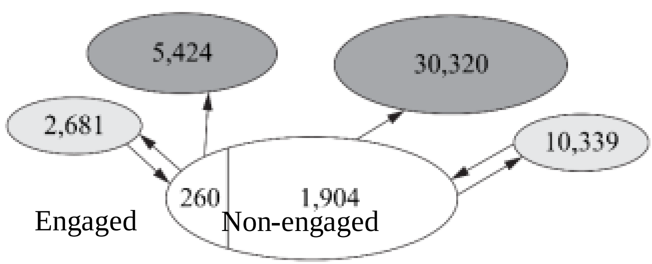


Figure 3.2. Dyadic Relationship by Engagement

3.3.3 *Summary Statistics*

Our dependent variable is a binary variable indicating whether a user engaged in #GivingTuesday 2015 (*Engage*). We have two key independent variables: one is the number of reciprocal followees who participate in #GivingTuesday 2015 (*R_Followees*), the other is the number of participated non-reciprocal followees (*NR_Followees*). We control for users' tenure (*Tenure*), which is the number of days since they joined Twitter. We also control for users' activity level using the number of tweets generated two weeks before the #GivingTuesday 2015 (*Two_Weeks*). The activity level needs to be controlled because some people are by nature less communicative than others. These people are less likely to generate charitable content, not from any failure to respond to the influence from followees, but from their own reticence. In a different specification we used the number of tweets generated ten days before we impose treatment (*Ten_Days*). We also calculated the number of tweets generated one month before #GivingTuesday (*One_Month*), and used it to select active users. We included a dummy variable (*Picture*) to indicate whether the user included a picture in her 2014 engagement. While the psychological reasons for selfies are not the focus of this work, we do notice that the inclusion of pictures likely reflects users' unobservable attributes, such as their needs for social acclaim. We further include users' gender (*Gender*). This is coded by our research assistants by looking at users' name and profile pictures. Another important user attribute we controlled for is users' number of followers (*Followers*). Since the number of followers measures the size of audience, it is likely to affect users' choice of charitable engagement.

We summarize the variables we used in Table 3.12. Within one month of #GivingTuesday 2015, users posted an average of 31 tweets. From their charitable content in 2014, we found that 16% of users included pictures in their posts. In the #GivingTuesday movement, uploading a picture to announce a donation is strongly encouraged. With a manual check, we found that most of these pictures are selfies, usually with a donor holding a card explaining how and why she gave. We further found that 36% of users are male. This shows that women are the dominant force of charitable content generation in this movement. On average, a user has 17 engaged non-reciprocal followees and 6 reciprocal followees that participated in this movement in 2015. Looking at the number of followers, we included users with extremely high numbers of followers. They are likely to be celebrities. We did not remove them from our main empirical analysis because we find it trivial to determine the threshold for celebrities. In our third robustness check, we kept only users with fewer than 1,500 followers to control for potential celebrity effects. Given the skewed distribution of the number of followers, the value of the mean is not a good representation of the distribution. The median for the number of followers in our data is 627.

Table 3.12. Summary Statistics

Variable	**No. Obs.**	**Mean**	**St. Dev.**	**Min**	**Max**
One_Month	2,164	30.84	35.95	1	552
Two_Weeks	2,164	14.50	24.64	0	534
Ten_Day	2,164	8.96	16.65	0	390
$Picture$	2,164	0.16	0.37	0	1
$Male$	2,164	0.36	0.48	0	1
$Followers$	2,164	84,411.32	1,546,271	9	54,258,936
$NR\_Followees$	2,164	16.52	17.40	0	163
$R\_Followees$	2,164	6.10	8.40	0	64
$Tenure$	2,164	1,911.73	639.37	365	3,532

3.4 EMPIRICAL ANALYSIS

3.4.1 *Theoretical Foundations*

We discuss the theoretical foundations behind our variables, and how our work extends existing theories.

3.4.2 *Reciprocal Followees*

Reciprocal ties in networks are usually associated with strong and steady relationships in terms of “time spending together, emotional intensity, intimacy, and reciprocal services that characterize the tie” (Granovetter 1973). From a dynamic point, reciprocal ties are stable. On the one hand, non-reciprocal relationships are likely to become reciprocal relationships because people like others who like them. On the other hand, non-reciprocal relationships are more likely to dissolve to avoid embarrassment or distress in traditional networks (Rivera et al. 2010). In our broadcasting context, people can subscribe to the broadcasts of nonprofit organizations and news media. Such a non-reciprocal relationship is not correlated with discomfiture, and should also be stable. Compared to such non-reciprocal relationships, people place more trust in information disseminated by their reciprocal followees (Aral and Walker 2011, Coleman 1988, Leider et al. 2009). In addition, people feel more pressure to conform to the actions of their reciprocal followees (Emerson 1962). As a result, individuals may be more responsive to the charitable content generated by their reciprocal followees.

We use the number of reciprocal followees who participated in #GivingTuesday as a measurement for the influence from strong ties. This variable is likely to be endogenous because compassionate users may follow each other, and their participation in #GivingTuesday can be caused by their shared interest to philanthropy instead of social influence. However, we only find the number of participated reciprocal followees marginally endogenous in our data, as will be detailed in estimation results. Shortly speaking, the reasons are twofold. First, we selected only noncommercial users who are individuals that do not seem to work in nonprofit organizations. Second, #GivingTuesday is an annual event which is a small part of Twitter activities, and users generally follow each other due to other reasons like their interpersonal relationships.

3.4.2.1 Non-reciprocal Followees

Davis (1970) considers non-reciprocal relationship (asymmetric relationship) to exist only among different hierarchies. Within each hierarchy, people either have mutual positive relationships or mutual nonpositive relationships (Davis 1970)**Error! Reference source not found.**. Following this idea, users' non-reciprocal followees generate charitable content and disseminate information between hierarchies. In Granovertter (1973), non-reciprocal relationship was regarded as a powerful bridge to connect network cliques and diffuse information. Shi et al. (2014) extended the theory to social exchange, and proposed that the novelty of information more likely drives non-reciprocal subscribers to share content than reciprocal subscribers in social broadcasting networks. One important distinction of our context is that we have controlled for the novelty of information. Since all related tweets are under the theme of #GivingTuesday, we can examine the intrinsic role of reciprocity.

We use the number of non-reciprocal followees who participated in #GivingTuesday to examine the role of weak ties in inducing users to generate charitable content. This variable is likely to be endogenous because users' subscription to nonprofit organizations or news media is likely to correlate with their propensity to participate in #GivingTuesday. We will discuss in detail how we handle the endogeneity issues in our model specifications.

3.4.2.2 Followers

The broadcasting system also brings a new measure that is potentially related to users' charitable content generation: the number of followers (Shriver et al. 2013). The number of followers is likely to correlate with users' content generation in two ways. First, individuals with more followers are likely to be more altruistic. These users are more likely to engage in charitable content generation because they want to help the victims. They have a large base of followers because they are helpful to others (Bekkers and Wiepking 2010). Second, individuals with more followers may have higher reputation-based incentives. Their high counts of followers are results of their continuous pursuit of followers. One important way to recruit followers is to post tweets. Generating charitable tweets can well serve the purposes of attracting followers because charitable engagement is considered pro-social behavior, and regarded highly (Shriver et al. 2013, Toubia and Stephen 2013).

We first investigated the role of followers as a potential moderator of the social influence from followees' engagement. We then designed a field experiment to disentangle the underlying mechanism behind the correlation between the number of followers and users' charitable content generation.

3.4.3 *Model Specification*

We devised three models to examine the role of social influence and reciprocity in affecting users' participation in #GivingTuesday: In Model (1), we examine the role of reciprocal followees ($R\_Followees$); In Model (2), we examine the role of non-reciprocal followees ($NR\_Followees$); In Model (3), we examine both measures of social influence ($R\_Followees$ and $NR\_Followees$).

Since the biggest challenge in our context is the endogeneity issue with participated followees ($R\_Followees$ and $NR\_Followees$), we apply a latent instrumental variable (LIV) approach to handle this problem. This approach was developed by Ebbes et al. (2005), and applied in marketing (Rutz et al. 2012). The original model was a linear regression setting where the outcome variable is continuous. We revised this model to suit our binary outcome (Newey 1987). The first two models are constructed based on an LIV model that handles one endogenous regressor, and the appropriate specification of the third model depends on the exogeneity conclusions from the estimation results of the first two models. If both $R\_Followees$ and $NR\_Followees$ are endogenous, we will need an LIV model that handles two endogenous regressors. It will turn out that only $NR\_Followees$ is significantly endogenous, and we just need the LIV model with one endogenous regressor for Model (3). $R\_Followees$ is only marginally endogenous, and we can treat it as exogenous. We did not use an observed instrumental variable because a causal instrumental variable is difficult to find (Shriver et al. 2013). Moreover, instrumental variables are untestable and weak instruments may even lead to more biased results than when no instrumental variables are used (Stock et al. 2012).

The observable outcome $Engage_i$ is a binary variable indicating whether user *i* created tweets including the keyword "giving tuesday" or "GivingTuesday".

$$Engage_i = \begin{cases} 1, & \text{if } Engage_i^* \geq 0; \\ 0, & \text{otherwise.} \end{cases} \tag{3.6}$$

A latent continuous variable $Engage^*$ determines whether individuals make the decision of participating:

$$Engage_i^* = \gamma R\_Followees_i + \mathbf{\beta x}_i + u_i \, ; \tag{3.7}$$

$$\mathbf{\beta x}_i = \beta_0 + \beta_1 Tenure_i + \beta_2 Two\_Weeks_i + \beta_3 Picture_i + \beta_4 Gender_i + \beta_5 Followers_i \, ; \tag{3.8}$$

$$R\_Followees_i = \sum_j \pi_j \delta_{ij} + v_i \, . \tag{3.9}$$

The above specification is Model (1), where the endogenous regressor is participated followees ($R\_Followees$). We can change $R\_Followees$ to $NR\_Followee_i$, and it becomes the specification for Model (2). By including $R\_Followees$ as an element of $\mathbf{x}_i$ on top of Model (2), we define the specification of Model (3).

The latent variable $Engage^*$ is determined by a vector of exogenous covariates $\mathbf{x}_i$, an endogenous variable $R\_Followees$, and an error term u_i. Following Ebbes et al. (2005), the endogenous variable $R\_Followees$ can be expressed as in equation , where π_j, $j = 1, \cdots, M$ is the categorical mean of category *j* and δ_{ij} is an indicator function which is 1 if *i* belongs to category *j* and 0 otherwise. The term $\sum_j \pi_j \delta_{ij}$ serves as the unobservable discrete instrument we construct. The probability of any *i* belonging to a category is λ_j, where $j = 1, \cdots, M$. It has been proved that the choice of *M* doesn't affect the unbiasedness of estimation (Ebbes et al. 2005).

From the above equations, we can see that the endogeneity issue arises if the error term of equation 3.7 is correlated with the error term of equation 3.9. This leads to the correlation between $R\_Followees_i$ and u_i. By explicitly modeling the data generating process, we control for this endogeneity problem. We assume that the error terms (u_i, v_i) follow bivariate Normal with mean zero and variance matrix $\begin{pmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{pmatrix}$. In the appendix, we derive the log likelihood function as below:

$$LL = \sum_i \ln\left(\sum_{j=1}^{M} \lambda_j \cdot \left(\begin{aligned} &(1 - Engage_i) + (2 \cdot Engage_i - 1) \\ &\times \Phi\left(\frac{\gamma R\_F_i + \boldsymbol{\beta}\mathbf{x}_i + \rho(\sigma_u / \sigma_v)(R\_F_i - \pi_j)}{\sqrt{1-\rho^2}\,\sigma_u} \right) \end{aligned} \right) \cdot \phi\left(\frac{R\_F_i - \pi_j}{\sigma_v} \right) \right) - \sum_i \ln \sigma_v, \quad (3.10)$$

where $\Phi(\cdot)$ is the cumulative density function of a standard normal distribution, $\phi(\cdot)$ is the probability density function of a standard normal distribution, and R_F_i is an abbreviation for $R\_Followees_i$. We derive a maximum likelihood estimator with the above likelihood function, and demonstrated in our simulated data that this approach can recover the coefficient for the endogenous variable unbiasedly. In our estimation, we cannot identify both σ_u and σ_v. Therefore, we estimate the ratio of them as $\sigma = \sigma_u / \sigma_v$.

We do not need an LIV model that handles two endogenous regressors for Model (3) because $R\_Followees$ can be treated as exogenous, but the idea to include one more endogenous variable is simple. With two endogenous regressors, we consider v_i as a vector. The variance-covariance matrix will become three-dimensional, and the likelihood for observing the outcome and two endogenous regressors will change based on the variance-covariance matrix.

3.4.4 *Results*

First we determined the number of categories of our latent instrumental variable (*M*). Following Ebbes (2005), we calculated the integrated classification likelihood (ICL) when *M* equals to 2, 3, and 4 for a model with one endogenous regressor $Participated\_Followees$ as the sum of $R\_Followees$ and $NR\_Followees$ (Biernacki et al. 2000). ICL is a modified version of BIC, and is suitable for selecting the number of components in mixture models. We report the fitness statistics in Table 3.13. As can be seen, the improvement in ICL is unnoticeable when *M* increases from 3 to 4. We choose $M = 3$ in our following analysis. Our results stay robust when we choose $M = 2$ or $M = 4$.

Table 3.13. Fitness Statistics

M	**ICL**
2	8093.970
3	8066.356
4	8065.462

The estimation results are reported in Table 3.14. We find from Model (3) that the overall effect from reciprocal followees is positively significant (0.311, $p < 0.001$), and the effect from non-reciprocal followees is negatively significant (-0.204, $p < 0.05$). The effect from participated reciprocal followees is consistently significant in the baseline specification of Model (1), where we only include reciprocal followees in the model (0.217, $p < 0.001$). From the baseline specification of Model (2), we see an insignificant overall effect from non-reciprocal followees ($NR\_Followees$). By combining the result of the interaction specification of Model (2), we see that the insignificance is a result of the negative main effect of $NR\_Followees$ (-0.321, *p* <

0.05) cancelling out with the positive crossover effect of $NR\_F \times Followers$ (0.038, $p < 0.05$). This positive interaction effect ($NR\_F \times Followers$) also shows the moderating role of $Followers$. At a higher level of $Followers$, greater influence is found from non-reciprocal followees. We also check the moderating role of $Followers$ on the influence from reciprocal followees ($R\_F \times Followers$). Again, we find a positively significant coefficient (0.06, $p < 0.01$). This indicates that at a higher level of $Followers$, the influence from participated reciprocal followees increases. In sum, the results of the three models consistently show that: 1) the number of participated reciprocal followees positively affects users' charitable content generation, 2) the number of participated non-followees negatively affects users' participation in #GivingTuesday, and 3) the number of followers positively moderates the influence from both types of participated followees.

Table 3.14. Regression Results for Empirical Analyses

	Model 1		**Model 2**		**Model 3**
	Baseline	Interaction	Baseline	Interaction	
Constant	-1.710^{***}	-1.077^{*}	-1.767^{***}	-1.187^{*}	-1.341^{**}
	(0.417)	(0.458)	(0.438)	(0.505)	(0.451)
$Tenure$	-0.018	-0.010	-0.003	-0.011	-0.011
	(0.058)	(0.058)	(0.056)	(0.056)	(0.058)
Two_Weeks	0.120^{***}	0.122^{***}	0.133^{***}	0.128^{***}	0.119^{***}
	(0.033)	(0.034)	(0.032)	(0.032)	(0.033)
$Picture$	-0.005	-0.011	-0.016	-0.019	-0.001
	(0.099)	(0.100)	(0.097)	(0.097)	(0.098)
$Female$	-0.073	-0.085	-0.075	-0.078	-0.061
	(0.075)	(0.075)	(0.074)	(0.074)	(0.076)
$Followers$	0.012	-0.094^{*}	0.072^{***}	0.002	0.004
	(0.022)	(0.041)	(0.019)	(0.038)	(0.023)
$R\_Followees$	0.217^{***}	-0.200			$\mathbf{0.311}^{***}$
	(0.059)	(0.142)			(0.051)
$R\_F \times Followers$		$\mathbf{0.060}^{**}$			
		(0.019)			
$NR\_Followees$			-0.050	-0.321^{*}	$\mathbf{-0.204}^{*}$

			(0.074)	(0.141)	(0.080)
$NR\_F \times Followers$				**0.038***	
				(0.017)	
AIC	7545.792	7535.553	7718.827	7713.962	7275.142
BIC	7630.988	7620.748	7804.023	7799.157	7359.831

Note: Standard errors in parentheses, + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We report the LIV coefficients in Table 3.15. The most important information is the significance of ρ. In Model (1), ρ is marginally significant at the p-value of 0.127. This means the number of participated reciprocal-followees shows evidence of endogeneity, but the extent of endogeneity doesn't affect our estimation statistically. Therefore, we can treat $R\_Followees$ as exogenous. The marginal significance in the endogeneity of $R\_Followees$ is different from the significant endogeneity found in previous literature (Bapna et al. 2011). This might be due to our special context where homophily is less likely to happen because #GivingTuesday is an annual event, and only represents a very small portion of people's activities on Twitter. In Model (2), ρ is positively significant (0.298, $p < 0.05$), showing that we need to handle the endogeneity issue to obtain unbiased estimations. Moreover, from the positive sign of ρ, we learn that the unobservable factors affecting people's charitable engagement is positively correlated with the unobservable factors affecting the engagement of people's non-reciprocal followees. These unobservable factors could be people's level of altruism. Since $R\_Followees$ does not show a statistically significant endogeneity, we only handle the endogeneity for $NR\_Followees$ in Model (3), where both $R\_Followees$ and $NR\_Followees$ are included.

The categorical means for the latent instrument are reported in π_1, π_2, and π_3. We can see that the point estimates of categorical means within each model have a wide spread. We also see a significant estimation for σ that is less than 1. This indicates that the variation of the error term in equation 3.7 is lower than the variation of the error term in equation 3.9.

Table 3.15. Estimation Results for LIV Parameters

		Model 1	Model 2	Model 3
LIV categorical means	π_1	1.606***	3.199***	3.206***
		(0.039)	(0.046)	(0.046)
	π_2	2.834***	1.979***	1.987***
		(0.041)	(0.065)	(0.066)
	π_3	0.378***	0.660***	0.670***
		(0.030)	(0.084)	(0.085)
LIV categorical probabilities	λ_1	0.444***	0.467***	0.762***
		(0.0211)	(0.0562)	(0.0352)
	λ_2	0.210***	0.421***	0.0498***
		(0.0190)	(0.0534)	(0.0107)
	λ_3	0.347***	0.111***	0.188***
		(0.014)	(0.0126)	(0.0274)
Covariance Matrix	ρ	0.134	0.298*	0.260**
		(0.087)	(0.113)	(0.122)
	σ	0.472***	0.589***	0.587***
		(0.0125)	(0.0215)	(0.0214)

Note: Standard errors in parentheses, + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.4.5 *Discussion*

The findings are very insightful. On the one hand, we confirm the theories proposed by Granovetter (1973) in a distinct way. We find that after controlling for the novelty of content, the positive role played by weak ties in information diffusion has disappeared. More strikingly, the effect of non-reciprocal followees' engagement even becomes negative, as shown in the Model 3 of Table 3.14. We speculate that this negative effect may source from compassion fatigue, crowding-out, or information overload. Whichever reason it is, our results suggest that the current allocation of fundraising resources in the charitable movement #GivingTuesday is not efficient. Specifically, it seems that too much involvement from nonprofit organizations and news media inhibits the information diffusion by noncommercial individual users.

On the other hand, the social influence of reciprocal followees has always been positive. This is in accordance with past literature (Bapna et al. 2011, Meer 2011), and confirms the important role of grassroots power in fundraising. By encouraging donors to share their donations or stories related to the acts of giving, they can mobilize their friends and family, and reinforce the norm of giving in their peer networks.

Overall, our findings suggest that an optimal allocation of fundraising resources is not achieved. The negative effect of non-reciprocal followees' engagement suggests information overload. To improve fundraising performance, the tweets generated by news media or nonprofit organizations (usually indicating non-reciprocal relationships) should be reduced in number to better stimulate individual charitable content generation. Opportunities to collaborate with broadcasting systems like Twitter may alleviate this problem at the root. For example, nonprofit organizations can utilize Twitter's recommendation system which pushes tweets that users might be interested in to their emails. By pushing charitable tweets generated by their reciprocal followees, Twitter can help engage users to generate charitable content.

3.5 FIELD EXPERIMENT

The goal of this experiment is to examine the underlying mechanism behind the moderating role of $Followers$. Two explanations can justify the moderating roles we find: 1) people with more followers are more altruistic, and 2) people with more followers are in a higher pursuit of reputation. If the first explanation is true, people's charitable content generation patterns do not change after we add synthetic followers to them. On the other hand, if the second explanation is true, people's charitable content generation patterns will alter after they obtain more followers (Toubia and Stephen 2013).

Understanding the true motivation is important because it helps with fundraising strategy designs. When the altruistic motivation dominates, nonprofit organizations should show the needs of the victims to solicit donations. When the reputation-based motivation dominates, nonprofit organizations should make donors' contributions more visible to enhance the reputation effect.

3.5.1 *Experimental Design*

We manipulated users' followers by adding synthetic users to their accounts in the last four days before #GivingTuesday 2015. This design was first used by Toubia and Stephen (2013) in examining motivations for generating tweets. Coppock et al. (2016) also conducted a field experiment asking people to support an online petition by sending them private messages. A major difference between Toubia and Stephen (2013) and our study is that they intended to examine general motivations, while we are interested in users' charitable content generation.

We selected an additional 100 users from the population as our treatment group, and imposed treatment on these users. In our study we created 16 synthetic users, with half looking compassionate. The reason to make the synthetic users look compassionate is to mimic the follower acquisition associated with charitable content generation. We wrote a program to generate random tweets that looked compassionate. Typical tweets are like "GivingTuesday is coming, are you ready?" We generated the names of these synthetic users using www.fakenamegenerator.com. We uploaded profile pictures for each user, and let them follow 20-30 other users. To make these users look more realistic, we added followers to their accounts. On average, we added 200 followers to each of these synthetic followers. In the last four days before #GivingTuesday 2015, we added followers to each treated user. Every day we added four followers. One follower is added in the morning, one in the afternoon, one is in the evening, and one at night. Each day, a combination of two ordinary users and two compassionate users were added. At the end of the four-day treatment period, we did a manipulation check, and found that 91 users received the same level of treatment. We left out the 9 users that did not pass the manipulation check due to operational problems.

The major variable of our interest is $Treatment$. This is a binary variable indicating whether we added synthetic followers to the subjects. While we used the number of tweets generated two weeks before #GivingTuesday 2015 to control for user activity level in the previous empirical analyses, we used the number of tweets generated ten days before our four-day treatment (Ten_Days) in this experimental study. This avoids potential confoundedness of our treatment because users' posting behavior may be affected by the increase in their followers in the last four days before the campaign. It is worth mentioning that results stay the same if we use Two_Weeks to control for user activity level.

We examine how the treatment affects the influence from users' numbers of followers ($Followers$), and the interactions with followers ($R\_Followees \times Followers$ and $NR\_Followees \times Followers$). A simple probit model is used without controlling for endogeneity because we employ exogenous shocks as identifying sources. In addition, we examine the heterogeneous treatment effects with respect to the level of activity (Ten_Days) and whether a picture was included in users' 2014 charitable content ($Picture$). This is done by including the interactions between these factors and the treatment. The significance of these interaction terms will have implications on whether these factors serve as proxies of users' reputation-based incentives.

3.5.2 *Propensity Score Matching*

Ideally, we can use all the data from the previous empirical study as our control group. However, some of the key variables, such the number of donated followees, are not known until the day of the event. Therefore, it is hard to guarantee that all characteristics of the treatment group and the control group are not statistically different. We use the Propensity Score Matching technique to select a subset of users from the 2,164 users in our previous study as control, and compare them with the treatment group. This approach selects similar users into the control group as the treatment group. The validation check results when the matching ratio is 1 are listed in Table 3.16. The matching ratio is the ratio between the number of observations in the control group and that of the treatment group. We use different ratios in our study. Given a treatment group of 91 users, the total observation is 182 if the ratio is 1, 546 if the ratio is 5, and 1001 if the ratio is 10. While we follow the convention to guarantee the pass of the validation check, such validation checks between control group and treatment group are usually not necessary (Mutz and Pemantle 2011).

Table 3.16. Validation Check

	Treatment Mean	**Control Mean**	**T-test**
Tenure	1,869.96	1,896.79	0.7829
	(625.22)	(685.40)	
Ten Day	22.92	18.65	0.447
	(32.79)	(42.28)	
Picture	0.26	0.29	0.7415
	(0.44)	(0.45)	
Male	0.27	0.29	0.8698
	(0.45)	(0.45)	
Followers	781.90	889.33	0.1693
	(541.62)	(507.99)	
$P\_Followees$	4.97	7.08	0.1002
	(5.73)	(10.73)	
$NP\_Followees$	20.66	22.03	0.6943

	(19.35)	(27.07)

Note: we report the original scale of these variables, and use the log transform in our estimation for skewed continuous variables.

3.5.3 *Results*

We report our findings in Table 3.17. We evaluate four models. Model (a) includes only the treatment dummy and other control variables. Model (b) includes interaction between treatment and all factors containing followers, and is of our major interest. Model (c) and (d) are simply checking the interactions between treatment and other variables to explore potential interesting findings. The models are specified in the rows, and the columns indicate different matching ratios. It can be seen that the results are consistent across all matching ratios, and consistent even when we do not use matching techniques at all.

First, we found that the treatment effect is insignificant in all specifications. This can be seen from the insignificant coefficient of *Treatment* in Table 3.17. Second, the treatment has no impact on either the direct effect of *Followers* or the moderation roles of *Followers*. This can be seen from Model (b) of Table 3.17. Third, the interaction between *Treatment* and *Picture* is significant in a consistent manner. This can be seen from Model (c) of Table 3.17. Last, while the interaction between *Treatment* and Ten_Days is statistically insignificant (Model (d)), it is marginally significant, showing that active Twitter users are either more likely to have reputation motivations or more aware of the treatment we imposed on them.

Table 3.17. Regression Results for the Experiment

		No Matching	Matching Ratio = 1	Matching Ratio = 5	Matching Ratio = 10
Model a	*Treatment*	-0.015 (0.179)	0.085 (0.275)	0.027 (0.195)	0.041 (0.186)
Model b	*Treatment*	-0.979 (3.932)	-1.262 (8.217)	-1.256 (4.372)	-1.247 (4.152)
	$Treatment \times Followers$	0.074 (0.610)	0.273 (1.256)	0.198 (0.682)	0.176 (0.647)

	$Treatment \times Followers \times P\_F$	-0.071	0.486	0.199	0.138
		(0.369)	(0.759)	(0.392)	(0.379)
	$Treatment \times Followers \times NP\_F$	-0.081	-0.634	-0.213	-0.186
		(0.290)	(0.716)	(0.324)	(0.305)
Model c	$Treatment$	-0.253	-0.322	-0.275	-0.261
		(0.228)	(0.336)	(0.247)	(0.238)
	$Treatment \times Picture$	**0.771****	**1.473****	**1.031****	**1.028****
		(0.378)	**(0.679)**	**(0.429)**	**(0.404)**
Model d	$Treatment$	-0.630	-0.339	-0.497	-0.468
		(0.478)	(0.712)	(0.521)	(0.501)
	$Treatment \times Ten\_Days$	0.224	0.153	0.189	0.184
		(0.154)	(0.239)	(0.170)	(0.164)
	Observations	2255	182	546	1001

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

We access the power of our estimations to show the sufficiency of the sample. The minimum detectable effect size (MDES) for Model (b) of Table 3.17 (when the matching ratio is one) is 0.343 standard deviation units, given a power of 0.8 and a significance level of 0.5. This demonstrated that we have an efficient experimental design.

