ABSTRACT

Title of Document:	INFORMATION TRANSPARENCY AND USER BEHAVIOR IN EMERGING ONLINE MARKETPLACES: EMPIRICAL STUDIES OF SOCIAL MEDIA AND OPEN INNOVATION MARKETS.
	Abrar Al-Hasan, Doctor of Philosophy, 2013
Directed By:	Professors Siva Viswanathan, Department of Decision, Operations and Information Technologies
	Professors Hank Lucas, Department of Decision, Operations and Information Technologies

Web 2.0 and social media have significantly increased the amount of information available to users not only about firms and their offerings, but also about the activities of other individuals in their networks and markets. It is widely acknowledged that this increased availability of information is likely to influence a user's behavior and choices. However, there are very few systematic studies of how such increased information transparency influences user behavior in emerging marketplaces. My dissertation seeks to examine the impact of increased information transparency – particularly, information about other individuals - in two emerging platforms. The first essay in my dissertation compares online "social" marketing on Facebook with "non-social" marketing and examines their relative impacts on the likelihood of

adoption, usage and diffusion of an “App”. While social marketing - wherein a user gets to see which of her other friends have also “liked” the product being marketed– is one of the fastest growing online marketing formats, there are hardly any studies that have examined the value of the social aspect of such marketing. I find that social marketing is associated with increased app adoption, usage, and diffusion as compared to non-social marketing. The study also uncovers interesting tradeoffs between the effects of different types of “social” information on user behavior outcomes. The second essay examines the behavior of contestants in an open innovation design marketplace, wherein firms seek solutions from a crowd through an online contest. The study examines how the availability of information about other contestants as well as the availability of feedback information provided to others by the contest holder, impacts a focal contestant’s behavior and outcomes. I find that contestants adopt different strategic behaviors that increase their odds of winning the contest under the different information-transparency regimes. The findings have interesting implications for the design of online contests and crowdsourcing markets. Overall, my dissertation provides a deeper understanding of how the visibility of different types of information in online platforms impacts individual behaviors and outcomes.

INFORMATION TRANSPARENCY AND USER BEHAVIOR IN EMERGING ONLINE MARKETPLACES: EMPIRICAL STUDIES OF SOCIAL MEDIA AND OPEN INNOVATION MARKETS.

By

Abrar Al-Hasan

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2013

Advisory Committee:
Professor Siva Viswanathan, Chair
Professor Hank Lucas
Professor Il-Horn Hann
Professor Guodong (Gordon) Gao
Professor Alan Neustadtl , (Dean's Representative)

Dedication

I would like to dedicate this work to my amazing Mom, Dad, Grandmother, and loved ones who made all of this possible, for their endless encouragement, patience and belief in me. And to the soul of my beloved nanny Aisha.

Acknowledgements

I would like to express my deepest gratitude for the valuable guidance I received from Professor Siva Viswanathan from the very start to the completion of my dissertation. I would also like to thank Professor Hank Lucas for his considerable mentoring and guidance. I would like to also thank Professor Il-Horn Hann for his constant support and direction. I am truly grateful to Professor Guodong (Gordon) Gao and Professor Alan Neustadtl for having served as committee members, and for their valuable insights in helping me to complete this dissertation. I also would like to give special thanks to Dr.Abhishek Pani for his extensive efforts and assistance.

Table of Contents

List of Tables

Tables for Essay 1

Tables for Essay 2

List of Figures

Figures for Essay 1

Figures for Essay 2

Chapter 1: Introduction and Overview

The ability of online markets to bring together individuals and businesses has transformed and redefined the ways business is conducted. In particular, the increased visibility of information enabled by Web 2.0 technologies has led to an explosion of new business models that seek to leverage this increased availability of information. Firms are now using information available in Web 2.0 not only for online word of mouth and marketing purposes, but also for tapping into the "wisdom" of the crowds. While prior studies have demonstrated that access to more information can lead to greater levels of innovation (Tjosvold & McNeely, 1988; Katz, 1982; Keller, 1986), there are hardly any studies of the consequences of such increased information transparency for individual behaviors and outcomes.

One type of information that has become very salient is information about "others" in an individual's network or market. This sudden increase in the amount and variety of information available to an individual about others in her network is likely to have significant impact on her behaviors and choices – an issue that serves as the primary focus of my dissertation. My dissertation examines the impact of increased information transparency, and in particular information on other users in the community, on an individual's behavior in two different online emerging marketplaces. The first essay compares and contrasts the value of "social" information with "non-social" information and how they impact different outcomes in one of the largest online social media platforms. The second essay examines the role of informational spillovers in open innovation design contests, in an online crowdsourcing market wherein firms seek design solutions from a crowd. A better understanding of the increase of visibility of information impact on consumer behavior in this emerging landscape has implications for academics, policymakers, and entrepreneurs who seek to leverage the power of online markets.

The first essay in my dissertation examines a firm's use of social media, in particular Facebook, as a marketing platform. Using a unique dataset from an established firm that sought to disseminate an App to users on Facebook, the goal of this essay is to understand the impact of information visibility of "others" on a focal users' behavior. Unlike traditional marketing channels, Web 2.0 and social media allow for the visibility and dissemination of information about other individuals in a person's network, "social information" that has been typically difficult or costly to obtain. In particular, "social marketing" seeks to leverage the availability of information about other users in an individual's social network to market or advertise products. My first essay examines the impact of traditional non-social online marketing methods that contain no social information (email and ads), and social online marketing methods that contain social information (social ads, friend invites, and Facebook newsfeeds) on a focal user's behaviors including the likelihood of adoption, use, and diffusion of the App.

I find that the visibility of "social" information is associated with greater App adoption, use and diffusion as compared to "non-social" information. I find that both the first degree as well as the second degree of friends in an individuals' network, have a significant influence on her adoption behavior. I also examine the role of two different types of media engagement (active versus passive) and their interactions with social and non-social sources of information, and how they impact App use and diffusion. I find that social information and engagement when taken together have interesting impacts on app use and diffusion. I also find that passive social information sources are better than active social information sources in fostering diffusion.

My findings are robust to a large number of alternate specifications and highlight the differential impacts of non-social versus social marketing methods, and the role of social information on user outcomes, and ultimately advertising effectiveness. My study takes both the marketing and information systems research a step forward by highlighting the significance of "social information," and how it can be used with media vehicles to better drive product adoption, use, diffusion and ultimately sales.

The second essay examines the behavior of contestants in an open innovation marketplace with a primary focus on the role of different information visibility regimes on the behavior of contestants as well as contest outcomes. Unlike traditional R&D efforts, where firms rely on internal firm solutions, the Internet enables a firm to seek solutions from a large number of people from geographically dispersed locations, and from a diverse set of skill sets, thereby increasing the potential for innovation. For example, leading organizations, such as NASA and Dell, have been engaging in crowdsourcing methods to successfully spark new innovations (McKendrick, 2012).

While such crowdsourcing marketplaces have been growing rapidly, there have been very few systematic studies of individual behaviors in these marketplaces. In particular, given the increased information transparency on "others" in such markets, where information about prior submissions in a contest and related feedback might be available to later participants, it is not clear how such information impacts a potential contestant's behavior and outcomes. Using a dataset from one of the early pioneers of online crowdsourcing contests, this essay seeks to understand the role of different information visibility regimes on the entry behaviors of contestants and their likelihood of winning the contest. Contestants not only view information related to them, but also view information on other contestants or "others" competing with them in the contest.

Drawing on theories of design of contests (Moldovanu & Sela ,2001; Liu et al. ,2007), time to entry (Lieberman & Montgomery, 1988; Urban & Star, 1991; Reinganum, 1981) and information spillover (Brynjolffson & Hitt 2000; Chang & Gurbaxani, 2012; Mun & Nadiri, 2002; Cheng & Nault, 2007), I examine the strategies that lead to higher probability of winning under different contest designs, and how different types of feedback information given to prior submissions in a contest impacts a focal contestant's submission behavior and ultimately her probability of winning the contest. I compare "open" contests possessing greater degrees of information transparency with "blind" contests that have limited information transparency, to examine how these different information regimes influence contestants' choices and outcomes.

I find that contest design, particularly relating to information visibility, significantly influences contestant's behavior as well as outcomes. Both early submissions as well as late submissions are more likely to win a contest compared to submissions at other times during the contest. In examining the impact of informational spillovers relating to the submissions, I find that late submissions are more likely to win in open contests. However, in blind contests where there are no informational spillovers and when no feedback is provided by the contest holder, there are no specific submission times that perform better than others. Only a contest holder's expertise, experience, and skill are associated with her likelihood of winning the contest. I also find evidence of informational spillovers relating to feedback given to others. Interestingly, in both open as well as blind contests, when feedback is provided by the contest holder to contestants, we find that early submissions are more likely to win, particularly when these early contestants make a resubmission.

I also find that the benefits from informational spillovers differ depending on the type of feedback provided by the contest holder. In open contests with feedback, the more specific the feedback given to others in a contest, the higher the chances of late submissions winning the contest. However, in blind contests where users cannot see each other's submissions, specific feedback given to other contestants does not benefit a focal user, while generic feedback increases the likelihood of late submissions winning the contest. These results highlight the importance of the role of different information regimes and information spillover in open innovation contests on time to entry and ultimately the probability winning the contest. Overall, I find that information spillover of certain types of visibility of information diminishes the well-known competitive strategy of first movers, and benefits late movers when it comes to time to entry.

Overall, the two essays of my dissertation investigate the impact of increased information transparency of "others" on user behavior in two different emerging online marketplaces that were enabled only through the emergence of Web 2.0. My dissertation contributes to the literature and practice in the following ways.

First, the two studies in my dissertation provide rich empirical evidence on how increased information transparency influences user behaviors. Essay 1 focuses on the role of friends – social information- in this process. Essay 2 focuses on the role of different information visibility regimes (different contest designs) on user behavior as well as contest outcomes.

Second, my dissertation contributes to a growing literature on diffusion of information, and visibility of information in online markets. The first essay highlights information diffusion in online social networks as a new mechanism for "social marketing" or "social advertising." The second essay directly examines the role of diffusion of information through feedback and different contest design regimes as information spillover and its impact on user behavior, or the "timing of entry". My dissertation shows both the benefits and the drawbacks of information transparency and visibility. For both academics and practitioners, understanding how consumers behave in these emerging landscapes is imperative. More importantly, my study sheds light on the overall impact of IT and Web 2.0 on not only firms, but also on user behavior in the context of the emerging online marketplaces.

Lastly, both studies foci have major implications for platform developers and policy makers in these emerging markets. For the first essay, my dissertation provides guidance to platform developers and policy makers so that they can enable and constrain the visibility of "social information", and can engineer the user experience to increase sharing, interaction, and virality. The second study provides useful guidelines for crowdsourcing platforms on the design of online contests and valuable insights into how the micro-structure of these contests can influence outcomes.

Chapter 2: The Value Of “Social” Marketing: A Comparison Of “Social" And “Non-Social” Online Marketing Of Facebook Apps

2.1 Introduction

With the advent of Web 2.0, consumers and their social relations have moved to the center of the internet stage (Mounier, 2005; Guillard, 2005). This phenomenon had led to a transformation of how firms engage with their customers. In particular, the rapid growth of social media has led to the increased use of social media platforms, such as Facebook, by firms for marketing their products and services. Not only are firms embracing social media, but a survey by the *Wall Street Journal* finds that 71 percent of U.S. adults who purchase online, use consumer product reviews for their purchases, of which 42 percent trust such a source (Spors, 2006). More recently, an article illustrated that Google is trying to integrate social signals from a users’ social network, into their search results to "deliver more relevant results" (Lynley, 2013).

The most prevalent social media platform for firms is Facebook. According to recent estimates, over 700 billion minutes a month are spent on Facebook.[1] As of 2011, there are 500 million active Facebook users, half of whom log in on any given day. Given the popularity of Facebook among its users, it is no surprise that firms are increasingly seeking to target Facebook users. An important aspect of social media platforms such as Facebook is the social connections among users and the potential of such connections to influence an individual’s behavior. Firms are increasingly seeking to leverage such social information to tailor their marketing and advertising

[1] Statistics retrieved from http://www.statisticbrain.com/facebook-statistics/

strategies. One of the primarily mechanisms that firms use to engage with customers on Facebook is “Apps” or applications. Facebook Apps are highly popular with its users, with over 20 million Apps being installed each day. App users also tend to be highly engaged. Facebook Apps not only allow firms to engage users, but also enable them to obtain information about the user as well as her social connections and use this information to design their marketing strategies. However, despite the growing popularity of social media and the use of social information for marketing activities, there is very little systematic understanding of how information about “others” in an individual’s social network influences her behaviors. Understanding how such increased information transparency – in particular, about “others” in one’s social network - affects an individual’s behavior and choices would be valuable to both firms seeking to leverage these emerging technologies, as well as to technologists seeking to devise new tools and artifacts.

While prior studies have demonstrated that access to more information can lead to greater levels of innovation (Tjosvold & McNeely, 1988; Katz, 1982; Keller, 1986), there are hardly any studies of the consequences of such increased information transparency for individual behaviors and outcomes especially in the social media marketplace. Firms seeking to engage consumers in social media platforms have a choice of two broad information provisioning strategies. Firms can continue to target potential consumers using traditional “non-social” formats such as display ads, email solicitations, or adopt emerging “social formats” such as “social ads, newsfeeds from friends, etc. While firms are actively experimenting with both “social” and “non-social” information-provisioning formats, there is a dearth of research to guide decision making in this context.

This study is among the first to empirically examine the value of “social information” in comparison to “non-social” information and investigate their differential impacts on user outcomes as well as their implications for firms’ social media strategies. In particular, this study seeks to answer the following questions.

1) *What is the impact of a users’ Facebook social network Adoption information on the likelihood of her App adoption?*

2) *What is the impact of different information sources (for example, Social Information (friend invitation, newsfeed, social ad) versus Non-Social Information (email, Traditional online ad) on App Use and Diffusion?*

Answers to these questions will enable firms to value the effectiveness of “social information,” and will identify the types of information or viral product design features that impact product adoption, use, diffusion and ultimately advertising effectiveness and sales.

My study examines detailed data from a Facebook App that was developed by a leading vendor in Facebook's Preferred Developer Consultant Program. I examine the impact of different sources of information (social and non-social) on user behavior and App outcomes. I find that the type of the source of information, particularly relating to information about users’ friends - or social information, significantly influences user behavior and App outcomes.

I find that users’ Facebook social network adoption information impacts a users’ likelihood of App adoption. In particular, by having higher number of friends adopting the App (both first or second degree friends) increases the likelihood of a user adopting the App.

Furthermore, I find that social information sources (friend invite, newsfeed, and social ad) positively impact a user's App use and App diffusion outcomes. However, non-social information sources (email and traditional ads), negatively impact App use outcomes, and are not significant predictors of App diffusion outcomes.

I also find that the benefits from social information sources differ depending on the type of information engagement (active versus passive). I find that active sources of information (friend invite) lead to higher App use outcomes that are user generated (uploading stories and photos - direct) as compared to passive sources of information (newsfeed, social ads). Also, active social sources of information (friend invite) lead to higher local diffusion (outdegree) as compared to passive social sources (newsfeed, social ads). However, passive social sources lead to higher App use outcomes that are indirect (passengering planes) as compared to active social sources of information. More importantly, I find that passive social sources (newsfeed, social ads) lead to higher beyond local diffusion (diameter) as compared to active social sources (friend invite). This finding implies that passive social sources are more effective in fostering diffusion.

Overall, I find that the visibility of different information sources (social versus non-social) and the interactivity of the information (active versus passive) generates identifiably different social contagion effects. My results have interesting implications for firms in managing consumer social interactions and advertising. In particular, my results show that traditional online advertising (non-social) may no longer be as effective, and underscore the importance of use of social information in online advertising. This study also sheds light on how viral products can be designed to generate social contagion not only by means of social information visibility but also by the interactivity of the information (active versus passive). These findings draw attention to the fact that with the increase in information visibility online, marketing methods should adapt to provide the right type of information to attract consumers.

My study makes a number of contributions. It is one of the first studies to examine the value of "social information" in the social media marketplace, and the consequences of differing social information. Using detailed data on different types of "social" and "non-social" information, my study is able to tease out the differing impacts of "social" and "non-social" information relating to App adoption, use and diffusion. Furthermore, my study is one of the first to identify the effects of the interactive nature of such information on App related outcomes. My study takes both the marketing and information systems research a step forward by highlighting the significance of "social information," and how it can be used with media vehicles to better drive product adoption, use, diffusion and ultimately sales. More importantly, my study sheds light on the overall impact of IT and Web 2.0 on not only firms, but also on user behavior in the context of the social media marketplace, and how a consumer can be used in the marketing campaign as a "co-creator" of value to the firm.

The rest of the paper is organized as follows. Section 2 reviews the literature. Section 3 describes the research context and Section 4 describes the data and methodology. The results are in Section 5. Section 6 discusses the implications of my study and concludes. Lastly, Section 7 describes the limitations of this study and future research.

2.2 Literature Review

This study draws upon two broad streams of literature. The first relates to research on advertising and media engagement and its impact on user behavior. The second relates to social information and its impact on user behavior.

2.2.1 Advertising and Media Engagement

It is well known that given increasing ad clutter, consumers are rather skeptical of advertising and wary of its influence (Dahlen, 2005; Friestad & Wright, 1995; Goodstein, 1993). The increasing ad clutter in traditional advertising media, such as TV and newspapers, has had negative effects on both the media and their advertising content. Ha and Litman (1997) find that increased advertising levels in magazines reduce the effectiveness of each individual ad, as well as circulation and profitability of the magazines. As a result, firms are constantly seeking novel ways to reduce this negative bias of consumers towards advertisements such as placing traditional ads in non-traditional environments. The idea behind certain placements is to blur the boundary between the advertised message and its surrounding content, making it more challenging for consumers to identify advertising as advertising, as it is made less irritating and disrupting of the content (Dahlen & Edenius, 2007). Social media platforms provide a novel mechanism for firms to engage customers. In particular, the use of social information (i.e., information about an individual's social network) has become a popular mechanism for firms to engage users on social media platforms. While prior research has largely focused on the attitude of consumers towards traditional ad formats, this is among the first study to compare the effectiveness of social ads with non-social ads. In doing so, this study contributes to the extant literature on advertising, particularly in the social media marketplace.

In addition to the sources of information (i.e., social or non-social), prior research has also examined how the engagement and interactivity of the information presented to users impacts outcomes (for instance see, Stern, 1997; Stern, 1994; MacInnis et al., 1991; Batra & Ray, 1983; Miniard & Cohen, 1983; Muncy & Hunt, 1984). Ad formats differ in the range of interaction they allow. The formats may range from a limited click in static online banner ad, to a full range of interactive features in rich media ads, such as those in Macromedia's Flash. Flash allows users to interact by providing search abilities, audio and video capabilities, capability to play games and enter contests, send e-mail and complete purchase transactions without ever leaving the publisher's site. Prior work has examined two main forms of online media engagement or interactivity: *passive* advertising exposure and *active* advertising exposure (Chatterjee, 1998; Chatterjee et al., 2003). Banner ads or display ads are considered to be passive advertising, while targeted communications are considered to be a form of active advertisement exposure (Chatterjee 1998). Active advertisement exposure is under the consumer's control, and passive advertisement exposure is under the marketer's control (Chatterjee, 1998). Empirical studies (Chatterjee et al., 2003; Aaker & Brown, 1972; Bronner & Neijens, 2006; Coulter, 1998; Cunningham et al., 2006; DePelsmacker et al., 2002; Feltham & Arnold, 1994; Gallagher et al., 2001; Nicovich, 2005; Wang, 2006; Calder et al., 2009) suggest that in general when consumers are more engaged with the advertising vehicle (active exposure), they are more responsive to advertising than when they are not as engaged (passive exposure).

Although there has been a strong interest in media engagement and advertising effectiveness, the question still remains as to what type of involvement is effective in social media advertising. Social media advertising not only includes different types of media engagement (social ads, friend invitations, newsfeed, etc.), but may also include social information that can impact advertising attitude differently. In the context of my study, I analyze behavior of users towards passive advertising exposure (traditional online ads, social ads, and newsfeed) and active advertising exposure (emails and friend invite) in the context of social media.

2.2.2 Social Information and Social Influence

Individuals are likely to be affected by the opinion and behaviors of others; this type of influence is termed social influence (Kelman, 1958). While there is a large and diverse body of work related to social information and social influence, I review two closely related streams of research. The first being social information processing theory, the second focuses on how word-of-mouth (WOM) and social contagion (i.e., how information is diffused in social networks), impacts user behavior.