3.5.4 *Discussion*

The results from this field experiment showed that a sudden increase in followers right before a fundraising event does not have a significant impact on users' engagement probabilities. This indicates that altruism, instead of reputation, is more likely to be the underlying motivation of users' charitable content generation in #GivingTuesday. For nonprofit organizations, a better fundraising strategy is to emphasize the welfare of the victims. By sharing more information about why to give and how the donations can help the beneficiaries, the fundraising can be more successful.

We also learned from the field experiment that users who included pictures in their past charitable engagement are more likely to respond to an increase in followers. While we are not sure about the underlying psychological reasons for the inclusion of pictures (usually selfies) in the tweet, we suspect that this is related to people's level of need for social acclaim. That is to say, users who included pictures in their previous charitable content are more likely to have reputational incentives. To these users, a better way to promote user content generation is to announce their donations. This finding can improve user targeting and engagement.

3.6 ROBUSTNESS CHECKS

We conducted three robustness checks to validate our findings and interpretations.

3.6.1 *Split-Sample Analysis*

To assess the explanation of information overload, we split our sample into two subsamples: one subsample of users who are exposed to fewer than 17 participated followees, the other subsample of users who are exposed to more than 17 participated followees. The number of participated followees is the sum of the number of reciprocal and non-reciprocal followees. We use the sum of these two measures to better reflect the overall exposures to charitable content. We chose the cutoff value of 17 because it is the median of *Participated Followees*. If information overload exists, we should see a greater negative effect in the subsample with a higher level of participated followees, compared with the subsample with a lower level of participated followees. Below, we report our results in Table 3.18.

Table 3.18. Robustness Check – Split Sample Analysis

	$Participated\_Followees < 17$	$Participated\_Followees \geq 17$
Participated_Followees	-0.052 (0.267)	-0.961* (0.567)

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

From Table 3.18 we can see that the effect of *Participated_Followees* is insignificant for the subsample with lower participated followees and significantly negative for the other subsample. The insignificant effect of *Participated_Followees* when

$Participated\_Followees < 17$ shows that the positive influence from reciprocal followees cancels out with the negative influence from non-reciprocal followees. Similarly, the negatively significant effect of $Participated\_Followees$ in the other subsample shows that the negative influence from non-reciprocal followees outperforms the positive influence from reciprocal followees. This result confirms our explanation, and implies that information overload occurs when users have more than 17 participated followees; it does not occur when they have fewer followees that engaged in #GivingTuesday.

3.6.2 *Pre-Participation of #GivingTuesday*

While most users in our sample only engaged in #GivingTuesday on 12/1/2015, we do find a small group of users who generated charitable content before the exact day of the movement. This group of users may have a higher probability to engage on the day of the movement, and their pre-participation is confounded with the level of activity (Two_Weeks).

In this robustness check, we excluded the 376 users who generated content including #GivingTuesday before the actual day of the movement. We conducted our evaluation following the specification of Model (3) in Table 3.14 with the subsample of users who never posted a single related message before #GivingTuesday 2015, and report the results in Table 3.19. As can be seen, the major results stay robust.

Table 3.19. Robustness Check – Pre-Participation

	Estimate	**St. Error**
Constant	-1.841***	0.495
Tenure	0.043	0.064
Two_Weeks	0.103**	0.035
Picture	0.059	0.101
Female	-0.049	0.080
Followers	0.017	0.024

$R\_Followees$	**0.275*****	**0.054**
$NR\_Followees$	**-0.199****	**0.080**

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.001$

3.6.3 *Celebrity Effect*

On Twitter, certain individuals are celebrities or micro-celebrities that put continuous efforts in gaining status, visibility, and attention (Page 2012). These users may have different patterns of charitable content generation compared to the grassroots users. In our previous analyses, we did not set a cap for the number of followers. In this robustness check, we follow Shi et al. (2014), and choose only users who have fewer than 1,500 followers to run Model (3) in Table 3.14. The estimation for ρ was significant at p-value of 0.083, showing a significant endogeneity. We report the point estimates in Table 3.20. As can be seen, $R\_Followees$ has a positively significant effect (0.292, p-value<0.001), and $NR\_Followees$ has a negatively significant effect (-0.337, p-value<0.1). This is consistent with our main findings.

In addition, we find that the coefficient of $Female$ becomes significant in this subsample (-0.157, p-value<0.1). Since all the users have participated in #GivingTuesday 2014, this result indicates that females are less likely to engage repeatedly in charitable content generation. This is despite the fact that females constitute 64% of all users who participated in #GivingTuesday 2014 from our sample.

Table 3.20. Robustness Check – Celebrity Effect

	Estimate	St. Error
Constant	0.005	0.620
$Tenure$	0.040	0.062
Two_Weeks	0.164***	0.042
$Picture$	-0.147	0.124
$Female$	-0.157*	0.088
$Followers$	-0.218***	0.062

$R\_Followees$	**0.292*** **	**0.069**
$NR\_Followees$	**-0.337* **	**0.176**

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.001$

3.7 Conclusions

User content generation has been of central interest to researchers because it is a matter of life and death to online communities. Users' participation in specific topics not only diffuses information, but also brings out the wisdom of the crowd to select important news and express opinions. In addition, trended topics in Twitter is of high advertising value because advertisers can pay to launch promoted trends to advocate for their products. By analyzing the social influence from strong ties (reciprocal relationships) and weak ties (non-reciprocal relationships), we discover that a critical aspect of topic stimulation is the level of information overload. This aspect is often overlooked. We find that overexposure to information from non-reciprocal ties likely discourages individual users' content generation. This impact from content generated by non-reciprocal nodes to individual users' content generation has never been discovered before.

Our work confirms the theory proposed by Granovetter (1973) and extends it in the social media environment where the cost of generating content is lower than ever before. We support the theory proposed by Granovetter (1973) because when the novelty of content is fixed, weak ties no longer show their power in information dissemination. Moreover, in an environment where the cost of generating information is extremely low, information overload likely occurs, making weak ties obstacles for information diffusion.

Practically speaking, nonprofit organizations should start their solicitation early to avoid information overload. They should not post a large quantity of tweets, but rather focus on the quality of the tweets. They should try their best to induce their followers to share charitable information in their cliques to reach the largest audience. To better solicit donations, nonprofits should tailor their solicitations to different users. For users who have altruistic motivations, usually those with more followers, it is best to show them the needs of the victims. For those who have reputation-based motivations, for example, those who included pictures in their previous posts, it is best to announce their giving. An example of announcing donations in social media is to mention donors' usernames in a tweet. Nonprofit organizations should employ tactics to minimize information overload and to recruit users based on their motivations.

Social media campaign organizers, like the management board of #GivingTuesday, should learn that it is not always advantageous to include a large number of nonprofit organizations. While it is good to have a special day of giving to increase people's awareness, the high density of fundraising on a single day leads to information overload. One way to alleviate this problem is to restrict the broadcasts from highly influential news media that are subscribed to by lots of users. Another way is to suggest that nonprofit organizations participate every other year, or carry out their campaigns at other times of the year. These tactics redistribute the fundraising demand, and reduce the possibility of information overload.

Our work has its limitations. First of all, charitable content generation is not equal to donations. Since the ultimate goal of fundraising is monetary, we believe that more valuable research can be done by combining the charitable content generation with donation decisions. Second, we did not process the user-generated charitable content. By extracting different features of the content using natural language processing, a deeper investigation can be carried out to better guide fundraising strategies. In addition, we considered a simple measure of tie strength: reciprocity. As shown in many studies, a more comprehensive measure of tie strength can be obtained by calculating the social embeddedness and summarizing the social interactions between users. Further, our analysis is based on a specific topic #GivingTuesday, which is an annual movement that is likely to drive a high intensity of engagement. We believe that our findings can be generalized to other campaigns because network structure is an inherent feature of the Twittersphere. Last but not least, our experiment is limited due to our inability to control the behaviors of nonprofit organizations. To better shed light on the strategy of nonprofit organizations, collaboration with nonprofit organizations to conduct field experiments is a promising area for future research to consider.

Despite all the limitations, we believe that our findings are of great theoretical and practical value. Social media fundraising as a new channel for nonprofit organizations to directly communicate with potential donors may alter the way people evaluate the quality of nonprofits, and make their choices of giving. However, the domain of social media fundraising is largely under-explored, and more systematic analyses are urgently needed to understand the role of social networks in reshaping fundraising. We look forward to more works in this domain.

Chapter 4. E-COMMERCE – MEMBERSHIP FREE SHIPPING

4.1 INTRODUCTION

U.S. e-commerce sales have grown from $69 billion in 2005 to over $340 billion in 2015. In the meantime, the percentage of e-commerce sales to total retail sales rose from 2% in 2004 to 7% in 2015 (U.S. Department of Commerce 2015). While the proliferation of e-commerce has greatly lowered consumers' transaction costs, it poses challenges to sellers due to fierce competition and heavy shipping costs. According to an operational expenses report, purchased transportation, shipping, and warehousing services account for 7.7% of the total operational cost for electronic shopping, while this percentage is as low as 1.2% for other retailing sectors (U.S. Department of Commerce 2012).

Consumers consider shipping speed one of the most important elements when evaluating the quality of their online shopping experience (Collier and Bienstock 2006). E-tailers have been putting continuous effort into speeding up shipping in order to gain advantage in the ever-competitive online market. About ten years ago, consumers usually needed to wait for a week or longer to receive products they ordered online. Nowadays, two-day delivery is pervasive due to Amazon's aggressive shipping offerings. Amazon pioneered the revolution by introducing its loyalty program, Amazon Prime, that offers free two-day shipping. In 2010, several e-tailers teamed up to form ShopRunner that also provides the same perk for orders placed with its affiliated e-tailers (Fowler 2010). Recently, Google Express started same-day delivery to sharpen the competition, pushing Amazon to add same-day shipping to its Prime program. In the future, Amazon plans to launch Prime Air that cuts the shipping time to 30 minutes by means of drones, making the delivery almost instantaneous (Mangalindan 2015).

While the shipping speed is accelerating, consumers have been paying less in shipping fees over the past several years. In the early phase of e-commerce, free shipping was an occasional enticement only offered on holidays. Now, free shipping has become a necessary strategy for e-tailers to stay competitive (Stevens and Banjo 2014). It was reported that nearly 90% of e-tailers offered some form of free shipping; this number was only 65% two years before (Stevens and Banjo 2014). So, customers expect free shipping and rank shipping cost the top reason to abandon their shopping carts.

Two major forms of free shipping policies are commonly used by e-tailers: contingent free shipping (CFS) and membership free shipping (MFS). Contingent free shipping has been considered an effective shipping schedule to increase order size. Under this schedule, consumers are eligible for free shipping if the total transaction amount of an order reaches a pre-announced threshold. For example, *Target* provides free shipping for orders exceeding $50 in value (for more examples, see the top of Figure 4.1). This shipping schedule serves as a quantity discount that entices consumers to place large orders and ship less frequently (Lewis 2006). For example, a *Target* shopper, who has $45 worth of products in her shopping cart, can avoid the shipping fee by adding extra items worth $5 or more. Although this consumer doesn't have an immediate use for the extra items, she can save them for future use. On average, two-thirds of all online orders from *Target* enjoyed free shipping. It was reported that 60% of online shoppers made efforts to place orders that are large enough to qualify for contingent free shipping (UPS 2015).

The other free shipping policy, membership free shipping, has become popular among e-tailers in the last decade. MFS is commonly advertised as a premium service included in an e-tailer's loyalty program. Under this policy, consumers pay a membership fee upfront, and enjoy free shipping along with other benefits throughout the membership period. The shipping service attached to a loyalty program is usually expedited shipping, and other benefits include free return shipping, cash back, etc. Under this shipping policy, consumers have high purchasing frequency. They order as needed because shipping is free once they subscribe to the membership program. Amazon pioneered membership free shipping with its loyalty program, Prime. By paying an annual membership fee of $99, consumers enjoy free two-day shipping for selected Prime items. Amazon Prime has acquired more than 20 million members by January 2014 since its launch in 2005.

Despite the surging popularity of MFS, the high cost associated with expedited shipping may hinder e-tailers' profitability. Unlike CFS, subscribers of MFS have no incentive to place large orders; instead, they order as needed, resulting in a higher purchasing frequency and a smaller order size. As stated by Amazon's CEO Bezos, "Two-day shipping becomes an everyday experience rather than an occasional indulgence.” This poses heavy financial burdens on e-tailers. It has been estimated that Amazon lost at least $11 for each member (Tuttle 2011). In 2013 alone, Amazon collected $3.1 billion in shipping payments, while it paid $6.64 billion in shipping expenditure (Stevens and Banjo 2014). To reduce such a huge financial deficit, Amazon in March of 2014 increased its Prime membership fee from $79 to $99. It was also reported that Drugstore.com lost between $10 and $15 per order; the loss is largely due to shipping and handling costs (Barsh et al. 2000).

Competition seems to have played a vital role in affecting e-tailers’ shipping choices. Following the success of Amazon Prime, Kmart and Sears offers Shop Your Way Max loyalty program that provides similar perks for an annual membership fee of $39. Walmart.com offers Shipping Pass at $50 a year for free three-day shipping. Newegg offers Premier membership at $49.99 for one year’s two-day free shipping. More examples are presented in the bottom of Figure 4.1. Given the dynamics in the online electronic market, a competitive framework is essential for a comprehensive understanding of free shipping policies.

(Figure adapted by author from Wall Street Journal)

Figure 4.1. Examples of Contingent Free Shipping and Membership Free Shipping

We consider two e-tailers that sell a homogeneous product to a group of customers with heterogeneous demands. The leading e-tailer chooses either MFS or CFS as his free shipping strategy, and then the following e-tailer makes his choice. Consumers choose the e-tailer that maximizes their utility. A fraction of consumers demands one unit of the product (low-demand consumers), and the rest have two-unit demands (high-demand consumers) in their lifespans of two periods (Friedman 1957). In each period, these high-demand consumers consume one unit of the product. High-demand consumers choose a purchase plan from the following options: in the case of *buying as needed*, consumers purchase one unit in each period; alternatively, as in the case of *buying and stocking*, they purchase two units at the beginning of the first period, consume one unit, and stock one unit for the second period. Consumers incur a *holding cost* when they purchase and stock.

We include both horizontal and vertical differentiation in our modeling framework. By offering faster shipping, e-tailers increase consumers' valuation of the product. Products delivered by expedited shipping and regular shipping are considered *vertically differentiated* products. Motivated by real-life practices, we assume that MFS offers expedited shipping whereas CFS provides standard shipping. We also allow consumers to have *horizontally differentiated* preferences over the e-tailers. Horizontal differentiation comes from the variety of other products sold by the e-tailers, website design, customer service, and so on.

With such a model, we examine the economic impacts of the two free shipping schedules, MFS and CFS. Our research questions are: *How is MFS different from CFS? Why is expedited shipping usually attached to MFS? When should e-tailers use which shipping schedule? What is optimal for society?* We first analyze a simple scenario where all consumers have two-unit demand (homogeneous model). We then relax the homogeneity assumption by allowing a fraction of consumers to have one-unit demand (heterogeneous model).

Our results show that under symmetric shipping configurations, when both-e-tailers use the same shipping schedule, MFS generates a higher profit than CFS solely because of the expedited shipping attached to MFS. Essentially, a higher shipping speed leads to a higher level of price discrimination. Price discrimination is mainly driven by two effects: the *shipping effect* and the *membership effect*. On the one hand, high-demand buyers may avoid the per-order shipping fee because they are either eligible for a fee waiver under CFS or offered free shipping under MFS. As a result, the shipping cost burden from high-demand consumers transfers into a higher per-unit price, which is faced by both high-demand and low-demand consumers. We refer to such a shift in per-unit price as the shipping effect. On the other hand, the membership payments collected from the high-demand buyers subsidize the low-demand buyers by driving down the per-unit price. This particular effect of membership fees on the per-unit price is referred to as the membership effect.

Although it generates higher profits when both e-tailers use MFS than when they both use CFS, e-tailers may both end up using CFS which leads to a suboptimal outcome, as in the prisoner's dilemma. This would be the case when the fraction of high-demand consumers is considerably large, and the high cost of frequent expedited shipping provides incentives for the following e-tailer to deviate from cooperation. However, under conditions when consumers highly value expedited shipping compared to standard shipping, cooperation can be guaranteed.

Another important point we illuminate concerns social welfare; we find that the level of holding cost is the key in determining the most efficient shipping policy for society. Even without offering any form of free shipping, consumers have an incentive to order and stock when the holding cost is low, in order to save on shipping trips. We refer to consumers' purchasing frequency without offers of free shipping as *original purchasing frequency*. We denote the *benchmark holding cost* to be the level that makes consumers indifferent to the choice of buying as needed and buying and stocking when neither CFS nor MFS is offered. We find that if the holding cost is above the benchmark level, society is better off when both e-tailers offer MFS than when they both offer CFS. If the holding cost is below the benchmark level, society is better off when both e-tailers offer CFS than when they both offer MFS. The introduction of a free shipping policy may alter consumers' purchasing frequency. With MFS, consumers are more likely to order as needed, and with CFS, consumers have a higher incentive to order and stock. Our results suggest that any free shipping policy that *alters* consumers' original purchasing frequency will harm social welfare.

To the best of our knowledge, this work is one of the earliest theoretical investigations of the economic impacts of MFS and CFS. Our findings carry significant theoretical and practical implications. Theoretically, MFS and CFS are new forms of nonlinear pricing strategies that have received scant attention. The pricing regime of MFS is a hybrid of a two-part tariff and a two-block tariff, on top of vertical product differentiation. The pricing regime of CFS, on the other hand, can be regarded as a decreasing-block tariff, where the decrease in price is at the expense of holding cost. The difference between these two nonlinear pricing strategies has never been examined in a competitive setting when consumer demand heterogeneity is taken into account. Practically speaking, shipping policies have long been regarded as complex managerial decisions. When Amazon first launched its Prime program, no one was able to predict consumers' response, and they chose $79 as the membership fee simply because it's a prime number (Stone 2010). When Amazon hiked its membership fee from $79 to $99, practitioners debated intensely over whether this was a good decision (Gustin 2014, Matthews 2014). Our findings bring economic insights to the above issues faced by e-tailers.

The remainder of the manuscript is organized as follows. In section 2, we review related work. In section 3, we introduce our model setup. In section 4, we describe the equilibrium for our homogeneous model. We analyze the equilibrium of our heterogeneous model in section 5. In section 6, we discuss limitations of our study, and possibilities for future work. Finally, we draw conclusions in section 7.

4.2 LITERATURE REVIEW

Our work is closely related to the economic study of price discrimination under duopoly/oligopoly. Specifically, we model second-degree discrimination, where firms offer a menu of options for consumers to self-select based on their types (Katz 1984, Stole 1995). Katz (1984) assumed a general quantity-based pricing schedule where a single-unit price and an r-unit price are offered. In our study, r takes the value of two, with MFS and CFS offering different forms of discounts. MFS resembles a two-part tariff, where the membership fee is the lump-sum entry fee. The general equilibrium of two-part tariff competition was widely discussed (e.g., Brown and Heal 1980 and Yin 2004). Compared with a uniform price, two-part tariffs result in a lower per-unit price, and higher profits for the firms (Yin 2004). The distinction of our model is the inclusion of higher quality attached to the high demand option in the menu offered to consumers.

The schedule of CFS is essentially a decreasing-block tariff, where the effective price for the second unit is lower than that of the first unit. While the specific schedule of CFS received little attention from economists, it is analyzed in operations literature (Leng and Becerril-Arreola 2010, Yang et al. 2005a). Leng and Parlar (2005) used a leader-follower game to analyze the best response functions of consumers to solve for the optimal threshold value for CFS. Leng and Becerril-Arreola (2010) showed that retailers face profit loss in accommodating heterogeneous consumers. Yang et al. (2005) investigated the direction and dispersion of price change as the threshold value for CFS changes. Our work contributes to the existing literature by explicitly comparing CFS with MFS.

Loosely related to our work, marketing literature has looked at the effectiveness of free shipping as a promotion and consumers' perception of the partitioned price (per-unit price and shipping fee, respectively). Overall free shipping policy was regarded as the most effective promotion strategy to increase order size (Lewis 2006, Lewis et al. 2006). It was argued that the diversity of shipping fees increases the difficulty for consumers to compare prices across different e-tailers, and improves profit margin (Frischmann et al. 2012). In terms of consumers' perceptions of partitioned prices, different results were found. On the one hand, Morwitz et al. (1998) found that consumers tend to overlook the shipping and handling fees (Morwitz et al. 1998). Schindler et al. (2005) followed up on this study and found that consumers' skepticism of the fairness of price explains such inattention (Schindler et al. 2005). On the other hand, Koukova et al. (2012) studied consumers' perceptions under a contingent free shipping schedule, and found that shipping cost causes customers to abandon their shopping carts (Koukova et al. 2012). By providing a theoretical analysis where consumers are assumed rational in their consideration of partitioned price, our work complements these empirical investigations.

4.3 Model Setup

We develop a game-theoretical model with three players: two competing e-tailers (M_1 and M_2) and a group of consumers with size being normalized to a mass of one. We model the two e-tailers' choices of shipping strategies in a sequential game. Without loss of generality, we assume M_1 to be the market leader who chooses a shipping policy before the follower M_2. Such a sequential setting captures the shipping policy dynamics in the e-tailing market (Alesci and

Smith 2015). The timeline of the game is as follows. M_1 moves first by choosing his free shipping policy and announcing it on the website, followed by M_2's choice. Next, e-tailers who use MFS determine the membership fee. Finally, both e-tailers set the per-unit price simultaneously. Each consumer chooses an e-tailer that maximizes her utility.

4.3.1 *Demand Structure*

The following assumptions rationalize our demand system.

Assumption 1. M_1 locates at position 0, and M_2 locates at position 1 of a unit segment. Consumers are distributed uniformly along the segment [0, 1]. The two e-tailers sell a single product. Consumers' distance to each e-tailer represents their horizontally differentiated preference, in the form of preference mismatch cost.

This horizontal differentiation setup follows the Hotelling model (d'Aspremont et al. 1979, Hotelling 1929). Practically, horizontal differentiation grants e-tailers with higher market power to conduct price discrimination with free shipping schedules (Li and Dinlersoz 2012).

The horizontal preference mainly comes from the variety of other products sold by the e-tailers. While we focus on a single product sold by both e-tailers, these e-tailers also sell other products that do not necessarily overlap. For example, one consumer may like Amazon better because it offers a wider range of products; another consumer may like Office Depot better because it offers a more dedicated supply of office products. Such a horizontal differentiation between e-tailers is discussed in Abhishek et al. (2015). This horizontal preference also captures other features that differentiate these two e-tailers: website design, customer service, security, product information, inertia, and imperfect information due to search costs (Abhishek et al. 2015, Blut et al. 2015, Wu et al. 2004). Evidence of the general existence of horizontal differentiation is the prevalence of price dispersion (Pan et al. 2002).

In our model, a consumer located at position x incurs a mismatch cost of xt if she purchases from M_1, and $(1-x)t$ if she purchases from M_2, where t is the unit measurement of the mismatch cost. The magnitude of t can be interpreted as the competitiveness or degree of similarity between the two e-tailers. For example, the value of t would be larger in the competition between Amazon and Office Depot than in the competition between Amazon and Walmart. This is because Amazon and Walmart sell similar ranges of products. Since the mismatch cost is at e-tailer level instead of at transaction level, it incurs only once for each consumer (Katz 1984).

Assumption 2. A fraction δ of consumers demand two units of product (high-demand consumers) in their lifespans of two periods, and consume one unit in each period. The rest $1-\delta$ consumers demand one unit of product (low-demand consumers) in their lifespans of one period.

E-tailers only know the distribution of consumers, but not the type of each consumer. They cannot practice first-degree price discrimination, and will rely on second-degree price discrimination.

In real life, demand heterogeneity exists in general. A family with children orders wipes more frequently than a family without children. Similarly, a family that prefers home cooking orders paper towels more frequently than a family that prefers to eat out. It was reported that Amazon Prime is the top brand for housewives because they order more frequently than other consumers.

Assumption 3. Consumers homogeneously value one unit of product v_h if it is delivered via expedited shipping, and v_l if it is delivered via standard shipping, where $v_h > v_l$. The shipping cost is s_h per parcel for expedited shipping, and s_l per parcel for standard shipping. We assume in our heterogeneous model that consumers are rational in shipping options, and $v_h - v_l = s_h - s_l$.

Assumption 3 reflects the fact that expedited shipping is vertically more attractive than standard shipping (Phlips and Thisse 1982). In our study we introduce two models: 1) a homogeneous model where all consumers are of the high-demand type, and 2) a heterogeneous model where both high-demand and low-demand consumers exist. We will discuss in our homogeneous model that e-tailers' optimal shipping strategy can be trivial based on the relationship between $v_h - v_l$ and $s_h - s_l$. In order to examine the intrinsic difference brought by the novel design of MFS, we tease out the influence of the higher value provided by expedited shipping with a rational shipping assumption. This assumption restricts that the valuation difference between expedited shipping and standard shipping to equal the shipping cost difference ($v_h - v_l = s_h - s_l$). This assumption is applied only in our heterogeneous model.

A low-demand consumer's utility net of mismatch cost is $v_h - p - s_h$ if she uses expedited shipping, and $v_l - p - s_l$ if she chooses standard shipping. Under the rational shipping assumption, this low-demand consumer receives equal utility from both cases, and the e-tailer gains equal profits regardless of the shipping speed the consumer chooses. Rational shipping assumption also amounts to saying that shipping is exogenous. In economics literature, shipping charge is often assumed to be exogenous when package-delivery market is competitive (Spulber 1981). The rationale is that shipping is not a result of production from the e-tailers. Rather, it is a service purchased from package-delivery companies like UPS, Fedex, and USPS. In addition, we assume shipping to be charged at a parcel level. It is a common assumption that the shipping cost is independent of the purchased value in the e-tailing environment because the order quantity is usually small (Yang et al. 2005b, Zhou et al. 2009).

Assumption 4. When a consumer purchases two units of the product at the beginning of the first period, she suffers from a holding cost h>0 for the extra unit she purchases but does not consume until the second period.

The concept of holding cost is similar to that used in operations management. It can result from the extra space needed to stock, the discount factor of cash flow, and the effort of planning ahead (Chiang et al. 1994). Consumers incur a low effort of planning ahead when they are certain about their future consumption and satisfaction from the product. Consumers incur a high effort of planning ahead when the product quality declines over time. For example, the clothes we purchased for next year may end up being out of fashion if we have no information about the future trend. For ease of exposition, we let all consumers have the same level of holding cost, so that the only varying part is the demand.

Conceptually, the holding cost can go negative when consumers enjoy positive utility from stocking commodities at home. However, given our deterministic demand setting, holding cost takes a positive value.

Remark 1. When neither MFS nor CFS is offered, the high-demand consumers will purchase two units of the product at the beginning of the first period to save one shipping trip fee if $0 < h < s_l$, and place a one-unit order in each period if $s_l \leq h < 2s_l$.

Proof. All proofs are in the appendix, unless indicated otherwise.

Remark 1 aims to provide a good understanding of the relationship between holding cost and purchasing frequency. We will revisit Remark 1 multiple times when we discuss holding cost in later analyses. From a modeling perspective, s_l is a critical threshold of holding cost, and the no free shipping setting in Remark 1 provides a useful reference point for us to interpret later results.

4.3.2 *Contingent Free Shipping*

With the policy of Contingent Free Shipping (CFS), consumers will be eligible for free standard shipping if they buy two units in one order. This quantity discount allows e-tailers to practice price discrimination based on consumers' demands. When a high-demand consumer places an order of two units, her utility net of the mismatch cost is $2v_l - 2p - h$. If she purchases one unit in each period, she will need to pay the shipping fee, and her utility becomes $2v_l - 2p - 2s_l$. The condition for any high-demand consumer to choose CFS rather than paying the shipping fee each time is therefore $h < 2s_l$. The low-demand consumers, on the other hand, are not eligible for contingent free shipping, and will always pay the shipping fee, assuming the per-unit price is higher than the shipping cost savings. Thus, a low-demand consumer derives utility of $v_l - p - s_l$. When $h \geq 2s_l$, the average utility net of the mismatch cost is the same across high-demand and low-demand consumers, and the heterogeneity will have no impact on e-tailers' profits. For ease of exposition, we only consider the case when $0 < h < 2s_l$.

4.3.3 *Membership Free Shipping*

With the policy of Membership Free Shipping (MFS), consumers pay a membership fee d upfront and enjoy expedited free shipping for every order they place. To avoid the holding cost, a high-demand consumer who joins the membership program will utilize MFS by placing a one-unit order in each period. Her per-unit utility net of the mismatch cost is $v_h - p - d/2$, which is strictly larger than $v_h - p - d$, the utility of a low-demand consumer who signs up for the loyalty program. From this point of view, high-demand consumers gain more benefits from MFS than low-demand consumers. E-tailers can leverage such a demand heterogeneity to conduct price discrimination. Clearly, e-tailers' choice of membership fee d has an impact on consumers' decisions of whether to subscribe. When d is too small ($0 \le d < s_h$), both types of consumers would join the membership program; when d is too large ($d \ge 2s_h - s_l + h$), no consumer would sign up for it. However, when d is moderate ($s_h \le d < 2s_h - s_l + h$), high-demand consumers would subscribe to the membership program whereas low-demand consumers will stay unsubscribed. We are particularly interested in the third case because only when d is moderate, the membership program can serve as a screening tool to sort consumers with respect to their demand types.

4.3.4 *E-tailer's Choice*

E-tailers choose shipping policy and product price to maximize their profits. Based on their choices, the market has three different shipping configurations:

1. *C* – Symmetric configuration where both e-tailers choose CFS.
2. *M* – Symmetric configuration where both e-tailers choose MFS.
3. *MC* – Asymmetric configuration where E-tailer M_1 chooses MFS, and e-tailer M_2 chooses CFS. We use *M.C* to represent M_1, and *C.M* to represent M_2.

The profit functions for different shipping configurations are presented below. When both M_1 and M_2 choose MFS (or CFS), their profits are equally Π_M (or Π_C) because the market structure is *symmetric*. For brevity, we only list the profit for M_1. When M_1 chooses MFS and M_2 chooses CFS, the shipping structure is *asymmetric*. M_1's profit is $\Pi_{M.C}$, and M_2's profit is $\Pi_{C.M}$. We note that although the general form of Π_M and $\Pi_{M.C}$ are the same, they correspond to different values of x_h, and will have different magnitudes. So is the case for Π_C and $\Pi_{C.M}$.

$$\Pi_M = \Pi_{M.C} = \overbrace{\underbrace{\delta}_{\text{fraction}} \underbrace{x_h}_{\text{demand}} \underbrace{(\underbrace{2p+d}_{\text{marginal revenue}} - \underbrace{2c-2s_h}_{\text{marginal cost}})}_{markup}}^{\text{profit from high-demand market}} + \overbrace{(1-\delta)x_l(p-c)}^{\text{profit from low-demand market}}; \tag{4.11}$$

$$\Pi_C = \Pi_{C.M} = \delta x_h(2p-2c-s_l) + (1-\delta)x_l(p-c).$$

We shall also stress that allowing e-tailers to simultaneously offer MFS and CFS in our model would not change the analytical results. This is because the holding cost is constant across consumers.

4.3.5 *Consumers' Choices*

Consumers choose either M_1 or M_2 to maximize their utility. Throughout the paper, we use *U* to denote consumer's utility, with superscripts representing consumers' type (*H* and *L*) and subscripts representing different shipping configurations (*M*, *C*, *M.C*, and *C.M*). For example, U_M^L denotes the utility a low-demand consumer derives from M_1's product when both e-tailers offer MFS. Under an asymmetric shipping configuration, $U_{M.C}^H$ denotes the utility if a high-demand consumer purchases from M_1 that offers MFS, with M_2 offering CFS. Similarly, $U_{C.M}^L$ is the utility if a low-demand consumer purchases from M_2 that offers CFS, with M_1 offering MFS. Accordingly, consumers' utility under two symmetric configurations (*M*, *C*) can be expressed as:

$$\begin{aligned} &U_M^H = 2v_h - 2p - xt - d; \\ &U_M^L = v_l - p - xt - s_l; \\ &U_C^H = 2v_l - 2p - xt - h; \\ &U_C^L = v_l - p - xt - s_l. \end{aligned} \tag{4.12}$$

In the two symmetric configurations, the marginal consumer in both high-demand and low-demand markets locates in the middle of the unit segment ($x_h = x_l = 1/2$). All consumers with $x < 1/2$ will purchase from M_1, and those with $x > 1/2$ will purchase from M_2.

In the asymmetric configuration when M_1 uses MFS, and M_2 uses CFS, a consumer who locates at x has utilities as follows:

$$
\begin{aligned}
U_{M.C}^{H} &= 2v_h - 2p - xt - d; \\
U_{M.C}^{L} &= v_l - p - xt - s_l; \\
U_{C.M}^{H} &= 2v_l - 2p - (1-x)t - h; \\
U_{C.M}^{L} &= v_l - p - (1-x)t - s_l.
\end{aligned} \tag{4.13}
$$

We illustrate the consumer utility in Figure 4.2. The high-demand market is at the top, and the low-demand market is at the bottom. As we can see, the utility of purchasing from each e-tailer decreases as the consumers locate farther away from the e-tailer. There is a marginal consumer (x_h for the high-demand market, and x_l for the low-demand market) who receives equal utility by purchasing from M_1 and M_2. All consumers on the marginal consumer's left will purchase from M_1, and those on her right will purchase from M_2.

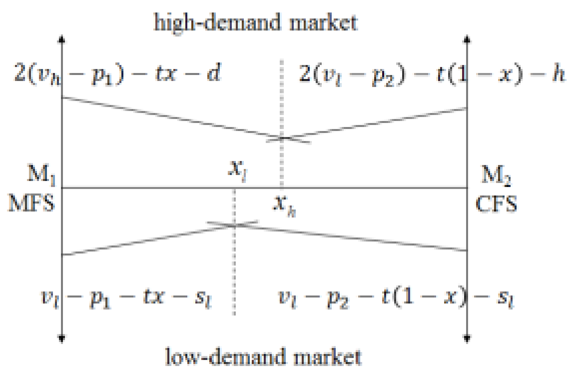


Figure 4.2. Utility Functions of Both Markets

In this paper, we are interested in the case when competition in both markets exists. This is the case when $0 \le x_h \le 1$ and $0 \le x_l \le 1$, which is guaranteed when $t > s_l$.

4.4 Equilibrium Analysis for Homogeneous Model

The homogeneous model assumes that all consumers are of high-demand type ($\delta = 1$). The goal of introducing this model is twofold. First, we aim to understand how our framework is extended from a classic Hotelling model. Second, we discuss the role of rational shipping assumption. Rational shipping assumption ($v_h - v_l = s_h - s_l$) is not imposed in the homogeneous model, but only in the heterogeneous model.

Under symmetric shipping configurations *M* and *C*, we have:

$$p_M = c + s_h - d/2 + t/2 \text{ and } p_C = c + s_l/2 + t/2. \quad (4.14)$$

The optimal profits are equal to that of the classic Hotelling model:

$$\Pi_M = \Pi_C = t/2. \quad (4.15)$$

Under the asymmetric configuration *MC*, when M_1 uses MFS and M_2 uses CFS, the optimal prices are:

$$p_{M.C} = c - d/2 + s_h + \sigma + t/2\,; \quad p_{C.M} = c + s_l/2 - \sigma + t/2\,, \quad (4.16)$$

where $\sigma = 1/6[2(v_h - v_l) - (2s_h - s_l - h)]$. The prices in (6) are similar to those in (4) except for an adjustment term σ. In terms of profit, one e-tailer will earn more than $t/2$, and the other will earn less than $t/2$:

$$\Pi_{M.C} = (t + 2\sigma)^2/2t\,, \text{ and } \Pi_{C.M} = (t - 2\sigma)^2/2t\,, \quad (4.17)$$

where $\sigma = 1/6[2(v_h - v_l) - (2s_h - s_l - h)]$.

The relativity of the profits under the asymmetric configuration depends on the relative strength of $2(v_h - v_l)$ and $2s_h - s_l - h$. When the former is greater than the latter, M_1 earns more than $t/2$, and M_2 earns less than $t/2$, and vice versa. Notably, $2(v_h - v_l)$ is two units of the valuation difference between MFS and CFS, and $2s_h - s_l - h$ is the cost difference between MFS and CFS. Specifically, $2s_h$ is the cost for two expedited shipping trips that would occur under MFS, and $s_l + h$ is the cost of one standard shipping trip plus the holding cost that would occur under CFS.

In terms of strategy equilibrium, both e-tailers will choose the same shipping schedule. They will both choose MFS if $2(v_h - v_l) > 2s_h - s_l - h$, and CFS otherwise. If we take $(v_h - v_l) - (s_h - s_l)$ as a constant, then the sign of $s_l - h$ becomes the determinant of shipping strategy. When holding cost h is greater than s_l, both e-tailers will choose MFS. Otherwise, they will choose CFS. This finding can explain the shipping schedules of markets like online pharmacies (e.g., CVS, Right Aid, and drugstore.com). These online drug stores face highly homogeneous customers whose demand is regular and large. In addition, the holding cost of medication is relatively small. In such a market, all e-tailers use CFS, and no one uses MFS. It is worth mentioning that the threshold for the holding cost (s_l) is the same as the threshold that changes customers' original purchasing frequency (see Remark 1). In other words, e-tailers' optimal free shipping choice preserves customers' original purchasing frequency. This guarantees social welfare maximization. Such a result resembles Bertrand price competition in a homogeneous duopoly, except for a positive profit caused by the product differentiation.

However, most e-tailing markets face heterogeneous demand (e.g., Best Buy and Office Depot) where price discrimination is feasible. We incorporate the factor of demand heterogeneity in our heterogeneous model. In order to focus on the novel design of MFS instead of the higher perceived value for expedited shipping, we impose the rational shipping assumption $v_h - v_l = s_h - s_l$ as described in Assumption 3.

4.5 Equilibrium Analysis for Heterogeneous Model

For the heterogeneous model, we first discuss the price equilibrium in section 4.5.1. We then investigate the shipping strategy equilibrium in section 4.5.2. Price equilibrium presents the optimal prices for each different combination of shipping policies, and strategy equilibrium is calculated using the payoffs at the optimal prices.

4.5.1 *Price Equilibrium Analysis*

Given the three market shipping configurations introduced in section 4.3.4, we discuss the optimal prices in section 4.5.1.1, shipping cost in section 4.5.1.2, and membership fee in section 4.5.1.3. We then compare the consumer surplus in section 4.5.1.4, and social welfare in section 4.5.1.5.

4.5.1.1 Per-Unit Price

The per-unit price is often considered an important factor in evaluating e-tailers. In reality, e-tailers usually change per-unit price dynamically based on various factors including competition and seasonal promotions. This makes it difficult to determine whether the per-unit price of MFS adopters is different from that of CFS adopters. Nevertheless, this topic is widely discussed by consumers. In the Amazon marketplace, consumers complain that the same items, especially inexpensive and small ones, are priced higher if included in the Prime program (Customers 2015). Furthermore, lawsuits have been filed against Amazon for charging Prime members higher prices than non-Prime users for the same item (Abel 2014). The phenomenon that per-unit price will be raised when free shipping is offered is demonstrated empirically (Yao and Zhang 2012). In this section, we compare the per-unit prices under different shipping configurations, and show how the shipping effect and membership effect redistribute surplus between the two types of consumers.

Proposition 1. *The per-unit prices at equilibrium are*

$$p_C = \underbrace{c}_{\text{Cost}} + \underbrace{2\delta^* s_l}_{\text{Shipping effect}} + \underbrace{\delta^* t}_{\text{Horizontal differentiation}}, \tag{4.18}$$

$$p_M = \underbrace{c}_{\text{Cost}} - \underbrace{2\delta^* d}_{\text{Membership effect}} + \underbrace{4\delta^* s_h}_{\text{Shipping effect}} + \underbrace{\delta^* t}_{\text{Horizontal differentiation}}, \tag{4.19}$$

$$p_{M.C} = \underbrace{c}_{\text{Cost}} - \underbrace{2\delta^* d}_{\text{Membership effect}} + \underbrace{4\delta^* s_h}_{\text{Shipping effect}} - \underbrace{2/3\delta^*(s_l - h)}_{\text{Vertial differentiation}} + \underbrace{\delta^* t}_{\text{Horizontal differentiation}}, \tag{4.20}$$

$$p_{C.M} = \underbrace{c}_{\text{Cost}} + \underbrace{2\delta^* s_l}_{\text{Shipping effect}} + \underbrace{2/3\delta^*(s_l - h)}_{\text{Vertical differentiation}} + \underbrace{\delta^* t}_{\text{Horizontall differentiation}}, \tag{4.21}$$

where $\delta^* = (1+\delta)/(1+3\delta)$.

In order to show how the shipping effect and the membership effect are reflected in the per-unit price, we list the optimal price as a function of membership fee d in Proposition 1 when at least one e-tailer chooses MFS. When we plug in the optimal membership fee, we will get $p_M = c + 2\delta^* s_h + \delta^* t$, and $p_{M.C} = c + (2/3)\delta^* t - (1/6)\delta^* h + (2/3)\delta^* s_l$.

We interpret components of the price as follows. Each price consists of a unit cost c, and a horizontal differentiation premium $\delta^* t$. The horizontal differentiation premium is the same across all cases, allowing the price difference to reflect the tension between membership effect and shipping effect. All prices have a positive shipping effect term, $2\delta^* s_l$ for CFS adopter, and $4\delta^* s_h$ for MFS adopter. This shipping effect illustrates how low-demand consumers bear part of the shipping cost of high-demand consumers. The shipping effect is higher in MFS because of the more expensive and more frequent shipping. In addition, MFS adopters have a negative membership effect term $2\delta^* d$. This arises because the membership fee collected from high-demand consumers will drive the per-unit price lower.

Further, under asymmetric configuration (*MC*), the vertical differentiation premium, $2\delta^*(s_l - h)/3$, appears in the optimal prices. Such a vertical differentiation premium is added to the CFS provider, and subtracted from the MFS provider. The sign of this term depends on the relative strength of s_l and h. When the holding cost is high, the MFS provider can charge a higher per-unit price because it is in consumers' best interest to avoid a higher holding cost. In the meantime, the CFS provider has to lower its price to offset the disadvantage of inconvenience.

Corollary 1. Under symmetric configurations, $p_M - p_C = 2\delta^*(s_h - s_l)$, where $\delta^* = (1+\delta)/(1+3\delta)$.

Corollary 1 reveals that under symmetric configurations, shipping speed is the essence of price discrimination. The per-unit price is strictly higher when both e-tailers use MFS than when they both use CFS only if $s_h > s_l$. When the same shipping speed is used for both CFS and MFS ($s_h = s_l$), the per-unit price of CFS becomes the same as that of MFS. The membership fee collected from MFS will be used to offset the additional shipping trip occurs in MFS; these two policies yield the same profit. Furthermore, the difference in per-unit price is proportional to the shipping cost difference $s_h - s_l$. By using a faster shipping speed, e-tailers conduct a higher level of price discrimination and yield higher profits.

Corollary 2. When $0 < h < s_l$, the per-unit price is the lowest when both e-tailers use CFS (C); when $s_l \le h < 2s_l$, the per-unit price is the lowest from the CFS adopter when the competing e-tailer chooses MFS (C.M).

We look at the lowest per-unit price because it is a direct indicator of the level of price discrimination. Under any shipping strategy configuration, low-demand consumers will pay a constant shipping cost s_l and the per-unit price to enjoy the product shipped by standard shipping[5]. Thus, a low per-unit price implies low level of price discrimination.

While Corollary 1 looks at the symmetric configuration, Corollary 2 examines the asymmetric configuration. It is found that the asymmetric competition drives the per-unit price of the CFS adopter even lower than the level in the symmetric configuration *C*. This is the case when the holding cost is relatively high, and consumers have higher willingness to pay for MFS.

---

[5] The assertion that per-unit price is an indicator of the level of discrimination stays even when low-demand consumers choose expedited shipping and pay s_h.

4.5.1.2 Shipping Cost

In this section, we examine shipping cost as an instrument for price discrimination. We conducted a comparative study of e-tailers' profits with respect to shipping. We discuss the differences between symmetric and asymmetric shipping configurations.

Proposition 2. (a) Under symmetric shipping configurations, $\frac{\partial \Pi_C}{\partial s_l} = \frac{\partial \Pi_M}{\partial s_h} > 0$;

(b) Under asymmetric shipping configuration, $\frac{\partial \Pi_{M.C}}{\partial s_l} > 0$ *if* $\delta < \frac{9h+3t-18s_l}{h-21t-10s_l}$ *, and* $\frac{\partial \Pi_{C.M}}{\partial s_l} > 0$ *for all values of* δ*;*

(c) $\frac{\partial \Pi_{M.C}}{\partial s_l} > \frac{\partial \Pi_{C.M}}{\partial s_l}$ *if* $0 < \delta < \frac{9h+19t-24s_l}{9h-13t-24s_l}$*, and vice versa.*

Under symmetric configurations, e-tailer's profit is $\Pi_C = \frac{\delta(1-\delta)}{2+6\delta}s_l + \frac{(1+\delta)^2}{2+6\delta}t$ when they both use CFS, and $\Pi_M = \frac{\delta(1-\delta)}{2+6\delta}s_h + \frac{(1+\delta)^2}{2+6\delta}t$ when they both use MFS. Such a profit structure consists of two components: horizontal differentiation gain and vertical differentiation gain. The horizontal differentiation gain is $\frac{(1+\delta)^2}{2+6\delta}t$ in both cases. The vertical differentiation gain is $\frac{\delta(1-\delta)}{2+6\delta}s_h$ when expedited shipping is used, and $\frac{\delta(1-\delta)}{2+6\delta}s_l$ when standard shipping is used. With faster shipping, e-tailers are able to gain higher profits because of a higher level of price discrimination. Although we assumed shipping cost to be exogenous, our model can be easily extended to account for a more general relationship between shipping valuation and shipping cost (like a quadratic cost structure). Then we can endogenize the decision of shipping cost, and evaluate the optimal shipping cost/speed.

Under the asymmetric shipping configuration, the role of shipping cost over profits depends on the level of δ. At a lower level of δ, a higher shipping cost boosts profits for both e-tailers.

At a higher level of δ, a higher shipping cost will reduce the profit for the MFS adopter and increase the profit for the CFS adopter. The intuition is that the MFS adopter suffers from a higher level of δ, due to the high frequency of expedited shipping. In addition, as we can see in part (c) of Proposition 2, under asymmetric shipping configuration, the MFS adopter has a higher marginal return from the shipping cost than the CFS adopter when δ is relatively low.

Practically, the e-tailer that adopts MFS needs to be cautious about the level of δ. When δ is high, the increase in shipping cost will harm the e-tailer's profitability. Signing a long-term contract with carriers, buying future options for fuel prices, and developing its own supply chain are strategies the MFS adopter can use to secure its profit.

4.5.1.3 Membership Fee

The membership fee attached to MFS has been a puzzle since its inception. When Amazon first launched its Prime program, it charged an annual fee of $79 simply because it is a prime number (Stone 2010). After Amazon hiked its membership fee to $99, some practitioners considered it a brilliant decision (Gustin 2014), while others regarded it as a major mistake (Matthews 2014). As a result, understanding the role of the membership fee is of significant importance. In this section, we investigate the composition of an optimal membership fee in two configurations *M* and *M.C*.

Before we present the optimal membership fee, we discuss the screening conditions. It is easy to show that when $0 \le d < s_h$, all consumers who purchase from the MFS adopter will subscribe to MFS; when $s_h \le d < 2s_h - s_l + h$, high-demand consumers will subscribe to MFS and low-demand consumers will not; when $d > 2s_h - s_l + h$, no consumer will subscribe to MFS. We further demonstrate that

Lemma 1. When all consumers subscribe to MFS, the e-tailer will earn a strictly lower profit than that when only high-demand consumers subscribe to MFS.

From Lemma 1, we learn that e-tailers will never set the membership fee below s_h. We also do not consider the case when $d > 2s_h - s_l + h$ because it provides no insights to the MFS schedule. Thus, our analysis is based on the membership fee range of $s_h \le d < 2s_h - s_l + h$.

Proposition 3. The optimal membership fee is $d_{M.C} = (h-t)/2 + 2s_h - s_l$ under the asymmetric configuration, and $d_M = s_h$ under the symmetric configuration.

The membership fee is an interior solution for the asymmetric configuration *MC*, and corner solution for the symmetric configuration *M*. When M_1 uses MFS and M_2 uses CFS, the profit function for M_1 is concave in the membership fee $d_{M.C}$, and reaches the maximum when $d_{M.C} = (h-t)/2 + 2s_h - s_l$. Note that this is guaranteed to satisfy the screening condition $s_h \le d < 2s_h - s_l + h$. When both e-tailers use MFS, the profit decreases as membership fee d_M increases. The optimal membership fee $d_M = s_h$ is binding at the lower bound of the screening condition. It is easy to show that $d_{M.C} > d_M$.

The difference is caused for the following reasons. When both e-tailers use MFS, the demand is split in half because the market is symmetric. Although higher membership fee results in additional income, it also leads to a decrease in price due to the membership effect, which ultimately harms e-tailers' profitability. Therefore, e-tailers want to price the membership fee as low as possible. On the other hand, in asymmetric configuration, the MFS adopter faces head-to-head competition for market share with the e-tailer adopting CFS. Thus, he wants to keep the price competitive to grab more consumers from his rival. Facing these opposing influences, the optimal membership fee is an interior solution of a concave profit function.

Our results shed light on the decision of membership fee. First of all, when both e-tailers use MFS, the optimal membership fee is just the lower bound of the screening constraint. Essentially, e-tailers want to set the membership fee as low as possible, as long as they can screen high-demand consumers from low-demand consumers. This finding resembles the equilibrium in a two-part tariff monopoly where the optimal entry fee is the fixed cost (zero in our context). Secondly, when only one e-tailer uses MFS, the optimal membership fee is $d_{M.C} = (h-t)/2 + 2s_h - s_l$. It consists of two parts: shipping cost part and adjustment part. The shipping cost part $2s_h - s_l$ is basically the cost of the two expedited shipping trips that would occur in MFS minus the cost of a regular shipping trip that would occur if CFS is chosen. Intuitively, consumers are paying their own shipping expenses. The other term $(h-t)/2$ can be regarded as an adjustment term which has very practical implications. When the holding cost h is high, the e-tailer can charge a higher membership fee for ordering as needed. When the competition is intense (t is small), the e-tailer who uses MFS has an incentive to increase the membership fee to strengthen the membership effect, in order to drive the unit price low and stay competitive.

4.5.1.4 Consumer Surplus

Consumer surplus reflects the benefits enjoyed by consumers in each shipping configuration. After the introduction of MFS, different voices have raised regarding its impact on consumer surplus. One opinion is that MFS makes consumers better off because it provides more value (Stone 2010). The other opinion is that, although MFS looks attractive, consumers end up paying more (Stevens and Banjo 2014, Wohlsen 2013). In this section we formally compare consumer surplus under symmetric configurations. We do not discuss the consumer surplus under the asymmetric configuration, mainly because it carries a similar message to our strategy equilibrium.

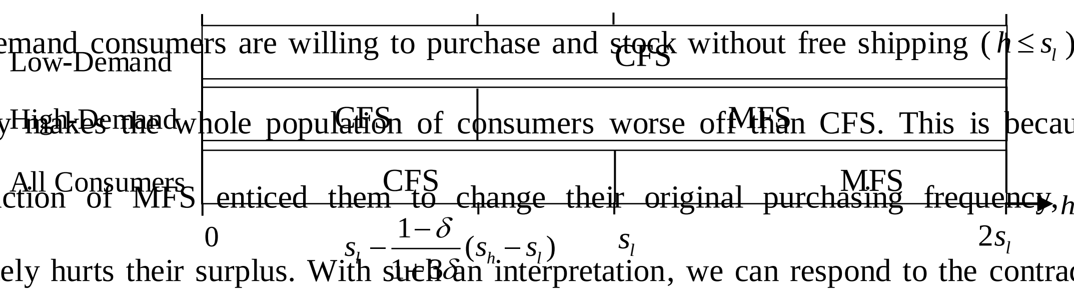
Proposition 4. (a) Low-demand consumers always have higher surplus when both e-tailers use CFS than when both e-tailers use MFS.

(b) High-demand consumers have higher surplus when both e-tailers use MFS than when they both use CFS if $h > s_l - \frac{1-\delta}{1+3\delta}(s_h - s_l)$, and lower surplus otherwise.

(c) The whole population of consumers has higher surplus when both e-tailers use CFS than when they both use MFS if $h < s_l$, and otherwise if $s_l < h < 2s_l$.

Part (a) of Proposition 4 can be easily inferred from section 4.5.1.1. Since the per-unit price is strictly lower in the symmetric configuration *C* compared to the symmetric configuration *M*, low-demand consumers prefer the former market configuration. However, the surplus for high-demand consumer is not that straightforward. We find that high-demand consumers' preferable market configuration depends on the magnitude of h (see Part (b) of Proposition 4). When h is low, MFS is less favorable to these high-demand consumers, and they prefer the symmetric configuration *C*. When h is relatively high, the high-demand consumers are better off in the symmetric configuration *M*. Part (c) of Proposition 4 examined the overall consumer surplus. When the holding cost is higher than s_l, consumers are generally better off in a market where both e-tailers use MFS. The threshold for the high-demand consumer to prefer the symmetric configuration *M*, $s_l - \frac{1-\delta}{1+3\delta}(s_h - s_l)$, is strictly lower than the threshold for all consumers to prefer the symmetric configuration *M*, s_l, due to low-demand consumers' strict preference over CFS. In Figure 4.3, we plot the shipping schedule that maximizes each consumer segment's surplus under symmetric configurations. It can be seen that CFS is preferable to both high-demand and low-demand consumers only when $h \le s_l - \frac{1-\delta}{1+3\delta}(s_h - s_l)$. In cases when $s_l - \frac{1-\delta}{1+3\delta}(s_h - s_l) \le 0$, the conflict of interests between high-demand and low-demand consumers can never reconcile because high-demand consumers strictly prefer MFS.