2.2.2.1 Social Information Processing Theory

According to established studies of organizational information systems, individual opinions of an information system are likely to be influenced by objective characteristics of the system, individual differences, and extent of use of the system (Rice & Aydin, 1991). The failure of individual attributes to sufficiently explain behavior, has brought theories of social influence, and more specifically, the social information processing model to the organizational setting (Salancik & Pfeffer, 1978). Social information processing theory proposes that individuals may be influenced by cues from others about what to attend to, how to value the important dimensions of a phenomenon, and how others evaluate the same phenomena (Salancik & Pfeffer, 1978). Different sources of information have different impacts on outcomes. Social sources of information are more likely to influence a users' attitude than non-social sources of information (Rice & Aydin, 1991; Ibarra & Andrews, 1993; Wellman, 1983; Dean & Brass, 1985; Hartman & Johnson, 2006). Examining the role of social influence on consumers' pre-purchase search efforts, Brown and Reingen (1987) show that information received from sources that have some personal knowledge about the consumer have more influence on the latter than sources that have no personal knowledge about the consumer.

My study seeks to take the social information processing theory a step forward by empirically analyzing how the increase in visibility of social information online influences user behavior in an online social media marketplace. I further analyze how the different types of social information impacts user behavior.

2.2.2.2 Social Contagion and WOM

Academics across several fields such as economics, marketing and sociology have been fascinated by the study of social interactions where individuals' actions or behavior depends upon the actions or choices of other actors (Granovetter, 1973; Banerjee, 1992; Ellison & Fudenburg, 1993; Manski, 1993; Bala & Goyal, 1998; Van den Bulte & Lilien, 2001). Prior research examined the conditions under which consumers are likely to rely on others' opinions to make a purchase decision (Godes & Mayzlin, 2004; Becker, 1991). Studies suggest that customers who were acquired through WOM and peer effects add more long term value to the firm than customers acquired through traditional marketing channels (Villanueva et al., 2008). Researchers have also examined the impact of WOM and peer effects on the diffusion of a new product (Zhang & Huang, 2011).

Though such social channels have influenced consumers for decades, current advances in technology have significantly increased the importance of consumer social interactions as a market force. More recently, studies have focused on targeting "influential" individuals who are likely to spread WOM most broadly (Katz & Lazersfeld, 1955; Watts & Dodds, 2007; Goldenberg et al., 2009). Biyalogorsky et al., (2001) use referral programs to create incentives for "influential" individuals to spread the word. Van der Lans et al. (2010) use observational evidence on viral campaigns to inform viral branching models of WOM diffusion. They define a viral marketing campaign as an online marketing message that stimulates consumers to forward the message to members of their social network. These friends are subsequently encouraged to forward the message to their social circle, and so on.

While viral marketing is a very popular concept, most current work has focused on viral marketing campaigns for existing products, there are hardly any studies examining the factors that drive virality. An exception is a recent study by Aral & Walker (2011) who study viral product design. Viral product design consists of including explicit characteristics and features into a product's design that brings about peer-to-peer influence to encourage adoption (Aral & Walker, 2011). Using a randomized experiment, Aral and Walker (2011) found that viral features (personalized referrals and automated broadcast notifications) generate identifiable and significant peer influence and social contagion effects. They also found that although personalized viral messages are more effective in encouraging adoption and are correlated with more user engagement and sustained product use, broadcast messaging is used more often, generating more total peer adoption in the network. Their findings highlight the importance of further examining how different types of online WOM and viral features are received by individuals.

There is growing evidence that social media allows firms to engage with customers more efficiently than do traditional marketing channels. Not only are consumers now better able to exchange information, but firms are also gaining the ability to directly initiate and manage consumer social interactions (Godes & Mayzlin, 2004; Chevalier & Mayzlin, 2006). While advances in technology are creating new opportunities for firms to directly facilitate and manage consumer social interactions, they also impose new challenges. Distinct strategic managerial actions are often necessary in such new social marketplaces. Deciding on the type of advertising and the information presented in the ad is a crucial decision that will impact a firm's profitability and positioning (Chen et al., 2010).

Thus far, most of the empirical studies on WOM and peer effects have focused on observing the distribution of WOM; they do not directly observe the reception of WOM. In this study, I take this stream of literature ahead a step, and directly observe the reception of WOM by an individual's social network. I compare the reception and effectiveness of online traditional non-social WOM to social WOM. Furthermore, I examine different types of social viral product designs and features, their impact on consumer behavior, and ultimately on advertising effectiveness.

2.2.3 Research Hypotheses

In this study, I examine the impact of different types of information sources (social versus non-social) on advertising effectiveness (App adoption, use and diffusion). I also explore the information engagement type (active and passive media), and how the engagement interacts with the type of information source to impact outcomes of interest.

Figure 2.3 shows the different categories I examine in this study. In particular, I examine social versus non-social sources of information; and active versus passive information engagements. Social sources are sources that provide information related to friends (social ads, newsfeed and friend invite). Non-social sources do not include such friend related information (emails and traditional ads). As for the engagement of the source, I use Chatterjee's (1998) definition of active versus passive media engagement. Traditional ads, social ads (similar to banner ads) are under the marketers control and are passive in nature. Similarly newsfeed sources pop-up in a user's Facebook newsfeed page, and are under the marketers' control, making them passive. On the other hand, emails and friend invites require a higher level of interaction by actually clicking on the email or invite, and thus are considered active engagements.

According to social influence theories and social information, individuals are influenced by social sources of information as they are perceived to be more trustworthy. I therefore propose that advertising with social information as compared to non-social information formats, influences consumers' attitudes more positively towards the advertised message as measured by her adoption, use and diffusion:

Proposition 1: *Ad effectiveness (adoption, use and diffusion) is higher when social information is presented in an ad compared with the non-presentation of social information.*

The concept of an active advertisement is a feature that differentiates Web advertising from advertising in traditional media. My study examines the two types of online advertising by comparing active advertising (emails, friend invite) and passive advertising (newsfeed, social ad, and traditional ads). Although studies have shown that media engagement improves advertising effectiveness, I argue that media engagement will not be effective under traditional non-social advertising; however, media engagement will be effective when social information is presented. Given that active information sources are under the consumer's control and have more media engagement as compared to passive information sources, I argue that active information sources will lead to higher effectiveness (use and diffusion) than passive information sources.

Proposition 2: *Media Engagement (active versus passive) of an ad will lead to higher advertising effectiveness when coupled with social information as compared to no social information. In particular, active information sources will lead to higher advertising effectiveness than passive information sources in the presence of social information.*

Overall, my study will contribute to an emerging area of research, namely the use and influence of online social and non-social information sources on consumers' decision-making processes. My study adds to the above mentioned streams of literature by examining how different sources of information (social versus non-social), and how different types of engagement (active versus passive) impacts user adoption, use, and diffusion in the social media marketplace. My study will also contribute to the WOM literature by providing empirical evidence on the reception and advertising effectiveness of different types of online WOM.

2.3 Research Context

I use App-related data that was developed by a leading vendor in Facebook's Preferred Developer Consultant Program. The App was launched at the end of March 2011, and was active till the middle of May 2011, for a period of around 6 weeks. The developer firm had the following channels to advertise the App: (a) Email, (b) Traditional online display ads, (c) Social ads – traditional online display ads that includes "Friends also Liked" button (see Figure 2.2), and (d) Other- users that adopted the App directly and not via online advertising. Lastly, given that the App was on Facebook, (e) some users learned about the App through Facebook's

Newsfeed, and (f) others adopted the App as a result of a direct invitation on Facebook from their friends.

The App used for this study is a game that consists of the following steps: First, a user "Like"-s the App. Second, the user gets to "pilot" a plane and invites 5 friends to "passenger" the plane. If all the friends accept, then the user is enrolled in the sweepstakes. The sweepstakes reward is a free ticket for the user and his 5 friends to go on a vacation. Users can become unlimited number of passengers, but can pilot a maximum of only 15 planes. Also, users are encouraged throughout the App lifetime to share stories and photos. Such activity will enroll them into another sweepstakes to win around $200.

Emails promoting the App were sent out in 3 periods: (a) early April, (b) mid-April, and (c) early-May. Prior to mid-April, the ads were all regular (non-social) display ads. Social ads only came into use after mid-April, creating a setting for a natural experiment that I use in this study. I collected data on App adoption, use and diffusion, along with Facebook related data of the users and their friends.

2.4 Data and Methodology

The data was compiled from various sources. Below, I describe the sample construction method and then provide the details of the empirical models and estimation methods.

2.4.1 Sample Construction

The vendor had collected data on users that "Like"-d the App which consists of a sample of 277,863 unique users. Of these 173,050 had Facebook URL links which I used to collect Facebook related data (I refer to this sample as the Outcome sample, henceforth). Table 2.1 shows the frequency distribution for the source of adoption of the users. Traditional online ads produced 31 percent of the sample, and social ads and email constitute 5 percent of the sample. Nearly 8 percent were invited by their friends, whereas approximately 30 percent came from Facebook newsfeeds. Table 2.2 shows further descriptive statistics, including the average App use and users' Facebook activity and some social network related metrics. On average, users have 15 Facebook "Like"-s, and around 300 Facebook friends.

As for users that saw the App but did not "Like" or adopt the App, I was able to obtain two sub-samples of such users. The first sub-sample consists of users that were "Fans" of the Vendors Facebook Page. These users were able to see the App, but some did not adopt or "Like" the App. I extracted users that were active fans starting 3 months prior start date of the App (from Dec 2010) up to the end of the App duration in May 2011. This sub-sample (henceforth, "FanPage") consisted of a total of 5,208 users of whom 3,914 users did not like the App. I constructed the second sub-sample (henceforth, "Top10CC") using the top 10 connected components[2] from the users that had "Like"-d the App (287 distinct users), and I collected their first degree Facebook friends network (103,493 distinct friends). These friends would have seen the App on the Facebook newsfeed through their friends that had "Like"-d the App. Of these 103,493 friends, 1,174 "Like"-d the App, and the rest did not. I then constructed the friendship network using day and time of adoption of the App.

---

[2] A connected component is basically a connected sub-graph of a graph to which no vertex can be added and still be connected. More formally, given a graph (G) with vertices (V) and edges (E) such that G= (V, E), a sub-graph S= (V', E') is a connected component if (a) S is connected, and (b) for all vertices u such that $u \in V$ and $u \notin V'$ there is no vertex $v \in V'$ for which $(u, v) \in E$.

Table 2.3 shows the adoption distribution of this sub-sample. I find that 58 percent of the Top10CC users had only first degree friends liking the App, 11 percent had only second degree friends liking the App, and 31 percent had both first and second degree friends liking the App. The third, and more robust, sample (henceforth, “Intersection”) is the intersection of the two sub-samples, which was composed of 387 users that saw the App but did not “Like” the App.

The number of users that “Like”-d the App and had Facebook URLs is 173,050 (outcome sample). The sub-samples, FanPage and Intersection, are much smaller in terms of the number of users that did not “Like” the App (3,914 and 387 respectively). Therefore, I used a random matched sample from the outcome sample of similar size for estimation (3,000 and 400 respectively) to generate a more balanced sample for estimation. The Top10CC sample was large enough (102,606) to include the whole outcome sample (173,050). Table 2.4 summarizes the sample constructions.

2.4.2 Empirical Model

I examine user behavior outcomes in terms of App adoption, use and diffusion as a function of the visibility of social information. I only observe App use and diffusion for users that “Like”-d the App. Thus, to draw conclusions about the larger population of all users that saw the App, not just the subpopulation of users that “Liked” the App, the Heckman (1979) two-stage estimation procedure for a continuous decision variable can be used to incorporate the App use and diffusion with the decision to “Like” or adopt the App. This method assumes the decision to adopt and use/diffuse are made simultaneously (that is, the error terms of the two equations are correlated).

The first step is the selection equation, which is estimated by maximum likelihood as an independent Probit model to determine the decision to adopt using information from the whole sample of adopters and non-adopters. A vector of inverse Mills ratios (estimated expected error) is generated from the parameter estimates (Greene, 1993).

App use and diffusion -- which is observed only when the selection equation equals 1 (that is, a user adopts the App) -- is then regressed on the explanatory variables, and the vector of inverse Mills ratios from the selection equation using ordinary least squares. Therefore, the second stage reruns the regression with the estimated expected error included as an extra explanatory variable, removing the part of the error term correlated with the explanatory variable and avoiding the selection bias.

I estimate a two-stage Heckman (1979) model where the first stage accounts for the selection (that is, a user's probability of adopting the App). The second stage examines the outcomes of interest (for users that have adopted the App) -- App use and App diffusion -- as a function of the source of adoption.

Selection Stage:

$$z_i^* = \omega_i \gamma + u_i \qquad (z_i = 1 \text{ if } z_i^* > 0 \text{ ; } z_i = 0 \text{ if } z_i^* <= 0)$$

Outcome Stage:

$$y_i = x_i\beta + \varepsilon_I \text{ if } z_i^* > 0$$

$$y_i = \text{------} \quad \text{if } z_i^* <= 0$$

z_i* identifies whether or not the user *i* "Like"-d the App or more formally adopted the App. The selection covariates that impact a user *i*'s App adoption are termed $\omega_{i.}$ The outcome covariates that impact a user *i*'s App use and App diffusion are termed $x_{i\,;}\, y_i$ and are the outcomes of interest for user *i*, namely, App use and App diffusion.

The Heckman model can be identified either through non-linearity intrinsic to selection models (for example, Uzzi, (1999)), or through exclusion restrictions. I consider exclusion restrictions, and use MediaExposure (days when the App received media exposure through Regis and Kelly's advertisement, New York Times,other emails that were sent out to advertise for the App, and introduction time of social ads) which increased traffic to the App and thereby adoption of the App. Yet, MediaExposure is not correlated with App use and diffusion as such metrics are post adoption. Therefore, for a user *i*, it will take some time from her MediaExposure, to actually use and diffuse the App. Empirically, this variable has an F-statistic exceeding 47, which is well above the cutoff point for a strong instrument, namely 10, as suggested by Staiger and Stock (1997).

2.4.3 Estimation

My study examines (a) App adoption, (b) App use, and (c) App diffusion as a function of the visibility of social information sources. Table 2.5 describes the variables and their description in detail. Below I describe the estimation method for each of the outcomes of interest.

2.2.3.1 App Adoption

In the App adoption stage, the outcome of interest is a binary variable as to whether or not the user adopted the App. To examine the impact of a users' Facebook social network adoption information on the likelihood of her App adoption, I investigate the social exposure of a user to the App in terms of her friends' adoption of the App. In particular, *Social Exposure to App,* entails the percentage of Facebook first degree and second degree friends that adopted the App at the time of user adoption decision. I control for user (a) Demographic information: gender and location, (b) Facebook Privacy Settings[3] (for example, "Add as Friend" button and "Send Message Button"), (c) Facebook Activity: Facebook "Like"-s and activities. In particular, the probability of a user *i*'s App adoption is the first stage of the Heckman model:

$$P\ (App\ Adoption_i) = f\ (Social\ Exposure\ to\ App_i,\ Demographic_i,\ Facebook\ Privacy\ Settings_i,\ Facebook\ Activity_i) + \varepsilon_i \qquad \textbf{Equation (2.1)}$$

To further control for possibility of omitted variable bias, I use Cox proportional hazard model with gamma frailty (Newman & McCullogh, 1984). This is a standard technique for assessing contagion in economics, marketing, IS, and sociology (Van den Bulte & Lilien, 2001; Iyengar et al., 2011; Aral & Walker, 2011; Nam et al., 2010). I estimate the Hazard of adoption for individual *i* at time *t* (time of adoption) as a function of individual characteristics and social influence:

$$h(t|x) = h_0(t)exp(x\beta) \qquad \textbf{Equation (2.2)}$$

[3] What I term Facebook Privacy Settings in this study is not actual Facebook Privacy settings. It is a measure of enabling and disabling certain buttons that might be indicative to privacy e.g. "Add as Friend" and "Send Message" buttons.

where $h_0(t)$ is the baseline hazard rate, x is a vector of covariates that impact the App adoption decision (both related and unrelated to social influence as in Equation 2.1), and β is the hazard ratio which is the exponentiated form of the coefficient β. To ensure vigor of results, I use all three samples (FanPage, Top10CC, Intersection) in estimating the likelihood of App adoption for user *i*.

2.2.3.2 App Use

Once users have adopted or "Like"-d an App, they can then use the App by becoming passengers of planes, uploading stories and uploading photos. Thus, the dependent variables for App use are (a) Log(Passenger count), (b) Log(Stories count), and (c) Log(Photos count)[4]. The main independent variable of interest is the visibility of social information in terms of the source of adoption. There are five main sources of adoption: email, regular online ads (traditional non-social information sources), friend invitations, Facebook newsfeed, and social ads (social information sources). I control for (a) Demographic information: gender and location, (b) Facebook Privacy Settings, (c) App Network : number of friends that like the App at time of adoption, number of invitations sent by user and accepted by friend, (d) Facebook Activity : Facebook "Like"-s and Facebook activity, total number of Facebook friends, (e) App Adoption Rate: day of adoption, Threshold lag, percentage of Facebook friend adopters, and (f) Friend App usage: average pilot count of friends, average passenger count of friends, average photo count of friends, and average story count of friends.

[4] While uploading stories and photos are user generated and considered direct App use, passengering is an indirect App use by only passengering a plane.

To examine the impact of visibility of social information sources of adoption on App use for user *i*, I estimate the second stage of the Heckman model with logarithmic App use outcome variables,

$$Log\ (App\ Use_i) = f\ (Adoption\ Source_i,\ App\ Network_i,\ Demographic_i,\ Facebook\ Privacy\ Settings_i,\ Facebook\ Activity_i,\ App\ Adoption\ Rate_i,\ Friend\ App\ Use_i) + \varepsilon_i$$

Equation (2.3)

2.2.3.3 App Diffusion

Once users have adopted or "Like"-d an App, they can not only use the App, but also send out invitations to their friends to join the App -- diffuse the App. I explore App diffusion outcomes on two main measures: (a) Diameter (maximum diffusion depth) and (b) Outdegree of a user. The extent to which the effect of the adoption source leads to adoption *beyond* a user's immediate local network is measured by the maximal diffusion depth (diameter) – the maximum network distance from a user to any peer adopter in a linked chain of adoptions. This is a measure of the breadth or reach of a user's network. The extent to which the effect of the adoption source leads to adoption in a user's immediate local network is measured by the outdegree of a user – number of invitations that were sent by the user and accepted by her friends. This is a measure of the width of a user's local network.

To examine the impact of visibility of social information sources of adoption on App diffusion, I use the same set explanatory and control variables as for the App use outcome. However, one concern is the argument that users that use the App more may be more likely to diffuse the App. Following Villas-Boas and Winer (1999), I use lagged total App use to instrument for potential App diffusion endogeneity. The total App use of a user is measured at time *t* whereas diffusion metrics are measured at time *t+1*, such that the total App activity of a user came before the friend's acceptance of the invitation. I estimate the second stage of the Heckman model with logarithmic App diffusion outcome variables,

$$Log(App\ Diffusion_{i(t+1)})^{5} = f\,(Adoption\ Source_i,\ Total\ App\ Use_{i\,(t)},\ App\ Network_i,\ Friend\ App\ Use_{\,i\,(t)},\ Demographic_i,\ Facebook\ Activity_i,\ App\ Adoption\ Rate_i) + \varepsilon_i$$

Equation (2.4)

2.5 Results

Table 2.6a provides the estimation results for the App adoption outcome. Table 2.6b displays the regression results of the App use outcomes, and Table 2.6c displays the results of the App diffusion model. The key research objective is to investigate the role of visibility of social information to users on App outcomes. I start by examining regression coefficients of the App adoption outcome, followed by App use outcomes, and lastly App diffusion outcomes. In particular, I refer to the regression coefficients of the Heckman model using the Intersection sample (since it is the most robust sample, and the Heckman model is most appropriate for my data). All samples and models provide consistent results in terms of sign and significance.

2.5.1 App Adoption

Table 2.6.a reports the likelihood of adoption estimates using both Heckman selection estimation method, and Cox proportional model with gamma frailty for all

[5] Lagging avoids endogeneity problems unless (a) people are forward looking not only about their own behavior but also (b) on the behavior of others' social ties, over which influence flows are symmetric (Lyengar et al., 2011). The first condition is quite unlikely in large networks as Facebook, and the second condition does not hold in my data as the Like invitation ties are directed ties, and I control for the exact time of adoption.

three samples. As for the Cox model, I report the exponentiated form of the coefficient β, which is called the *hazards ratio*. A hazards ratio greater than 1.0 for variable *x* indicates that it increases the probability of App adoption, while a ratio less than 1.0 indicates that it decreases the probability of App adoption.

Column 1 of Table 2.6a reports results of the first stage Heckman model and column 2 reports results of the Hazard model for the FanPage sample. Columns 3 and 4 report the estimation results of the first stage Heckman and Hazard model for the Top10CC sample respectively. Columns 5 and 6 report the estimation results of the first stage Heckman and Hazard model of the more robust Intersection sample.