Figure 4.3. Shipping Configuration Maximizing Consumer Surplus



Recall that F[illegible]g cost. If high-demand consumers are willing to purchase and stock without free shipping ($h \le s_l$), MFS actually makes the whole population of consumers worse off than CFS. This is because the introduction of MFS enticed them to change their original purchasing frequency, which ultimately hurts their surplus. With such an interpretation, we can respond to the contradicting statements from the industrial report: for products that induce high-demand consumers to place large orders without the free shipping, CFS offers higher consumer surplus than MFS. For products that induce consumers to place small orders without the presence of free shipping, MFS provides higher consumer surplus than CFS.

4.5.1.5 Social Welfare

It was reported that heavy residential shipping enabled by internet shopping leads to social waste. It drives fuel-surcharge high, and makes e-tailers at the border of life (Stevens 2015). In this section, we formally compare social welfare under different shipping configurations, and investigate the most efficient solution for society.

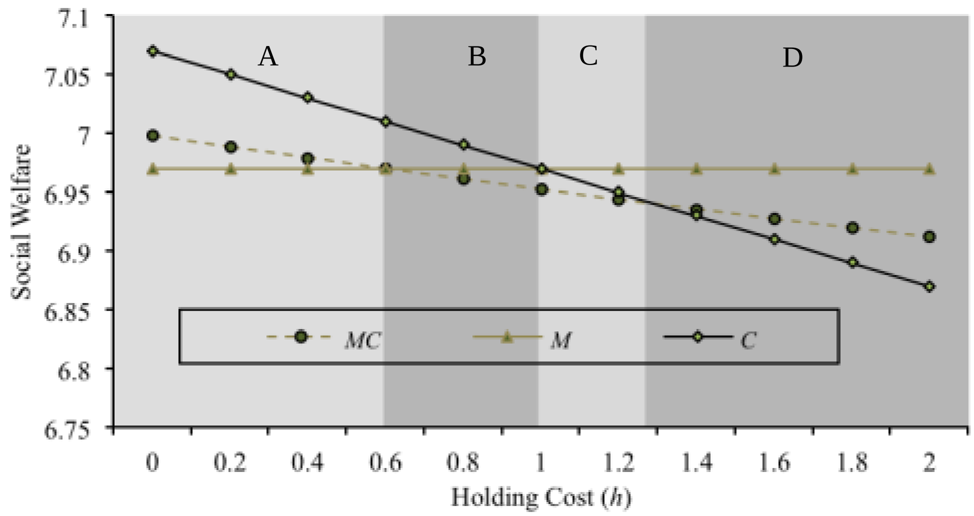
Proposition 5. (a) It is socially optimal to let both e-tailers use CFS if $0 < h < s_l$.

(b) It is socially optimal to let both e-tailers use MFS if $s_l < h \leq 2s_l$.

(c) It is never socially optimal to let one e-tailer use MFS, and the other use CFS.

The result from Proposition 5 reinforces the impact of holding cost on social welfare. When consumers are incentivized by MFS to place one-unit orders for products that they would have placed two-unit orders for in the absence of free shipping policy, social waste is generated. Similarly, when consumers are incentivized by CFS to place two-unit orders for products that they would have placed one-unit orders for in the absence of free shipping, social waste is produced. In other words, free shipping policy, whether in the form of MFS or CFS, creates social waste when it changes consumers' original purchasing frequency in the absence of free shipping. This is one of the key findings of this paper.

We plot the value of social welfare against the level of h in Figure 4.4. It is intuitive that at a low level of h, the symmetric configuration *C* creates the highest social surplus, the symmetric configuration *M* produces the lowest surplus, and the asymmetric configuration *MC* generates a medium level of social surplus (area A). The opposite is the case at a high level of h, as in area D. Surprisingly, when h is at a medium level, the asymmetric configuration *MC* corresponds to the lowest level of social welfare (area B and C). This is due to the inefficiency from preference mismatch. Such an inefficiency is caused when the marginal consumer does not locate at the center of the horizontal line under an asymmetric configuration. Under such a condition, some consumers purchase from the e-tailer that they would not patronize if the market were symmetric.



$(t=6, \delta=0.1, v_h=10, v_l=9.5; s_h=1.5; s_l=1, c=0.8)$

Figure 4.4. Social Welfare under Different Shipping Configurations

4.5.2 *Shipping Strategy Equilibrium*

In this section, we present the optimal shipping strategy given the payoffs we derive from our heterogeneous model. In section 4.5.2.1, we describe the dilemma faced by e-tailers. In section 4.5.2.2 and 4.5.2.3, we present the equilibria for the sequential-move game when $h<s_l$ and $s_l \le h < 2s_l$, respectively.

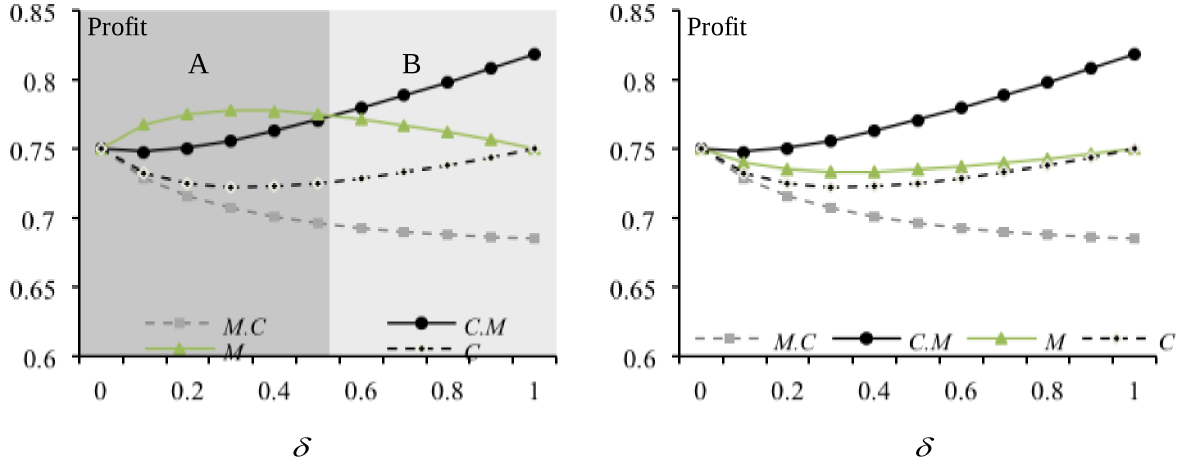
4.5.2.1 Prisoner's Dilemma

The goal of this section is to compare profits to understand the prisoner's dilemma faced by e-tailers in the sequential-move game. This is done by showing 1) the attractiveness of cooperation, 2) the threat of the rival's deviation, and 3) the incentive for the rival to deviate. We discuss the condition when $h<s_l$ and $s_l \le h < 2s_l$, separately. We leave the mathematical formulation of profit functions Π_M, Π_C, $\Pi_{M.C}$ and $\Pi_{C.M}$ in the appendix to increase readability.

Lemma 2. When $h<s_l$, (a) $\Pi_M > \Pi_C$, (b) $\Pi_{M.C} < \Pi_C$ if $\hat{\delta} < \delta < 1$, (c) $\Pi_{C.M} > \Pi_M$ if $v_h - v_l < \tilde{\hat{v}}$, *where* $\hat{\delta} = \frac{-9h^2 - 30ht - 9t^2 + 36hs_l + 48ts_l - 36s_l^2}{7h^2 + 66ht - 9t^2 + 4hs_l - 48ts_l - 20s_l^2}$ *and*

$$\tilde{v} = \frac{8h\left(-h\delta + 3t\left(1+\delta\right)\right) + \left(-3t\left(11+5\delta\right) + h\left(-9+25\delta\right)\right)s_l + \left(18-26\delta\right)s_l^2}{18t\left(-1+\delta\right)}.$$

Lemma 2 reveals both the attractiveness and the risk faced by the first-moving e-tailer when $h < s_l$. From Lemma 2 (a), we immediately learn that when both e-tailers use MFS, they will gain higher profits than when they both use CFS because $\Pi_M - \Pi_C = \frac{\delta(1-\delta)}{2+6\delta}(s_h - s_l) > 0$, as illustrated in the *M* and *C* curves in both panels of Figure 4.5. This gives e-tailers motivation to move from CFS to MFS. However, the first mover faces the threat that if the second mover doesn't cooperate and keeps using CFS, the first mover may end up gaining less profit than in the case of symmetric configuration *C* as illustrated in Lemma 2 (b) and the *M.C* dashed line at the bottom of Figure 4.5. The following e-tailer has the incentive to deviate from cooperation when it gives him a higher profit, as presented in Lemma 2 (c) and the *C.M* curves in Figure 4.5. Also from Lemma 2 (c), we learn that this suboptimal result can be avoided when expedited shipping is highly valued by consumers and the proportion of high-demand consumers is low. We plot the profit functions when $v_h - v_l$ is high in the left panel, and when $v_h - v_l$ is low in the right panel of Figure 4.5. As can be seen, the left panel has a unique region A. In this region, the level of δ is low enough (compared to region B), and using MFS will give both e-tailers the highest profits. In Region A, the second-mover is guaranteed to follow the first mover in using MFS.



(The figures are illustrated when $h = 0.8$, $s_l = 1$, and $t = 1.5$. In the left panel, $s_h = 2$, and in the right panel, $s_h = 1.2$.)

Figure 4.5. Profit Comparison

These findings provide an explanation for why Amazon aggressively accelerated its speed of delivery from two-day to one-day, and even to half-an-hour by means of drones. From our analysis, doing so can prevent the second mover from deviating. After all e-tailers are forced to use MFS, the per-unit price can be raised, and all e-tailers benefit from such a cooperation.

When the holding cost is high ($s_l \le h < 2s_l$), such a prisoner's dilemma no longer exists.

Lemma 3. *When* $s_l \le h < 2s_l$, $\Pi_M > \Pi_C$ *and* $\Pi_{M.C} > \Pi_C$.

From Lemma 3, we learn that the suboptimal outcome when both e-tailers use CFS will never happen. When the holding cost is higher than s_l, the first moving e-tailer will gain higher profit (than when they both use CFS) by introducing MFS regardless of the second mover's action.

4.5.2.2 Strategy Equilibrium When $0 < h < s_l$

The sequential-move game accentuates the realistic setting where e-tailers choose free shipping policy in a strict order. We let the leader M_1 choose his shipping strategy before the follower M_2, based on his anticipation of M_2's behavior, as illustrated in the game tree below.

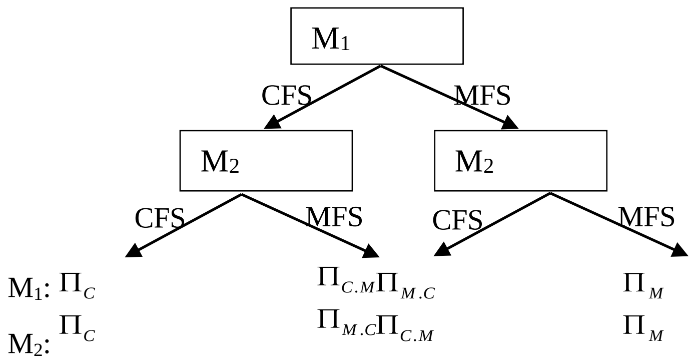


Figure 4.6. Sequential-Move Game Tree

Proposition 6. *When* $0 < h < s_l$, *the equilibrium is as follows:*

(a) When $v_h - v_l > \tilde{\tilde{v}}$, M_1 chooses MFS, and M_2 follows up by choosing MFS.

(b) When $v_h - v_l \le \tilde{\tilde{v}}$, there are three conditions based on the level of δ.

i) If $0 \le \delta < \underline{\delta}$, M_1 chooses MFS, and M_2 chooses CFS. M_1 gains higher profit than M_2.

ii) If $\underline{\delta} \le \delta < \bar{\delta}$, M_1 chooses CFS, and M_2 chooses MFS. M_1 gains higher profit than M_2.

iii) If $\bar{\delta} \le \delta \le 1$, both e-tailers choose CFS.

where $$\tilde{\tilde{v}} = \frac{8h\left(3t(1+\delta) - h\delta\right) + \left(h(-9+25\delta) - 3t(11+5\delta)\right)s_l + (18-26\delta)s_l^2}{18t(-1+\delta)},$$

$$\underline{\delta} = \frac{3h^2 + 26ht + 3t^2 - 18hs_l - 38ts_l + 24s_l^2}{3h^2 - 38ht + 3t^2 - 18hs_l + 26ts_l + 24s_l^2} \text{ and } \bar{\delta} = \frac{-9h^2 - 30ht - 9t^2 + 36hs_l + 48ts_l - 36s_l^2}{7h^2 + 66ht - 9t^2 + 4hs_l - 48ts_l - 20s_l^2}.$$

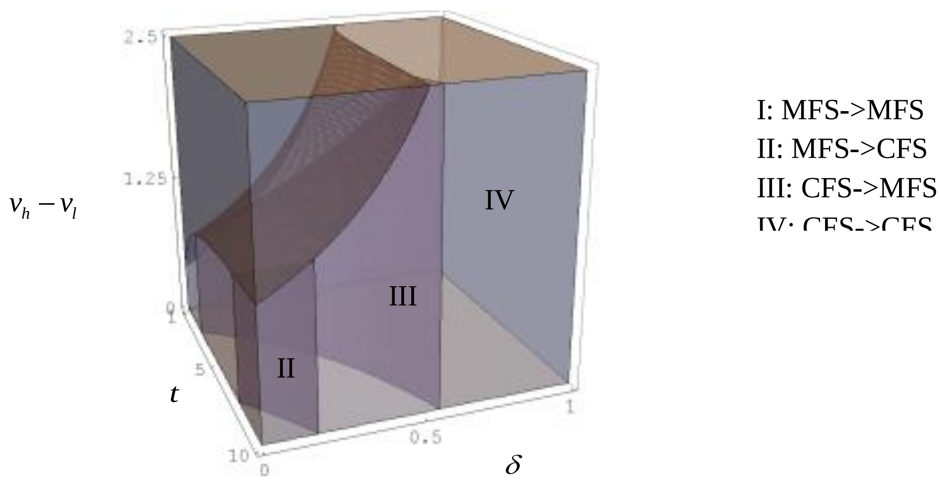
We interpret

Figure 4.6. Sequential-Move Game Tree

Proposition 6 as follows. First, when expedited shipping is highly valued, the first-moving e-tailer can guarantee that the second-moving e-tailer will follow his shipping strategy MFS, leading to a cooperative outcome. This can be seen from part (a) of

Figure 4.6. Sequential-Move Game Tree

Proposition 6 and area I of Figure 4.7.



(This graph is illustrated by setting $s_l = 1$, and $h = 0.5$)

Figure 4.7. Sequential-Move Game ($0 < h < s_l$)

Next, we find that the fraction of high-demand consumers δ plays an important role in determining the equilibrium status, as in part (b) of

Figure 4.6. Sequential-Move Game Tree

Proposition 6. When δ is small, the two competing e-tailers differentiate from each other (area II and III). One e-tailer will choose MFS, and the other will choose CFS (Case *i* and *ii*, respectively). Because of the first mover advantage, M_1 will always make a decision so that he gains higher profit than M_2 under this differentiated condition. When δ is big enough, both e-tailers will choose CFS (Case *iii*). This is because the cost of being betrayed is very high due to the high frequency of expedited shipping. This condition is like the prisoner's dilemma where two rational agents end up with a suboptimal outcome. However, only under this condition, society obtains the highest social welfare. For other conditions where social welfare is not maximized, the social planner can use taxes to achieve higher efficiency. Specifically, the regulator can impose higher taxes for expedited shipping or subsidize standard shipping. In the real world, shipping is still taxed uniformly as part of the product price, regardless of the shipping speeds. One reason is that expedited shipping is used in different environments, and the social planner cannot tell when it causes social waste. However, our finding is still applicable to online marketplaces. Owners of online marketplaces like Amazon work as regulators for all e-tailers that sell in their platforms. Amazon has grasped the efficiency loss due to excessively frequent expedited shipping, and allows consumers to choose "no rush shipping" on items for which they don't need two-day shipping. By choosing "no rush shipping", consumers experience a longer waiting time, and are compensated with Amazon credit. This policy works as a subsidy and preserves consumers' original purchasing frequency.

4.5.2.3 Strategy Equilibrium When $s_l \leq h < 2s_l$

Proposition 7. When $s_l \leq h < 2s_l$, the equilibrium status is as follows:

a) When $v_h - v_l > \ddot{v}$, M_1 will choose MFS, and M_2 will follow up by choosing MFS.

b) When $v_h - v_l \leq \ddot{v}$, there are three conditions based on the range of δ.

If $0 \leq \delta < \underline{\delta}$, M_1 chooses CFS, and M_2 chooses MFS. M_1 gains higher profit than M_2.

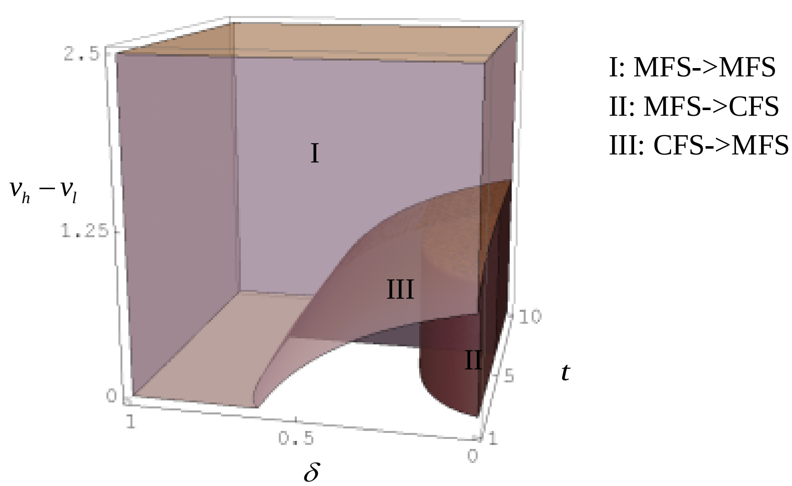
If $\underline{\delta} \le \delta < \bar{\delta}$ *, M₁ chooses MFS, and M₂ chooses CFS. M₁ gains higher profit than M₂.*

If $\bar{\delta} \le \delta < 1$*, no pure strategy equilibrium exists.*

where $\tilde{v} = \dfrac{8h\left(-h\delta + 3t(1+\delta)\right) + \left(-3t(11+5\delta) + h(-9+25\delta)\right)s_l + (18-26\delta)s_l^2}{18t(-1+\delta)}$,

$\underline{\delta} = \dfrac{3h^2 + 26ht + 3t^2 - 18hs_l - 38ts_l + 24s_l^2}{3h^2 - 38ht + 3t^2 - 18hs_l + 26ts_l + 24s_l^2}$ *and* $\bar{\delta} = \dfrac{24ht - 9hs_l - 33ts_l + 18s_l^2}{8h^2 - 24ht - 25hs_l + 15ts_l + 26s_l^2}$.

Similar to the equilibrium status when $0 < h < s_l$, both e-tailers should use MFS if expedited shipping is highly valued, as presented in part (a) of Proposition 7 and area I of Figure 4.8. However, there are two major differences between equilibria under $s_l \le h < 2s_l$ and $0 < h < s_l$. First, the symmetric configuration *C* will never be at equilibrium when $s_l \le h < 2s_l$. We can see the reasons in Lemma 3. Second, in the asymmetric equilibrium, the first-moving e-tailer will choose CFS when δ is large in the case of $0 < h < s_l$, and will choose MFS when δ is large in the case of $s_l \le h < 2s_l$. This is because the first-moving e-tailer puts more weight on the preference of the high-demand market when it has a larger share. High-demand consumers would prefer CFS when $0 < h < s_l$, and MFS when $s_l \le h < 2s_l$.



(This graph is illustrated by setting $s_l = 1$, and $h = 1.2$)

Figure 4.8. Sequential-move Game ($s_l \le h < 2s_l$)

4.6 FUTURE WORK

Offer of a free shipping schedule is a complicated system of decisions, and we did not incorporate all factors that influence the pricing. For tractability reasons, we used a simple model to illustrate the tension between the shipping effect and the membership effect in the presence of demand heterogeneity. Based on this model, we are able to draw normative conclusions on the distinctions between CFS and MFS. While we do not explicitly model the following factors, we believe these factors hold interest for future exploration.

First, it will be interesting to look at e-tailers' dynamic pricing decisions. A loyalty program can be regarded as a strategy to increase consumers' switching cost, in order to "lock in" consumers (Chen and Hitt 2002). Theoretically, e-tailers can strategically increase product price after consumers subscribe to the membership program. We don't consider such a dynamic pricing in our paper because such an increase in price may hinder subscriptions from new members. A dynamic model that capture all these trade-offs is beyond the investigation of this work.

Second, we don't model the fact that e-tailers practice price discrimination based on consumers' heterogeneous preference for high quality shipping (Li and Dinlersoz 2012). It was found that e-tailers strategically charge a higher markup for faster shipping options (Li and Dinlersoz 2012, Yao and Zhang 2012). We did not model this dimension of vertical product differentiation because we assumed consumers have homogeneous willingness to pay for any unit of product, given the shipping speed. Such a modeling setting allows us to focus on the effect of demand heterogeneity.

Third, we don't take into account the economy of scale based on the number of processed orders. In the real world, package-delivery companies offer discounts to e-tailers who ship more heavily (Stevens 2016). This makes the shipping cost per order related to the e-tailer's market share. We don't explicitly model this because it will only make MFS more favorable due to the higher frequency of shipping. We also don't include the role of package-delivery companies like UPS, Fedex, and USPS. We assume that e-tailers are price takers who face a competitive package-delivery market. This is primarily because transportation companies don't differ much from each other, and hence have low market power, which is likely to result in a competitive market.

Fourth, we only allowed two shipping speeds: standard shipping and expedited shipping. In reality, multiple shipping speeds are available. When both e-tailers use MFS, they can choose different shipping speeds and set different membership fees. Such product differentiations likely follow traditional vertical differentiation equilibria. Since our focus is the comparison between MFS and CFS, we simplify the competition when both e-tailers choose MFS/CFS to a symmetric case.

Last, e-tailers are vertically differentiated in aspects other than shipping service. Amazon provides video streaming, unlimited photo cloud storage, Kindle book borrowing, and other benefits to its Prime members, which creates a vertical advantage over its rivals. However, it also charges the highest membership fee at $99. If the higher membership fee is used to offset the additional cost for the extra benefit, we still argue that the inclusion of these benefits becomes part of horizontal differentiation. Therefore, the simplification does not affect the generality of our results. Nevertheless, it can be studied as a separate topic.

4.7 CONCLUSION

Over the last two decades, contingent free shipping is massively used as a promotion strategy to induce larger orders. After Amazon's successful introduction of membership free shipping, a tough decision of whether to follow and use MFS is faced by all e-tailers. While some retailing giants like Walmart have started to offer MFS in the hope of winning consumers back from Amazon, many others remain hesitant. In this work, we compare the two different shipping policies MFS and CFS, to shed light on the decision makings for e-tailers.

We find that under symmetric shipping configurations, CFS and MFS are equivalent in profitability if the same shipping speed is attached to both. MFS outperformed CFS solely because expedited shipping is attached to MFS. Thus, e-tailers have an incentive to use MFS with expedited shipping. However, the leading e-tailer faces the threat of the following e-tailer deviating from cooperation. When the fraction of high-demand consumers is high, the cost of being betrayed is so high that both e-tailers end up using CFS, as in the prisoner's dilemma. We show that when the expedited shipping attached to MFS is highly valued compared to standard shipping, the leader can guarantee that the follower will cooperate. We also find that preserving consumers' original purchasing frequency is the key in social welfare maximizing. These findings open the black box of MFS and carry significant importance to both researchers and practitioners.

We end our paper by suggesting that the following e-tailers examine 1) the level of holding cost, 2) the valuation difference between expedited shipping and standard shipping, and 3) the fraction of high-demand consumers, while making shipping policy decisions. When expedited shipping is highly valued, or considered as a necessary part of online shopping experience, e-tailers should follow the trend and use MFS. On the other hand, when most consumers have high demand, blindly following the leader to use MFS will result in a loss of profit.

Chapter 5. E-COMMERCE - LIVE CHATTING TOOLS

5.1 INTRODUCTION

Unlike brick-and-motor stores, online stores face the disadvantage of information asymmetry, lack of trust, and lack of adaptive assistance to help and guide customers (Resnick and Zeckhauser 2002). In physical store, customers can judge the trustworthiness of the retailer by the appearance and display of the stores. They can touch and feel the products to evaluate their qualities. Further, customers are offered adaptive assistance by salespersons. Salespersons can recognize customers' decision styles, and provide tailored service and recommendations (Perrault and Brousseau 1989). In online marketplaces, customers interact with a system. They face a different set of information, including sellers' feedback scores, product reviews, and prices. Sellers' feedback profiles and product reviews are essential parts of the online reputation system, which makes it possible for strangers to trust each other and conduct business transactions without contracts (Dellarocas 2003). Considerable research has found that reputation is positively correlated with purchase (Ba and Pavlou 2002, Chevalier and Mayzlin 2006, Livingston 2005).

In the past, feedback score (at seller level) and product review count (at seller-product level or product level) have usually been studied separately. This is because most prior analyses of online markets are conducted on eBay or Amazon, where the website design restricted the availability of data. To be specific, eBay allows buyers to rate and leave text feedbacks to both sellers and products. However, sellers can opt out from displaying customers' reviews, leading to missing data in product reviews. Amazon uniquely identifies each book: the same book sold by different sellers has a distinct rating. Therefore, the analysis of the relationship between book reviews and sales only focuses on product-level feedback (Mayzlin et al. 2003). Seller-level reputation, while available, is usually not used in researches because it is not clear the share of book sales at each seller. In this study, we analyze a different online market where both seller-level feedback score and seller-product-level product review are available, allowing us to control for their different roles.

More importantly, the online marketplace we study provides live chatting functions to allow communication between buyers and sellers. In recent years, the use of live chatting tools has changed the way customers interact with sellers. With live chatting tools, sellers display a "chat" button on the website for consumers to click away and initiate a live chat. Sellers will get notifications either from their mobile phone or computers, and respond to the buyer in real-time. A few online businesses use proactive live chatting tools that pop up to ask whether any assistance is needed. Live chatting tools have shown great ability to improve customer satisfaction and boost conversion rates. According to a survey conducted by Forrester Research (Strothkamp et al. 2010), "Around 44% of online consumers say that having questions answered by a live person while in the middle of an online purchase is one of the most important features a Web site can offer." AT&T claimed to have reached the highest satisfaction when the chat function is used (Power 2013). Abt Electronics experienced conversion rates as high as 20% after introducing live chat (Abt 2010). Wells Fargo claimed to see both high satisfaction and a double-digit increase in conversion rates with the use of live chatting tools (Strothkamp et al. 2010).

With the adoption of live chatting tools, the role of reputation may change. As is known, customers base their decisions on prior shoppers' decisions, leading to information cascade, where products with high reputation attract more customers (Li and Hitt 2008). As a result, the rich get richer, and new sellers have a hard time surviving even when they sell high quality products (Muylle and Basu 2004). The introduction of live chatting tools adds human-based interaction to customers' online shopping experience. This may change the way consumers acquire knowledge to reduce information uncertainty in searching for products. Therefore, it may provide an unprecedented opportunity for sellers with relatively low reputation to attract customers and increase turnover in the competitive online marketplace.