Referring to the column 5 of Table 2.6a, I find that the coefficient on *FirstDeg_FrndAdop* is positive and significant ($\beta_{FirstDeg\_FrndAdop}$= 0.4731), indicating that higher number of first degree friends adopting the App (at time of adoption decision of user *i)*, the greater the likelihood of focal user *i*'s adoption. Furthermore, the coefficient on *SecondDeg_FrndAdop* is also positive and significant ($\beta_{SecondDeg\_FrndAdop}$= 0.1047), implying that higher number of second degree friends adopting the App (at time of adoption decision of user *i*), the greater the likelihood of focal user *i*'s adoption. The coefficient on the interaction of the first and second degree friends is positive and significant ($\beta_{FirstDeg*SecondDeg}$ = 0.5027), implying that having higher number of first and second degree friends adopting the App (at time of adoption decision of user *i)*, the greater the likelihood of focal user *i*'s adoption. Overall, these findings suggest that social exposure of a user to the App in terms of her friends' adoption information of the App has a positive and significant impact on a user's adoption decision- supporting proposition 1 for App adoption. More interestingly, this is the case not only for the exposure of first degree friends' adoption information, but also for the exposure of second degree friends' adoption information.

In terms of economic interpretations, a one unit increase in social information on the number of first degree friends adopting results in 17.03% increase in the probability of a user adopting as indicated by the average marginal effect. A one unit increase in social information on the second degree friends adopting results in 3.87% increase in the probability of the user adopting as indicated by the average marginal effect. A one unit increase in social information on the number of first and second degree friends adopting results in 18.10% increase in the probability of the user adopting as indicated by the average marginal effect. My findings support Proposition 1a that advertising effectiveness in terms of adoption is higher when there is social information in the advertising medium rather than when there is no social information in the advertising medium.

Results for the control variables align well with expectations. The coefficient for *Female* is positive and significant, indicating that on average females are more likely to adopt than males. This finding is not surprising; as noted by the vendor that develops Facebook Apps on average for most Facebook Apps, females are predominant. The coefficient for *likes_count* is positive and significant, implying that users that have a higher number of Facebook Likes have a higher probability of App adoption. This is in accordance with what one would expect. Users that have a high number of "Like"-s or adoptions of Apps, are more prone to adopting an App.

2.5.2 App Use

Table 2.6.b reports the second stage of the Heckman regression model for the App use outcomes (a) Log Passenger count, (b) Log Stories count, and (c) Log Photos count. Columns 1-3 of Table 2.6b report the estimation results for the FanPage sample. Columns 4-6 report the estimation results for the Top10CC sample. Lastly, columns 7-9 report the estimation results for the more robust Intersection sample. Estimation results from all three samples are similar in terms of sign and significance.

Referring to columns 7-9 in Table 2.6b, the signs of the coefficients reported for the social and non-social sources of adoption variables are in accordance with what one would expect as compared to the baseline. The regression coefficients for non-social sources are negative and significant (Passengering: $\beta_{source\_ad} = -0.0109$, $\beta_{source\_email} = -0.1251$; Stories: $\beta_{source\_ad} = -0.1607$, $\beta_{source\_email} =-0.0313$; Photos: $\beta_{source\_ad} =-0.1363$, $\beta_{source\_email} = -0.0081$), implying that non-social sources of information have a negative impact on App use. Whereas, the regression coefficients for social sources are positive and significant (Passengering: $\beta_{source\_invited} = 0.2449$, $\beta_{source\_nf} = 0.4224$, $\beta_{source\_socialad} =0.3196$; Stories: $\beta_{source\_invited} = 0.2561$, $\beta_{source\_nf} = 0.1228$, $\beta_{source\_socialad} = 0.1015$; Photos: $\beta_{source\_invited} = 0.2079$, $\beta_{source\_nf} = 0.1375$, $\beta_{source\_socialad} = 0.0729$), indicating that social sources of information have a positive impact on a users' App use. These findings support Proposition 1 that effectiveness of the information, as measured by App use, is higher when social information is presented in the ad as compared to when no social information is presented.

More interestingly, while the regression coefficients of social sources variables are positive; the regression coefficients of non-social sources are negative.

These findings indicate that media engagement is effective under social sources of information supporting Proposition 2 for App use. To further analyze how social information interacts with media engagement to impact App use, I examine the effect sizes of the different types of engagement (active versus passive) regression coefficients on App use outcomes.

For user generated content (stories and photos) App use outcomes, the regression coefficients of active social sources (friend invite) rank higher (Stories: $\beta_{source\_invited} = 0.2561$; Photos: $\beta_{source\_invited} = 0.2079$) than the regression coefficients of passive social sources (newsfeed and social ads - Stories: $\beta_{source\_nf} = 0.1228$, $\beta_{source\_socialad} = 0.1015$; Photos: $\beta_{source\_nf} = 0.1375$, $\beta_{source\_socialad} = 0.0729$), implying that active social sources are more effective than passive social sources when it comes to direct App use of a user (for example, user generated content). Similarly, the regression coefficients of active non-social sources (email) on user generated content App use outcomes, although negative, have a less negative impact than passive non-social sources (traditional ad). As for passengering App use outcome, the regression coefficients of passive social sources ($\beta_{source\_nf} = 0.4224$; $\beta_{source\_socialad} =0.3196$) rank higher than the regression coefficients of active social sources ($\beta_{source\_invited} = 0.3196$), indicating that passive social sources are more effective towards the App use of the overall system rather than to the user herself (for example, more passengers on planes). Non-social sources, although estimates are negative, also show a less negative impact on passive sources than active sources for passengering App use outcome.   These findings show that while active social sources are more effective in direct App use (supporting proposition 2); passive social sources are more effective in indirect App use.

I also find several interesting controls to be significant, particularly relating to user gender, adoption time, Facebook “Like”-s and activities, Facebook privacy settings, and friend adoption information.

As for gender, parallel to the intuition that females tend to use Apps more than males, the coefficient for *Female* is positive on all App use outcomes, implying that females are associated with higher App use as compared to men. In terms of adoption time, I find that users that adopted earlier have a higher App activity *(day_lvl_adop)*, largely because they had more time to use the App.

I also find a users’ Facebook “Like”-s and activities to be significant. The coefficient *likes_count* is positive, implying that the higher the total Facebook “Like”-s the higher the App use. This finding is consistent with the notion that Facebook “Like”-s can be considered as a measure of Facebook activity; and the more the user is active on Facebook, the higher the App activity. Furthermore, certain specific “Like” categories on Facebook have significant impacts as well (for example, *fbc_interests, fbc_activities*).

With regards to Facebook privacy settings, the coefficient (*ln_ttl_privacy*) is positive on the passengering outcome (an indirect use); and negative on the stories and photos outcomes (user generated content – direct use). This finding denotes that the higher the privacy settings of a user, the higher the passengering activity; yet the more public the higher the users’ activity in writing stories or posting photos. This finding makes sense in that users that do not want their information to be visible to all would only be passengers (a type of App use that does not present information about the user). Less private users, that don't mind their information going public, would post stories and pictures.

Lastly, in terms of friend adoption information, the coefficients *FirstDeg_FrndAdop* and *SecDeg_FrndAdop* are both positive and significant, showing that the number of friends that have "Like"-d the App at time of adoption, being first or second degree friends, is positively associated with App use. Yet the actual total number of Facebook friends (*num_ofFB_friends*) although the estimate is significant, has a coefficient of 0, meaning that the number of Facebook friends does not really impact App use. On a similar tone, the negative and significant coefficients on *threshold_percFriendAdop* show that the smaller the lag between the last friend that adopted and a users' adoption time, the higher the App use.

2.5.3 App Diffusion

Table 2.6c reports the second stage of the Heckman regression model for the App diffusion outcomes (a) Log Diameter, and (b) Log Outdegree. Columns 1 and 2 of Table 2.6c report the estimation results for the FanPage sample. Columns 3 and 4 report the estimation results for the Top10CC sample. Lastly, columns 5 and 6 report the estimation results for the more robust Intersection sample.

The signs and significance of the coefficients from the three samples are consistent, particularly the source of adoption estimates relative to the baseline. Referring to columns 5 and 6 in Table 2.6c, the estimation coefficients reported for *source_invited*, *source_nf* and *source_socialad* are positive and significant, implying that social sources of information have a positive impact on a users' App diffusion.

Interestingly, non-social sources - *source_ad* and *source_email* - are not significant predictors of App diffusion as compared to the baseline. Taken together, these findings support Proposition 1 for App diffusion.

Once again, the findings that media engagement has no significant impact on App diffusion when information sources are non-social; and that the regression coefficients of social source variables are positive and significant, indicate that media engagement is effective under social sources of information supporting Proposition 2 for App diffusion. To further analyze how social information interacts with media engagement to impact App diffusion, I examine the effect sizes of the different engagement (active versus passive) regression coefficients on App diffusion outcomes.

In examining diffusion with respect to the depth of the network (diameter), the coefficients on passive social sources of information ($\beta_{source\_nf} = 0.3339$; $\beta_{source\_socialad} = 0.2369$) rank higher than the coefficients on active social sources of information ($\beta_{source\_invited} = 0.1725$), implying that passive social sources of information are more effective than active social sources of information when it comes to diffusion. On the other hand, in examining diffusion with respect to the width of the network reach (outdegree) or local diffusion, the coefficients on active social sources of information ($\beta_{source\_invited} = 0.4673$) rank higher than the coefficients on passive social sources of information ($\beta_{source\_nf} = 0.1588$; $\beta_{source\_socialad} = 0.0496$), indicating that that active social sources of information are more effective than passive social sources of information when it comes to local diffusion. These findings are consistent and similar to App use findings on media engagement interaction. Active social sources are more effective than passive social sources when it comes to direct user impact (for example, local diffusion), supporting proposition 2. However, passive social sources are more effective towards the overall system (for example, beyond local diffusion).

I also find several interesting controls to be significant, particularly relating to user gender, adoption time, Facebook “Like”-s and activities, Facebook privacy settings, and friend adoption information.

As for gender, I find that the coefficient on *Female* is positive, indicating that on average females are associated with higher App diffusion as compared to males. Again this is in line with the reasoning that females tend to use Apps more than males. As for time of adoption, I find that users that adopted earlier have a higher App diffusion (*day_lvl_adop*), which is reasonable since they had more time to send out invitations to friends. I also find that certain specific “Like” categories on Facebook have significant impacts as well (for example, *fbc_interests*, *fbc_fav_athletes*), highlighting the fact that Facebook activities can be a good predictor of App diffusion.

In terms of Facebook privacy settings, the coefficient on *ln_ttl_privacy* is negative for both App diffusion outcomes, showing that the lower the privacy of the user the higher the App diffusion. This result is in line with the fact that users who are private may not want to share the App and send invitations.

With regards to friend adoption information, the coefficient *FirstDeg_FrndAdop* and *SecDeg_FrndAdop* are both positive and significant, showing that the number of friends that have “Like”-d the App at time of adoption, being first or second degree friends, is positively associated with App diffusion outcomes. The coefficient on *num_ofFB_friends* although significant is 0, denoting that the number of Facebook friends does not have any real impact on App diffusion.

2.5.4 Robustness Checks

To ensure the robustness of my findings, I conducted a series of additional tests. In this section, I present different robustness checks that address potential concerns relating to endogeneity, homophily versus social influence, and sample representativeness.

2.5.4.1 Potential Endogeneity

The higher the user's App use, the more likely is the user to send out invitations, resulting in higher diffusion. Therefore, App use may be correlated with some unobservable user-specific characteristics that might influence App diffusion. To control for this potential problem, I use a Two Stage Least-Squares (2SLS) regression with IV. Under the 2SLS approach, in the first stage, each endogenous variable is regressed on all valid instruments, including the full set of exogenous variables in the main regression. Since the instruments are exogenous, these approximations of the endogenous covariates will not be correlated with the error term. In the second stage, each endogenous covariate is replaced with its approximation estimated in the first stage and the regression is estimated as usual. The slope estimator thus obtained is consistent (Wooldridge, 2001). The intuition behind the use of IV's is that they are likely to be correlated with the relevant independent variables but uncorrelated with unobservable characteristics that may influence the dependent variable. In particular, a valid IV for App use will be correlated with App use but not with the second stage error.

I instrument for total App use with Facebook "Like"-s count, which was highly significant in predicting App use outcomes. In theory a user that has more "Like"-s on Facebook (a type of Facebook activity) will be more likely to be active on Apps. Although Facebook Likes Count might impact her App use behavior, it is unlikely to impact App diffusion, as sending out invitations and having friends accepting them is an act beyond one's direct Facebook "Like"-ing activity. In terms of statistical correlation, I find that users Facebook Likes Count is -0.2147 correlated with total App use, and 0.0001 with diameter diffusion outcome, and -0.0235 with the outdegree diffusion outcome. I conduct a 2SLS regression, where I instrument for total App use in the first stage by using a user's Facebook Likes Count as an IV. Moreover, the total App use of a user is measured at time *t* whereas diffusion metrics are measured at time *t+1*, such that the total App activity of a user came before the friend's acceptance of the invitation. Table 2.7 shows the regression estimates of the 2SLS model for App diffusion outcomes. The first-stage F statistic for both App diffusion outcomes is highly significant and much higher than the minimum value of 10, alleviating weak instrument concerns (Staiger & Stock, 1997). More importantly, the regression coefficients in Table 2.7 for both types of App diffusion outcomes are consistent with my results in terms of sign and significance.

2.5.4.2 Homophily versus Social Influence

To distinguish homophily driven diffusion from social influence, several approaches for identifying peer effects have been proposed, including peer effects models and extended spatial autoregressive models (Kelejian & Prucha,1998; Oestreicher-Singer & Sundararajan, 2012; Trusov et al., 2009, Bramoulle et al., 2009), actor-oriented models (for example, Snijders et al., 2007), dynamic matched sample estimation (Aral et al., 2009), structural models (for example, Ghose & Han, 2010), and ad hoc approaches (Christakis & Fowler, 2007). Natural experiments are by far the best approach (for example, Sacerdote, 2001; Tucker, 2008). Therefore, to distinguish homophily driven diffusion from social influence, I analyze data from a natural experiment, on the move from traditional ads to social ads.

Initially, only traditional ads were presented to users; however, after mid-April, social ads were introduced. In particular, after mid-April, if a user had friends that "Like"-d the App, the user was presented with a social ad as opposed to a traditional ad. However, after mid-April, if a user had no friends that had "Like"-d the App, then the user was presented with a traditional ad. In sum, before mid-April only traditional ads were presented to users regardless of whether or not they had friends that "Like"-d the App; after mid-April social ads were presented if users had friends that "Like"-d the App. Table 2.8 shows the frequency distribution of the four categories of the before social ads and after social ads groups ((1)Before Social Ads & Friends "Like"-d (2) Before Social Ads & No Friends "Like"-d (3) After Social Ads & Friends "Like"-d (4) After Social Ads & No Friends "Like"-d). Table 2.9 shows the averages and paired two sample t-tests of the before and after groups to ensure that groups are not inherently different. Both averages and t-tests show that there is no significant difference between the before and after groups.

To examine the impact of traditional non-social ads versus social ads on App use and App diffusion, I conduct an ordinary least squares (OLS) regression for App use outcomes, and a 2SLS regression for App diffusion outcomes. Table 2.10a reports

OLS regressions results of users that adopted through the various groups of ads on App use outcomes ((a) Log(Passenger), (b)Log(Stories), and (c)Log(Photos)). Table 2.10b reports 2SLS regression model for App diffusion outcomes ((a) Log(Diameter) and (b) Log(Outdegree)) using Facebook Likes count as an IV for total App use; the total App use is also lagged one period to control for potential endogeneity.

The regression coefficients of social ads are all positive and significant for both App use and App diffusion outcomes. These estimates imply that users that come from social ads are associated with higher App use and App diffusion as compared to users that came from traditional ads and had friends that liked the App. The coefficients for *before_nofriendsliked- Traditional Ad* are negative and significant for both App use and App diffusion outcomes. This finding indicates that users that came through a traditional ad with no friends that “Like”-d the App are negatively associated with App use and App diffusion as compared to users that came through a traditional ad with friends that liked the App. Lastly, I find no significant difference between the before and after users that came from traditional ads and had no friends that liked the App for both App use and App diffusion outcomes. Overall, these results show that the App use and App diffusion of a user is impacted by the knowledge that their friends “Like” the App at time of adoption, highlighting the role of visibility of social information.

2.5.4.3 Representative Sample

Selection effects could occur when users in the study sample are not representative of the overall Facebook population. Obtaining official demographic statistics on Facebook’s user base from Facebook is somewhat challenging because

Facebook does not publish such information. I therefore use recently released statistics on Facebook demographics by socialbakers.com and istrategylabs.com, both being social targeting advertisement services that have a focus on Facebook. Figures 2.4a and 2.4b show the overall comparison of my study sample to the Facebook population. Specifically, Figure 2.4a displays the geographic distribution and Figure 2.4b displays the gender distribution. In general, I find that although my sample has a slightly higher percentage of women and of users on the American continent than the Facebook population, the demographics of my study population are comparable to those of the broader Facebook population. Also, the published Facebook demographics fall within less than two standard deviations of my study sample means.

2.5.4.4 Additional Specifications

I also estimated several alternative model specifications. Using all three samples, I ran a Probit and a Logit model for predicting App adoption (Equation 2.1). I further looked at Probit with IV by using the MediaExposure variable as an instrumental variable. All the three alternative models provided highly consistent results in terms of sign and significance.

App use and App diffusion outcomes are count data and might have a distribution that is similar to the Poisson distribution. I therefore ran Poisson regression models for App use and App diffusion (Equations 2.3 and 2.4 respectively), on all three samples. Regression coefficients are highly consistent with my findings in terms of sign and significance. Furthermore, since the App use and App diffusion outcomes are count data, the data may have many zero values for some users (there might be over-dispersion of zeros in the data), thus a negative binomial model might also be appropriate. I estimated a negative binomial regression model for App use and App diffusion on all three samples. I obtained consistent findings in terms of sign and significance of the estimates of interest.

2.6 Conclusions and Implications

Advances in Web 2.0 technologies have significantly transformed the way in which firms interact and market to their consumers. Most noticeably, there has been a significant growth in the number of firms using social media advertising, and, in particular, Facebook as a main platform to target their marketing efforts. Such social media advertising mechanisms provide and make visible information on what "others" in a consumer's social network are doing. Prior to Web 2.0, such "social" information was difficult to obtain. While advances in technology are creating new opportunities for firms to directly facilitate and manage consumer social interactions, they also impose new challenges. Distinct strategic managerial actions are often necessary in such new social marketplaces. Although there have been many studies on the significance of peer effects, my study is one of the first attempts to understand how such increased information visibility on "others" affects an individual's behavior and ultimately advertising effectiveness in the social media marketplace.

From a theoretical perspective, my study complements and extends the literature in IS that has studied social influences in technology adoption, use and diffusion behavior. In particular, I empirically test and extend social influence theories to the online social media marketplace. I show the role of social information in alleviating the negative impacts of the advertisements, and its essential role as a new mechanism for advertising effectiveness.

My findings not only ventures to gain a richer theoretical understanding of social influence, but also to find ways through which one might ultimately increase the effectiveness of social media marketing and viral product diffusion. The results of this study offer some interesting implications for firms in managing consumer social interactions and advertising. My work also sheds light on how viral products can be designed to foster social contagion.

I conduct my study on a Facebook App, and present important new results on how social information in advertisements impacts user adoption, use and diffusion, and, ultimately, the advertising effectiveness. Overall, I find that social information has econometrically identifiable positive impacts on social influence, product diffusion and ultimately advertising effectiveness. Interestingly, it seems that with the increase of ad clutter online, advertisements with no social information have a negative impact on App use. This finding provides managers and technology developer's compelling evidence on the importance of emphasizing social information in designing their communication strategies. Furthermore, since the combination of social features seems to drive a positive feedback loop in which product use drives peer adoption, and peer adoption drives product use. If so, managers should seek to enable this feedback loop by designing both viral features and social features into their products.

I also examine how media engagement interacts with social information. My findings show that media engagement that lacks social information may no longer be effective in terms of use and diffusion. Media engagement has a significant positive impact only when social information is visible. Studies have shown that consumer avoidance of advertising in a traditional advertising medium increases with consumption of the medium by way of repeated exposures (Elliott & Speck 1998; Speck & Elliott 1997). My findings highlight how this perception is now evident in Web 2.0, in that traditional online advertising and traditional media engagement may no longer be as effective, and they underscore the importance of use of social information in online advertising. This result provides important information to managers and technological developers, given that technology is evolving at a fast pace, their marketing methods and strategy design should reflect this evolution.

Furthermore, I not only examine the social information on first degree friends, but also second degree friends. Thus, I am able to also assess the influence of the behavior of those with whom the focal individual may only come into contact infrequently, such as at shopping malls or social events. My results underscore the salience of these more digital "influences" in driving adoption, use and diffusion.

I also look closely at the interaction effect of social information and media engagement. I find that for user generated content (direct user impact) active social sources have a stronger impact than passive social sources. Yet passive social sources are more effective than active social sources towards the overall system - in terms of the indirect use outcome (passengering). As for diffusion, I find that active social sources of information are more effective than passive social sources of information when it comes to local diffusion. More interestingly, and in contrast to the general notion that WOM is considered to be more effective at promoting product diffusion when it is active, I find that in terms of network reach and beyond-local diffusion, passive social information sources have a stronger impact than active social information sources. In general, active social advertising is more locally effective but less globally effective than passive social advertising, therefore managers should optimally design their viral features accordingly. In particular, to foster diffusion marketers should invest in passive social advertising methods. These findings highlight the importance of understanding the different features in social media marketing, and could enable firms to optimally create and manage social contagion when it comes to new product diffusion.

Findings from my study indicate that firms would do best to optimize their investment in advertisements such that they provide different levels of information to induce different outcomes. For example, to encourage App virality, ads should include sharing information such as a link to share the App with their friends, and provide information that “User X” has shared the App with “User Z.” This type of social information might further induce virality for the App and the overall success of the ad.

Another interesting finding is that the Facebook privacy settings of a user have a significant impact on user behavior in terms of App use and diffusion. This finding will allow managers and marketers to target users more precisely by factoring in the Facebook privacy settings of a user in their marketing algorithms. This will reduce marketing costs and will enable firms to be more efficient.

Overall my study provides empirical evidence on how Web 2.0 has changed the role of the consumer, and it emphasizes the importance of using consumers and their social information as an essential marketing medium. Firms can use the findings here to develop guidelines for optimal investments in various social marketing methods. The findings of my study would guide platform developers to enable and constrain the visibility of information that operates in their ecosystem, and to engineer the user experience to increase sharing, interaction and virality.

2.7 Limitations and Future Work

As for my study, although I was able to collect a subsample of the users that saw the advertisement for the App and did not “Like” the App, I could not observe all of the users. Future studies can replicate my study and control for this selection bias by use of randomized experiments. In addition, I was only able to collect gender and location as demographic information, and although I used hazard models to account for the omitted variable bias, other demographic information, such as education and work, might have an impact on outcomes.

Also, I was not able to directly measure the strength of the social tie between users. A fine grained analysis could be useful to provide additional insight into social information of different “strengths,” and their impact on user behavior. Also, a natural question that arises is whether some social information from certain friends is more salient and “influential” than others. Rogers (1995) refers to such individuals in his identification of opinion leaders, as fundamental forces in the diffusion of innovations. The identification of such leaders of social information would be an interesting and important area for future research.

Although it is abundantly clear that social influence is present, the precise mechanism through which social influence exerts itself in this specific context is less understood. I identified social information as a possibility, but there may be others, and my analysis does not allow me to distinguish between them. More qualitative data via interviews or surveys may shed further light on this issue. Some studies on the adoption over the Internet have shown that non-adoption was driven by a different set of factors (for example, fear of obsolescence, mistrust), from those driving adoption (Brown & Venkatesh, 2005). It would be interesting to examine what role social information plays in non-adoption of Apps, and whether social information can relieve or compound such fears.

Research on advertising has shown that repeated exposure to traditional advertising media triggers the ad schema, thus leading to a reduction in the effectiveness of the ad (Elliott & Speck 1998; Speck & Elliott 1997). My study provides evidence on how social information can alleviate such triggers. However, with time, this notion could also take form in the context of social information. With the increased use of social media advertising, online consumers might be bombarded with social advertising which might ultimately trigger their ad schemas. Therefore, future studies should examine if users, with time, change their ad schemas towards social media advertising, and how the type of social information presented can help mitigate this effect.

Lastly, because my detailed study was limited to a single App, corroboration of these novel findings by subsequent research would be useful. This is especially so as both the amount of adoption, use and diffusion of the product is likely to be contingent on the nature of the product. Overall, the findings of my study can be generalized to most Facebook Apps, and to social media advertising that functions in a similar way to Facebook (for example, Twitter, Google plus).

Chapter 3: Information Spillover and Strategic Behaviors in Open Innovation Crowdsourcing Contests: An Empirical Investigation

3.1 Introduction

The ability of online markets to efficiently bring together individuals and businesses has redefined and transformed traditional ways of conducting business. More recently, there has been an explosion of new business models that leverage online interactions and the "wisdom of the crowds". In particular, firms have increasingly begun to leverage online crowdsourcing marketplaces to seek solutions to business problems as well as to undertake research – activities that were traditionally performed within the boundaries of the organization. Crowdsourcing markets use a distributed problem solving and production model and seek to tap into the wisdom of crowds to provide solutions or products for firms (Kleemann et al., 2008). Lego, for instance, encourages its most fanatical customers to redesign its famous sets (Rodgers, 2011). Other big corporations such as Dell, have turned customer complaints into increased profit margins by tapping the crowd for solutions to their problems (Bensen, 2013). An important objective of these crowdsourcing markets is to attract high quality solvers, and to obtain good, diverse solutions (Terwiesch & Ulrich, 2009; Terwiesch & Xu, 2008). The effectiveness and success of a crowdsourcing marketplace depends largely on the market's ability to not only incentivize participants to submit high quality solutions, but also deter strategic gaming by the participants.

Most online crowdsourcing markets use “contests”, with anonymous users (“the crowd”) submitting solutions to the contest holder’s problem and competing for prize money. Contest design plays a crucial role in the success of the marketplace. It is widely recognized that contestants’ incentive to exert effort depends largely on the competitive environment as defined by the rules of the contest. Contests are often characterized by a number of parameters such as the number of players, heterogeneity of players, the number and amount of prizes, and the information available to participants, among others. All of these factors play an important role in influencing the behavior of individual players. While a number of these factors play a critical role in offline contests as well, an important development in online contests is the increased availability of information – in particular, information relating to other contestants. Two broad categories of contests are popular in online crowdsourcing markets – open contests, wherein information about other contestants are made visible to all participants, and closed (also known as blind) contests, wherein the visibility of such information is limited. Open contests in particular, such as the one used in the marketplace I study, make information about a contestant as well as her submissions visible to all other contestants. While this encourages greater participation, open contests also suffer from a number of drawbacks.

Participants in these online crowdsourcing markets are faced with a number of strategic choices – an important choice being the order of submission or timing of entry in a contest. While early movers (i.e., contestants who submit early) may enjoy some benefits and also be able to deter entry of later contestants, late movers might benefit from their ability to learn from early entrants. However, another critical feature of online innovation contests is the provision of feedback to contestants by the contest holders. Given the uncertainties surrounding the contest holder's requirements and tastes, feedback provided by the contest holder on the submissions can be very helpful to the contestant in refining their solutions. In particular, early submissions have a higher likelihood of obtaining valuable feedback that can help early entrants to refine and resubmit their solutions. While early entrants are more likely to benefit from feedback provided on their submission, the visibility of such feedback information to all contestants can provide later entrants valuable information that increases their chances of winning the contest. Poorly designed contests can dissuade potential submissions, while well-designed contests that deter strategic gaming by participants and can stimulate participation and growth of online crowdsourcing markets. Understanding how informational spillovers in open innovation contests impact the entry behaviors of contestants and their likelihood of winning can provide valuable insights into the effectiveness of open innovation contests.

My study analyzes data from one of the largest online crowdsourcing markets for design to examine the strategic behavior of contestants and their impact on outcomes. The marketplace I study allows individuals or businesses to setup contests for the design of logos, graphics, and websites. Contestants are usually individual designers who compete to provide solutions, with the winner being financially rewarded. The marketplace uses two types of contests (a) an open contest, where all submissions by contestants as well as the associated feedback provided by the contest holder are visible to all the participants, and (b) a blind contest, where information about a submission is not available to other contestants. The presence of these different types of contests enables us to study the role of informational spillovers relating to submissions as well as feedback on the strategic entry decisions of contestants as well as its impact on their likelihood of winning.

More specifically, this study seeks to investigate the following questions.

1) *How do informational spillovers relating to submissions (i.e., the ability to see earlier submissions) influence a focal contestant's behavior (timing of entry) and her likelihood of winning?*

2) *How do informational spillovers relating to feedback (i.e., the ability to see the feedback provided by the contest holder to earlier submissions) influence a focal contestant's behavior (timing of entry) and her likelihood of winning? Further, how does the type of feedback (specific versus generic)*[6] *impact these outcomes?*

I find that contest design, particularly relating to information visibility, significantly influences contestant's behavior as well as outcomes. Both early submissions as well as late submissions are more likely to win a contest compared to submissions at other times during the contest. In examining the impact of informational spillovers relating to the submissions, we find that late submissions are more likely to win in open contests. However, in blind contests where there are no informational spillovers and when no feedback is provided by the contest holder, there are no specific submission times that perform better than others. Only a contest holder's expertise, experience, and skill are associated with her likelihood of winning the contest. I also find evidence of informational spillovers relating to feedback. Interestingly, in both open as well as blind contests, when feedback is provided by the

[6] A definition of specific versus generic feedback will be provided later. See Figure 3.11**.**

contest holder to contestants, I find that early submissions are more likely to win, particularly when these early contestants make a resubmission.

I also find that the benefits from informational spillovers differ depending on the type of feedback provided by the contest holder. In open contests with feedback, the more specific the feedback given to others in a contest, the higher the chances of late submissions winning the contest. However, in blind contests where users cannot see each other's submissions, specific feedback given to other contestants does not benefit a focal user, while generic feedback increases the likelihood of late submissions winning the contest.

Finally, in examining the role of skill, experience, and expertise, I find that in the case of open contests, contestants with high skill, experience, or expertise that submit late are more likely to win. However, in the case of blind contests with feedback, I find that high skilled contestants who submit early are more likely to win.

My findings suggest that contestants in these markets strategically time their submissions to increase their chances of winning the contest. Interestingly, it is the contestants with higher skills, experience, and expertise that are more likely to act strategically and win the contest. My results have interesting implications for the design of innovation contests. While feedback provided by the contest holder to a contestant could benefit that contestant, I find that when this feedback is very specific, the informational spillovers are higher and other contestants that submit later tend to benefit from such feedback. Such information spillovers might not be detrimental to the contest holder; however they could discourage contestants from participating in such open contests. While informational spillovers relating to submissions as well as feedback could help later contestants converge to a winning solution more quickly, they could have an unintended side effect of reducing variety. Blind contests, on the other hand, could promote greater variety by reducing information spillovers from earlier submissions.

My study makes a number of contributions. It is one of the first studies to examine the role of informational spillovers in open innovation contests and the consequences of such informational spillovers. Using data on open as well as blind contests with and without feedbacks, my study is able to tease out the differing impacts of informational spillovers relating to submission and feedbacks. While a few recent studies have examined the role of the feedback process on the idea generation (Wooten et al., 2011), this is the first study to identify the effects of informational spillovers that occur when such feedback is made visible to other contestants. My study also contributes to the vast stream of research in marketing, economics, and strategy on the role of timing of entry and its implications for market outcomes. Most of the studies examining the timing of entry of market participants examine the strategic behavior of firms. This study is among the first studies to examine strategic entry behavior of individuals in a decentralized marketplace. Finally, and most importantly, my study contributes to the emerging literature on online crowdsourcing markets and the design of online contests.

The rest of the paper is organized as follows. Section 2 reviews the literature. Section 3 describes the research context and Section 4 describes the data and methodology. The results are in Section 5. Section 6 discusses the implications of my study and concludes. . Lastly, Section 7 describes the limitations of this study and future research.

3.2 Literature Review

This study draws upon research relating to a number of contexts including the optimal design of contests, timing of entry, the role of information externality, as well as the role of feedback in innovations.

3.2.1 Design of Contests

Most of the work relating to contest design is analytical in nature. Prior work has examined a number of contest-related factors to examine their impact on outcomes. One stream of literature has focused on the “prize structure”, including how many prizes should be offered and how the total award should be allocated among them. Kalra and Shi (2001) study a multiple-player model of sales contests and find that the number of prizes should be increased and the spread should be decreased when salespeople are more risk averse. Glazer and Hassin (1988) and Moldovanu and Sela (2001) study the design of prize structures in contests in which contestants differ in skills. Glazer and Hassin show that the winner-take-all design is optimal if the skill distribution is uniform and the performance function is linear. Moldovanu and Sela (2001) show that the winner-take-all design is optimal under general distributions, as long as the performance function is linear. Lastly, Liu et al. (2007) study a one period model of a consumer contest in which consumers’ performance is a multiplicative function of their skill and consumption, and winners are determined by rank order of performance. They find that both skill distribution and number of contestants play an important role in determining the optimal prize structure in consumer contests.

Researchers have examined the design of an optimal contest that generates the highest revenue, as well as contest structures that maximize the effort exerted by contestants. The "contest structure" involves whether and how to segment the consumer population and how to choose a performance evaluation criteria. Gradstein and Konrad (1999), for example, provide a rationale for a multi-stage contest design by endogenizing the choice of contest structure. They demonstrate that, depending on a return to scale parameter of the contest success function, a multi-stage contest may encourage higher effort by the participants than a one-stage contest. Similarly, Baik and Lee (2000) study a two-stage contest with effort carryovers. They show that in the case of player-specific effort carryovers, the rent-dissipation rate (defined as the ratio of the expended total effort to the value of the prize) increases in the carryover rate and the rent is fully dissipated with carryover rate equal to one. Fu and Lu (2006) investigate the optimal structure of a multistage sequential-elimination contest with pooling competition in each stage. They show that the optimal contest excludes one contestant at each stage until the finale in which a single winner takes the entire prize. Liu et al. (2007) employ a game-theoretical approach to investigate consumer contest design issues, including segmentation, and handicapping. They find that increasing contest size is beneficial to the marketer and that the marketer may achieve less dispersive skill distributions by segmenting or screening contestants according to their skill levels, and by adopting a performance evaluation scheme that handicaps high-skilled contestants.

While there is a growing body of theoretical research examining the impact of different contest parameters, empirical research on the impact of contest design is scant, and has been limited to examining the impact of prize on effort exerted. For example, Maloney and McCormick (2000), analyze responses of individual runners to different prizes. They find a significant relation between the performance and the prize value and that higher prize values are associated with higher effort levels. Lynch and Zax (2000) examine data on road races in the United States. They find that the performance increases in response to larger prize spreads. However, when controlling for ability, the impact of the prize spread disappears. The authors conclude that the larger prize spreads produce better performance not because they encourage all runners to run faster but because they attract faster runners. A recent experimental study by Shermeta (2011) compares the performance of four simultaneous lottery contests and the effort levels exerted. Most of the empirical studies of contest design focus on contests in offline settings. An exception is a recent study by Huang et al. (2012), where the authors examine the effect of incentive prize structure design of online crowdsourcing contests on the solutions produced by the crowd. They use data from a crowdsourcing marketplace Threadless, and find that participants exert less effort as competition for the prize increases, indicating that the prize may adversely affect the quality of the solutions produced by the crowd.

While the "prize structure" as well as the "contest structure" are important determinants of a number of outcomes of interest, there are hardly any studies that have examined the differences in "information structures" – in particular, information visibility – within contest and their impacts on outcomes. This study is among the first to empirically examine the role of different information visibility regimes on the behavior of contestants as well as contest outcomes in online settings.

3.2.2 Timing of Entry

There is substantial theoretical and analytical research that highlights the importance of timing of entry for market participants (for instance see, Lieberman & Montgomery, 1988; Urban & Star, 1991). First movers, for instance, have been shown to deter entry of later entrants by locking in consumers (Lieberman & Montgomery, 1988; Klemperer, 1987; Dewan et al., 2003; Lee & Grewal, 2004). Other studies have explored additional benefits of early entry including the evidence of scale effects (Rao & Rutenberg, 1979), experience effects (Smiley & Ravid, 1983), asymmetric information about product quality and risk averse buyers (Conrad, 1983), and reputational effects (Bain, 1956; Krouse, 1984), among others. In contrast to these studies of early mover advantages, other studies find that late movers have an advantage when they can lower their uncertainty and costs by learning from the experiences of early movers (Reinganum, 1981; Fudenberg & Tirole, 1985; Dutta et al., 2005; Hoppe, 2000; Hoppe & Lehmann-Grube, 2004). The main disadvantage of pioneers is that it is generally more costly to be a pioneer than to be an early follower or a late entrant since product innovation tends to be more costly than product imitation (Mansfield et al., 1981; Levin et al., 1987). Other studies have explored how a later entrant can diminish the impact of the first movers by moving away from the first mover, and by developing a more desirable position (Carpenter & Nakamoto, 1989; Hauser & Shugan, 1983).

Findings of empirical studies examining the impacts of timing of entry are mixed (for reviews of empirical findings see, Kerin et al., 1992; Robinson et al., 1994; Kalyanaram et al., 1995; Zahra et al., 1995; Mueller, 1997; Lieberman & Montgomery, 1998). For instance, Robinson and Fornell (1985) analyzed 371 consumer goods firms and found that first movers had around 20 percent higher market shares than later entrants. Similarly, Robinson (1988) studied 1209 mature industrial goods businesses and found that order of entry alone explained 8.9% of the variation in market share and that first movers had higher market shares than later entrants. Other studies have also found support for first mover advantages (for example, Lillien & Yoon, 1990; Cooper, 1979; Lieberman, 1989; Parry & Bass, 1990). Other empirical studies have found that later movers have higher market share and better performance than early movers. Shankar, Carpenter and Krishnamurthi (1998) examine brand sales in the prescription drug market and find than an innovative late mover can create sustainable advantage. Huff and Robinson (1994) also find that pioneer's relative market share declines over time with competition. Other studies examining firms have also found negative effects of early entry (for example, Kalyanaram & Wittink, 1994; Brown & Lattin, 1994; Sullivan, 1992; Mascarenhas, 2006). Most of the empirical research on timing of entry to market has been limited to offline settings and have focused on the firm as the primary unit of analysis.

More recently, there has been growing number of empirical studies on timing of entry of individual participants in online auction marketplaces. Studies on online bidding behavior suggest that early and late bidding could affect outcomes in opposite ways leading to opposing conclusions. On the one hand studies have found that a significant amount of bidders bid early in the auction (Bajari & Hortacsu¸ 2003; Hasker et al., 2004), while on the other, studies have found that a substantial fraction of bidders submit their bids towards the end of an auction (Ockenfels & Roth, 2002, 2006; Roth & Ockenfels, 2002). Early bidding can lower a bidder's cost of searching for alternatives, at the same time making other competitors less interested in competing (Vadovic, 2009). Studies have also shown that entry deterrence incentive leads bidders to bid early to discourage other bidders from entering the auction (Nekipelov, 2007). Although a late online bid is at a large risk of not being successfully transmitted due to network traffic (Bajari & Hortaçsu, 2004) and is discouraged by online auction sites, late bidding softens competition compared to early bidding (Bajari & Hortaçsu, 2004) and is a deliberate strategy meant to avoid incremental bidding or 'price war' behaviors (Schindler, 2003; Bajari & Hortaçsu, 2004; Roth & Ockenfels, 2002). In addition, late bidding is considered to be a best response strategy at times for informed bidders to protect their information (Roth & Ockenfels, 2002) and prevent learning (Nekipelov, 2007).

Despite the recent growth of crowdsourcing marketplaces, little attention has been paid to timing of entry in online crowdsourcing contests. The findings from a few recent studies have shown mixed results. Archak (2010) examines a crowdsourcing software development website TopCoder.com, and finds that high rated contestants face tougher competition in the contest phase. Yet they strategically play Stackelberg leaders in the registration phase of the contest by committing early to particular projects, thus deterring entry of opponents to that contest. Yang et al.

(2011) use Taskn.com to show that winners are more likely to be those who submit early or later during the submission period as opposed to those submit in the middle. They also find that intentionally waiting to submit solutions later is associated with higher winning probability. Yet they find evidence of such behavior only in the context of pure ideation based projects such as naming projects and creative writing projects.

The multiplicity of explanations provided in the literature regarding the evidence and causes of timing of entry enables us to appreciate the richness and complexity of markets, and in particular the importance of entry time as a competitive strategic choice in different market structures. These studies highlight the importance of understanding the role of different market structures and "information structures" on the timing of entry. Thus far most of the empirical research on time to entry has been conducted on the firm level of analysis and in offline settings. This study is among the first to empirically examine the impact of different information visibility regimes on contestants' timing of entry and their related outcomes in an online crowdsourcing marketplace.

3.2.3 Information Spillover

The impact and value of information spillover or information externalities have been examined in wide variety of fields, including economics, IT and finance. There are various strands of research relating to information spillover in different literatures; however, the main idea is that the lack of information about some essential variable that is of public interest can be compensated for, at least partially, by looking at what other similar agents do. For example, if the information that is privately available to agent A to form his decision has some value for agent B (a neighbor of A) the observation of A's actions can help B make a better decision since A's actions will partly reveal his information.

In general information spillover occurs when each agent has some private piece of information which, if combined with the others' would increase the information available to each about some relevant common variable. If pooling is ruled out, each agent's private information will be embedded in his decisions. The other agents' choices become an alternative source of information. As a consequence, individual agents' decisions will be affected both by their private information and by other agents' actions. Specifically, private information spills over through individual actions. In the open innovation contest, the lack of information on the taste of the contest holder is the main variable of interest for all contestants. Visibility of certain information can provide insights into the taste of the contest holder which can be harvested by each agent.