This paper aims to understand: 1) how live chat affects customers' online shopping experience, 2) how seller-level reputation and product-level reviews affect customers' online shopping behavior differently, and 3) whether live chat serves as substitutes for reputation. Besides exploring answers to these questions, we also examine the power of our model in predicting customers' browsing, chatting, and purchasing behavior, in the hope of aiding in the design of web recommendation and proactive live chatting.

To achieve these goals, we used clickstream data from Alibaba, the biggest online selling platform in China. The data were collected in the tablet market where two brands, Apple and Samsung, dominate the market. Our data consist of 4000[6] customers' trajectories of product browsing, live chatting, and purchasing histories. In the first stage of viewing products, customers form their choice sets. In the second stage of live chatting, customers interact with sellers to ask questions and allay their concerns. In the final stage of purchasing, customers select products from their choice sets or leave the platform without purchasing (Gensch 1987). By

[6] 3000 users were used in the training set, and 1000 users were used in the validation set.

analyzing customers' decisions at each stage, we uncovered how live chats affect customers' online shopping patterns.

We obtained interesting results from our analyses. First, we found that live chat positively affects a customer's repeat visit to the same seller only when the seller has low feedback scores or the product has low counts of reviews. Second, live chats improve the purchase probability significantly for all sellers. Third, the positive effect of live chat in converting sales is mitigated by the level of feedback scores. Fourth, customers are more likely to request live chats when the seller has a low feedback score or when the product has a low count of reviews. Fifth, customers usually start their search with products with high products review counts[7] but low seller feedback scores, and gradually move on to products with lower product review counts and higher feedback scores. This is likely to be the most efficient search path because price is found to be positively correlated only with feedback score, but not with product review count. In other words, customers can obtain high quality products without suffering high prices if they use product review count as their primary search criterion, at least in our context. Lastly, our predictive model performed very well in the holdout sample, indicating potential opportunities for real-time proactive live chat or better recommendation system design.

Our work makes significant contributions to the literature of information systems and marketing. First, this work is the first to examine the role of live chat in consumers' browsing and purchasing decisions. By analyzing determinants of these two decisions, our work sheds light on consumers' decision making in online shopping, and adds to the literature of consumer choice theory. Second, we analyze factors to reduce seller-level uncertainty (feedback scores) and product-level uncertainty (review count) simultaneously, allowing for a deeper

[7] We note that product review count is likely to be highly correlated with product sales volume, and these two criteria are likely to be equivalent.

understanding of their distinctions, which is under-explored in prior research. Third, we obtain insightful findings about when consumers need help from human-based web assistants. We discover the substitution effect of live chat for reputation, showing that sellers with low reputation gain higher benefits from live chatting tools. Furthermore, the good performance of our predictive model indicates great potential for platform designers to improve product recommendation and facilitate real-time proactive live chatting.

This manuscript is organized as follows. In Section 2 we review related literature. In Section 3 we introduce our research context. In Section 4 we describe our data. We conduct the empirical study in Section 5, and present our predictive analysis in Section 6. We provide managerial implications, and conclusions in Section 7.

5.2 Literature Review

5.2.1 *E-Commerce*

Our work is related to the literature of electronic commerce in information systems. In electronic commerce literature, business-to-consumer communication media are studied to improve consumers' shopping experience (Åberg and Shahmehri 2000). Åberg and Shahmehri (2000) proposed human-based web assistants, and developed a prototype system to evaluate its performance. They found that web assistants can improve the usability of online shops in terms of relevance, efficiency, attitude, and learnability. Basso et al. (2001) used an experiment to compare an online store with no audio or real-time interpersonal communication, with other stores with different forms of communication media. They found that online stores with web-based instant messenger create more trust than those without such instruments. The trust will in return lead to customers' greater willingness to share the information to their friends, and repeated purchases (Basso et al. 2001).

Further, our work is related to online word of mouth, or reputation systems (Resnick and Zeckhauser 2002). Information technologies harness a large word-of-mouth network, allowing sellers to build trust and customers to cooperate (Dellarocas 2003). The online feedback mechanism is especially important for online marketplaces, due to the lack of trust between sellers and buyers (Kollock 1999, Overby and Jap 2009). Most literature in information systems investigate factors to reduce seller uncertainty, and a few literature addressed product uncertainty (Dimoka et al. 2012). At seller level, factors like feedback scores are used to reduce uncertainty (Ba and Pavlou 2002, Dewan and Hsu 2002, Kauffman and Wood 2006). In addition, text comments left by customers to sellers are also found to significantly influence price premium (Ghose et al. 2006, Pavlou and Dimoka 2006). At product level, auction features are used to reduce uncertainty (Bapna et al. 2008, Dimoka et al. 2012, Li et al. 2009). Money-back guarantee and certification of product are also used to control for product quality (Li et al. 2009). In addition, pictures uploaded by sellers and product descriptions are regarded important factors to reduce product uncertainty (Andrews and Benzing 2007, Kauffman and Wood 2006, Ottaway et al. 2003). Since most prior studies are conducted on eBay, we do not see the use of product review counts as a factor to reduce product uncertainty. This is because on eBay, sellers can opt out from displaying product review.

In another online marketplace Amazon, product reviews have been widely studied with respect to their impacts on sales. Mayzlin and Chevalier (2006) examined the effect of book reviews on sales, and found positive evidence. In their study, they not only considered the number of reviews, but also other factors including the valence of the reviews (Chevalier and Mayzlin 2006). Park et al. (2007) examined the moderating role of involvement in the relationship between reviews and purchase using an experimental approach, and found that a higher count of reviews positively affects sales, and will be moderated by user involvement (Park et al. 2007). However, sellers' ratings are usually not considered in the above studies because no information is provided on the volume sold by each seller.

Another problem related to our work is the relationship between reputation and price. Ambiguous conclusions have been reached over this question (Houser and Wooders 2000, Resnick and Zeckhauser 2002). Ba and Pavlou (2002) found positive relationship between feedback score and price in the categories of music, software, and electronics. Houser and Wooders (2000) found positive correlation between feedback and price in the product category of Pentium chips. On the other hand, Resnick and Zeckhauser (2002) found no significant correlation between feedback and price in the market of MP3 players.

In our work, we include both feedback scores at seller level and review counts at product level. The inclusion of both measurements allows us to reconcile the different findings of reputation, and better understand their differences.

5.2.2 *Marketing*

Our work is also related to the choice theories in marketing literature. A two-stage choice process is usually used to model how consumers first select a subset of products to view, and then choose one or more from the subset to purchase (Bettman 1979, Gensch 1987). It has been proposed that consumers' decision rules in determining which products to view and which products to purchase differ because of their cost-benefit trade-offs. In general, simpler rules are used in early stages because consumers face a larger set of choices, and more effortful rules are used in later stages to achieve the maximum of benefit (Gensch 1987). While most models assume the first stage to be latent, Moe (2006) addressed the limitation with clickstream data, and explicitly modeled the first stage decisions. Our paper extends the two-stage choice framework, and incorporates consumers' choice of live chat as a second stage after the first stage of browsing and before the last stage of purchasing.

Path data analysis in general is of considerable interest to marketing researchers. Hui et al. (2009) reviewed previous works in path analysis in marketing, and pointed out the under-exploration of path data. Montgomery et al. (2004) used a dynamic multinomial probit model to predict web browsing and purchase, and found that users' paths imply their shopping goals, and this information can be utilized to predict their future movements (Montgomery et al. 2004). Along this line, Olbrich and Holsing (2011) used a logit model to understand consumer's purchasing behavior in a social shopping community (Olbrich and Holsing 2011).

Loosely related to our work is another stream of marketing literature that used a structural search model to explicitly model consumers' searching behavior. With the availability of additional identification sources, they recover the underlying search cost. Hong and Shum (2006) developed a structural approach to recover consumers' search cost with aggregate data (Hong and Shum 2006). De los Santos et al. (2012) also used a structural approach to recover the search costs for automobiles, and they found significant search costs which will result in a limited number of dealer search attempts (De los Santos et al. 2012). Chen and Yao (2015) advanced the search model to incorporate search refinement with sorting and filtering (Chen and Yao 2015). Our work only focuses on determinants of customers' decisions at each stage of their shopping experience, and does not concern recovery of search cost.

5.3 RESEARCH CONTEXT

We obtained our individual-level clickstream data from Alibaba, the largest online marketplace in China. Alibaba Group was founded in 1999. By August 2016, it has 434 million registered users. Every year, 12.7 billion orders are placed on Alibaba. The Alibaba platform had over 8.5

million active sellers by August 2014 (Smith 201). Alibaba sells enormous varieties of products, and our data are collected from the section of tablet markets. The online tablet market flourishes in China because Apple used to price its tablets higher in China than in Hong Kong and many other areas and countries. In addition, Apple used to release its new tablets at different times in different countries, and China is usually among the last to release the new product. The price difference and time gap stimulated the secondary market because people cannot wait till the official date of release, or else want the tablets at lower prices. The secondary market largely operates in online marketplaces, like Alibaba.

Alibaba is a typical two-sided market where sellers and customers both need to register in order to sell or buy. Each seller owns an online shop with a distinct link. Products are displayed in each online shop for consumers to browse and purchase. Different sellers can sell the same product (e.g., the same iPad model) and price them differently.

5.3.1 *Shopping Process*

Our data document three key activities: product page view, live chat, and purchase. In a consumer's shopping session, she can conduct multiple shopping trips. A shopping trip starts with a detailed product page view, and ends with a new product view. To show the concept of shopping trips, we illustrate an example in Table 5.21 following the example from Moe (2006). In this example, a consumer viewed three distinct product pages (three shopping trips), conducted one live chat, and purchased one item. We further describe customers' choices at each stage of their shopping trips.

Table 5.21. Example of Shopping Process

Time (t)	**View (Stage 1)**	**Choice Set**	**Chat (Stage 2)**	**Enhanced Choice Set**	**Purchase (Stage 3)**
1	Shop I - A	I-A	No chat	I-A	No purchase

2	Shop II – A	I-A, II-A	No chat	I-A, II-A	No purchase
3	Shop I - B	I-A, II-A, I-B	Chat with Shop I	I-A’, II-A, I-B’	Purchase I-A
4	STOP				

The first stage is product page view. A product page is specifically designed to introduce the product (Figure 5.1 (c)), but not other products. It is not an administration page or catalogue page. On each product page, there’s a button to initiate “live chat,” and two buttons for “purchasing” and “adding to carts” (Figure 5.1 (c)). In the example in Table 5.21, the consumer viewed model A from shop I, then browsed the same model in shop II, and then went back to shop I to check another model B. After seeing model B in shop I, she initiated a live chat, and finally chose model A from seller I, which she viewed in her very first visit to shop I. A customer may be directed to the product page multiple ways. For example, she may search the key word “iPad” in the search engine, and choose the product that most interests her from the search results (Figure 5.1 (b)). The customer may also learn the link of the product from a friend who is a loyal customer of that seller. In addition, a user may click onto advertisements while browsing through related pages, or simply follow the best sellers at the landing page (Figure 5.1 (a)). In our data, we do not have information about how a customer managed to arrive at the product page. Therefore, we allow all products to be options in the viewing stage. We believe the potential bias due to the algorithm of the search engine is of little concern because customers usually use refinement tools to sort the results based on price or reputation after receiving the search results. We explicitly include evolving factors to control for potential sorting refinement. Also, the results returned by a search engine are by default in the form of a matrix instead of a list (Figure 5.1 (b)). This makes the ranking influence even more ambiguous.

The second stage is live chat. The live chatting tool in Alibaba was developed by the selling platform, and each seller is encouraged to use it. The live chatting tool has its own name, “Ali Wangwang”. Ali Wangwang is not simply an instant messenger; it collects information about users’ web browsing history, and displays the product just visited by the customer or mentioned by the customer on the right panel of her chat dialogue (Figure 5.1 (d)). To improve live chatting experiences, Taobao developed two different versions of live chatting portals, the web-based version and the instant messenger version. Both versions are integrated in terms of data storage and management. Due to such designs, we can collect all the live chat logs from Alibaba. We processed the timestamp for a page view as the starting time of a shopping trip, and the timestamp for the next page view as the ending time of a shopping trip. If a live chat is requested between the starting and ending time, we claim that a live chat is conducted during this shopping trip. After a live chat, the customer learns new information about this seller, and her choice set gets updated to an enhanced choice set. As in the example of Table 5.21, the shopper requested a live chat in her third trip with shop I, and both items from shop I in her choice set get updated. Specifically, I-A is updated to I-A’, and I-B is updated to I-B’. It is notable that in our data we only keep users’ first visits to the same product and their first chatting communication with the same seller because the first visit plays the most critical role. The follow-up visits are usually for the purposes of refreshing users’ memories, and the follow-up live chats are sometimes concerns that have not yet been resolved.

The third stage we document is the purchasing stage. In each purchasing decision, a customer can purchase anything she has browsed before. Note that we do not capture users’ clicking behavior putting products into the shopping carts, but capture the timestamps when the system received their payments. In the same order, they may purchase more than one item. Each product has a list price, but this price is negotiable, and can be adjusted before checking out. Therefore, we use the effective price received in the payment system as product price.

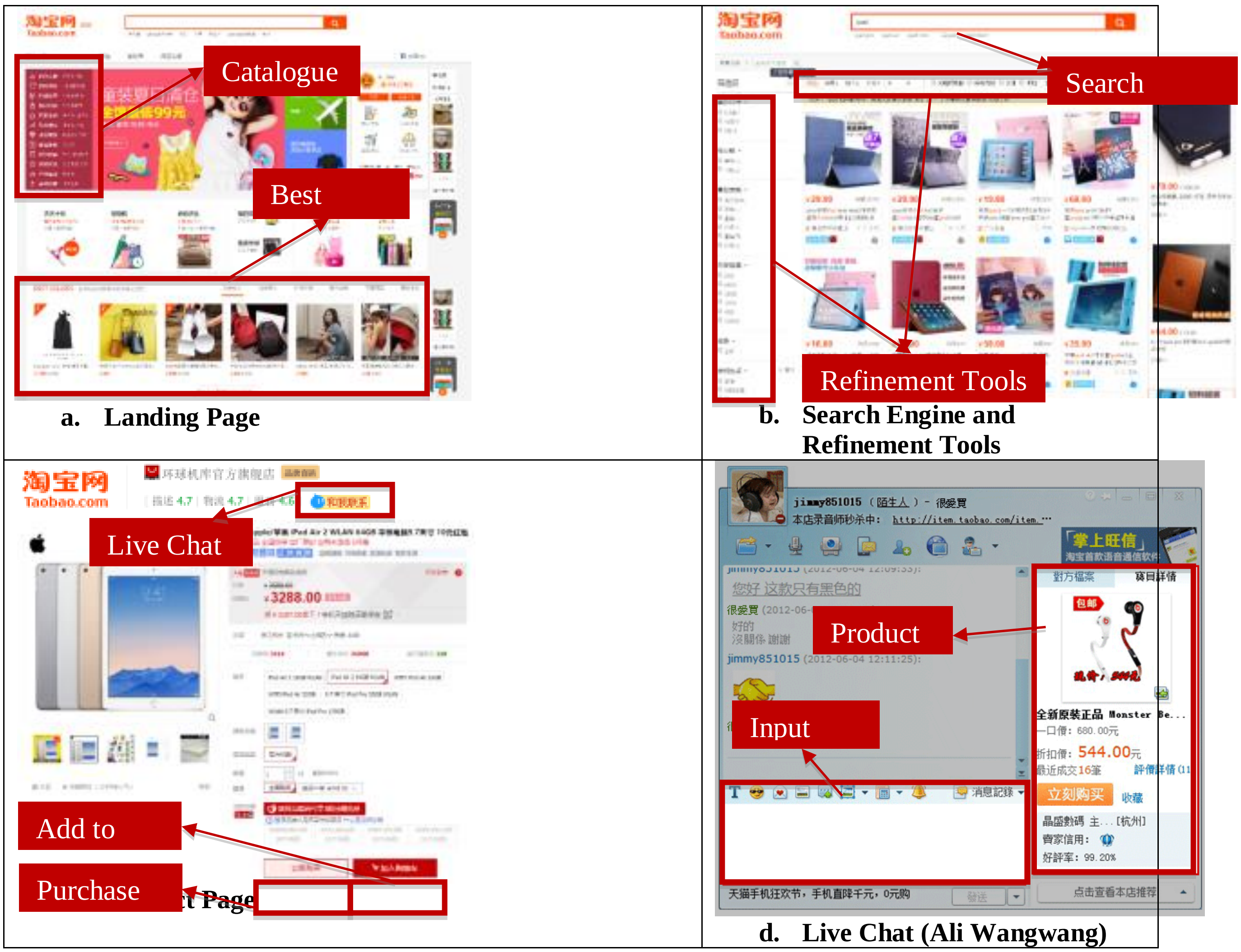


Figure 5.1. Taobao Platform Screenshots

After each shopping trip, customers can stop the current shopping session by leaving the platform. Otherwise, a new shopping trip will start when they click to visit a new product page.

5.3.2 *Reputation System*

The reputation system in this online marketplace provides two levels of key information: seller level and product level. After competing each transaction, customers can voluntarily provide feedback information. The feedback information consists of two major parts: feedback scores and product reviews. Feedback scores will be aggregated at seller level, and product reviews will be listed in each product page. Note that a seller's feedback score is the sum of feedbacks from all his historical transactions. Customer's feedback for each transaction takes three values: a positive rating corresponds to +1 point in the feedback score, a neutral rating leads to no point change, and a negative rating translates into -1 point in the feedback score. The feedback score is the same as the one used in eBay, and is usually displayed next to the seller's name (Figure 5.2 (b)). EBay developed icons of stars with different colors to indicate the level of feedback score, and Taobao uses icons of hearts, diamonds, and crowns to reflect different reputation levels.

At the product level, product reviews carry customers' evaluations of the specific product (Figure 5.2 (a)). Unlike in Amazon where each distinct product has the same review despite its supply by different sellers, Alibaba treats a product sold by different sellers as different products. As mentioned above, when consumers complete a transaction, they can voluntarily provide feedback information including the feedback score and the review. The product reviews are displayed at each product page (Figure 5.2(a)). The review count is likely a measure of sales volume. This is because Alibaba has an automatic mutual reviewing system. Both the seller and the buyer can give each other a feedback score within 15 days of a transaction. If only one party gives a positive feedback and the other party remains silent, a positive feedback will be given at the end of the 15-day period to the other party by the system. In the meantime, the automatic review "the customer did not write a review, and we assume it was a positive experience" will be posted by the system to the specific product page. If both parties remain silent, the feedback will be canceled. Since sellers usually give a positive feedback to gain a reciprocal positive feedback, the review count is highly correlated with the sales volume.

Figure 5.2. Reputation System

Although tablets are highly homogeneous products with a low level of information uncertainty, customers face potential loss due to the dishonesty of sellers. Specifically, some sellers label refurbished tablets as new, and earn high profit margins by cheating. Some sellers claim to have stock when they do not, and customers experience a longer waiting time than necessary. Therefore, the reputation of sellers and reviews of products are important measures valued by customers.

5.4 DATA

We obtained product visit histories for 20 days from March 10, 2013, to March 30, 2013, for all customers who browsed at least one tablet product in Alibaba before March 15, 2013. We also have customers' product visit daily summaries for four months between March 10, 2013 and July 10, 2013, which is used to validate that our sample does not suffer from truncation. We have live chatting logs and transaction records for four months between March 10, 2013, and July 10, 2013. In terms of reputation data, we have sellers' feedback scores at the weekly level, and all product review histories. From this data, we sampled a sufficiently large subset of shoppers to analyze their shopping paths.

5.4.1 *Sampling Technique*

Purchasing is a rare event which has an occurrence rate of around 8%. This low occurrence leads to the difficulty in prediction because a binary outcome model usually underestimates rare events (King and Zeng 2001). Therefore, we follow the literature to oversample the rare event of purchase. To build our training set, we randomly selected 1000 customers who eventually made a purchase and 2000 customers who ended up buying no tablets during this four-month period. To construct our holdout sample, we randomly selected 1000 users from the population with no restriction on their purchasing frequency. Note that this oversampling approach is used mainly for the purpose of prediction. In terms of empirical analysis, using oversampled data or the original data should yield the same results.

5.4.2 *Evolving Factors*

To model the evolvement of customers' browsing, chatting, and purchasing paths, we construct two variables: $Evolving\_FS$ (evolving factor for feedback score) and $Evolving\_RC$ (evolving factor for review count). These variables are constructed using the difference between the focal product and users' dynamic browsing path information. The construction approach extended the dispersion measurement used by Rao (2000) and Moe (2006). While they were interested in the dispersion, we are interested in the direction of evolvement (Bradlow and Rao 2000)**Error! Reference source not found.**. For example, we are interested in whether a customer started from a low feedback score and gradually searched for products with high reputation scores. Because of our different goals, we do not calculate the absolute value of the difference, but the relative difference between the focal product's feature ($Factor_{ijt}$) and the average feature of users' browsing histories $\sum_{k=1}^{t-1} Factor_{ik} / (t-1)$. The formula is presented below:

$$Evolving\_Factor_{ijt} = Factor_{ijt} - \frac{1}{t-1}\sum_{k=1}^{t-1} Factor_{ik} , \tag{5.22}$$

where the factor can be feedback score, or review counts. We use i to denote customer, j to denote product, and t to denote time. In our estimation we can interpret the coefficient of $Evolving\_Factor_{ijt}$ to determine the direction of evolvement. When the coefficient is positive, customers search from low factor level to high factor level, and vice versa.

5.4.3 *Data Overview*

In our training sample, 1,000 customers purchased at least one tablet, and 2,000 customers did not purchase any tablet during our four months of observation. In the viewing stage, the 2,000 non-purchasing customers visited 7,009 unique product pages, and the 1,000 purchasing customers visited 6,945 unique product pages. This indicates that purchasing customers on average visit twice the number of product pages as non-purchasing customers. This is consistent with prior literature (Moe and Fader 2004). On average, each customer visited 4.65 product pages. In total, non-purchasing customers viewed 41 distinct tablet models, and purchasing customers viewed 46 distinct models. In our context, products of the same model sold by different sellers are considered different products. Based on this definition, non-purchasing customers viewed 415 products sold by 256 sellers, and purchasing customers viewed 539 products sold by 330 sellers. Out of all the product visits, 2,772 are made to Samsung products and 11,232 are made to Apple products. We select these two brands because they dominant the tablet market. By June 2013, Apple and Samsung tablets accounted for 55.4% of all tablet sales in Alibaba. In the chat stage, we capture 350 first-time live chat conversations between buyers and sellers. In the purchasing stage, we observe 1,282 purchase decisions from the 1,000 purchasing customers.

We introduce the variables inTable 5.22. . It is worth mentioning that there are two different variables $Chatted_{ijt}^{v}$ and $Chatted_{ijt}^{p}$. The dummy variable $Chatted$ is used to denote whether a customer has chatted with a seller before. This variable is different for the view stage ($Chatted_{ijt}^{v}$) and the purchase stage ($Chatted_{ijt}^{p}$) because there is time gap between product view and purchase. Referring to our example in Table 5.21, $Chatted_{j=I-A,t=3}^{v}=0$ in the third trip because the customer has not chatted with shop I until after viewing the product. However, $Chatted_{j=I-A,t=3}^{p}=1$ in the third trip because a live chat conversation has just been conducted.

Table 5.22. Variable Description

Variable	**Description**
$View_{iji}$	whether individual *i* viewed product *j* at time *t*
$Purchase_{ijt}$	whether individual *i* purchased product *j* at time *t* after viewing it
$Chatted_{ijt}^{v}$	whether live chat has been conducted between individual *i* and the seller of product *j* before the page view
$Chatted_{ijt}^{p}$	whether live chat has been conducted between individual *i* and the seller of product *j* before the purchase
$Price\_Diff_{jt}$	the price difference between the focal product and the average price of the same model
$Feedback\_Score_{j}$	logarithm of sellers' feedback score
$Evolving\_FS_{ijt}$	evolving factor of feedback score
$Product\_Review\_Cnt_{jt}$	logarithm of product review count
$Evolving\_RC_{ijt}$	evolving factor of the review count

$Apple_j$	whether a product is from Apple

Our sample comprises 7,615,021 observations. This is the result of all subjects (*i*) choosing all alternative products (*j*) at all their shopping trips (*t*). The number of shopping trips for each individual is the number of pages she viewed. We eliminated visited product pages from customers' subsequent choice sets because we only count the first product view. As we discuss in detail in later sections, we constructed a binary choice for each alternative instead of applying a multinomial probit model in the consideration of computation efficiency. The drawback is that the predicted choice probabilities of all alternatives may not add up to 1. However, this does not affect our inference from the estimation results. This also does not affect our predictive model, where we deal with the relativity of probabilities.

We report the summary statistics in Table 5.23. As can be seen, the number of observations for the purchase stage is 13,936, much less than the number of observations for the first stage. This is because purchase can only be made after viewing the product page. For the choice of $Chat$, there are 13,283 observations. This figure is lower than the number of viewed pages because we excluded cases when a live chat has been conducted between a seller and a buyer in previous shopping trips.

Table 5.23. Summary Statistics

Variable	**Mean**	**St. Dev.**	**Min**	**Max**	**No. Obs.**
$View$	0.0018	0.0428	0.0000	1.0000	7,615,021
$Chatted^v$	0.0005	0.0233	0.0000	1.0000	7,615,021
$Chatted^p$	0.0006	0.0247	0.0000	1.0000	7,615,021
$Price$	2876.1260	1022.8991	100.0000	7999.0000	7,615,021
$Evolving\_FS$	-1.9853	2.7638	-14.3796	12.4747	7,615,021
$Feedback\_Score$	7.9999	2.4508	0.0000	13.6865	7,615,021
$Evolving\_RC$	-3.1404	2.7970	-9.7009	9.1498	7,615,021
$Review\_Cnt$	3.1518	2.1016	0.0000	9.1498	7,615,021

Purchase	0.0916	0.2884	0.0000	1.0000	13,936
Chat	0.0371	0.1891	0.0000	1.0000	13,283

5.4.4 *Reputation and Price*

We conducted a model-free analysis to examine the relationships between price and the two factors, seller-level feedback score and product-level review count. While the relationship between price and reputation is not the focus of this work, it lays a foundation for us to interpret our findings. It is worth noting that the relationship between reputation and price has been widely studied, and evidence showed positive or null relationship (Ba and Pavlou 2002, Resnick and Zeckhauser 2002).

We plot the feedback score and product review count over price in Figure 5.3 to show the correlation between these two factors and price. As can be seen in Figure 5.3, higher prices are correlated with higher feedback scores, but not necessarily with product review counts. Such a result is surprising because it indicates that there exists intrinsic difference between these two measures. We will discuss potential reasons for their differences in the discussion of our empirical analysis.