Most of the work relating to information spillovers is analytical in nature. Prior theoretical work has examined how information spillover occurs, and how it affects related outcomes of interest. Knowledge spillovers are external benefits dissipated to non-innovating firms from innovating firms due to the non-rival property of knowledge. Knowledge spillovers occur because accumulated knowledge at one firm can be transmitted to other firms due to its public good characteristics (Griliches 1979, 1998). In studying information spillovers, Brynjolfsson and Hitt (2003) find that a firm's IT-related knowledge can diffuse to rival firms through consulting firms, employee transfer, and educational institutions. Knowledge spillovers can result in learning, observation, and replication of others' innovations (Chang & Gurbuxani, 2012). Studies have also shown that knowledge spillovers can in turn become the source of long-run productivity growth (Romer 1986, 1994; Coe & Helpman, 1995). Other studies have argued that information spillovers diminish the incentive of innovators to undertake information production in the first place (Benveniste et al., 2002; Hoffmann-Burchardi, 2001). In innovation contests one firm's R&D effort may spillover and benefit its rival (D'Aspremont & Jacquemin, 1988). Kamien et al. (1992) show how positive input spillovers can affect R&D decisions when firms are engaged in a two-stage game of innovation.

Given the nature of information spillover, it is hard to empirically isolate and test for information spillovers (Chang & Gurbaxani, 2012). Consequently, empirical studies have been scarce and most are experimental studies conducted in offline settings. Cheng and Nault (2007) focus on industry-level spillover benefits that result from IT investments made by upstream industries. Chang and Gurbuxani (2012) examine the effects that result from IT related spillovers on firm level productivity. Experimental studies, mostly in the finance, have found that the disclosure of information about a firm presented in different ways affects the valuations and trades of investors and even experienced financial analysts (Hirst & Hopkins, 1998; Dietrich et al., 2001; Hopkins et al., 2000). There is also evidence that individuals fail to make use of all publicly available information (Lipe, 1998).There is other evidence suggesting that investors' and analysts' assessments are influenced by the format and salience with which public signals are presented (Hand,1990).

Thus far research on information spillover has been largely theoretical, and the few empirical studies have been conducted in offline settings at industry and firm level of analysis. There are hardly any studies that have empirically examined information spillovers at the individual level in online settings. This study is among the first to empirically examine the role of information spillovers on the behavior of contestants and related outcomes in online crowdsourcing markets for innovation.

3.2.4 Feedback in Innovation Processes

Several motivation theories attest that feedback is effective for motivating goal pursuit because it increases outcome expectancy of the goal and perceived self-efficacy of the pursuer (Atkinson, 1964; Bandura & Cervone, 1983; Lewin, 1935; Weiner, 1974; Zajonc & Brickman, 1969). Feedback on successful and failed actions allows individuals to adjust and direct their efforts to match the challenge they are facing (Bandura, 1991; Dweck & Leggett, 1988; Festinger, 1954; Locke & Latham, 1990). The fundamental idea that information can lead to learning is explored in the literature of mental models. Mental models activate when new information is incorporated into one's base of knowledge, resulting in conceptual change. Enhancement occurs when consistent information reinforces the existing framework, and revision occurs when the new information is inconsistent with prior beliefs (Vosniadou, 1994). Vosniadou points out that learning failures are more probable when revisions are needed, which can produce inconsistencies. This suggests that feedback schemes that increase the amount of accurate information will reduce misunderstandings, enrich learning related to the quality function, and thus improve the average quality of submissions.

A significant amount of empirical research has been conducted on the role of feedback and innovation. Tjosvold and McNelly (1988) study organizations and demonstrate that the quality and type of communication, rather than its frequency, improves organizational innovation. Their findings supports the theory that interaction and feedback will lead to increased levels of innovation. At the group level, several empirical studies link higher levels of information gathering and both internal and external group communication with better performance (Katz, 1982; Keller, 1986). Highsmith (1978) suggests that a lack of meaningful, positive feedback largely reduces the rate of idea generation in group sessions. He conducts a simulated study in an organizational context, and finds that the average number of ideas combined goes down in the absence of communication. However, when there is variability in communication, there is no effect on the average number of proposals combined or on the variance (Seshadri & Shapira, 2003). A recent field experiment conducted on online contests by Wooten and Ulrich (2011) find that the type of feedback (star ratings) given to a particular user is associated with differences in performance in the idea generation process.

The increase in interest on the role of feedback has been mainly due to Web 2.0, and due to the emergence of online feedback mechanisms. These mechanisms have proved to be viable for fostering cooperation among strangers by ensuring that the behavior is publicly visible, and may therefore, affect the behavior of users in the community. To date, most empirical studies have been conducted in offline settings. The unit of analysis thus far has been on firm and group level, with the exception of Wooten and Ulrich’s field experiment. This study is one of the first to empirically examine the role of feedback, at the individual level, on the behavior of contestants and contest outcomes in the online setting of open innovation crowdsourcing marketplace.

In open innovation crowdsourcing contests, contest holders can give direct feedback to users in the form of a star rating, textual feedback, or both. This feedback (star ratings or text reviews) is not only visible to the contestant given the feedback to but also to the rest of the contestants. While online feedback is visible not only to the user to whom feedback is given but also to the rest of the community, to date most studies have analyzed the direct impact of feedback on a user; what has not yet been studied is how different types of feedback given to a user impacts not only her behavior but also the behavior of other contestants in the contest. This study seeks to fill an important research gap, namely, the role of feedback provided to others on the behavior of contestants and contest outcomes in an open innovation crowdsourcing marketplace.

3.2.5 Research Hypotheses

Although submitting early will lead to higher chance of getting feedback, and being able to work upon the feedback for a higher probability of winning, I argue that information spillovers through the visibility of information in the contest benefits late entry since imitation costs are lower than production costs. In particular, when the design submissions are visible, design information spills over and contestants are able to see each other's designs. The visibility of designs in a contest, benefits contestants since imitation costs are lower than production costs; thus moving later will lead to a higher probability of winning. Thus, I hypothesize that in open innovation contests, moving later will lead to a higher probability of winning.

Proposition 1: *Visibility of designs of other contestants in a contest will benefit a contestant, such that she will "strategically wait"*[7] *to "imitate"*[8] *and submit later; leading to a higher probability of winning.*

Winning a design contest is mainly based on the tastes of a contest holder (Terwiesch & Xu, 2008). These tastes cannot be readily measured. However, feedback given in a contest by a contest holder can be used to signal the taste of a contest holder. Since feedback is visible to all in a contest, contestants are able to benefit from the information spillover of the contest holders' tastes from the feedback given to others in a contest when designs are visible; and build upon the feedback to their own designs. I therefore propose:

Proposition 2a: *Visibility of feedback to other contestants in a contest will benefit a contestant, such that she will "strategically wait" to build upon the feedback; leading to a higher probability of winning.*

Moreover, different types of feedback send out different forms of information. Text feedback in online contests is visible to all. Whether the text feedback is more general or more specific drives different levels of information spillover. Text feedback can be specific (for example, "I like submission X but please try a different shade of color.") or more generic (for example, "I like the submissions thus far but I would like to see more flowers."). When information about a specific feedback provided to a contestant spills over, other contestants can make better use of the

[7] I define "strategically waiting" as intentionally submitting later to benefit from information spillover.
[8] *Imitate* henceforth is defined as copying and building upon others' work.

feedback as compared to generic feedback. Therefore, if a contestant is given specific feedback when designs are visible, later entrant benefit more, than if the feedback is generic. I therefore propose:

Proposition 2b: *The higher the specificity of the feedback to other contestants in a contest, the higher the information spillover to a contestant; thus later submission will lead to a higher probability of winning.*

3.3 Research Context

In this study I use one of the most dominant crowdsourcing contest design websites to study the role of information externalities and their impact on the strategic behavior of participants. The community is a design crowdsourcing website where contests are held and users submit logos or website designs.

Figure 3.1 describes the 4 steps of the contest process from start to finish. The main steps are (1) Contest Launch, (2) Submissions, (3) Feedback, and (4) Announcing the Winner. Below I describe the steps more extensively.

3.3.1 Contest Launch

To launch a contest, a contest holder provides the following information:

- Contest Prize (the online community used in this study has a minimum prize of $299).
- Design description. The contest holder needs to provide project details which usually include objective, slogans, and any other information the contest holder is seeking.

- The contest type, whether open or blind. Blind contests are contests where the submitted solutions (logos, for this study) are not visible to anyone in the contest except for the contest holder. Open contests are contest where the submissions are visible to everyone. Also blind contests are hidden from search engines like Google, whereas open contests are visible in search engines.

The contest can then be launched and displayed in the design marketplace. Logo designs contests are open for 7 days. The contests are displayed based on the ending dates.

3.3.2 Submissions

Designers (contestants) who wish to enter a contest view and evaluate the contest. Figure 3.2 displays how contests are displayed in an open innovation marketplace. Designers can view several contests, the type of the contest (open or blind), the title of each contest, the end date, the number of entries, the prize amount, the contest holder, and details about the contest and contest holder. Figure 3.3 displays the information made public on the contest holder. Users can not only view current open contests of a contest holder, but also contest holders' past activity in terms of total contests held, total prizes awarded, and average feedback. Users can get more information on the contest requirements by viewing the design brief as shown in Figure 3.4. In addition, users are able to see the current submissions in a contest. Figure 3.5 displays submission entries of contestants along with the star feedback given by the contest holder. Users are able to obtain further information on contestants by clicking on their profiles. Figure 3.6 displays a contestant profile, which lists the total contests entered and won by the contestant, as well as her portfolio of designs.

Designers can submit a solution at any time before the end of the contest. If the contest is open, designers can see submission of other contestants; however, if the contest is blind, designers cannot see submission of other contestants. A key decision for a contestant is when to submit, (that is, early or late in the contest). Figure 3.7 shows the submission order frequency on the lifetime of the contests- note that while there is a high number of early entrants there is also a high number of late entrants.

3.3.3 Feedback

After contestants begin submitting solutions, the contest holder is encouraged by the community to send feedback to solvers about their solutions by communicating her average feedback % score. A contest holders average feedback (%) score depends on the amount of feedback given in previous contests. Feedback by a contest holder to contestants is given in two ways - (a) star rating of one to five stars on a submission, and (b) text feedback. I observe that contestants usually prefer to submit improved solutions after receiving feedback. Although the community encourages feedback there are quite a large number of contest holders that do not provide any feedback. Figure 3.8 displays the text feedback frequency, and Figure 3.9 shows the star feedback frequency. On average, both figures show that feedback is mainly given at the early stages of the contest timeline, and decreasing in frequency towards the end of the contest.

3.3.4 Announcing the Winner

Once the contest duration ends, the contest holder then announces a winner who is awarded the prize.

3.4 Data and Methodology

I collect data in 3 time periods, namely (a) Nov 2010, (b) April 2011, and (c) Nov 2011- Oct 2012. The sample is composed of 3,893,221 designer participations (or submissions) in 16,645 contests. 13,225 are open contests (of which 6091 have feedback and 7134 do not), and 3420 are blind contests (of which 2417 have feedback and 1003 do not). Table 3.1 lists the variables and their descriptions. Table 3.2a reports the descriptive statistics of the variables used in this study. On average, contest holders have held around 3 contests, have awarded around $718, and have provided an average of 70% feedback. Contestants on average have been on the crowdsourcing community for around 422 days, have an average of 13 wins, and have entered an average of 287 contests. Contests on average have 211 entries, 69 designers, and a prize of $517.

Table 3.2b reports the descriptive statistics of the variables in terms of the different contest designs and information regimes. Namely, (a) open contests with feedback, (b) open contests with no feedback, (c) blind contests with feedback, and (d) blind contests with no feedback. Table 3.2b displays that on average, there are more submission entries for open contests with feedback (267) than for open with no feedback (228), and even fewer for blind contests (197 with feedback and 151 with no feedback). A similar pattern is also reported in terms of the number of designers. Furthermore, in all contest designs, on average early submitters (defined as the first

20 percentile of submissions) have the highest number of resubmissions. In addition, for open contests, the averages reported for late submitters in terms of skill is higher than the averages reported for early and middle submitters. Yet for blind contests, the averages reported for early submitters in terms of skill is higher than the averages reported for late and middle submitters. The reported averages suggest that contestants with different skills are likely to have different entry times. Figure 3.10a briefly describes the benefits and drawbacks to a contest holder on the choice of the contest design, and Figure 3.10b briefly describes the benefits and drawbacks of a contestant on the choice of time to entry.

3.4.1 Variables

The following sections describe the variables used in this study. The data is in panel form and the unit of analysis is the contestant *i* in contest *t*.

3.4.1.1 Dependent Variable

The main dependent variable in this study is a binary variable *winner_dummy*, indicating whether or not the contestant *i* won contest *t*.

3.4.1.2 Independent Variables

To examine the impact of information spillover on contestant time of entry and contest outcome, I explore the following categories of explanatory variables related to the contestant *i* in contest *t*: (1) Submission order, (2) Feedback to others in a contest, (3) Direct feedback to a contestant, (4) Resubmissions, and (4) Expertise and past experience.

To capture timing of entry behavior I measure c_*suborder*, the order of the first submission of a contestant *i* in contest *t*, along with submission order squared *c_subOrder2* to capture any late strategic behavior.

Feedback by a contest holder to a contestant is either given through stars or through text. Both types of feedback (stars and text) are visible to others in a contest. To capture information spillover in terms of star feedback given to others, I measure the following metrics (a) *c_maxstar_prior:* the maximum star rating given in a contest *t* at the time of entry of a contestant *i*, (b) *c_avgstar_prior:* the average star rating given in a contest *t* at the time of entry of a contestant *i*, and (c) *c_ttlnumuserswithstars_prior:* the total number of users that were given stars in a contest *t* at the time of entry of a contestant *i*. In terms of textual feedback, I employed a Linguistics Inquiry and Word Count (LIWC) program to categorize the type of feedback given in a contest. I find that feedback is either given to a specific submission, for example, the feedback contains a callout to a submission number or a username, or is general and given to no particular submission (see Figure 3.11 for examples). I construct *contest_SpecificFdbk* to measure the total count of the specificity of the feedback given to others in the contest *t* at the time of contestant *i*'s entry, and *contest_GenericFdbk*, to measure the total count of generic feedback given in a contest *t* at the time of a contestant *i*'s entry.

To examine the impact of direct feedback given to a contestant *i*, I construct two metrics to quantify both types of feedback. The first metric c_*feedback_dum*, measures whether or not the contestant was given text feedback in contest *t*. The second metric c_*max_star*, measures the maximum star rating given to the contestant in contest *t*.

I further explore total submission behavior of a contestant. I construct two metrics to measure resubmissions of a contestant *i* in contest *t. c_ttl_resubs,* measures the total number of resubmissions of a contestant *i* in contest *t*, and *c_timeToResubmit* measures the lag between one resubmission and the next of a contestant *i* in contest *t*.

Lastly, I construct three different metrics to measure contestants *i*'s expertise and past experience: (a) *c_skill* –measures of past wins of contestant *i* divided by the total number of contest participations, (b) *c_experience*- measures the total number of contest participations of contestant *i* in the community, and (c) *c_expertise_dum*- a binary variable that indicates whether contestant *i* participates in more than logo contests within the community (for example, website designs, banner design, t-shirt design).

3.4.1.3 Controls

I include several additional controls in my analysis. In particular, I control for (a) Contestant-Related Factors: Total number of days in the community, total number of contests currently participating in (b) Contest-Related Factors: total number of entries, and amount of contest prize, contest design description length, and (c) Contest Holder - Related Factors: total number of matches held, total prize amount awarded, and average feedback. (See Table 3.1 for further description of control variables).

3.4.2 Empirical Model

The main objective of this study is to measure the effect of information spillovers on the contestant's timing of submission and the probability of her winning the contest. I examine different contest designs in particular open contests and blind contests. To test my propositions, I examine the probability of winning a contest based on the submission order, the feedback given in a contest whether to the contestant herself or to other contestants in a contest, along with other variables such as contestant expertise, and resubmissions in a contest. I also control for contestant variables, contest controls and contest holder controls as described in section 3.4.1. My data is in panel form, and the unit of analysis is the contestant (*i*) in contest (*t*), in particular, in its simplest form I estimate the following model:

$$Prob(Winning_{it} = 1 \mid x) = F(x, \beta)$$

where x entails the full set of explanatory and control variables and β are the coefficients of interest, in particular,

$$P(Winning_{it} = 1 \mid x) = F(Contestant\ Submission_{it}\ Order_{it}, Feedback\ to\ Others\ in\ Contest_{it}, Direct\ Feedback\ to\ Contestant_{it}, Contestant\ Expertise_{it}, Contestant\ Resubmissions_{it}, Contestant\ Controls_{it}, Contest\ Controls_{t}, Contest\ Holder\ Controls_{t}) + \varepsilon_{it} \qquad \textbf{Equation (3.1)}$$

Given the complexity and difficulty in controlling for all omitted variable bias in terms of contestants choice of participation in contests, I need to control for individual effects. A standard modeling choice when faced with bias caused by missing variables is fixed effects (Agarwal et al., 2009). More specifically, I observe the submission order and whether the contestant won at the individual contestant level, for 16,645 different contests. This panel data allows me to estimate a model that controls for omitted bias with individual fixed effects,

$$Win_{it} = X_{it}\beta + \gamma S_{it} + \alpha_{it} + \varepsilon_{it} \qquad \textbf{\textit{Equation (3.2)}}$$

where i indicates the contestant, t denotes the time in contests, Win_{it} is the observed winning dummy for contestant i in contest t such that variable Win_{it} is 1 if contestant i is the winner, while Win_{it} is 0 for all others in the contest, X_{it} is a vector of time varying explanatory variables that includes feedback and the above specified variables and controls in equation (3.1), S_{it} is the submission order for individual i in contest t, α_{it} is the unobserved individual effects, and ε_{it} is the disturbance.

The above equation (3.2) can be estimated using a simple panel data fixed effects model. However, one concern with this strategy is that the time of entry or submission order may be correlated with some unobservable contestant-specific characteristics that may influence outcome of winning the contest. If some explanatory variables are correlated with errors, then ordinary least-squares regression gives biased and inconsistent estimates. The Wald test of exogeneity (chi2 (1) = 5.48; $p<0.05$) for submission order points to the presence of endogeneity, implying that the parameter of interest γ will be estimated with a positive bias and underscoring the need for an instrumental variable (IV) approach. To control for this potential problem, I use a Two Stage Least-Squares (2SLS) regression with IV's. Under the 2SLS approach, in the first stage, each endogenous variable is regressed on all valid instruments, including the full set of exogenous variables in the main regression. Since the instruments are exogenous, these approximations of the endogenous covariates will not be correlated with the error term. So, intuitively they provide a way to analyze the relationship between the dependent variable and the endogenous covariates. In the second stage, each endogenous covariate is replaced with its approximation estimated in the first stage and the regression is estimated as usual. The slope estimator thus obtained is consistent (Wooldridge, 2001). Therefore, I need to instrument for S_{it} in a two-stage least squares model to obtain consistent estimates.

The intuition behind the use of IV's is that they are likely to be correlated with the relevant independent variables but uncorrelated with unobservable characteristics that may influence the dependent variable. Thus, valid instruments Z_{it} predict the submission order but are uncorrelated with the second stage error ε_{it}. Using Hausman and Taylor (1981) estimation method for finding an IV, I instrument for submission order with the average submission order in the other contests for user *i* at contest *t*. In theory, a user's past submission behavior would be a good predictor of her current submission behavior. Feedback given to a user in contest *t* can also be endogenous to the probability of winning in contest *t*. Therefore, I also instrument for Feedback of contestant *i* with the average star rating received for contestant *i* in other contests. Users' past receipt of feedback will be a good predictor of her current feedback. Thus the regression model extends to a two-step fixed effects with IV:

(1) $S_{it} = X_{it}\beta + Z_{it} + \alpha_{1it} + \varepsilon_{1it}$ ***First Step Control Equation (3.3)***

(2) $Win_{it} = X_{it}\beta + \gamma S'_{it} + \alpha_{2it} + \varepsilon_{2it}$ ***Second Step Primary Equation (3.4)***

where Z_{it} is my full set of instruments for submission order and feedback. As denoted earlier, X_{it} entails the rest of the explanatory and control variables (for example, contestant expertise, contestant resubmissions, contestant controls, contest controls and contest holder controls). The first stage estimation is the submission order and the second stage is the probability of winning the contest. Using this two-step fixed effects model, I estimate the probability of winning the contest for the various types of contests: (a) open contests with feedback (b) open contests with no feedback (c) blind contests with feedback and (d) blind contests with no feedback. I conduct a two stage (2SLS) Fixed Effect and find that for all the models, the first-stage F statistic is highly significant and much higher than the minimum value of 10, alleviating weak instrument concerns (Staiger & Stock, 1997). Variance inflation factors across all models range from 1.18 to 3.37, suggesting that the estimates obtained are not biased because of multicollinearity. The main drawback of the Fixed Effects estimator is that it prevents the use of any explanatory variables that are time invariant. I therefore also conduct a two stage Random Effects model to mitigate this drawback.

3.5 Results

Table 3.3 summarizes the main findings. In this section, I discuss the estimation results in Table 3.4 for each explanatory parameter of interest (namely, submission behavior, feedback to others, direct feedback to user, and expertise). The key research objective is to investigate the role of different information visibility regimes on the behavior of contestants as well as contest outcomes. Therefore, for each parameter of interest, I discuss the Fixed Effects coefficients for the different contest design regimes visibility (open with feedback, open with no feedback, blind with feedback, blind with no feedback).Note that both fixed effects and random effects are consistent in terms of sign and significance.

3.5.1 Submission Behavior

Columns 1 and 2 of Table 3.4 show the results for open contests with feedback using the fixed effects model and random effects model respectively. In this type of contest**,** both the design submissions and feedback is visible. I can make several inferences from the regression coefficients; in particular, I refer to the fixed effects model estimates in column 1. The coefficient on *c_subOrder* is negative and significant ($\beta_{(c\_subOrder)}$ = -0.1223) and the coefficient on *c_subOrder^2* is positive and significant ($\beta_{(c\_subOrder\wedge 2)}$ = 0.1667), implying that winners are more likely to submit either early or late. The coefficients on the interaction of the expertise variables and *c_subOrder* is positive, implying that users with higher skill, expertise or experience and who submit late have a higher probability of winning. These findings suggest that later contestants benefit from information spillovers in such a contest regime.

I next examine the behavior of contestants in open contests with no feedback where only the design submissions are visible. Columns 3 and 4 of Table 3.4 show the results for open contests with no feedback using the fixed effects model and random effects model respectively. Referring to column 3, the coefficients on *c_subOrder* and *c_subOrder^2* ($\beta_{(c\_subOrder)}$ = 0.1634; $\beta_{(c\_subOrder\wedge 2)}$ = 0.0076) are positive and significant, implying that the later the submission the higher the probability of winning. This finding shows that although there is no feedback from the contest holder, users are able to benefit from the information spillover of the design visibility of their competitors- supporting Proposition 1.

Columns 5 and 6 in Table 3.4 report the regression coefficients for blind contests with feedback, where design submissions are not visible, and only feedback is visible. Note that the signs of the coefficients of the submission order variables are in accordance with what one would expect for both fixed effects model (column 5) and random effects model (column 6). Exploring column 5, the coefficients on both *c_subOrder* and *c_suborder^2* ($\beta_{\text{(c\_subTime)}}$ = -0.2544; $\beta_{\text{(c\_subOrder^2)}}$ = -0.0302) are negative and significant, implying that earlier submission entries increase the probability of winning the contest. The coefficients on the interaction of the expertise variables and *c_subOrder* is negative, implying that users with higher skill, expertise or experience and who submit early have a higher probability of winning. Taken together, these findings suggest that when design submissions are not visible, users do not benefit from late submissions (no information spillover), and instead submit early and build upon their own feedback.

Lastly, columns 7 (fixed effects model) and 8 (random effects model) of Table 3.4 report the regression coefficients for blind contests with no feedback, where there is no visibility of design or feedback. Referring to column 7, the coefficients on both *c_subOrder* and *c_suborder^2* are insignificant, showing no evidence of strategic behavior related to timing of entry and probability of winning. Only the coefficients of *c_skill*, *c_expertise_dum*, and *c_experience* are positive and significant, indicating that when there is no information spillover, users are thus not able to behave strategically and only benefit from their own skill and expertise.

In conclusion, I find that in open contests with feedback (with design and feedback visibility), winners either submit early or late – a U-shaped behavior. In open contests with no feedback (with only design visibility), winners submit late. In contrast, I find that winners seem to submit early in blind contests with feedback (with only feedback visibility), and I find no strategic submission behavior in blind contests with no feedback (no information spillover). These findings put together all support proposition 1. I find that when there is information spillover, users tend to use such information and act strategically. When there is information spillover on the design visibility, users strategically wait to "imitate," build upon others' work and submit late, thereby increasing their likelihood of winning the contest. However, when there is no information spillover, users are limited to their own skills and abilities and to submitting early and building upon the contest holders' feedback directed to them.

3.5.2 Feedback to Others

I study the impact of the feedback to others on a users' submission behavior and probability of winning a contest. I examine the impact of feedback in two ways**.** First, I explore the visibility of feedback through stars and text, next I further analyze the type of textual feedback provided by the contest holder.

3.5.2.1 Visibility of Feedback

I examine the impact of feedback visibility given to others on one's submission behavior and the probability of winning in open and blind contests. In open contests where design submissions of others are visible, feedback can have information spillover effects, whereas in blind contests feedback has no information spillover effects since the design submissions of others are not visible, and, thus, feedback may not have much value.

Referring to the regression coefficients for open contests with feedback, in particular column 1, I can make several interpretations from the regression coefficients. The coefficients on *c_maxstar_prior*, *c_avgstar_prior*, and *c_ttlnumuserswithstars_prior* ($\beta_{c\_maxstar\_prior}$= -0.0007; $\beta_{c\_avgstar\_prior}$ = -0.0035; $\beta_{c\_ttlnumuserswithstars\_prior}$ = -0.0050) are negative and significant, implying that higher the feedback given to others at the time of a contestant's first submission, the lower the probability of a focal contestant winning the contest. The more interesting results, however, are related to the coefficients for the interaction of feedback given to others variables, and *c_subOrder*. The coefficients of these interactions are positive and significant ($\beta(_{c\_maxstar\_prior*subOrder)}$=0.1284; $\beta(_{c\_avgstar\_prior*subOrder)}$= 0.1174; $\beta(_{c\_ttlnumuserswithstars\_prior*subOrder)}$ = 0.1055), implying that the higher the feedback given to others and the later the submission of a focal contestant, the higher the probability of winning the contest- supporting Proposition 2a. This result indicates that the visibility of feedback to other contestants will benefit a contestant as it provides information about the contest holders' tastes. By leveraging this information and submitting late, a contestant can increase her probability of winning. In particular, when a contestant delays her submission time by one unit, and the maximum star rating of other contestants at the time of submission is increased by one unit, the contestant will increase her probability of winning the contest by 12.84%. Furthermore, by delaying submission of a design by one time unit, and the average star rating of others in a contest is increased by one unit, the contestant will increase her probability of winning the contest by 11.74%. Lastly, by deferring submission by one time unit, and the total number of users given star feedback is increased by one unit, the contestant will increase her probability of winning the contest by 10.55%.

Regression coefficients for blind contests with feedback in column 5 provide anticipated estimates. The regression coefficients for feedback to others are negative and significant ($\beta_{c\_maxstar\_prior}$= -0.0896; $\beta_{c\_avgstar\_prior}$ = -0.0507; $\beta_{c\_ttlnumuserswithstars\_prior}$ = -0.1126), implying that the higher the feedback given to others at time of a contestants first submission the lower the probability of winning the contest. Yet, in contrast to open contests with feedback, I observe that in blind contests, the coefficients for the interaction of feedback given to others variables and *c_subOrder* are negative and significant ($\beta_{(c\_maxstar\_prior*subOrder)}$ = -0.0054; $\beta_{(c\_avgstar\_prior*subOrder)}$= -0.0139; $\beta_{(c\_ttlnumuserswithstars\_prior*subOrder)}$ = -0.0729), implying that the higher the feedback given to others and the later the submission, the lower the probability of winning the contest. This finding provides further support for Proposition 2a. Since later contestants are not able to benefit from the information spillover of feedback given to other contestants, early submitters have a higher chance of winning than late submitters by building upon their own work.

3.5.2.2 Type of feedback

I further analyze how the different types of textual feedback impact outcomes. I compare the specificity of the text feedback in open contests and blind contests. In open contests, I find that the regression coefficient in column 1 for *contest_SpecificFdbk* is positive and significant ($\beta_{(contest\_SpecificFdbk)}$ = 0.2421), implying that the higher the specificity of the feedback to other contestants in a contest, the higher the likelihood of winning the contest. The coefficient on the interaction of submission order and the specificity of the feedback is positive and significant, ($\beta_{(c\_subOrder * contest\_SpecificFdbk)} = 0.2841$), indicating that the later the submission and the higher the specificity of the feedback, the higher the likelihood of winning the contest - supporting Proposition 2b. Interestingly, the regression coefficient for generic feedback is not significant predictor of the outcome of interest.

In blind contests, I observe that the regression coefficient in column 5 for *contest_SpecificFdbk* is negative and significant ($\beta_{(contest\_SpecificFdbk)} = -0.1052$), showing that the higher the specificity of the feedback given to others, the lower the probability of winning. In addition, the coefficient on the interaction of submission order and the specificity of the feedback is also negative and significant, ($\beta_{(c\_subOrder*contest\_SpecificFdbk)} = -0.1931$), denoting that the later the submission and the higher the specificity of the feedback, the lower the likelihood of winning. These findings show that specific feedback to other contestants is not very useful to a contestant when the designs are not visible. However, the coefficient for generic feedback is positive and significant ($\beta_{(contest\_GenericFdbk)} = 0.0344$), implying that the higher the generic feedback given to others, the higher the likelihood of winning for a focal contestant. The coefficient for the interaction of submission order and generic feedback is negative ($\beta_{(c\_subOrder*contest\_GenericFdbk)} = -0.2152$) suggesting that the earlier the submission and the higher the generic level of the feedback, the higher the likelihood of winning.

When design submissions of others are visible, the more specific the feedback, the greater the benefits from information spillovers relating to the feedback; thus, later submissions have a higher probability of winning. In particular, a unit increase in the specificity of the feedback to other contestants at the time of a users' first submission in an open contest will lead to a 24.21% increase in her probability of winning the contest. However, when design submissions of others are not visible, the specific feedback is less valuable to others, and I find no benefit of late submissions. In other words, early submissions have a higher likelihood of winning the contest. In such blind contests, I find that a one unit increase in generic feedback to other contestants at the time of a users' first submission will lead to a 3.44% increase in her probability of winning the contest. These findings highlight the interaction between the information spillovers relating to the design and the specificity of feedback and how they impact outcomes.

3.5.3 Direct Feedback to User

I next examine the impact of direct feedback to a contestant and resubmission in both open and blind contests with feedback.

In open contests with feedback, I find that the regression coefficients in column1 on direct feedback are both positive and significant ($\beta_{(c\_maxstars)} = 0.0751$; $\beta_{(feedback\_dum)} = 0.0624$). Indicating that the higher the star rating feedback of a contestant, the higher the probability of winning the contest; and all else being equal, contestants that get feedback are more likely to win the contest as compared to contestants that do not get feedback. The positive and significant coefficient on the total number of resubmission variable ($\beta_{(c\_ttl\_resubs)} = 0.0557$), shows that the higher the number of resubmissions, the higher is the likelihood of being a winner. Further examining the impact of direct feedback; the regression coefficient for the interaction of star feedback and time to resubmit is negative and significant

($\beta_{(c\_maxstars*timeToResubmit)}$ = -0.0976), denoting that the higher the star rating feedback a user gets and the earlier the resubmission, the higher is the likelihood of being a winner. In addition, the coefficient for the interaction of total resubmissions and feedback dummy is positive and significant ($\beta_{(c\_ttl\_resubs*feedback\_dum)}$ = 0.0835), indicating that users that get feedback and resubmit are more likely to win the contest as compared to users that do not get feedback and resubmit

Results for blind contests are also similar to open contests. The coefficients in column 5 on direct feedback are also both positive and significant ($\beta_{(c\_maxstars)}$ = 0.2011; $\beta_{(feedback\_dum)}$ = 0.1319), indicating that the higher the feedback a user gets, the higher is the likelihood of her winning. Similarly, the positive significant coefficient on total resubmissions ($\beta_{(c\_ttl\_resubs)}$ = 0.1043), implies that the higher the number of resubmissions, the higher is the likelihood of being a winner. The interaction coefficient for star feedback and time to resubmit is negative ($\beta_{(c\_maxstars*timeToResubmit)}$ = -0.0339), showing that the higher the star rating feedback a user gets and the earlier the resubmission, the higher is the likelihood of being a winner. Lastly, the interaction coefficient for the total resubmissions and feedback dummy is positive and significant ($\beta_{(c\_ttl\_resubs*feedback\_dum)}$ = 0.1721), denoting that users that get feedback and resubmit are more likely to win as compared to users that do not get feedback and resubmit

I also find that users resubmit when there is no feedback in the contest. In open contests with no feedback, the coefficient in column 3 on the total number of resubmissions is positive and significant ($\beta_{(c\_ttl\_resubs)}$ =0.0089), implying that higher number of resubmissions increase likelihood of winning. In blind contests with no feedback, the number of resubmissions is not a significant predictor winning the contest. Although there is no information spillover in terms of feedback in open contests with no feedback, users are still able to benefit from information spillovers relating to the design visibility. Therefore, contestants may resubmit a better design at a later stage and increase their likelihood of winning the contest. However, in blind contests information spillover relating to the design visibility is not available, and thus there is no benefit of resubmission.

3.5.4 Skill, Expertise, and Experience

In all types of contests**,** the regression coefficients for all expertise variables are positive and significant, implying that winners are more likely to have high skill (or high rate of past wins), experience and have expertise in more than one type of contest design. Interestingly, in open contests (with and without feedback) where designs are visible, the coefficients for the interaction of *c_subOrder* and expertise variables are positive and significant, showing that contestants that submit late and have high skill, design expertise, or experience, are more likely to be winners. However, in blind contests with feedback, the coefficients for the interaction of *c_subOrder* and expertise variables are negative, indicating that contestants that submit early and have high skill, design expertise or experience are more likely to be winners. However, when there is no feedback spillover in blind contests (blind with no feedback), I do not find any significant results for the interaction of *subOrder* and expertise variables on the probability of winning the contest. Taken together, when there is design and feedback spillover, the more skilled and experience contestants are better able to capitalize on the information spillover and submit late are more likely to win. However, in the case of blind contests with feedback, when there is no design visibility, high skilled contestants who submit early are more likely to win. These results show that it is the contestants with higher skills, experience, and expertise that are more likely to act strategically and win the contest.

3.5.5 Controls

Results for the control variables align very well with expectations and are consistent with the results. For all types of contests, the regression coefficients for total contest entries, contest prize amount, contest description length, contest holders average feedback and contest holder total prizes awarded are negative and significant. More entries, a higher prize, a more detailed description, a more attractive contest holder all imply competition, and therefore negatively impact the probability of winning a contest. As expected, the coefficient of a member's age is positive and significant, showing that contestants that have been in the online marketplace for longer have a higher probability of winning the contest.

3.5.6 Robustness Checks

To ensure the robustness of my findings, I conducted a series of additional tests. In this section, I present multiple robustness checks that address selection bias, qualitative surveys to contest holders, brief informal contestant interviews, and additional specifications to ensure the vigor of my findings.

3.5.6.1 Selection Bias

An issue of concern is whether there is selection bias in terms of contestants' choice in selecting an open versus a blind contest. I use a t-test to test the difference of means on contestants' skill, experience, expertise and membership age in terms of choice of open contest and blind contests. I find no significant difference in either skill p =0.203; experience p = 0.186; expertise p = 0.843; or member age p = 0.102, indicating no self-selection bias in choice of open versus blind contests relating to these variables.

In addition, I examine whether or not contest holders choose open versus blind contests in terms of whether they have observed any type of participation selection by designers. For example, whether or not contest holders have observed that more "effortful" or more "skilled" designers participate in blind contests as opposed to open contests. Given the information spillover and imitation findings, by blinding the contest, one would think that it would attract designers that would put more "effort," and, thereby, might require a higher reward. Therefore, I use a t-test to test the difference of means of contest prize amount for contest holders that held both open and blind contests. On average open contests have a lower prize amounts than blind contests ($open_{\mu}$ = \$537.6704, $open_{\sigma}$ = \$304.9263; $blind_{\mu}$ = \$615.1873, $blind_{\sigma}$ = \$344.3936), however I accept the null hypothesis that the means are the same and find no significant difference (p = 0.1405). This findings shows that contest holders do not provide different rewards for open versus blind contests.

3.5.6.2 Contest Holder Surveys and Informal Contestant Interviews

I conducted an online survey of contest holders and an informal brief online interview of a few designers to better understand their behaviors and choices in this marketplace.

To better understand why contest holders would choose to hold an open contest versus a closed contest, and why they provide feedback, I conducted a survey on the contest holders. Figure 3.12a shows the survey email sent, and Figure 3.12b displays the actual survey.

As for the sample I surveyed, I initially went through the contest holders in my sample and gathered any email information they had made public. I also went on the design community Facebook Page and Twitter account and collected contact information on users that had made comments as contest holders. In all, I collected contact information on 270 contest holders.

I contacted contest holders via email, Facebook private message or Twitter private message. I got responses from 32 contest holders, 64% of which had conducted both open and blind contests, and the rest had only held open contests. I find that contest holders sometimes have an idea of the design they want at the time, but at other times they do not.

In terms of understanding why contest holders chose one type of contest over another, I asked them about the advantage of holding one type of contest over another. As for open contests versus blind contests, most respondents (63%) indicated that there is more participation and visibility in open contests. One respondent indicated that it attracts designers from outside of the community. Interestingly some pointed out that designers learn from each other or "*Designers piggy back on each other when I give feedback, makes it simpler for me and we get to the design faster*." This is a thought provoking comment, that some contest holders not only have noticed that designers imitate via feedback, but also they actually prefer the imitation/learning that occurs because it is easier and faster for them. One respondent even indicated why he prefers the imitation: "*... a lot of copy cats, which is sometimes* *good when I have a pretty good idea of the logo design I want.*" This comment suggests that some contest holders prefer open contests when they have an idea in mind, such that designers build upon each other through the feedback.

As for the advantage of blind contests over open contests, most of the contest holders denoted that the privacy of the logo designs seems to be the main advantage (78%). Some of the contest holders specified that there is "*no design imitation and more variety of designs"* in blind contests as opposed to open contests. This comment shows that blinding the designs resulted in more varied submissions.

Remarkably, all respondents noticed that in open contests later submissions were sometimes similar to earlier submissions. However, this did not seem to happen in blind contests. This supports the finding that users that submit later benefit from the information spillovers.

All respondents provided feedback in their contests. While 33% of them wait for a certain number of submissions before providing feedback, and 19% wait a few days before providing feedback, 48% provide feedback as soon as they see a potentially good design. This indicates that feedback can be generalized to all (when they wait for a certain number of submissions) or can be more specific such as when they see a potentially good design. Fifty-six percent of respondents provide feedback to the top 2-3 submissions, 26% to the best submission only, and 18% to the top 4-5 submissions. This shows that while some contest holders focus on a particular best submission, a lot of contest holders prefer keeping their options open by providing feedback to more submissions. The main reason behind providing feedback is to suggest improvements to an existing design (63% of the time), and a few contest holders suggested the reason was to attract more submissions (26%). All respondents indicated that (a) they saw more submissions from other contestants upon providing feedback, (b) they got a better design from the contestants they provided feedback to and (c) feedback also benefits other contestants who submitted after the feedback. These findings strongly support our finding that users that submit later benefit from feedback given to others. Overall, the survey provided further support for my findings relating to information spillover and timing of submissions.

To better understand who the designers are and their submission strategy, I informally interviewed a few designers. I contacted seven designers that had a high number of wins via private message and was able to get response from four of them. Figure 3.13 displays the responses. Three out of the four designers were located outside of the US. Only one designer considered the design community to be that designer's sole source of income; the rest used the community as a hobby and as a way to augment their income.

In terms of choice of what contests to join, most of them seem to want a contest that is "interesting" to them. One designer preferred open contests; another preferred contests above $300. The clarity of the design brief was also important to one. A couple of designers also noted that an active contest holder that provides feedback attracted them.

Lastly, to get an understanding of their submission strategy, I asked them if they had a preference as to when to submit their first design. While two designers preferred early submission to get feedback, the other two preferred to wait and submit late. In fact, one said that it would be a "waste of time" to enter the contest during the first two days. In particular, both designers that preferred late submission wanted to wait for "participation from others" and wait for the contest holder to give a rating or feedback to get a better idea of what "he is after".

3.5.6.3 Additional Specifications

To further validate my findings, I estimated alternative model specifications. I ran a two stage random effects model as shown in Table 3.4 using the same fixed effects IV's as instruments. Most applications in economics have made the choice between the random effects and fixed effects estimators based on the standard Hausman test (1978). The null hypothesis of the Hausman is that the preferred model is random effects versus fixed effects. This test basically tests whether the disturbances are correlated with the regressors, and the null hypothesis is they are not. Rejecting the null hypothesis means that the fixed effects model is consistent and the random effects model is inconsistent. I therefore run the Hausman test and find that the null hypothesis is rejected (chi2 = 26.34; Prob>chi2 = 0.0005). Therefore I am more confident in my fixed effects models.