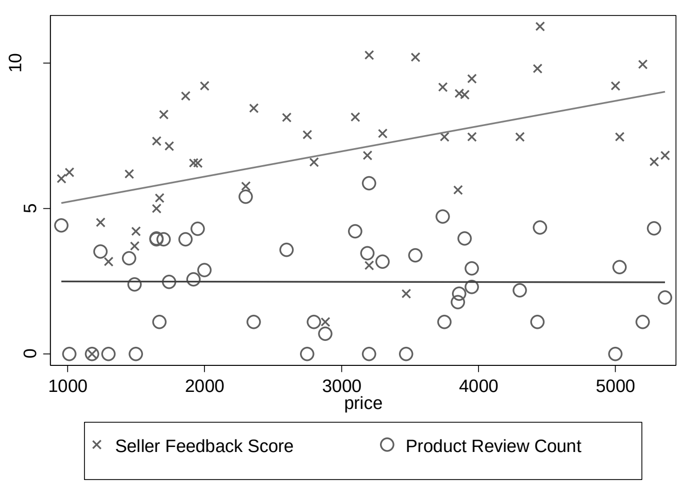


Figure 5.3. Scatterplot of Feedback Score and Product Review Count vs. Price

5.5 EMPIRICAL ANALYSIS

5.5.1 *Stage 1 – View*

The dependent variable in the viewing stage is the binary variable $View_{ijt}$ which takes the value of 1 if a customer *i* visits product page *j* at time *t*, and 0 otherwise. As introduced before, the dimension of time is defined by customers' shopping trips.[8] If a customer visits 3 distinct product pages, then her maximum value of *t* is 3. She faces decisions at times 1, 2, and 3. The dimension of *j* is determined by all the alternatives available to the customer. In total the 3000 users in our sample browsed through 555 distinct pages, and we allow each product page to be an alternative for customers to choose. In the following formulation we omit the subscripts because there is no ambiguity. We use superscript *v* to indicate coefficients for the viewing stage. The determination of the binary decision $View$ is based on a latent continuous variable $View^*$:

$$\begin{aligned} View^* &= \left(\mathbf{X}^v\right)^T \boldsymbol{\beta}^v + \varepsilon_1^v \\ &= \beta_0^v + \beta_1^v Chatted^v + \beta_2^v Price\_Diff + \beta_3^v Evolving\_Price\_Diff \\ &\quad + \beta_4^v Feedback\_Score + \beta_5^v Evolving\_Feedback\_Score \\ &\quad + \beta_6^v Product\_Review\_Cnt + \beta_7^v Evolving\_Review\_Cnt + \beta_8^v Apple \\ &\quad + \beta_8 Hot\_Product + \varepsilon_1^v, \end{aligned} \tag{5.23}$$

and

$$View = \begin{cases} 1, & View^* > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{5.24}$$

Our major interest is in the binary variable $Chatted^v$, which indicates whether the buyer has chatted with the seller on previous shopping trips. To further examine the effect of $Chatted^v$ at different levels of price, feedback score, and review count, we include the interaction term

---

[8] We do not model the choice of stopping because it does not affect the relative choice preference among available products.

between $Chatted^v$ and $Price\_Diff$, $Feedback\_Score$, and $Product\_Review\_Cnt$ in the Interaction Model. These interactions investigate the role of these factors as moderators of $Chatted^v$.

The dummy variable $Chatted^v$ is very likely to be an endogenous regressor because it might correlate with unobservable factors that also correlate with the choice of viewing the product. For example, if a store has a limited-time special offer which is not observed by us researchers, customers may initiate live chat to ask about the details of the promotion. Customers are also likely to browse through many products sold by the same seller simply because they want to redeem the promotion offer.

To control for such endogeneity, we simultaneously estimate an additional equation for the endogenous variable $Chatted^v$ (equation 5.23 and 5.25). As is well documented, this can be achieved with a biprobit model (Wilde 2000).

$$\begin{aligned} Chatted^{v*} &= \left(\mathbf{Z}^v\right)^{\mathbf{T}} \boldsymbol{\alpha}^v + \left(\mathbf{X}^v\right)^T \boldsymbol{\gamma}^v + \varepsilon_1^v \\ &= \alpha_0^v + \alpha_1^v IV\_Battery + \alpha_2^v IV\_Storage + \alpha_3^v IV\_Screen\_Size + \left(\mathbf{X}^v\right)^T \boldsymbol{\gamma}^v + \varepsilon_1^v, \end{aligned} \tag{5.25}$$

and

$$Chatted^v = \begin{cases} 1, & Chatted^{v*} > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{5.26}$$

In equation 5.25, $\mathbf{Z}$ are excluded instruments, and $\mathbf{X}$ are included instruments. The excluded instruments $\mathbf{Z}$ need to be correlated with the endogenous variable $Chatted^v$ but not correlated with the outcome variable $View$. We constructed three instrumental variables: 1) the average of battery life of other models with the same brand (Apple or Samsung), 2) the average storage (from 8GB – 128 GB) of other models with the same brand, and 3) the average screen size of other models of the same brand (Berry et al. 1995). These variables are likely to be correlated with $Chatted^v$ because a customer is usually interested in several models of the same brand. For example, a customer who viewed iPad 4 16G WIFI is also likely to be interested in iPad 4 32G WIFI Therefore the attributes of iPad 4 32G WIFI are likely related to her probability of chatting with the seller. However, these characteristics should not be correlated with the customer's viewing decision for the focal product because they are not features of the focal product. We conducted over-identification test, and the inclusion of three instruments does not lead to over identification.

By allowing the error term ε_1^v of equation 5.23 to be correlated with the error term ε_2^v of equation 5.25, we account for the endogeneity problem. For example, if a limited-time promotion that is unobservable to researchers is affecting both the choice of chatting and viewing, this should be taken care of by the correlation in the error term. We assume that $Var(\varepsilon_1^v) = Var(\varepsilon_2^v) = 1$, and $Cov(\varepsilon_1^v, \varepsilon_2^v) = \rho^v$. The significance of ρ^v will indicate whether we have encountered endogeneity issue.

From the regression results presented in Table 5.24, we have fruitful findings. Frist, we notice that ρ^v is significant for both the IV Model and the IV + Interaction Model. This verifies the endogeneity, and indicates that we should look at the IV models to interpret our results.

Table 5.24. Regression Result for Stage 1 - View

Variable	**Basic**	**Interaction**	**IV**	**IV + Interaction**
$Chatted^v$	1.368***(0.032)	3.516***(0.144)	0.0249 (0.172)	2.026***(0.547)
$Price\_Diff$ (x10^{-4})	-0.416***(0.055)	-0.431***(0.055)	-0.420***(0.054)	-0.433***(0.055)
$Feedback\_Score$	-0.0334***(0.003)	-0.0324***(0.003)	-0.0325***(0.003)	-0.0318***(0.003)
$Evolving\_FS$	0.0317***(0.003)	0.0318***(0.003)	0.0316***(0.003)	0.0317***(0.003)
$Review\_Cnt$	0.235***(0.003)	0.236***(0.003)	0.234***(0.003)	0.236***(0.003)
$Evolving\_RC$	-0.0119***(0.003)	-0.0124***(0.003)	-0.0122***(0.003)	-0.0126***(0.003)
$Apple$	-0.0616***(0.008)	-0.0634***(0.008)	-0.0587***(0.008)	-0.0612***(0.008)
$Chatted^v \times PD$ (x10^{-4})		0.357 (0.686)		0.255 (0.643)
$Chatted^v \times FS$		-0.162***(0.019)		**-0.132*****(0.024)
$Chatted^v \times RC$		-0.158***(0.021)		**-0.152*****(0.020)

Constant	-3.614***(0.018)	-3.633***(0.018)	-3.616***(0.018)	-3.635***(0.018)
IV_Battery (x10^{-4})			2.728*** (0.536)	2.755* (0.536)
IV_Storage			-0.0563***(0.006)	-0.0564***(0.006)
IV_Screen_Size			0.877***(0.124)	0.869***(0.123)
Constant			-11.090***(0.792)	-11.02***(0.789)
atanh(ρ^v)			0.453***(0.073)	0.387* (0.154)
N	7615021	7615021	7615021	7615021
Pseudo R^2	12.77%	12.9%	-	-

Standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, included instruments are not reported

From the IV Model, we see that $Chatted^v$ has an insignificant effect overall. However, when we look at the IV + Interaction Model, we see that $Chatted^v$ has a positively significant main effect, mitigated by the levels of feedback score (we use FS for readability in Table 5.24) and review count (we use RC for readability in Table 5.24). This crossover interaction effect has two indications. First, revisits to products with low counts of reviews or products sold by sellers with low feedback scores are positively affected by past live chats. Second, revisits to products with high reputation and reviews are negatively affected by past live chats. This finding shows the substitution effect of live chats for reputation. This finding is one of the key contributions of this paper.

We also notice that the main effect of $Feedback\_Score$ is negative, and the main effect of $Review\_Cnt$ is positive across all model specifications. This says that the reputation at seller level works differently from the reviews at product level. Customers prefer to visit sellers with relatively low reputation and with high product review counts. Recalling our model-free evidence in Figure 5.3 where review count did not show a correlation with price but feedback score showed a positive correlation with price, we speculate that the preference pattern we find reflects customers' optimal choice in balancing trust and price. Following the criterion of product review counts, customers can effectively search for products with high quality and low price. If the criterion is the measure of feedback score, customers will have to pay a higher price for better sellers. However, when they focus on review counts, they do not have to pay a higher price to get a product with many reviews.

The coefficients before $Evolving\_FS$ and $Evolving\_RC$ uncover users' browsing patterns. We find that users start from products with high review counts, and gradually explore products with low review counts. On the other hand, they start by looking at sellers with low feedback scores, and gradually move to sellers with higher feedback scores. This can be seen from the positively significant coefficient before $Evolving\_FS$ and the negatively significant coefficient before $Evolving\_RC$. Again, we can combine this result with our model-free evidence from Figure 5.3. By putting a higher priority on the dimension of review count, customers can get credible products without suffering higher prices.

Last, we see that price difference has a significantly negative effect, in accordance with economic theories. However, the interaction between $Price\_Diff$ and $Chatted^v$ is insignificant, meaning that the effect of chatting is not influenced by price. We also see a negatively significant effect from the brand of Apple.

5.5.2 *Stage 2 – Chat*

We use a probit model as our main model in stage 2, with the dependent variable being the binary choice $Chat_{ijt}$ indicating whether customer *i* initiated a live chat with the seller of product *j* for the first time.[9] We address the potential endogeneity issue with regard to price difference using the instrumental variable approach. Note that the instrumental approach is different from that used in the viewing stage because price is continuous. It will turn out that there is no

[9] Repeat live chat with the same seller is left out in this study so that we do not need to create another index for the seller.

endogeneity problem for price difference, and a simple model is sufficient. Below we specify our main model:

$$
\begin{aligned}
Chat^{*} &= \left(\mathbf{X}^{c}\right)^{T} \boldsymbol{\beta}^{c} + \varepsilon_{1}^{c} \\
&= \beta_{0}^{c} + \beta_{1}^{c} Price\_Diff + \beta_{2}^{c} Feedback\_Score \\
&\quad + \beta_{3}^{c} Evolving\_Feedback\_Score + \beta_{4}^{c} Product\_Review\_Cnt \\
&\quad + \beta_{5}^{c} Evolving\_Review\_Cnt + \beta_{6}^{c} Apple + \beta_{7}^{c} Hot\_Product + \varepsilon_{1}^{c},
\end{aligned}
\tag{5.27}
$$

and

$$
Chat = \begin{cases} 1, & Chat^{*} > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{5.28}
$$

We use the same three instrumental variables as we used in Stage 1 because these are cost shifters that are not affected by the demand, and have been used widely as instruments for price (Berry et al. 1995). It is notable that although we used the same instrumental variables, the arguing reason for the validity is different. With the help of these instruments, we simultaneously estimate a second equation:

$$
Price\_Diff = \lambda_{0} + \lambda_{1} IV\_Battery + \lambda_{2} IV\_Storage + \lambda_{3} IV\_Screen\_Size + \left(\mathbf{X}^{c}\right)^{T} \boldsymbol{\gamma}^{c} + \varepsilon_{2}^{c}, \tag{5.29}
$$

where the error terms of equation 5.27 and 5.29 are related, with the covariance matrix:

$$
Var(\varepsilon_{1}^{c}, \varepsilon_{2}^{c}) = \Sigma = \begin{pmatrix} \sigma_{1} & \rho^{c}\sigma_{1}\sigma_{2} \\ \rho^{c}\sigma_{1}\sigma_{2} & \sigma_{2} \end{pmatrix}. \tag{5.30}
$$

Since we cannot identify both σ_1 and σ_2, we normalize σ_1 and estimate their ratio as σ.

The result of the Main Model and the IV Model are reported in Table 5.25. As can be seen, the correlation ρ^c is insignificant, meaning that there is no endogeneity issue. Therefore, the findings are interpreted based on the estimation of the Main Model, and we do not handle the endogeneity issue with price difference in all other model specifications of other stages. There are two reasons for the exogeneity of price difference. First, we calculate price difference between the current product and the average price of the same model, and thus there is no direct reason for the price difference to be affected by demand. Second, we have accounted for seller-level feedback score and product-level review count which are the major potential reasons for higher prices.

Table 5.25. Regression Results for Stage 2 - Chat

	Main		IV	
Price_Diff (x10^{-3})	-0.0369	(0.039)	-0.163	(0.106)
Feedback_Score	**-0.104*****	(0.016)	-0.105***	(0.016)
Evolving_FS	0.0184	(0.017)	0.0196	(0.017)
Review_Cnt	**-0.118*****	(0.017)	-0.116***	(0.017)
Evolving_RC	0.0468**	(0.018)	0.0451*	(0.018)
Apple	-0.00188	(0.058)	-0.0277	(0.061)
Constant	-0.172	(0.101)	-0.160	(0.102)
IV_Battery			0.464***	(0.047)
IV_Storage			312.0***	(8.107)
IV_Screen_Size			-1236.8***	(108.459)
Constant			-1804.6**	(660.249)
atanh(ρ^c)			**0.0635**	(0.050)
$\ln\sigma$			6.062***	(0.006)
N	13283		13283	
R^2	9.38%		-	

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$;
Inclusive instruments are not reported.

From the Main Model results in Table 5.25, we learn that customers are more likely to initiate a live chat when sellers do not have a high feedback score (-0.104, with $p < 0.001$), and when the product does not have a high count of reviews (-0.118, with $p < 0.001$). This is likely to be because with a low level of reputation or review, customers cannot establish enough trust to conduct purchases, and seek to learn more about the sellers to identify their quality. This indicates that live chatting tools may work as substitutes for seller level reputation and product level review information. This finding is one of the key findings in our work.

We also find that the evolving review count ($Evolving\_RC$) has a positively significant coefficient (0.0468, with $p < 0.01$). This is in contrast with the insignificant effect from evolving feedback score ($Evolving\_FS$). These results indicate that customers are more likely to initiate live chats when the focal product has more product reviews than the previous products they viewed. This might be because they discover negative comments or concerns raised by other customers, and want to resolve their uncertainties. However, in our robustness check and holdout sample analysis later, the evolving factor of review count does not show a significant effect. Thus the effect of $Evolving\_RC$ is ambiguous in the chat stage.

Last, we notice that price difference has an insignificant impact, this might be because price is a straight forward characteristic that does not require any explanation by web assistants.

5.5.3 *Stage 3 – Purchase*

We developed a probit model to examine customers' purchase choices. The dependent variable is a binary variable $Purchase_{ijt}$ indicating whether the customer *i* eventually purchased the product *j*. Note that the customer does not have to purchase right after viewing the pate. Our major interest is the effect of past live chat ($Chatted^p$) on the eventual purchase. As discussed before, the dummy variable $Chatted^p$ is likely to be endogenous because unobservable factors like customers' urgency of use likely affect both their purchase decisions and their live chat decisions. We used the same approach as in the first stage to handle endogeneity:

$$\begin{aligned} Purchase^* &= \left(\mathbf{X}^p\right)^T \boldsymbol{\beta}^p + \varepsilon_1^p \\ &= \beta_0^p + \beta_1^p Chatted^p + \beta_2^p Price\_Diff + \beta_3^p Feedback\_Score \\ &\quad + \beta_4^p Evolving\_Feedback\_Score + \beta_5^p Product\_Review\_Cnt \\ &\quad + \beta_6^p Evolving\_Review\_Cnt + \beta_7^p Apple + \beta_7^p Hot\_Product + \varepsilon_1^p, \end{aligned} \tag{5.31}$$

and $$Purchase = \begin{cases} 1, & Purchase^* > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{5.32}$$

$$\begin{aligned} Chatted^{p*} &= \left(\mathbf{Z}^p\right)^{\mathbf{T}} \boldsymbol{\alpha}^p + \left(\mathbf{X}^p\right)^T \boldsymbol{\gamma}^p + \varepsilon_2^p \\ &= \alpha_0^p + \alpha_1^p IV\_Battery + \alpha_2^p IV\_Storage + \alpha_3^p IV\_Screen\_Size + \left(\mathbf{X}^p\right)^T \boldsymbol{\gamma}^p + \varepsilon_2^p, \end{aligned} \tag{5.33}$$

and $$Chatted^p = \begin{cases} 1, & Chatted^{p*} > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{5.34}$$

As can be seen from the regression results listed in Table 5.26, the correlation ρ^p is significant, indicating the existence of endogeneity. Therefore, we interpret the results using the IV models. It is interesting to note that ρ^p is negatively significant in the purchase stage but ρ^v is positively significant in the view stage. A negative ρ^p indicates that the unobservable factors impact chatting choice and purchasing choice differently. Since users usually chat with sellers who have low reputation but generally would prefer to purchase products with high quality, we speculate that the unobservable factor is related to customers' private information about the quality of the product. On the other hand, the positive ρ^v identified in the view stage indicates that the unobservable factor affects chat and viewing in the same direction. One possible explanation might be users' urgency of use of the product.

Table 5.26. Regression Results for Stage 3 - Purchase

	Basic	**Interaction**	**IV**	**IV + Interaction**
$Chatted^p$	1.604^{***}(0.052)	2.662^{***}(0.200)	1.816^{***}(0.060)	2.692^{***}(0.200)
Price_Diff (x10^{-4})	-0.513 (0.337)	-0.600 (0.378)	-0.494 (0.337)	-0.592 (0.377)
$Feedback\_Score$	0.0462^{***}(0.014)	0.0693^{***}(0.014)	0.0497^{***}(0.014)	0.0681^{***}(0.014)
$Evolving\_FS$ (x10^{-2})	0.164 (1.456)	-0.469 (1.474)	0.122 (1.454)	-0.378 (1.472)
Review_Cnt	-0.0945^{***}(0.013)	-0.104^{***}(0.014)	-0.093^{***}(0.013)	-0.100^{***}(0.014)
$Evolving\_RC$	0.0603^{***}(0.014)	0.0640^{***}(0.014)	0.0595^{***}(0.014)	0.0623^{***}(0.014)
Apple	0.152^{**}(0.050)	0.161^{**}(0.051)	0.161^{***}(0.050)	0.166^{***}(0.051)
$Chatted^p \times PD$ (x10^{-4})		0.378 (0.855)		0.431 (0.859)

$Chatted^p \times FS$		-0.151^{***}(0.028)		-0.124^{***}(0.029)
$Chatted^p \times RC$		0.0402 (0.028)		0.026 (0.028)
Constant	-1.465^{***}(0.093)	-1.641^{***}(0.100)	-1.523^{***}(0.093)	-1.665^{***}(0.099)
IV_Battery (x10^{-3})			0.546 (0.354)	0.554 (0.354)
IV_Storage			-0.055 (0.044)	-0.054 (0.044)
IV_Screen_Size			-0.483 (0.708)	0.497 (0.710)
Constant			2.148 (4.361)	2.186 (4.376)
atanh(ρ^p)			-0.234^{***}(0.042)	-0.269^{***}(0.045)
N	13936	13936	13936	13936
R^2	13.54%	13.93%	-	-

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Included instruments are not reported.

From the regression results of the purchase stage, we discover interesting findings. First of all, we find that conducting live chats greatly increases the overall purchase probability, after accounting for endogeneity (Model IV of Table 5.26). Second, we find that when making the final decision of purchase, customers choose sellers with high reputation ($Feedback\_Score$) and products with relatively low review counts. This can be seen from the positive coefficient of $Feedback\_Score$ and negative coefficient of $Review\_Count$. Finally, we find that price difference has a marginally negative effect, with a p-value of 0.143 for the IV Model and a p-value of 0.1117 for the IV + Interaction Model in Table 5.26.

5.5.4 *Robustness Checks*

We conducted two robustness checks to validate our findings. These robustness checks also provide additional information that complements our main analysis.

5.5.4.1 Individual Heterogeneity

We account for individual heterogeneity by incorporating fixed effect at individual level. We estimate a fixed effect model for choices in each stage, and report the results in Table 5.27. Note that in our data, 973 customers out of the 3000 customers had only one shopping trip. Imposing a fixed effect may produce different results because we lose some information. We find that the effects of *Chatted*, *Review_Cnt*, and *Evolving_FS* are the same as our previous estimations (Basic Model for stage 1, Main Model for stage 2, and Basic Model for stage 3). The marginal significance of *Price_Diff* in the purchase stage of our previous model becomes significant in the fixed effect model ($p = 0.036$).

We do notice that there is a slight inconsistency in the effect of *Evolving_RC* and *Feedback_Score*. Specifically, *Evolving_RC* is only significant in the view stage and not significant in the latter two stages after we control for individual heterogeneity, unlike as suggested by our main estimation where it is significant across all stages. Similarly, *Feedback_Score* is no longer significant in the purchase stage after we control for individual heterogeneity. Despite the discrepancy in terms of significant level, the signs of the coefficients stay the same as reported by our previous estimations. The difference in results is possibly due to our short panel. In addition, the interpretation is different between our previous analyses and this fixed effect model. The fixed effect model looks at within subject variation, and showed that customers' preference concerning feedback score is mainly reflected in the viewing stage. On the other hand, our previous analysis looks at average patterns, and found that people prefer to purchase products with higher feedback score in general. This robustness check enriches our main analyses, and indicates that the cognitive processing of feedback score information is mainly at the view stage.

Table 5.27. Robustness Check – Fixed Effects Model

	FE + View	**FE + Chat**	**FE + Purchase**
Chatted	3.565^{***}(0.080)		2.332^{***}(0.131)
Price_Diff ($\times 10^{-3}$)	-0.123^{***}(0.017)	-0.0488 (0.117)	-0.182^{*} (0.087)

Feedback_Score	-0.157***	(0.009)	-0.0891	(0.107)	0.0589	(0.068)
Evolving_FS	0.159***	(0.010)	-0.107	(0.103)	-0.0140	(0.064)
Review_Cnt	0.756***	(0.010)	-0.383***	(0.115)	-0.164*	(0.070)
Evolving_RC	-0.0611***	(0.010)	0.196	(0.111)	0.0818	(0.068)
Apple	-0.185***	(0.024)	-0.286	(0.298)	0.0988	(0.224)
N	7614470		2543		7812	
Pseudo R^2	13.83%		9.65%		11%	

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.5.4.2 Brand Heterogeneity

To account for the difference between the Samsung market and the Apple market, we split our sample into two parts, one part containing only Samsung products and the other part containing only Apple products. We ran basic models for each stage to check the consistency of data patterns. Running separate regressions for different product categories is commonly used in marketing literature (Moe 2006).

We report the results in Table 5.28. As can be seen, the effect of *Chatted* (for simplicity, we use *Chatted* to denote $Chatted^v$ or $Chatted^p$, depending on the stage we refer to) is positively significant for both brands. The chat effect is slightly stronger for Samsung products, possibly because Apple has a higher reputation for quality. By checking the results from the chatting stage, we find that customers are more likely to request live chats when *Feedback_Score* and *Review_Cnt* are low for Apple products, consistent with our earlier findings. For Samsung products, customers are more likely to request live chats only when they face sellers with low feedback score, but not when they see products with low counts of product reviews. This might be due to the fact that Samsung products have much lower sales volumes and review counts in the first place. This might also be due to the fact that Samsung shoppers are more knowledgeable about the tablet qualities. In terms of the evolving factors, the submarkets of Apple and Samsung showed different patterns, indicating that we need to be cautious while generalizing our results. Overall, the estimations based on different brands showed the robustness of our major conclusions, and provide information about the differences across brands.

Table 5.28. Robustness Check – Brand Heterogeneity

	Apple			Samsung		
	View	Chat	Purchase	View	Chat	Purchase
Chatted	1.309^{***}		1.500^{***}	1.681^{***}		1.833^{***}
	(0.036)		(0.064)	(0.075)		(0.091)
Price_Diff ($\times 10^{-3}$)	-0.00966	-0.0699	-0.126^{**}	0.00388	-0.0845	0.0772
	(0.00801)	(0.0651)	(0.048)	(0.00846)	(0.0490)	(0.0549)
Feedback_Score	-0.819×10^{-5}	-0.0945^{***}	0.0703^{***}	-0.0500^{***}	-0.117^{***}	0.0155
	(0.003)	(0.023)	(0.017)	(0.004)	(0.024)	(0.024)
Evolving_FS	0.00348	-0.0254	-0.0296	0.0557^{***}	0.0864^{***}	0.0425
	(0.003)	(0.024)	(0.018)	(0.005)	(0.026)	(0.027)
Review_Cnt	0.276^{***}	-0.169^{***}	-0.0970^{***}	0.103^{***}	0.00513	-0.138^{***}
	(0.003)	(0.022)	(0.015)	(0.005)	(0.031)	(0.029)
Evolving_RC	-0.0373^{***}	0.0808^{***}	0.0623^{***}	0.0388^{***}	-0.00920	0.0778^{**}
	(0.003)	(0.023)	(0.016)	(0.005)	(0.030)	(0.029)
Constant	-4.289^{***}	0.0191	-1.553^{***}	-2.825^{***}	-0.513^{***}	-1.088^{***}
	(0.028)	(0.183)	(0.139)	(0.024)	(0.141)	(0.154)
N	5352937	10734	11218	2262084	2549	2718
R^2	15.39%	9.45%	10.13%	6.70%	3.24%	27.65%

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.5.4.3 Discussion

We summarize our results, and compare how the key factors affect consumers' decisions differently at different stages. We create a billboard to present our results in Table 5.29. First of all, consistent evidence shows that live chats have a positive effect in inducing repeat visits for sellers with low reputation or products with low counts of reviews. Live chats also significantly improve conversion rate. In terms of price, we find that the difference between product price and the average market price for the same model has a negative effect only in the first stage. This indicates that consumers process price information during the viewing stage. This is to say, after viewing a product page, customers no longer consider price as a dominant factor in their subsequent choices. It is worth pointing out that price is usually considered a simple measurement that does not take much cognitive effort to comprehend.

Table 5.29. Key Factor Billboard

	View	Chat	Purchase
Chatted	↑	NA	↑
Price_Diff	↓	-	-
Feedback_Score	↓	↓*	↑*
Evolving_FS	↑	-	-
Review_Cnt	↑	↓	↓
Evolving_RC	↓	↑*	↑*
$Chatted^{p} \times PD$	-	NA	-
$Chatted^{p} \times FS$	↓	NA	↓
$Chatted^{p} \times RC$	↓	NA	-

Notes: "↓" indicates a negatively significant effect and "↑"a positively significant effect. Insignificant effect is designated by "-". The symbol * indicates an effect becoming insignificant when adding fixed effects in our robustness check.