I also estimated a Probit and a Logit model for the probability of winning the contest. Furthermore, I estimated at a two stage Probit with IV using the same instrumental variables (average submission order and average star rating). All three models provided highly consistent results in terms of sign and significance.

3.6 Conclusions and Implications

The massive growth and consumption of Web 2.0 has transformed and redefined the roles of traditional ways of conducting business by proficiently bringing together individuals and firms. Firms are no longer limited to resources within their own corporation. Web 2.0 connects firms with new emerging online marketplaces that allow them to tap into the wisdom of crowds for resources and solutions. Recently, firms have been using such emerging crowdsourcing open innovation marketplaces at an escalated rate (for example, Mountain Dew, AOL, Starbucks, Dell). According to data from fourteen large crowdsourcing firms, their revenues grew 53% between 2009 and 2010, and 74% between 2010 and 2011(Silverman, 2012). These open innovation marketplaces have made visible, information that was traditionally not only costly, but also impossible to provide in contests prior Web 2.0. Such information relates to contestants' submissions and feedback from the contest holder - not only to the user but to others in the contest. While advances in Web 2.0 are creating new opportunities for firms to find solutions they also entail new challenges. Traditional managerial actions are no longer appropriate in these new emerging marketplaces. My study seeks to be one of the first attempts to understand the role of information spillover in different contest designs on the behavior of contestants as well as contest outcomes.

From a theoretical perspective, my study complements and extends the emerging literature on online crowdsourcing markets and in particular on the design of online contests. This study is among the first to empirically examine the role of different information visibility regimes on the behavior of contestants as well as contest outcomes in online settings. My study also contributes to the vast stream of research on the role of timing of entry and its implications for market outcomes. Most of the studies examining the timing of entry of market participants examine the strategic behavior of firms. This study is among the first studies to examine strategic entry behavior of individuals in a decentralized marketplace. My study also complements and extends the literature in IS that has studied the impact of IT-related information spillovers on firms and industries (for example, Brynjolffson & Hitt, 2000; Chang & Gurbaxani, 2012; Mun & Nadiri, 2002; Cheng & Nault, 2007). In particular, I empirically extend literature on timing of entry, and information spillover in the context of open innovation marketplaces at the individual level. In this study, I show that information spillover is a crucial factor in determining the optimal time of entry. First mover advantages have been known to benefit firms in terms of market share, profitability and long terms effects. However, when it comes to the open innovation contest marketplace, I find that information spillover in certain contest designs diminishes the well-known competitive strategy of first movers, and thus benefits late movers instead.

My findings not only give access to a richer theoretical understanding of the role of information visibility regimes, but also ways through which one might ultimately increase the effectiveness contest designs in open innovation marketplaces. The results of this study offer some interesting implications for firms in managing open innovation marketplaces.

This study is conducted on a leading online open innovation marketplace for designs, and present important new results on how different information visibility regimes impacts user behavior and contest outcomes. I compare “open” contests possessing greater degrees of information transparency with “blind” contests that have limited information transparency, and examine how these different information regimes influence contestants’ choices and outcomes.

I find that contest design, particularly relating to information visibility, significantly influences contestant's behavior as well as outcomes. Information spillover on design visibility has econometrically identifiable impacts on contestant's behavior and contest outcome. Technology providers in such crowdsourcing marketplace should be aware of such behavior under different information transparency regimes and design the market accordingly. I also find evidence of informational spillovers relating to feedback given to others. Technology providers of such open innovation marketplaces should be mindful of the potential impact of making feedback visible on user behavior. Technology is evolving at a fast pace, and thus it is crucial to take appropriate action in terms of designing the market. In particular, information spillover induces imitation, and imitation lowers the diversity of ideas, which is associated with lower innovation (Terwiesch and Ulrich, 2009). I recommend "blind"-ing contests for when the contest holder is looking for diversity of designs and innovative designs. Yet when the contest holder has a fair idea in mind, then an "open" contest would allow contestants to learn from each other and deliver the best solution to the contest holder. Another recommendation would be a multi-stage contest. The contest could start as "blind" to deliver the most diverse ideas. Once the contest holder likes a submission, she can then turn the contest to an "open" contest for designers to learn from each other and deliver the optimal solution. In order to maintain the attractiveness of the market, operators should provide such options to the contest holder to maintain the effectiveness of the marketplace.

Another novel finding is the benefits from informational spillovers differ depending on the type of feedback provided by the contest holder. In open contests with feedback, the more specific the feedback given to others in a contest, the higher the chances of late submissions winning the contest. However, in blind contests where users cannot see each other's submissions, specific feedback given to other contestants does not benefit a focal user, while generic feedback increases the likelihood of late submissions winning the contest. In order to encourage delivery of optimal submissions, market operators should therefore manage the content in the text feedback. Market operators should encourage contest holders to provide specific feedback in open contests, and generic feedback in blind contests to facilitate better communication between the contest holder and the contestants and improve the overall performance of a contest.

These results highlight the importance of the role of different information regimes and information spillover in open innovation contests on time to entry and ultimately the probability winning the contest. The findings of this study draw attention to the fact that with the increase in information visibility online, technology providers and contest designers should alter their ways to manage the visibility of information more proficiently. The findings of my study would guide platform developers to enable and constrain the visibility of information that operates in their ecosystem and would engineer the user experience to increase the efficiency of the marketplace.

3.7 Limitations and Future Research

My study has certain limitations. Although this study controlled for experience and expertise of the designers, I was not able to collect designer demographic information. Demographics such as gender, age, education, and profession may potentially impact outcomes. While the fixed effects models used in this study accounted for omitted variable bias, future studies can replicate this study and control for such demographic information.

Although it is abundantly clear that strategic behavior is present, the precise mechanism through which strategic behavior exerts itself in this specific context is less understood. I identified information spillover and its different types as a possibility, but there may be others, and my analysis does not allow me to distinguish between them. More qualitative data via interviews or surveys may shed further light on this issue. In addition, future studies can also survey designers to further investigate whether designers learn specific types of strategies with time and whether their strategies change overtime.

This study identified when winners are more likely to submit in different contest design regimes. However, I cannot suggest that a solver should take a late or early approach in particular, as it entails confounding effects on contestants such as the amount of time spent to deliver a solution. Yet, from a contest holder point of view, knowing when they will receive the best solutions is what they are most interested in. Future studies can explore the factors that will expedite the arrivals of best solutions.

My analysis of the textual comments is one of the first efforts to study the effect of written language feedback on strategic behavior of contestants. Though I identified the main feedback categories of text (specific versus generic), future studies could apply advanced text mining techniques to possibly identify more advanced text feedback classifications, and identify how a change in one of the classes can potentially change how contestants react in different contest visibility regimes. These are certainly interesting dynamic interactions that can be explored in future.

Lastly, since this study was limited to a single open innovation marketplace for a specific logo contest, additional corroboration of these novel findings by subsequent research that examine multiple crowdsourcing markets would be useful. This is especially so as the imitation costs of different types of contests and the amount of information visibility is likely to be contingent on the nature of the contest and marketplace accordingly.

Chapter 4: Conclusion

As the digitization of online markets continues to grow, it also presents novel opportunities and challenges in terms of designing marketplaces. In particular, the increased visibility of information enabled by Web 2.0 technologies has led to an explosion of new business models that seek to leverage this increased availability of information. One type of information that has become very salient is information about “others” in an individual’s network or market. This sudden increase in the amount and variety of information available to an individual about others in her network is likely to have significant impact on her behaviors and choices.

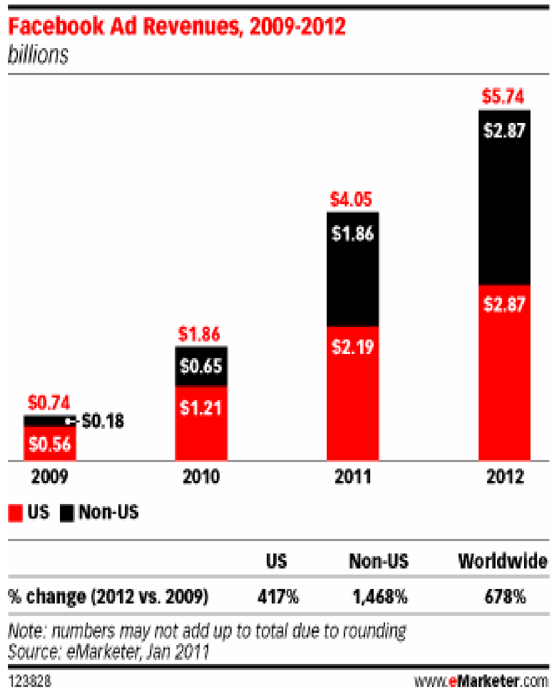
My dissertation examines the impact of increased information transparency, and in particular information on other users in the community, on an individual’s behavior in two different online emerging marketplaces. The first essay compares and contrasts the value of “social” information with “non-social” information and how they impact different outcomes in one of the largest online social media platforms--Facebook. The second essay examines the role of informational spillovers in open innovation design contests, in an online crowdsourcing market wherein firms seek design solutions from a crowd. A better understanding of the increase of visibility of information impact on consumer behavior in this emerging landscape has implications for academics, policymakers, and entrepreneurs who seek to leverage the power of online markets.

The findings of my dissertation add to the streams of literature that study social influence, information spillover and time-to-market. Essay 1 provides evidence of the value of social influence or “social information” on consumers’ decision making process, in particular, in alleviating the negative impacts of the advertisements, and in its essential role as a new mechanism for advertisements. The study also empirically shows that it is fruitful to examine how different types of "social information" interact with media interactivity, to impact marketing effectiveness at the consumer level. In contrast to the general notion that marketing methods are more effective at promoting product diffusion when the interactivity is active; I find that passive interactivity coupled with "social" marketing fosters virality more globally than active "social" methods. This essay takes a step towards understanding how a consumer can be used in the marketing campaign as a "co-creator" of value for the firm. Understanding optimal viral marketing design strategies enables firms to optimally create and manage social contagion. Essay 2 provides empirical evidence on the role of information spillover in determining the timing of entry at the consumer level. While first mover advantages have been known to benefit firms, the study finds that information spillover in certain contest designs diminishes the well-known competitive strategy of first movers, and benefits late movers instead. The study also highlights how transparency of information in different contest design regimes may lead to imitation, which may potentially limit the innovativeness and competitiveness of the marketplace. In order to maintain the attractiveness of the market for innovation seekers as well as designers, market operators must understand the implications of different contest designs and information regimes to enable the design of an optimal innovative marketplace.

In conclusion, the two essays of my dissertation provide a better understanding of how the increase in information visibility enables consumers to be “co-creators” of value to the firm, and how this role impacts user behavior and outcomes in emerging online marketplaces. The essays from this dissertation seek to inform firms and policy makers in designing better ways to interact with consumers through managing the transparency of information on “others”. Marketplaces that learn to utilize the power of information in an optimal way will be better positioned to succeed. Further developments in internet technologies and Web 2.0 will continue to change the salience of information visibility in emerging marketplaces, creating a rich area for future empirical research. While this dissertation takes an initial step, future research should examine how consumers respond to changes in information visibility of online communities. Lastly, my dissertation contributes to the growing IS literature of empirical studies in online markets, especially those related to online social networks and crowdsourcing marketplaces.

Appendices for Essay 1

Figure 2.1 Facebook Ad Revenues



Notes: Print screen adapted from:
http://vator.tv/news/2011-01-18-facebook-ad-revenue to-top-4-billion-in-2011

Figure 2.2 Social Ad

Notes: This is an adapted print screen of Facebook's social ad.

Figure 2.3 Advertising Information Sources

	Active	**Passive**
Non-Social	Email	Traditional Ad
Social	Friend Invite	• Newsfeed • Social Ad

Figure 2.4a Comparisons of Sample and Population Demographic Characteristics: Geography

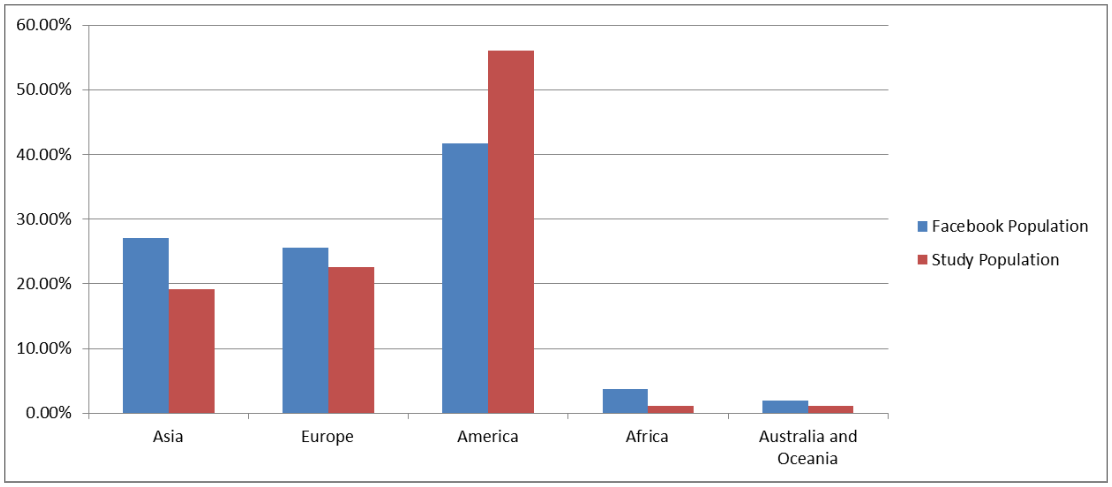


Figure 2.4b Comparisons of Sample and Population Demographic Characteristics: Gender

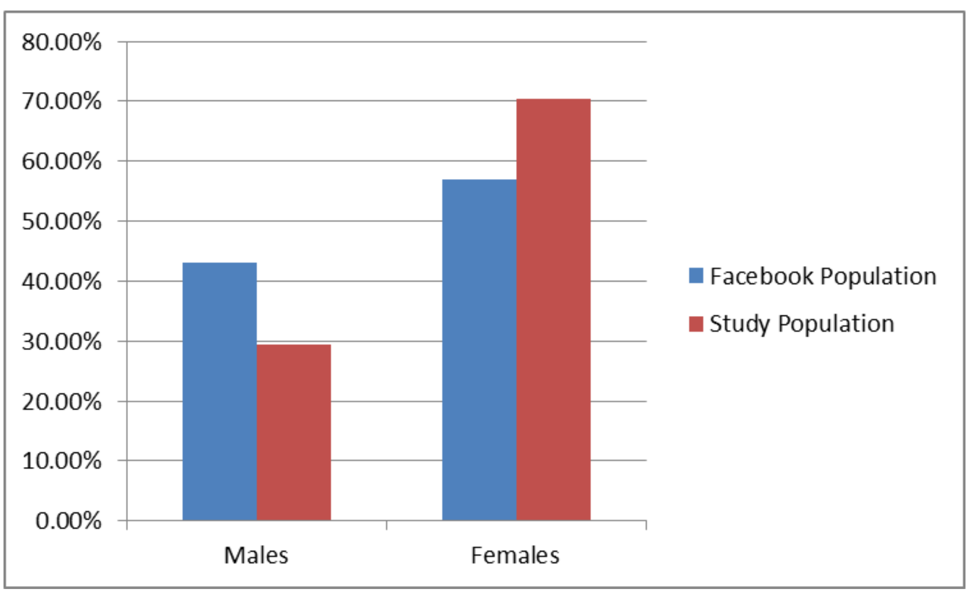


Table 2.1 Adoption Source Frequency Distribution

Adoption Source	Frequency	Percent
Traditional Ad	86,057	30.97
Social Ad	13,211	4.75
Email	12,222	4.40
Invited by Friend	21,108	7.60
Newsfeed	81,704	29.40
Other	63,561	22.87

Table 2.2 Overall Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
App Activity				
ln_passenger_count	0.3803	0.6016	0	5.6904
ln_pilot_count	0.4792	0.5263	0	3.2958
ln_stories	0.1687	0.3740	0	5.2781
ln_photos	0.0295	0.2147	0	5.3375
App Diffusion				
Diameter	0.5179	0.9635	0	9
Outdegree	0.2128	0.6962	0	24
Size_of_CC	1.7920	2.1205	1	35
Network App Controls				
friend_count	1.4014	3.7665	0	227
thresholddaily_percOfFriendsLike	0.0159	0.3356	0	39
App Adoption Controls				
day_lvl_adop	20.5326	11.2807	1	44
Facebook Privacy Controls				
ln_ttl_privacy	1.5385	0.1550	0.6931	1.6094
Facebook Activity Controls				
likes_count	15.3578	16.1432	0	1438
num_of_friends_onFB	299.9410	309.2102	1	4985
Facebook Likes/Interests Categories Count				
fbc_interests	0.6187	1.4241	0	9
fbc_activities	0.6154	0.8653	0	4
fbc_television	1.6549	2.0781	0	10
fbc_sports	0.0563	0.2938	0	3
fbc_movies	1.1886	1.7609	0	10
fbc_other	7.7900	7.8954	0	1413
fbc_music	1.8198	2.1521	0	10
fbc_favorite_teams	0.2763	0.8895	0	9
fbc_favorite_athletes	0.2879	0.9575	0	7
fbc_games	0.1692	0.7710	0	5
fbc_books	0.7340	1.4821	0	10

Table 2.3 Top 10 CC Sub-Sample Adoption Distribution

Distribution of users that liked the App	Count	Percentage
User & 1st degree Friend only	680	58%
User & 2nd degree Friend only	130	11%
User & 1st & 2nd degree Friend	364	31%
Total Likes	1174	100%

Table 2.4 Sample Construction Summary

Sample	Total users that "Like"-d App	Total users that did not "Like" App	Additions to sample from the Outcome Sample-users that "Like"-d App (total = 173,050)	Sample Size (N)
FanPage	1294	3914	3000	8202
Top10CC	1174	102606	173,050 (All)	275082
Intersection	----	387	400	787

Table 2.5 Variables and Description

	Category	Variable	Description
Outcome Variables	***App Adoption***	user_Adopted	Dummy variable – whether or not user adopted
	App Use	ln_passengered_count	Log of the number of instances of the App where user is a passenger.
		ln_photos_uploaded	Log of the number of photos uploaded to App by user. User Generated Content (UGC).
		ln_stories_uploaded	Log of the number of stories uploaded to App by user. User Generated Content (UGC).
	App Diffusion	ln_diameter	Log of the Diameter of the connected component of the user. A measure of beyond local diffusion or diffusion reach.
		ln_outdegree	Log of the Number of invitations sent and accepted. A measure of local diffusion.
Independent Variables	***Social Exposure***	FirstDegree	Total First Degree Friends /Total number of Facebook Friends that had adopted the App at time of user adoption decision
		SecondDegree	Total Second Degree Friends /Total number of Facebook Friends that had adopted the App at time of user adoption decision
	Adoption Source	source_email	The source the user came to like the App from: Email Source
		source_ad	Traditional ad
		source_socialad	Social ad
		source_invited	Facebook friend invitation
		source_nf	Facebook newsfeed
			Other – came directly to the App, not through online advertising. This is the baseline and not included in the analysis.
Controls	***Facebook Privacy Level***	ln_ttl_privacy	Whether the user allows such information to be visible (by

Control		Facebook default policy all of the below should be allowed). Log of the Total count of : ● Ability to “Send a Message” ● Ability to “Add user as Friend”
Friend App Activity Control	frnd_avg_passenger	Average number of Passengers for users’ friends.
	frnd_avg_pilot	Average number of Pilots for users’ friends.
	frnd_avg_photos	Average number of Photos for users’ friends.
	frnd_avg_stories	Average number of Stories for users’ friends.
Network App Controls	friend_count	Number of friends who have liked the App at time of adoption for the user.
	Percoffrndadopters	Percentage of Facebook Friends that liked the App.
	thresholddaily_percOfFriendsLike	(Day of Adoption – Max Friend Adoption Day)* (number of friends that liked the App) / (Total Number of facebook friends)
Facebook Activity	likes_count	Total Number of Likes and Activities on Facebook.
	num_of_FBFriends	Total number of Facebook Friends
	fbc_	Count of likes on Facebook on the following Categories : Interests, Activities, Television, Sports, Movies, Music, Favorite Teams, Favorite Athletes, Games, Books, Other
Adoption Rate Controls	day_lvl_adop	Day when the user “Liked” the App, according to the lifetime of the App.
Demographic Controls	Gender	Facebook listed Gender
	location_	Facebook listed geographic location