In terms of feedback score, we see a negative effect in the first two stages, and a positive effect in the third stage. In our robustness check where we controlled for individual level heterogeneity, it is only significant in the viewing stage. The effect from feedback score in the viewing stage is consistently negative and significant, meaning that people choose to look at products with low feedback scores. This is likely to be because products with lower feedback scores are usually cheaper. When we look at the second stage of live chat, we see that on average (our previous model), people likely initiate live chats with sellers whose feedback scores are low. However, within a specific user's choice set (fixed effect model), the user does not show such a preference for all the sellers he visited. Finally, we look at the third stage of purchase. When we look at the average purchase pattern from our main model, we find that people prefer to purchase from sellers with high feedback scores. However, when we look at within subject patterns (fixed effect model), customers do not show a preference to purchase from sellers with high feedback scores. Overall we can see from the fixed effect model that customers process the information of feedback scores as early as in the viewing stage. This is probably because this information is easy to comprehend. Alibaba has developed different icons to indicate sellers' feedback score level, just like the star icons used by eBay. The use of these visual cues makes it less costly to process this information (Fader and McAlister 1990).

The interpretation of product review count is closely related to feedback score and price. Product review count has a high correlation with feedback score (correlation = 0.5208), but it still carries a significantly different message. A reputable seller may introduce new products that have a low product review count. Similarly, a seller with a low reputation level may have one popular product with many reviews. From our scatterplot in Figure 5.3, we find that product review count is irrelevant to price (correlation = –0.0066). This result is probably due to two opposing effects: 1) the high count of product reviews itself might be a result of low prices; 2) with a signal of high quality from the high counts of reviews, sellers can charge a higher price. As a matter of fact, customers can leverage the fact that products with high review counts are not necessarily more expensive, and start their searching by looking at products with more reviews but relatively lower prices. This explanation is confirmed by the changes in the sign of the coefficients before $Review\_Cnt$.

The negative moderating effects of feedback score (on the first and third stages) and review count (on the first stage) on live chat have revealed fascinating features of live chat in substituting for reputation. In contrast to feedback scores, the moderating effect from review count is only significant in the viewing stage. This result implies that customers are more serious when making their purchasing decisions because chatting cannot substitute for product review count in the purchase stage. Talking is useful in removing concerns with a low feedback score, possibly because customers can understand that new sellers cannot have a high reputation score. However, talking does not make up for the low product review count, possibly because review count (sales volume) is a stronger criterion that better represents the quality of the product.

5.6 PREDICTIVE ANALYSIS

In our simple predictive analysis, we add another 1,000 customers from the true population as our validation set, and use the model we developed to score the outcome. The goal for this predictive analysis is to confirm the robustness of our main model in new data and to check the potential for calibrating this model into a predictive model.

5.6.1 *Validation Set*

Our validation set is randomly sampled from the true population data, and constructed separately from our training set. It contains 1,000 users having no overlap with the 3,000 users used in our unbalanced training set. These 1,000 users conducted 2,522 unique product visits with 29 unique tablet models and 115 different products provided by 80 different sellers. Note that products of the same model sold by different sellers are considered different products. Since customers in the validation set viewed a smaller set of products than did customers in the training set, their choice sets for the viewing stage is also reduced accordingly. In other words, if a product is browsed by a customer in the training set but not browsed by any customer in the validation set, this product will not be included in the validation set as a choice for the customers. The way we constructed our validation set guaranteed its independence from the training set. Out of these 1,000 users, 84 of them ended up purchasing a product. The conversion rate is 8.4%, lower than the reported benchmark conversion rate of 24.1% in electronics category of the U.S market (Burstein 2014).

We first ran a base model in this validation set as a robustness check to our previous models, and report the results in Table 5.30. The major findings for the effect of *Chatted* are consistent across all stages, and the effects for both *Feedback_Score* and *Review_Cnt* are consistent with our main findings. When we look at the effect from the evolving factors, the effect of *Evolving_FS* is exactly the same as in our main model, while the effect of *Evolving_RC* is slightly different. Specifically, the effect of *Evolving_RC* becomes insignificant in the second and third stages. Notably, the difference in significance reflects the same pattern as in our fixed effect robustness check in Table 5.27. Overall, this model shows almost the same pattern as in our previous estimations, demonstrating the validity of our analysis.

Table 5.30. Regression Results in Holdout Sample

	view	chat	purchase
Chatted	1.425***(0.122)		1.852***(0.179)
Price_Diff (x10^{-3})	-0.199***(0.051)	0.203 (0.366)	0.224 (0.271)
Feedback_Score	-0.0252** (0.008)	-0.0407 (0.061)	0.102* (0.051)
Evolving_FS	0.0431***(0.008)	-0.0849 (0.064)	-0.0759 (0.056)
Review_Cnt	0.256***(0.008)	-0.195**(0.060)	-0.203***(0.048)
Evolving_RC	-0.0184* (0.009)	0.109 (0.065)	0.0605 (0.053)
Apple	-0.118***(0.026)	-0.181 (0.174)	0.277 (0.174)
Constant	-3.506***(0.054)	-0.359 (0.347)	-1.864***(0.332)
N	272714	2478	2522
R^2	12.54%	15.34%	21.68%

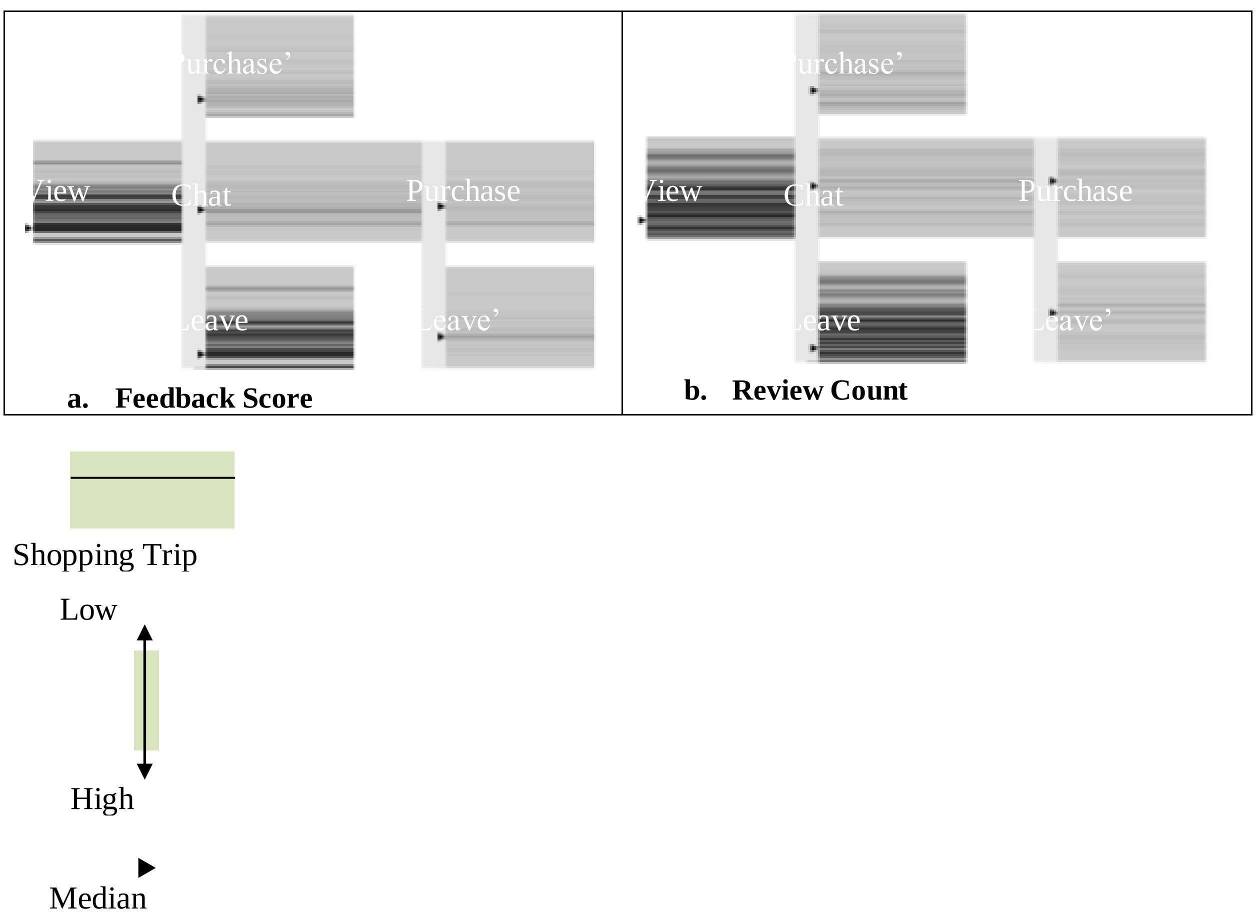
Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

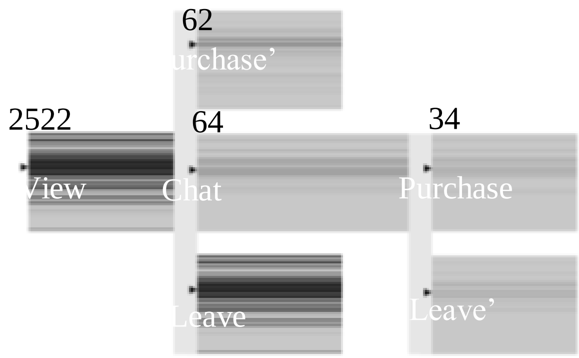
5.6.2 *Path Visualization*

We visualized customers' paths from the validation set to better learn customers' decision patterns in the population. We did not use the training set because it is oversampled, and does not reflect the general pattern. In Figure 5.4, we plotted the path based on three dimensions: feedback score, review count, and price difference. Each figure illustrates customers' choices with six rectangular boxes. Each rectangular box shows the decisions made by customers at a specific stage. Each *horizontal segment* inside the boxes represents one stage choice from a shopping trip. The relative location of the horizontal segment in the vertical space (with respect to the *vertical bar*) indicates the magnitude of the product characteristic. The small *triangle* to the left of the rectangular box presents the median level of this measurement.

We observed a total of 2,522 shopping trips: 64 of these trips had live chat communication before customers faced the purchase decision. After live chats, 34 trips ended up with purchases. We also observe 62 shopping trips in which customers purchased without conducting live chats.

From Figure 5.4 (a) and (b), we can see that compared to the *view set*, the *chat set* has a lower median level (the median level is lower when the small triangle is higher vertically). When customers choose to purchase without conducting live chats (*purchase' set*), the products are usually sold by sellers with higher feedback scores, compared with the level of feedback scores for products sold after live chats (*purchase set*). Another interesting pattern can be found in Figure 5.4 (b) by comparing the *purchase set* and the *leave' set* after conducting live chats. It is clear that the purchase set has a lower median level, indicating the substitution effect of live chat for product review count. From Figure 5.4 (c), we can find that although there is a large range of price difference, people who end up purchasing (*purchase set* and *purchase’ set*) select products concentrated in a small range of price difference.





c. Price Difference

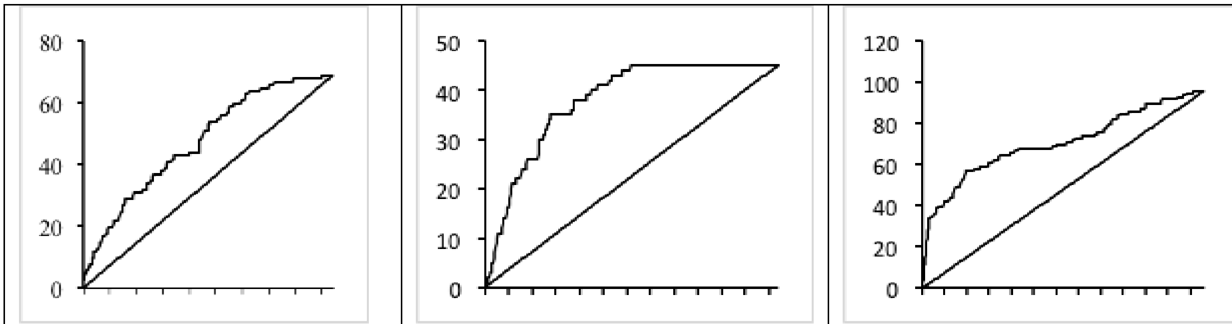
Figure 5.4. Path Data

5.6.3 *Performance Evaluation*

To generate the predicted probability with this true data in the validation set, we need to undo the oversampling. This is done by (Deutsch 2010):

$$correct\_prob = \frac{1}{1+\left(\frac{1}{original\_fraction}-1\right)/\left(\frac{1}{oversampled\_fraction}-1\right)\times\left(\frac{1}{scoring\_result}-1\right)}. \tag{14}$$

We report the lift charts in Figure 5.5. Lift charts are used to evaluate the effectiveness of predictive models. It shows the improvement of our model from a random guess. From all the three lift charts, we see a consistently high predictive performance compared to the naïve rule. The predictive power is especially good for the chat stage and purchase stage because the lift is high in the beginning.



a. Stage 1 - View b. Stage 2 - Chat c. Stage 3 - Purchase

Figure 5.5. Lift Charts

5.7 CONCLUSIONS

This paper leveraged clickstream data to understand online shoppers' decisions of product page view, live chat, and purchase. We have a rich set of findings centering on the differences between seller-level feedback scores, product-level review counts, and the role of live chat in customers' online shopping experiences.

Our results have important implications for sellers. New sellers can learn from our work that the best way to obtain feedback score is to promote one flagship product, instead of promoting multiple products with less effort on each product. This is because customers are most interested in relatively cheap products with high count of reviews. Despite their preference for viewing such products, feedback scores still play a critical role in the purchase stage. Therefore, sellers can increase the price after they gain sufficient feedback scores. Mature sellers can also learn that while live chatting may not help them with customers' repeat visits, they should still use live chatting tools because of its ability to improve conversion rates. Since mature sellers have well-established reputation, they do not need to prove their quality in the chatting conversations as much as new sellers. Instead, they should leverage live chatting opportunities to conduct cross selling and enhance long-term customer relationships to induce repeat visits.

Platform designers can also obtain insights from our work. First, they can leverage our results to optimize their search engines and recommendation systems. According to our results, a product with high review counts and relatively low price is the optimal choice for customers because these products are the most cost-effective. The quality of these products is demonstrated by the high review counts, and the price is reasonable because the sellers are still in pursuit of a higher feedback score. As is known, it is optimal for the online marketplace to charge customers less and charge sellers more as a two-sided platform. Therefore, the platform should recommend products satisfying the above conditions to customers to increase customer surplus. Second, other platforms can learn from our findings that allowing the coexistence of seller-level and product-level reputation will change the dynamics of the platform. Unlike the feedback score which is a life aggregation of the sellers' reputation, the product-level reviews may be easier to earn. Allowing customers to leave reviews for each product provides small sellers potential opportunities to survive and grow. Third, platform owners can consider improving the function of live chat. Instead of using a passive live chatting tool, platform owners can develop proactive live chats, which utilize real-time customer behavior to predict customers' need to talk to a web assistant. This is likely to improve customer satisfaction and alleviate the-rich-get-richer problem in the online marketplace.

With the introduction of live chatting tools, the way people interact with sellers may undergo a dramatic change. As a first attempt, this paper used clickstream data to understand the role of live chat. We look forward to more works in this domain.

REFERENCES

Abel 2014Abel, J. (2014), ‘Lawsuit alleges amazon charges prime members for "free" shipping’. https://www.consumeraffairs.com/news/lawsuit-alleges-amazon-charges-prime-members-for-free-shipping-031414.html

Åberg and Shahmehri 2000Åberg, J. and Shahmehri, N. (2000), ‘The role of human web assistants in e-commerce: an analysis and a usability study’, Internet research 10(2), pp. 114–125.

Abhishek et al. 2015Abhishek, V., Jerath, K. and Zhang, Z. J. (2015), ‘Agency selling or reselling? channel structures in electronic retailing’, Management Science .

Abt 2010Abt (2010), ‘Growth curve’. http://www.abt.com/about/news/intret_030110, retrieved on Sep. 20, 2016.

Adar and Huberman 2000Adar, E. and Huberman, B. A. (2000), ‘Free riding on gnutella’, First Monday 5(10).

Agrawal et al. 2015Agrawal, A., Catalini, C. and Goldfarb, A. (2015), ‘Crowdfunding: Geography, social networks, and the timing of investment decisions’, Journal of Economics & Management Strategy 24(2), 253–274.

Alesci and Smith 2015Alesci, C. and Smith, A. (2015), ‘Walmart challenges amazon prime’. http://money.cnn.com/2015/05/14/news/companies/walmart-amazon-shipping/

Alpizar et al. 2008Alpizar, F., Carlsson, F. and Johansson-Stenman, O. (2008), ‘Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in costa rica’, Journal of Public Economics 92(5), 1047–1060.

Andreoni 1988Andreoni, J. (1988), ‘Why free ride?: Strategies and learning in public goods experiments’, Journal of public Economics 37(3), 291–304.

Andreoni 1989Andreoni, J. (1989), ‘Giving with impure altruism: Applications to charity and ricardian equivalence’, The Journal of Political Economy pp. 1447–1458.

Andreoni 1990Andreoni, J. (1990), ‘Impure altruism and donations to public goods: A theory of warm-glow giving’, The economic journal 100(401), pp. 464–477.

Andreoni 1990Andreoni, J. (1990), ‘Impure altruism and donations to public goods: a theory of warm-glow giving’, The economic journal pp. 464–477.

Andreoni 2006Andreoni, J. (2006), ‘Leadership giving in charitable fund-raising’, Journal of Public Economic Theory 8(1), 1–22.

Andreoni and Petrie 2004Andreoni, J. and Petrie, R. (2004), ‘Public goods experiments without confidentiality: a glimpse into fund-raising’, Journal of public Economics 88(7), 1605–1623.

Andreoni and Scholz 1998Andreoni, J. and Scholz, J. K. (1998), ‘An econometric analysis of charitable giving with interdependent preferences’, Economic inquiry 36(3), 410–428.

Andrews and Benzing 2007Andrews, T. and Benzing, C. (2007), ‘The determinants of price in internet auctions of used cars’, Atlantic Economic Journal 35(1), pp. 43–57.

Aral and Walker 2011Aral, S. and Walker, D. (2011), ‘Creating social contagion through viral product design: A randomized trial of peer influence in networks’, Management science 57(9), pp. 1623–1639.

Aral and Walker 2014Aral, S. and Walker, D. (2014), ‘Tie strength, embeddedness, and social influence: A large-scale networked experiment’, Management Science 60(6), pp. 1352–1370.

Aral et al. 2009Aral, S., Muchnik, L. and Sundararajan, A. (2009), ‘Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks’, Proceedings of the National Academy of Sciences 106(51), 21544–21549.

Aral et al. 2009Aral, S., Muchnik, L. and Sundararajan, A. (2009), ‘Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks’, Proceedings of the National Academy of Sciences 106(51), pp. 21544–21549.

Ariely et al. 2009Ariely, D., Bracha, A. and Meier, S. (2009), ‘Doing good or doing well? image motivation and monetary incentives in behaving prosocially’, The American Economic Review pp. 544–555.

Ba and Pavlou 2002Ba, S. and Pavlou, P. A. (2002), ‘Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior’, MIS quarterly pp. 243–268.

Bakshy et al. 2011Bakshy, E., Hofman, J. M., Mason, W. A. and Watts, D. J. (2011), Everyone’s an influencer: quantifying influence on twitter, in ‘Proceedings of the fourth ACM international conference on Web search and data mining’, ACM, pp. 65–74.

Bapna and Umyarov 2015Bapna, R. and Umyarov, A. (2015), ‘Do your online friends make you pay? a randomized field experiment on peer influence in online social networks’, Management Science 61(8), pp. 1902–1920.

Bapna et al. 2008Bapna, R., Jank, W. and Shmueli, G. (2008), ‘Consumer surplus in online auctions’, Information Systems Research 19(4), pp. 400–416.

Bapna et al. 2011Bapna, R., Goes, P., Wei, K. K. and Zhang, Z. (2011), ‘A finite mixture logit model to segment and predict electronic payments system adoption’, Information Systems Research 22(1), 118–133.

Bapna et al. 2011Bapna, R., Gupta, A., Rice, S. and Sundararajan, A. (2011), Trust, reciprocity and the strength of social ties: An online social network based field experiment, in ‘Conference on Information Systems and Technology (CIST)’.

Barabási and Albert 1999Barabási, A.-L. and Albert, R. (1999), ‘Emergence of scaling in random networks’, science 286(5439), 509–512.

Barsh et al. 2000Barsh, J., Crawford, B. and Grosso, C. (2000), ‘How e-tailing can rise from the ashes’, McKinsey Quarterly pp. 98–109.

Basso et al. 2001Basso, A., Goldberg, D., Greenspan, S. and Weimer, D. (2001), First impressions: emotional and cognitive factors underlying judgments of trust e-commerce, in ‘Proceedings of the 3rd ACM conference on Electronic Commerce’, ACM, pp. 137–143.

Bekkers 2004Bekkers, R. H. (2004), Giving and volunteering in the Netherlands: Sociological and psychological perspectives, Utrecht University.

Bekkers and Wiepking 2010Bekkers, R. and Wiepking, P. (2010), ‘A literature review of empirical studies of philanthropy: Eight mechanisms that drive charitable giving’, Nonprofit and Voluntary Sector Quarterly .

Berenson 2015Berenson, T. (2015), ‘Facebook raised more than 10 million in 2 days for nepal’. http://time.com/3842365/facebook-money-nepal-earthquake/

Bernheim 1994Bernheim, B. D. (1994), ‘A theory of conformity’, Journal of political Economy pp. 841–877.

Berry et al. 1995Berry, S., Levinsohn, J. and Pakes, A. (1995), ‘Automobile prices in market equilibrium’, Econometrica: Journal of the Econometric Society pp. 841–890.

Bettman 1979Bettman, J. R. (1979), Information processing theory of consumer choice, Addison-Wesley Pub. Co.

Biernacki et al. 2000Biernacki, C., Celeux, G. and Govaert, G. (2000), ‘Assessing a mixture model for clustering with the integrated completed likelihood’, IEEE transactions on pattern analysis and machine intelligence 22(7), pp. 719–725.

Blut et al. 2015Blut, M., Chowdhry, N., Mittal, V. and Brock, C. (2015), ‘E-service quality: A meta-analytic review’, Journal of Retailing 91(4), 679–700.

Bøg et al. 2012Bøg, M., Harmgart, H., Huck, S. and Jeffers, A. M. (2012), ‘Fundraising on the internet’, Kyklos 65(1), 18–30.

Bowman 2004Bowman, W. (2004), ‘Confidence in charitable institutions and volunteering’, Nonprofit and Voluntary Sector Quarterly 33(2), 247–270.

Bradlow and Rao 2000Bradlow, E. T. and Rao, V. R. (2000), ‘A hierarchical bayes model for assortment choice’, Journal of Marketing Research 37(2), pp. 259–268.

Brown and Heal 1980Brown, D. J. and Heal, G. (1980), ‘Two-part tariffs, marginal cost pricing and increasing returns in a general equilibrium model’, Journal of Public Economics 13(1), 25–49.

Burstein 2014Burstein, D. (2014), ‘Ecommerce research chart: Industry benchmark conversion rates for 25 retail categories’.https://www.marketingsherpa.com/article/chart/conversion-rates-retail-categories, retrieved on Sep. 20, 2016.

Burtch et al. 2013Burtch, G., Ghose, A. and Wattal, S. (2013), ‘An empirical examination of the antecedents and consequences of contribution patterns in crowd-funded markets’, Information Systems Research 24(3), 499–519.

Burtch et al. 2015Burtch, G., Ghose, A. and Wattal, S. (2015), ‘The hidden cost of accommodating crowdfunder privacy preferences: a randomized field experiment’, Management Science 61(5), 949–962.

Cai et al. 2007Cai, H., Chen, Y. and Fang, H. (2007), Observational learning: Evidence from a randomized natural field experiment, Technical report, National Bureau of Economic Research.

Carman 2003Carman, K. G. (2003), ‘Social influences and the private provision of public goods: Evidence from charitable contributions in the workplace’.

Castillo et al. 2014Castillo, M., Petrie, R. and Wardell, C. (2014), ‘Fundraising through online social networks: A field experiment on peer-to-peer solicitation’, Journal of Public Economics 114, 29–35.

Chen and Hitt 2002Chen, P.-Y. and Hitt, L. M. (2002), ‘Measuring switching costs and the determinants of customer retention in internet-enabled businesses: A study of the online brokerage industry’, Information systems research 13(3), 255–274.

Chen and Yao 2015Chen, Y. and Yao, S. (2015), ‘Sequential search with refinement: Model and application with click-stream data’, Available at SSRN with paper ID 2130646.

Chevalier and Mayzlin 2006Chevalier, J. A. and Mayzlin, D. (2006), ‘The effect of word of mouth on sales: Online book reviews’, Journal of marketing research 43(3), pp. 345–354.

Chiang et al. 1994Chiang, W.-C., Fitzsimmons, J., Huang, Z. and Li, S. X. (1994), ‘A game-theoretic approach to quantity discount problems*’, Decision Sciences 25(1), 153–168.

Clark 2002Clark, J. (2002), ‘Recognizing large donations to public goods: an experimental test’, Managerial and Decision Economics 23(1), 33–44.

Coleman 1988Coleman, J. S. (1988), ‘Social capital in the creation of human capital’, American journal of sociology pp. 95–120.

Collier and Bienstock 2006Collier, J. E. and Bienstock, C. C. (2006), ‘How do customers judge quality in an e-tailer?’, MIT Sloan Management Review 48(1), 35–40.

Customers 2015Customers, A. (2015), ‘Higher prices on items with prime shipping?’. http://www.amazon.com/forum/amazon%20discussions%20feedback?_encoding=UTF8&cdForum=FxLCMW9UN2N6QR&cdThread=Tx13IW5V355AY70

d'Aspremont et al. 1979d'Aspremont, C., Gabszewicz, J. J. and Thisse, J.-F. (1979), 'On hotelling's "stability in competition"', Econometrica 47(5), 1145–1150.

Davis 1970Davis, J. A. (1970), 'Clustering and hierarchy in interpersonal relations: Testing two graph theoretical models on 742 sociomatrices', American Sociological Review pp. 843–851.

De los Santos et al. 2012De los Santos, B., Hortaçsu, A. and Wildenbeest, M. R. (2012), 'Testing models of consumer search using data on web browsing and purchasing behavior', The American Economic Review 102(6) 2955–2980.

Dellarocas 2003Dellarocas, C. (2003), 'The digitization of word of mouth: Promise and challenges of online feedback mechanisms', Management science 49(10), pp. 1407–1424.

Dempster et al. 1977Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977), 'Maximum likelihood from incomplete data via the em algorithm', Journal of the royal statistical society. Series B (methodological) pp. 1–38.

Deutsch 2010Deutsch, G. (2010), 'Overrepresentation - "sas"-oversampling'. http://www.data-mining-blog.com/tips-and-tutorials/overrepresentation-oversampling/, retrieved on Sep. 20, 2016.

Dewan and Hsu 2002Dewan, S. and Hsu, V. (2002), 'Adverse selection in reputations-based electronic markets: Evidence from online stamp auctions', Research Paper, University of Irvive.

Dewan et al. 2016Dewan, S., Ho, I. and Ramaprasad, J. (2016), 'Popularity or proximity: Characterizing the nature of social influence in an online music community', Information System Research .

Dimoka et al. 2012Dimoka, A., Hong, Y. and Pavlou, P. A. (2012), 'On product uncertainty in online markets: Theory and evidence', MIS quarterly 36.

Duan et al. 2008Duan, W., Gu, B. and Whinston, A. B. (2008), 'Do online reviews matter?—an empirical investigation of panel data', Decision support systems 45(4), pp. 1007–1016.

Ebbes et al. 2005Ebbes, P., Wedel, M., Böckenholt, U. and Steerneman, T. (2005), 'Solving and testing for regressor-error (in) dependence when no instrumental variables are available: With new evidence for the effect of education on income', Quantitative Marketing and Economics 3(4), pp. 365–392.

Einolf et al. 2013Einolf, C. J., Philbrick, D. M. and Slay, K. (2013), 'National giving campaigns in the united states entertainment, empathy, and the national peer group', Nonprofit and Voluntary Sector Quarterly 42(2), 241–261.

Einolf et al. 2013Einolf, C. J., Philbrick, D. M. and Slay, K. (2013), 'National giving campaigns in the united states entertainment, empathy, and the national peer group', Nonprofit and Voluntary Sector Quarterly 42(2), pp. 241–261.

Emerson 1962Emerson, R. M. (1962), ‘Power-dependence relations’, American sociological review pp. 31–41.

Fader and McAlister 1990Fader, P. S. and McAlister, L. (1990), ‘An elimination by aspects model of consumer response to promotion calibrated on upc scanner data’, Journal of Marketing Research pp. 322–332.

Feldstein and Clotfelter 1976Feldstein, M. and Clotfelter, C. (1976), ‘Tax incentives and charitable contributions in the united states: A microeconometric analysis’, Journal of Public Economics 5(1), 1–26.

Fowler 2010Fowler, G. A. (2010), ‘Retailers team up against amazon’. [Online; posted 06-Oct-2010]. http://www.wsj.com/articles/SB10001424052748703843804575534062509989530

Frazier et al. 2004Frazier, P. A., Tix, A. P. and Barron, K. E. (2004), ‘Testing moderator and mediator effects in counseling psychology research.’, Journal of counseling psychology 51(1), 115.

Frey and Meier 2004Frey, B. S. and Meier, S. (2004), ‘Social comparisons and pro-social behavior: Testing" conditional cooperation" in a field experiment’, The American Economic Review 94(5), 1717–1722.

Friedman 1957Friedman, M. (1957), A Theory of the Consumption, princeton university press Princeton, NJ.

Frischmann et al. 2012Frischmann, T., Hinz, O. and Skiera, B. (2012), ‘Retailers’ use of shipping cost strategies: Free shipping or partitioned prices?’, International Journal of Electronic Commerce 16(3), 65–88.

Gensch 1987Gensch, D. H. (1987), ‘A two-stage disaggregate attribute choice model’, Marketing Science 6(3) 223–239.

Ghose et al. 2006Ghose, A., Smith, M. D. and Telang, R. (2006), ‘Internet exchanges for used books: An empirical analysis of product cannibalization and welfare impact’, Information Systems Research 17(1), pp. 3–19.

Granovetter 1973Granovetter, M. S. (1973), ‘The strength of weak ties’, American journal of sociology pp. 1360–1380.

Gu et al. 2009Gu, B., Huang, Y., Duan, W. and Whinston, A. B. (2009), ‘Indirect reciprocity in online social networks-a longitudinal analysis of individual contributions and peer enforcement in a peer-to-peer music sharing network’, McCombs Research Paper Series No. IROM-06-09 .

Gustin 2014Gustin, S. (2014), ‘Amazon’s prime price hike is actually crazy brilliant’. http://time.com/23744/wall-street-amazon-prime/

Haaijer et al. 2000Haaijer, R., Kamakura, W. and Wedel, M. (2000), ‘Response latencies in the analysis of conjoint choice experiments’, Journal of Marketing Research 37(3), 376–382.

Harper et al. 2010Harper, Yan Chen, F. M., Konstan, J. and Li, S. X. (2010), ‘Social comparisons and contributions to online communities: A field experiment on movielens’, The American economic review pp. 1358–1398.

Hiltz and Turoff 1985Hiltz, S. R. and Turoff, M. (1985), ‘Structuring computer-mediated communication systems to avoid information overload’, Communications of the ACM 28(7), pp. 680–689.

Hong and Shum 2006Hong, H. and Shum, M. (2006), ‘Using price distributions to estimate search costs’, The RAND Journal of Economics 37(2), pp. 257–275.

Hotelling 1929Hotelling, H. (1929), ‘Stability in competition’, The Economic Journal 39(153), 41–57.

Houser and Wooders 2000Houser, D. and Wooders, J. (2000), ‘Reputation in internet auctions: Theory and evidence from ebay (working paper)’, Tucson, AZ: The University of Arizona.

Isaac 2016Isaac, M. (2016), ‘Twitter will offer selected tweets to keep users coming back’. http://www.nytimes.com/2016/02/11/technology/twitter-feed-information-overload.html, retrieved on Oct. 9, 2016.

Jones and Linardi 2014Jones, D. and Linardi, S. (2014), ‘Wallflowers: Experimental evidence of an aversion to standing out’, Management Science 60(7), 1757–1771.

Kamakura and Russell 1989Kamakura, W. A. and Russell, G. (1989), ‘A probabilistic choice model for market segmentation and elasticity structure’, Journal of Marketing Research 26, 379–390.

Katz 1984Katz, M. L. (1984), ‘Price discrimination and monopolistic competition’, Econometrica: Journal of the Econometric Society pp. 1453–1471.

Kauffman and Wood 2006Kauffman, R. J. and Wood, C. A. (2006), ‘Doing their bidding: An empirical examination of factors that affect a buyer’s utility in internet auctions’, Information Technology and Management 7(3) 171–190.

King and Zeng 2001King, G. and Zeng, L. (2001), ‘Logistic regression in rare events data’, Political analysis 9(2) 137–163.

Kollock 1999Kollock, P. (1999), ‘The production of trust in online markets’, Advances in group processes 16(1) 99–123.

Koning and Model 2013Koning, R. and Model, J. (2013), ‘Experimental study of crowdfunding cascades: When nothing is better than something’, Available at SSRN 2308161 .

Koukova et al. 2012Koukova, N., Srivastava, J. and Steul-Fischer, M. (2012), ‘The effect of shipping fee structure on consumers’ online evaluations and choice’, Journal of the Academy of Marketing Science 40(6), 759–770.

Kuppuswamy and Bayus 2014Kuppuswamy, V. and Bayus, B. L. (2014), ‘Crowdfunding creative ideas: The dynamics of project backers in kickstarter’, UNC Kenan-Flagler Research Paper (2013-15).

Kwak et al. 2010Kwak, H., Lee, C., Park, H. and Moon, S. (2010), What is twitter, a social network or a news media?, in ‘Proceedings of the 19th international conference on World wide web’, ACM, pp. 591–600.

Lee et al. 2015Lee, Y.-J., Hosanagar, K. and Tan, Y. (2015), ‘Do i follow my friends or the crowd? information cascades in online movie ratings’, Management Science 61(9), 2241–2258.

Lee et al. 2015Lee, Y.-J., Hosanagar, K. and Tan, Y. (2015), ‘Do i follow my friends or the crowd? information cascades in online movie ratings’, Management Science 61(9), pp. 2241–2258.

Leider et al. 2009Leider, S., Möbius, M. M., Rosenblat, T. and Do, Q.-A. (2009), ‘Directed altruism and enforced reciprocity in social networks’, The Quarterly Journal of Economics pp. 1815–1851.

Leng and Becerril-Arreola 2010Leng, M. and Becerril-Arreola, R. (2010), ‘Joint pricing and contingent free-shipping decisions in b2c transactions’, Production and Operations Management 19(4), 390–405.

Lewis 2006Lewis, M. (2006), ‘The effect of shipping fees on customer acquisition, customer retention, and purchase quantities’, Journal of Retailing 82, 13–23.

Lewis et al. 2006Lewis, M., Singh, V. and Fay, S. (2006), ‘An empirical study of the impact of nonlinear shipping and handling fees on purchase incidence and expenditure decisions’, Marketing Science 25(1), 51–64.

Li and Dinlersoz 2012Li, H. and Dinlersoz, E. (2012), ‘Quality-based price discrimination: evidence from internet retailers’ shipping options’, Journal of Retailing 88(2), 276–290.

Li and Hitt 2008Li, X. and Hitt, L. M. (2008), ‘Self-selection and information role of online product reviews’, Information Systems Research 19(4), pp. 456–474.

Li et al. 2009Li, S., Srinivasan, K. and Sun, B. (2009), ‘Internet auction features as quality signals’, Journal of Marketing 73(1), pp. 75–92.

Livingston 2005Livingston, J. A. (2005), ‘How valuable is a good reputation? a sample selection model of internet auctions’, Review of Economics and Statistics 87(3), pp. 453–465.

Lu et al. 2014Lu, C.-T., Xie, S., Kong, X. and Yu, P. S. (2014), Inferring the impacts of social media on crowdfunding, in 'Proceedings of the 7th ACM international conference on Web search and data mining', ACM, pp. 573–582.

MacLaughlin 2015MacLaughlin, S. (2015), 'Givingtuesday trends a closer look at the online fundraising impact in the united states'. https://www.blackbaud.com/files/resources/downloads/givingtuesday-trends.pdf, retrieved on Oct. 9, 2016.

Manchanda et al. 2008Manchanda, P., Xie, Y. and Youn, N. (2008), 'The role of targeted communication and contagion in product adoption', Marketing Science 27(6), 961–976.

Mangalindan 2015Mangalindan, J. (2015), 'Inside amazon prime'. http://fortune.com/2015/02/03/inside-amazon-prime/

Manski 1993Manski, C. F. (1993), 'Identification of endogenous social effects: The reflection problem', The review of economic studies 60(3), 531–542.

Manski 2000Manski, C. F. (2000), Economic analysis of social interactions, Technical report, National bureau of economic research.

Matthews 2014Matthews, C. (2014), 'Amazon prime price hike could be a major mistake'. http://time.com/5685/amazon-prime-price-hike-could-be-a-major-mistake/

Mayzlin et al. 2003Mayzlin, D., Chevalier, J. A. et al. (2003), The effect of word of mouth on sales: Online book reviews, Technical report, Yale School of Management.

Meer 2011Meer, J. (2011), 'Brother, can you spare a dime? peer pressure in charitable solicitation', Journal of Public Economics 95(7), 926–941.

Meer 2011Meer, J. (2011), 'Brother, can you spare a dime? peer pressure in charitable solicitation', Journal of Public Economics 95(7), pp. 926–941.

Meer 2014Meer, J. (2014), 'Effects of the price of charitable giving: Evidence from an online crowdfunding platform', Journal of Economic Behavior & Organization 103, 113–124.

Moe 2006Moe, W. W. (2006), 'An empirical two-stage choice model with varying decision rules applied to internet clickstream data', Journal of Marketing Research 43(4), pp. 680–692.

Moe and Fader 2004Moe, W. W. and Fader, P. S. (2004), 'Capturing evolving visit behavior in clickstream data', Journal of Interactive Marketing 18(1), pp. 5–19.

Montgomery et al. 2004Montgomery, A. L., Li, S., Srinivasan, K. and Liechty, J. C. (2004), 'Modeling online browsing and path analysis using clickstream data', Marketing Science 23(4), pp. 579–595.

Morwitz et al. 1998Morwitz, V., Greenleaf, E. and Johnson, E. J. (1998), 'Divide and prosper: consumers' reaction to partitioned prices', Journal of Marketing Research 35, 453–463.

Mutz and Pemantle 2011Mutz, D. and Pemantle, R. (2011), The perils of randomization checks in the analysis of experiments, in ‘annual meeting of the Society for Political Methodology’.

Muylle and Basu 2004Muylle, S. and Basu, A. (2004), ‘Online support for commerce processes and survivability of web retailers’, Decision Support Systems 38(1), pp. 101–113.

Narayanan and Nair 2013Narayanan, S. and Nair, H. S. (2013), ‘Estimating causal installed-base effects: A bias-correction approach’, Journal of Marketing Research 50(1), 70–94.

Newey 1987Newey, W. K. (1987), ‘Efficient estimation of limited dependent variable models with endogenous explanatory variables’, Journal of Econometrics 36(3), pp. 231–250.

Olbrich and Holsing 2011Olbrich, R. and Holsing, C. (2011), ‘Modeling consumer purchasing behavior in social shopping communities with clickstream data’, International Journal of Electronic Commerce 16(2), pp. 15–40.

Online+Chat/fulltext/-/E-RES57154?objectid=RES57154, retrieved on Sep. 20, 2016.

Ottaway et al. 2003Ottaway, T. A., Bruneau, C. L. and Evans, G. E. (2003), ‘The impact of auction item image and buyer/seller feedback rating on electronic auctions’, Journal of Computer Information Systems 43(3), pp. 56–60.

Overby and Jap 2009Overby, E. and Jap, S. (2009), ‘Electronic and physical market channels: A multiyear investigation in a market for products of uncertain quality’, Management Science 55(6), pp. 940–957.

Page 2012Page, R. (2012), ‘The linguistics of self-branding and micro-celebrity in twitter: The role of hashtags’, Discourse & Communication 6(2), pp. 181–201.

Pan et al. 2002Pan, X., Ratchford, B. T. and Shankar, V. (2002), ‘Can price dispersion in online markets be explained by differences in e-tailer service quality?’, Journal of the Academy of Marketing Science 30(4), 433–445.

Park et al. 2007Park, D.-H., Lee, J. and Han, I. (2007), ‘The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement’, International Journal of Electronic Commerce 11(4), pp. 125–148.

Pavlou and Dimoka 2006Pavlou, P. A. and Dimoka, A. (2006), ‘The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation’, Information Systems Research 17(4), pp. 392–414.

Perrault and Brousseau 1989Perrault, M. R. and Brousseau, K. R. (1989), ‘Selling to different styles’, Decision Dynamics Corporation, Los Angeles, CA.

Phlips and Thisse 1982Phlips, L. and Thisse, J.-F. (1982), ‘Spatial competition and the theory of differentiated markets: an introduction’, The Journal of Industrial Economics pp. 1–9.

Power 2013Power, J. (2013), ‘Overall customer care satisfaction rises as wireless customers use online chat feature more frequently due to improved capabilities’. http://www.jdpower.com/press-releases/2013-us-wireless-customer-care-full-service-performance-study-volume-2-and-2013-us, retrieved on Sep. 20, 2016.

Reinstein and Riener 2012Reinstein, D. and Riener, G. (2012), ‘Reputation and influence in charitable giving: An experiment’, Theory and decision 72(2), 221–243.

Resnick and Zeckhauser 2002Resnick, P. and Zeckhauser, R. (2002), ‘Trust among strangers in internet transactions: Empirical analysis of ebay’s reputation system’, The Economics of the Internet and E-commerce 11(2), pp. 23–25.

Rivera et al. 2010Rivera, M. T., Soderstrom, S. B. and Uzzi, B. (2010), ‘Dynamics of dyads in social networks: Assortative, relational, and proximity mechanisms’, annual Review of Sociology 36, pp. 91–115.

Rutz et al. 2012Rutz, O. J., Bucklin, R. E. and Sonnier, G. P. (2012), ‘A latent instrumental variables approach to modeling keyword conversion in paid search advertising’, Journal of Marketing Research 49(3), pp. 306–319.

Satow 1975Satow, K. L. (1975), ‘Social approval and helping’, Journal of Experimental Social Psychology 11(6), 501–509.

Saxton and Wang 2013Saxton, G. D. and Wang, L. (2013), ‘The social network effect: The determinants of giving through social media’, Nonprofit and Voluntary Sector Quarterly p. 0899764013485159.

Schervish and Havens 1997Schervish, P. G. and Havens, J. J. (1997), ‘Social participation and charitable giving: a multivariate analysis’, Voluntas: International Journal of Voluntary and Nonprofit Organizations 8(3), 235–260.

Schindler et al. 2005Schindler, R. M., Morrin, M. and Bechwati, N. N. (2005), ‘Shipping charges and shipping-charge skepticism: Implications for direct marketers’ pricing formats’, Journal of Interactive Marketing 19(1), 41–53.

Schultz et al. 2007Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J. and Griskevicius, V. (2007), ‘The constructive, destructive, and reconstructive power of social norms’, Psychological science 18(5), 429–434.

Schwartz 1974Schwartz, S. H. (1974), ‘Awareness of interpersonal consequences, responsibility denial, and volunteering.’, Journal of Personality and Social Psychology 30(1), 57.

Shang and Croson 2006Shang, J. and Croson, R. (2006), ‘The impact of social comparisons on nonprofit fundraising’, Research in Experimental Economics 11, 143–156.

Shang and Croson 2009Shang, J. and Croson, R. (2009), ‘A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods’, The Economic Journal 119(540), 1422–1439.

Shi et al. 2014Shi, Z., Rui, H. and Whinston, A. B. (2014), ‘Content sharing in a social broadcasting environment: Evidence from twitter’, MIS Quarterly 38(1), pp. 123–142.

Shriver et al. 2013Shriver, S. K., Nair, H. S. and Hofstetter, R. (2013), ‘Social ties and user-generated content: Evidence from an online social network’, Management Science 59(6), pp. 1425–1443.

Silverman et al. 1984Silverman, W. K., Robertson, S. J., Middlebrook, J. L. and Drabman, R. S. (1984), ‘An investigation of pledging behavior to a national charitable telethon’, Behavior Therapy 15(3), 304–311.

Sina 2015Sina (2015), ‘Weibo philanthropy raised 240 million rmb’. http://gongyi.sina.com.cn/gyzx/2015-01-14/115951212.html

Smith 2010Smith, B. G. (2010), ‘Socially distributing public relations: Twitter, haiti, and interactivity in social media’, Public Relations Review 36(4), 329–335.

Smith 2012Smith, A. (2012), ‘Real time charitable giving’. http://www.pewinternet.org/2012/01/12/real-time-charitable-giving/, retrieved on Oct. 9, 2016.

Smith 201Smith, C. (201), ‘By the numbers: 83 amazing alibaba statistics’. http://expandedramblings.com/index.php/alibaba-statistics/, retrieved on Sep. 20, 2016.

Smith et al. 2015Smith, S., Windmeijer, F. and Wright, E. (2015), ‘Peer effects in charitable giving: Evidence from the (running) field’, The Economic Journal 125(585), 1053–1071.

Soetevent 2005Soetevent, A. R. (2005), ‘Anonymity in giving in a natural context—a field experiment in 30 churches’, Journal of public Economics 89(11), 2301–2323.

Spulber 1981Spulber, D. F. (1981), ‘Spatial nonlinear pricing’, The American Economic Review 71(5), 923–933.

Stein 2015Stein, C. (2015), ‘The 2015 online fundraising report’. http://www.fundraisingschool.it/wp-content/uploads/2015/05/2015-Online-Fundraising-Report.pdf, retrieved on Oct. 9, 2016.

Stevens 2015Stevens, L. (2015), ‘Fuel prices fall, but fedex and ups boost surcharges’. http://www.wsj.com/articles/fuel-prices-fall-but-fedex-and-ups-boost-surcharges-1444088938

Stevens 2016Stevens, L. (2016), ‘‘free’ shipping crowds out small retailers’. http://www.wsj.com/articles/for-online-shoppers-free-shipping-reigns-supreme-1461789381

Stevens and Banjo 2014Stevens, L. and Banjo, S. (2014), ‘Free shipping is going to cost you more’. http://www.wsj.com/articles/free-shipping-is-going-to-cost-you-more-1414003507

Stock et al. 2012Stock, J. H., Wright, J. H. and Yogo, M. (2012), 'A survey of weak instruments and weak identification in generalized method of moments', Journal of Business & Economic Statistics.

Stole 1995Stole, L. A. (1995), 'Nonlinear pricing and oligopoly', Journal of Economics & Management Strategy 4(4), 529–562.

Stone 2010Stone, B. (2010), 'Amazon prime delivery service hooks shoppers'. http://www.sfgate.com/business/article/Amazon-Prime-delivery-service-hooks-shoppers-3244233.php

Strothkamp et al. 2010Strothkamp, B., Ensor, B. and McGowan, B. (2010), 'Case study: Wells fargo proves the business case for online chat'. https://www.forrester.com/Case+Study+Wells+Fargo+Proves+The+Business+Case+For+

Sugden 1984Sugden, R. (1984), 'Reciprocity: the supply of public goods through voluntary contributions', The Economic Journal pp. 772–787.

Tan 2015Tan, W. (2015), 'China's charities confront public trust crisis'. http://www.shanghaidaily.com/feature/news-feature/Chinas-charities-confront-public-trust-crisis/shdaily.shtml

Toubia and Stephen 2013Toubia, O. and Stephen, A. T. (2013), 'Intrinsic vs. image-related utility in social media: Why do people contribute content to twitter?', Marketing Science 32(3), pp. 368–392.

Townsend 2014Townsend, L. (2014), 'How much has the ice bucket challenge achieved?'. http://www.bbc.com/news/magazine-29013707, retrieved on Oct. 9, 2016.

Tucker 2008Tucker, C. (2008), 'Identifying formal and informal influence in technology adoption with network externalities', Management Science 54(12), pp. 2024–2038.

Tucker and Zhang 2011Tucker, C. and Zhang, J. (2011), 'How does popularity information affect choices? a field experiment', Management Science 57(5), 828–842.

Tuttle 2011Tuttle, B. (2011), 'Amazon prime loses $11 annually per member. and it's a huge success'. http://business.time.com/2011/11/14/amazon-prime-loses-11-annually-per-member-and-its-a-huge-success/

U.S. Department of Commerce 2012U.S. Department of Commerce (2012), 'Estimated detailed operating expenses of u.s. retail, accommodations, and food services firms1 by kind of business: 2012'. https://www2.census.gov/retail/releases/current/arts/bes.xls

U.S. Department of Commerce 2015U.S. Department of Commerce (2015), 'Quarterly retail e-commerce sales', Web. http://www.census.gov/retail/index.html

UPS 2015UPS (2015), 'Ups online shopping study: Empowered consumers changing the future of retail'. https://www.comscore.com/Insights/Press-Releases/2015/6/UPS-Online-Shopping-Study-Empowered-Consumers-Changing-the-Future-of-Retail

Vesterlund 2003Vesterlund, L. (2003), 'The informational value of sequential fundraising', Journal of Public Economics 87(3), 627–657.

Wiepking 2008Wiepking, P. (2008), 'For the love of mankind: A sociological study on charitable giving'.

Wikipedia 2014Wikipedia (2014). https://en.wikipedia.org/wiki/Ice_Bucket_Challenge

Wilde 2000Wilde, J. (2000), 'Identification of multiple equation probit models with endogenous dummy regressors', Economics letters 69(3), pp. 309–312.

Wohlsen 2013Wohlsen, M. (2013), 'Why amazon prime could soon cost you next to nothing'. http://www.wired.com/2013/03/amazon-prime-could-soon-cost-next-to-nothing/

Wooldridge 2010Wooldridge, J. M. (2010), Econometric analysis of cross section and panel data, MIT press.

Wu 2015Wu, R. (2015), 'How to launch a #givingtuesday campaign ╟ webinar recording'. https://www.causevox.com/givingtuesday-campaign-planning/, retrieved on Oct. 9, 2016.

Wu et al. 2004Wu, D., Ray, G., Geng, X. and Whinston, A. (2004), 'Implications of reduced search cost and free riding in e-commerce', Marketing Science 23(2), 255–262.

Wu et al. 2011Wu, S., Hofman, J. M., Mason, W. A. and Watts, D. J. (2011), Who says what to whom on twitter, in 'Proceedings of the 20th international conference on World wide web', ACM, pp. 705–714.

Xia et al. 2012Xia, M., Huang, Y., Duan, W. and Whinston, A. B. (2012), 'Research note-to continue sharing or not to continue sharing? an empirical analysis of user decision in peer-to-peer sharing networks', Information Systems Research 23(1), 247–259.

Yang et al. 2005aYang, Y., Essegaier, S. and Bell, D. R. (2005a), 'Free shipping and repeat buying on the internet: theory and evidence', Manuscript, Wharton School, University of Pennsylvania.

Yang et al. 2005bYang, Y., Essegaier, S. and Bell, D. R. (2005b), 'Free shipping and repeat buying on the internet: theory and evidence', Manuscript, Wharton School, University of Pennsylvania.

Yao and Zhang 2012Yao, Y. and Zhang, J. (2012), 'Pricing for shipping services of online retailers: Analytical and empirical approaches', Decision Support Systems 53(2), 368–380.

Yin 2004Yin, X. (2004), 'Two-part tariff competition in duopoly', International Journal of Industrial Organization 22(6), 799–820.

Zhou et al. 2009Zhou, B., Katehakis, M. N. and Zhao, Y. (2009), ‘Managing stochastic inventory systems with free shipping option’, European Journal of Operational Research 196(1), 186–197.

APPENDIX

The error terms in equations 3.7 and 3.9, (u_i, v_i), are assumed to follow bivariate Normal with mean 0 and variance matrix:

$$\begin{pmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{pmatrix}.$$

Therefore, the joint probability of $\left(Engage_i^*, R\_Followees_i\right)$ conditional on $\mathbf{x}_i$ and category j $prob\left(Engage_i^*, R\_Followees_i \mid \mathbf{x}_i, j\right)$, is bivariate Normal with the mean:

$$\begin{pmatrix} \gamma\pi_j + \boldsymbol{\beta}\mathbf{x}_i \\ \pi_j \end{pmatrix},$$

and variance matrix:

$$\begin{pmatrix} \gamma^2\sigma_v^2 + 2\gamma\rho\sigma_u\sigma_v + \sigma_u^2 & \gamma\sigma_v^2 + \rho\sigma_u\sigma_v \\ \gamma\sigma_v^2 + \rho\sigma_u\sigma_v & \sigma_v^2 \end{pmatrix}.$$

Using R_F_i as the abbreviation of $R\_Followees_i$, we can write:

$$prob\left(Engage_i^*, R\_F_i \mid \mathbf{x}_i, j\right) = prob\left(Engage_i^* \mid R\_F_i, \mathbf{x}_i, j\right) \times prob\left(R\_F_i \mid j\right)$$

where

$$prob\left(R\_F_i \mid j\right) = \frac{1}{\sigma_v}\phi\left(\frac{R\_F_i - \pi_j}{\sigma_v}\right)$$

and ϕ is the PDF of standard Normal. $prob\left(Engage_i^* \mid R\_F_i, \mathbf{x}_i, j\right)$ follows a Normal distribution with the mean:

$$E\left(Engage_i^* \mid R\_F_i, \mathbf{x}_i, j\right) = \gamma\pi_j + \boldsymbol{\beta}\mathbf{x}_i + \left(\gamma + \rho\frac{\sigma_u}{\sigma_v}\right)\left(R\_F_i - \pi_j\right),$$

and the variance

$$Var\left(Engage_i^* \mid R\_F_i, \mathbf{x}_i, j\right) = \left(1 - \rho^2\right)\sigma_u^2.$$

Now we evaluate:

$$prob\left(Engage_i = 1 \mid R\_F_i, \mathbf{x}_i, j\right) = prob\left(Engage_i^* \geq 0 \mid R\_F_i, \mathbf{x}_i, j\right)$$

$$= \Phi\left(\frac{\gamma\pi_j + \boldsymbol{\beta}\mathbf{x}_i + \left(\gamma + \rho\left(\sigma_u / \sigma_v\right)\right)\left(R\_F_i - \pi_j\right)}{\sqrt{1-\rho^2}\,\sigma_u}\right).$$

where Φ is the CDF of standard Normal. The log likelihood function as expressed in equation 3.10 follows.

VITA

Xue Tan is originally from China. She holds a B.Eng. in Management Information Systems from Renmin University of China. She has three years' working experience as a modeler for data warehousing design in the Chinese banking industry. Her research interests include social media fundraising and online markets.