Table 2.6a App Adoption Regression Models

Variable	(1) Expedia Fan Page Heckman First Stage	(2) Expedia Fan Page Hazard Ratio	(3) Top 10 CC Heckman First Stage	(4) Top 10 CC Hazard Ratio	(5) Intersection Heckman First Stage	(6) Intersection Hazard Ratio
FirstDeg_FrndAdop	0.5283***	1.1581***	0.5032***	1.2374***	0.4731***	1.1921***
	(0.0934)	(0.1461)	(0.0562)	(0.1112)	(0.1301)	(0.2102)
SecondDeg_FrndAdop	0.1003***	1.1001***	0.0617***	1.0847***	0.1047***	1.0619***
	(0.0173)	(0.1065)	(0.0032)	(0.9383)	(0.0254)	(0.1106)
FirstDeg*SecondDeg	0.5301***	1.1613***	0.5918***	1.1389***	0.5027***	1.1898***
	(0.1145)	(0.2012)	(0.0953)	(0.4172)	(0.1343)	(0.2312)
Female	0.3598***	1.2566***	0.3806***	1.1981*	0.2307***	1.1692***
	(0.0371)	(0.0431)	(0.0281)	(0.0968)	(0.0567)	(0.0627)
numofFBFriends	0.0028	1.0040	0.0053	1.0067	0.0014	1.0026
	(0.1039)	(0.0180)	(0.1748)	(0.0991)	(0.1276)	(0.0176)
location_GB	0.3754	1.2031	0.3330	1.2507	0.4538	1.2712
	(0.1923)	(0.2312)	(0.7213)	(0.1717)	(0.2109)	(0.6658)
location_US	0.2728	1.2071	0.4523	1.2993	0.1029	1.2018
	(0.1803)	(0.1692)	(0.6377)	(0.1577)	(0.2130)	(0.3109)
ttl_privacy	1.0226	1.0971	-0.5171	0.9677	-0.3021	0.8291
	(0.7515)	(0.1738)	(0.4170)	(0.1610)	(0.4321)	(0.2610)
likes_count	0.0005***	1.0153***	0.0009***	1.0773***	0.0007***	1.0020***
	(0.0001)	(0.0135)	(0.0000)	(0.0178)	(0.0002)	(0.0125)
fbc_interests	0.0077	1.0059	-0.1635	0.8541	-0.0054	0.9102
	(0.0800)	(0.0634)	(0.3169)	(0.5288)	(0.0042)	(0.0167)
fbc_activities	-0.1072	0.9158	-0.9324	0.6604	-0.1025	0.8994
	(0.1109)	(0.0760)	(0.6543)	(0.1647)	(0.1175)	(0.0925)
fbc_television	0.1399	1.0205	-0.2824	0.9204	0.1053	1.0179
	(0.0901)	(0.0673)	(0.1961)	(0.0766)	(0.0844)	(0.0572)

fbc_movies	-0.0357	0.9984	0.9247	1.0595	-0.0401	0.9711
	(0.1005)	(0.0659)	(0.6963)	(0.0463)	(0.1023)	(0.0463)
fbc_other	-0.0183	0.9899	-0.2168	0.2176	-0.2299	0.4291
	(0.0151)	(0.0523)	(0.1745)	(0.1172)	(0.1039)	(0.0739)
fbc_music	-0.1868	0.9464	0.8218	1.0160	-0.2031	0.8721
	(0.1652)	(0.0612)	(0.5243)	(0.0389)	(0.3145)	(0.0892)
fbc_favorites	0.0273	1.0297	-0.7908	0.8417	0.0207	1.0077
	(0.1275)	(0.0848)	(0.8190)	(0.1471)	(0.1069)	(0.0947)
fbc_sports	-0.4300	0.9486	-0.4722	0.8393	-0.5572	0.7638
	(0.2846)	(0.1499)	(0.3293)	(0.1155)	(0.4419)	(0.2311)
fbc_games	-0.3239	0.9834	-0.7349	0.8585	-0.6213	0.8183
	(0.1952)	(0.0782)	(0.1004)	(0.1028)	(0.1736)	(0.1736)
fbc_books	-0.0995	0.9700	-0.9612	0.8876	-0.0319	0.9321
	(0.1110)	(0.0626)	(0.8508)	(0.4306)	(0.4109)	(0.2377)
MediaExposure	0.0108***	1.0037***	0.0130***	1.0510***	0.0132***	1.0238***
	(0.0029)	(0.0060)	(0.0018)	(0.0470)	(0.0033)	(0.0232)
_cons	-0.1221		-1.1704		-0.2837	
	(0.2343)		(0.6088)		(0.5091)	
N	8208	8208	275082	275082	787	
Censored	3914		102032		387	
Uncensored	4294		173050		400	
Ll		-4685.0873		-4386.9960		-4619.4312

Notes: This table reports parameter estimates and standard errors of the first stage of the Heckman model and the Hazard model as specified in Section 2.4 for App adoption for all three samples (1) FanPage ,(2)Top10CC, and (3) Intersection. A Heckman first stage is a Probit model predicting the probability of App adoption. The table also reports hazards ratio estimates of a Cox proportional hazards model of the time to App adoption. A hazards ratio greater than 1.0 for variable *x* indicates that it increases the probability of App adoption, while a ratio less than 1.0 indicates that it decreases the probability of App adoption.
*p< 0.10, **p< 0.05, ***p< 0.005. Standard errors shown in parentheses.

Table 2.6b App Use Regression Models

Variable	(1) FanPage Log (Passenger)	(2) FanPage Log (Stories)	(3) FanPage Log (Photos)	(4) Top 10 CC Log (Passenger)	(5) Top 10 CC Log (Stories)	(6) Top 10 CC Log (Photos)	(7) Intersection Log (Passenger)	(8) Intersection Log (Stories)	(9) Intersection Log (Photos)
source_ad	-0.0063***	-0.1791***	-0.1534***	-0.0124***	-0.2164***	-0.1623***	-0.0109***	-0.1607***	-0.1363***
	(0.0012)	(0.0209)	(0.0295)	(0.0005)	(0.0234)	(0.0317)	(0.0028)	(0.0324)	(0.0377)
source_email	-0.0816***	-0.0268*	-0.0051***	-0.0843***	-0.0290***	-0.0086***	-0.1251***	-0.0313**	-0.0081***
	(0.0179)	(0.0142)	(0.0012)	(0.0162)	(0.0014)	(0.0006)	(0.0284)	(0.0151)	(0.0023)
source_invited	0.3530***	0.2559***	0.1214***	0.3300***	0.2549***	0.2301***	0.2449***	0.2561***	0.2079***
	(0.0477)	(0.0274)	(0.0419)	(0.0332)	(0.0491)	(0.0044)	(0.0603)	(0.0712)	(0.0427)
source_nf	0.5492***	0.1761***	0.0759***	0.3628***	0.1396***	0.1760***	0.4224***	0.1228**	0.1375**
	(0.0593)	(0.0457)	(0.0194)	(0.0120)	(0.0421)	(0.0310)	(0.1242)	(0.0467)	(0.0618)
source_socialad	0.4636***	0.1390***	0.0680***	0.3357***	0.0252***	0.0683***	0.3196***	0.1015***	0.0729***
	(0.0357)	(0.0122)	(0.0101)	(0.0511)	(0.0083)	(0.0211)	(0.0719)	(0.0257)	(0.0161)
ln_ttl_privacy	0.2092***	-0.0616***	-0.0569***	0.3264***	-0.1408***	-0.0277***	0.1866***	-0.0851***	-0.0114***
	(0.0175)	(0.0168)	(0.0115)	(0.0587)	(0.0417)	(0.0068)	(0.0485)	(0.0176)	(0.0034)
Female	-0.0370***	0.1505***	0.0650	-0.0512***	0.1563***	0.0448***	-0.0501***	0.0666***	0.0628***
	(0.0058)	(0.0135)	(0.0649)	(0.0091)	(0.0317)	(0.0108)	(0.0068)	(0.0146)	(0.0134)
location_GB	0.1897	0.1821	0.0004	0.4231**	0.2967***	0.1307***	0.1045**	0.0492	0.0150
	(0.1634)	(0.1824)	(0.1783)	(0.1519)	(0.0457)	(0.0425)	(0.0503)	(0.0467)	(0.1206)
location_US	0.1014***	0.1713***	0.1351***	0.1493***	0.1429***	0.1436***	0.1604***	0.1327***	0.1324***
	(0.0259)	(0.0252)	(0.0332)	(0.0132)	(0.0185)	(0.0122)	(0.0272)	(0.0333)	(0.0379)
FirstDeg_FrndAdop	0.4192***	0.3393***	0.0468***	0.4783***	0.3279***	0.0384***	0.3901***	0.3137***	0.0417***
	(0.0861)	(0.0926)	(0.0107)	(0.0348)	(0.0467)	(0.0032)	(0.0928)	(0.0937)	(0.0121)
SecDeg_FrndAdop	0.0072***	0.0021***	0.0023***	0.0089***	0.0057***	0.0029***	0.0088***	0.0024***	0.0022***

	(0.0015)	(0.0004)	(0.0006)	(0.0011)	(0.0012)	(0.0009)	(0.0016)	(0.0005)	(0.0006)
First*Second_FrndAdop	0.2809***	0.2523***	0.2450***	0.1604***	0.1369***	0.1660***	0.2310***	0.2181***	0.1807***
	(0.0314)	(0.0401)	(0.0496)	(0.0431)	(0.0476)	(0.0203)	(0.0445)	(0.0507)	(0.0533)
num_ofFB_friends	0.0000***	-0.0001****	0.0001***	-0.0008***	-0.0000***	-0.0001***	-0.0001***	-0.0001**	-0.0001*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0001)
threshold_percFriendAdop	-0.2927***	-0.0776***	-0.1297***	-0.2448***	-0.2168***	-0.1250***	-0.2153***	-0.1024***	-0.1229***
	(0.0322)	(0.0228)	(0.0311)	(0.0163)	(0.0532)	(0.0315)	(0.0589)	(0.0258)	(0.0328)
likes_count	0.0052***	0.0033***	0.0052***	0.0003***	0.0001***	0.0001***	0.0052***	0.0028***	0.0041***
	(0.0015)	(0.0008)	(0.0007)	(0.0000)	(0.0000)	(0.0000)	(0.0016)	(0.0008)	(0.0012)
fbc_interests	0.0137***	0.0132***	0.0078***	0.0145***	0.0300***	0.0074***	0.0154***	0.0130***	0.0075***
	(0.0029)	(0.0034)	(0.0014)	(0.0013)	(0.0086)	(0.0003)	(0.0042)	(0.0045)	(0.0015)
fbc_activities	0.0108**	0.0255***	0.0236***	0.0867***	0.2013***	0.0202***	0.0250***	0.0346***	0.0268***
	(0.0049)	(0.0057)	(0.0031)	(0.0064)	(0.0612)	(0.0006)	(0.0075)	(0.0100)	(0.0072)
fbc_tv	0.0491	-0.0150	-0.0318	0.1924***	0.0670***	0.0031***	0.0053	0.0163**	0.0385*
	(0.0461)	(0.0492)	(0.0479)	(0.0432)	(0.0206)	(0.0008)	(0.0044)	(0.0068)	(0.0202)
fbc_sports	-0.0781	-0.0943	-0.0284	-0.1961	-0.1474***	0.006	-0.0156	-0.0087	-0.0496
	(0.0929)	(0.0987)	(0.0962)	(0.1410)	(0.0268)	(0.0130)	(0.0107)	(0.0219)	(0.0567)
fbc_movies	-0.0049	-0.0163	-0.0138	-0.1847***	-0.0536***	-0.0003	-0.0594**	-0.0086***	-0.0022
	(0.0436)	(0.0464)	(0.0452)	(0.0640)	(0.0108)	(0.0004)	(0.0215)	(0.0018)	(0.0032)
fbc_other	0.0283***	-0.0233***	-0.0209***	0.0081***	-0.0075***	-0.0022***	0.0052**	-0.0075**	-0.0051***
	(0.0076)	(0.0048)	(0.0039)	(0.0012)	(0.0022)	(0.0001)	(0.0019)	(0.0027)	(0.0012)
fbc_music	0.0286***	0.0089***	0.0177***	0.0244***	0.0088***	0.0164***	0.0190***	0.0080**	0.0156***
	(0.0033)	(0.0021)	(0.0024)	(0.0030)	(0.0026)	(0.0053)	(0.0043)	(0.0029)	(0.0028)
fbc_fav_teams	0.0258***	0.0231**	0.0213***	0.0278***	0.0296***	0.0188***	0.0207**	0.0260**	0.0191***
	(0.0068)	(0.0098)	(0.0039)	(0.0059)	(0.0061)	(0.0045)	(0.0089)	(0.0093)	(0.0042)
fbc_fav_athletes	0.0378***	0.0250***	0.0855***	0.0343***	0.0257***	0.0815***	0.0383***	0.0117**	0.0833**

	(0.0093)	(0.0044)	(0.0259)	(0.0083)	(0.0084)	(0.0207)	(0.0073)	(0.0045)	(0.0299)
fbc_games	0.0112**	-0.0145***	-0.0527**	0.0145**	-0.0181***	-0.0522***	0.0185**	-0.0104*	-0.0589**
	(0.0042)	(0.0041)	(0.0221)	(0.0053)	(0.0037)	(0.0154)	(0.0070)	(0.0057)	(0.0239)
fbc_books	-0.0212	0.0250	-0.0556	0.0672***	-0.0193***	-0.0268***	-0.0052	0.0043	-0.0212
	(0.0444)	(0.0472)	(0.0459)	(0.0130)	(0.0056)	(0.0073)	(0.0047)	(0.0059)	(0.0179)
day_lvl_adop	-0.0065***	-0.0101***	-0.0084***	-0.0078***	-0.0114***	-0.0071***	-0.0067***	-0.0117***	-0.0072***
	(0.0004)	(0.0005)	(0.0016)	(0.0004)	(0.0001)	(0.0000)	(0.0007)	(0.0022)	(0.0021)
frnd_avg_pilot	0.0124*	-0.1757	-0.0179	0.0159***	0.0080*	-0.0038	0.0191**	-0.0246	-0.0014
	(0.0065)	(0.1541)	(0.1504)	(0.0051)	(0.0041)	(0.0023)	(0.0076)	(0.0332)	(0.0219)
frnd_avg_passenger	0.0025***	0.1395	0.2627	0.0020***	0.0006	-0.0003	0.0024**	0.0055	0.0369
	(0.0005)	(0.1636)	(0.1596)	(0.0007)	(0.0008)	(0.0005)	(0.0009)	(0.0401)	(0.1210)
frnd_avg_photo	0.0023	0.0017	0.0093*	0.0017	0.0013	0.0139*	0.0064	0.0151	0.0125*
	(0.0358)	(0.0372)	(0.0055)	(0.0098)	(0.0037)	(0.0075)	(0.0189)	(0.0168)	(0.0072)
frnd_avg_stories	0.0421	0.0201*	0.0002	-0.0094	0.0173*	-0.0006	0.0155	0.0217*	-0.0003
	(0.0301)	(0.0105)	(0.0024)	(0.0106)	(0.0098)	(0.0040)	(0.0250)	(0.0111)	(0.0039)
Inverse Mills Ratio	0.0469***	0.0731***	0.0273***	0.0167***	0.0314***	0.0461***	0.0313***	0.0954***	0.0332***
	(0.0085)	(0.0102)	(0.0079)	(0.0052)	(0.0054)	(0.0012)	(0.0109)	(0.0147)	(0.0097)
_cons	0.2320***	0.0704***	0.4757***	0.3411**	0.3074***	0.0338**	0.2873***	0.1891***	0.3800***
	(0.0360)	(0.0216)	(0.0338)	(0.1329)	(0.0254)	(0.0123)	(0.0590)	(0.0563)	(0.0386)
N	8208	8208	8208	275082	275082	275082	787	787	787

Notes: This table reports parameter estimates and standard errors of the second stage of the Heckman model as specified in Section 2.4 for App use outcomes (1) Log(Passenger), (2) Log(Stories), and (3) Log(Photos) for all three samples (1) FanPage , (2)Top10CC, and (3) Intersection.
*p< 0.10, **p< 0.05, ***p< 0.005. Standard errors shown in parentheses.

Table 2.6c App Diffusion Regression Models

Variable	(1) FanPage Log (Diameter)	(2) FanPage Log (Outdegree)	(3) Top 10 CC Log (Diameter)	(4) Top 10 CC Log (Outdegree	(5) Intersection Log (Diameter)	(6) Intersection Log (Outdegree
ln_total_activity	0.1698***	0.2398***	0.2275***	0.1795***	0.1832***	0.2282***
	(0.0180)	(0.0219)	(0.0226)	(0.0206)	(0.0199)	(0.0230)
source_ad	-0.0817	-0.0313	-0.0882	-0.0255	-0.0831	-0.0382
	(0.0580)	(0.0418)	(0.0593)	(0.2156)	(0.0595)	(0.0203)
source_email	0.0662	0.0586	-0.0639	0.0512	-0.0538	0.0531
	(0.0826)	(0.0906)	(0.1578)	(0.0466)	(0.0321)	(0.0336)
source_invited	0.1565***	0.4959***	0.0903***	0.5086***	0.1725***	0.4673***
	(0.0134)	(0.0891)	(0.0061)	(0.0460)	(0.0142)	(0.0912)
source_nf	0.3138***	0.1277***	0.3596***	0.3603***	0.3339***	0.1588***
	(0.0370)	(0.0294)	(0.0257)	(0.0736)	(0.0516)	(0.0366)
source_socialad	0.2104***	0.0598***	0.2299***	0.0640***	0.2369***	0.0496***
	(0.0558)	(0.0146)	(0.0399)	(0.0034)	(0.0595)	(0.0151)
ln_ttl_privacy	-0.0505***	-0.0942***	-0.0474***	-0.0828***	-0.0488***	-0.0808***
	(0.0113)	(0.0203)	(0.0041)	(0.0124)	(0.0123)	(0.0227)
Female	0.0289***	0.0525***	0.0348***	0.0587***	0.0365***	0.0551***
	(0.0101)	(0.0111)	(0.0047)	(0.0128)	(0.0121)	(0.0157)
location_GB	0.0849***	0.1644***	0.0748***	0.1877 ***	0.0796*	0.1415*
	(0.0131)	(0.0563)	(0.0109)	(0.0415)	(0.0411)	(0.0712)
location_US	0.0887***	0.0088***	0.0932***	0.0072***	0.0880***	0.0073**
	(0.0169)	(0.0027)	(0.0147)	(0.0008)	(0.0295)	(0.0028)
FirstDeg_FrndAdop	0.1843***	0.2041***	0.1751***	0.2587***	0.1542***	0.1965***
	(0.0156)	(0.0374)	(0.0147)	(0.0301)	(0.0287)	(0.0523)
SecondDeg_FrndAdop	0.0057***	0.0166***	0.0105***	0.0069***	0.0066***	0.0115***
	(0.0012)	(0.0027)	(0.0021)	(0.0021)	(0.0018)	(0.0032)

FirstDeg*SecondDeg	0.1091***	0.1482***	0.1439***	0.1827***	0.0937***	0.1203***
	(0.0244)	(0.0317)	(0.0189)	(0.0312)	(0.0246)	(0.0326)
num_ofFB_friends	0.0001***	0.0000***	0.0002***	0.0002***	0.0000***	0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
threshold_percFriendAdop	0.0870	0.0883	0.1898	0.0767	0.1215	0.1025
	(0.1242)	(0.1223)	(0.1187)	(0.1054)	(0.2342)	(0.2217)
likes_count	-0.0098	0.0094	0.0002	0.0001	0.0022	0.0090
	(0.0154)	(0.0172)	(0.0002)	(0.0001)	(0.0019)	(0.0058)
fbc_interests	0.0329***	0.0079***	0.0214***	0.0063***	0.0281***	0.0073***
	(0.0059)	(0.0009)	(0.0015)	(0.0006)	(0.0070)	(0.0022)
fbc_activities	0.0186	0.0341	0.0354	-0.0191	0.0253	0.0538
	(0.0329)	(0.0366)	(0.0696)	(0.0536)	(0.0183)	(0.0423)
fbc_tv	0.0065	-0.0130	0.0318	0.0317	0.0116	-0.0128
	(0.0229)	(0.0255)	(0.0435)	(0.0335)	(0.0078)	(0.0094)
fbc_sports	-0.0278	-0.0591	-0.2298***	0.2143***	-0.0477	-0.0344
	(0.0563)	(0.0627)	(0.0622)	(0.0434)	(0.1182)	(0.0272)
fbc_movies	0.0289	-0.0121	-0.0804*	0.0264	-0.0094	-0.0024
	(0.0260)	(0.0289)	(0.0420)	(0.0294)	(0.0083)	(0.0081)
fbc_other	0.0052	-0.0256	-0.0016	-0.0090	-0.0025	-0.0086
	(0.0168)	(0.0187)	(0.0092)	(0.0071)	(0.0021)	(0.0139)
fbc_music	0.0038***	0.0242***	0.0053***	0.0270***	0.0045***	0.0205**
	(0.0011)	(0.0068)	(0.0010)	(0.0060)	(0.0012)	(0.0081)
fbc_fav_teams	0.0105***	0.0042***	0.0160***	0.0058***	0.0117***	0.0047**
	(0.0014)	(0.0011)	(0.0029)	(0.0018)	(0.0029)	(0.0017)
fbc_fav_athletes	0.0097***	0.0245***	0.0170***	0.0269***	0.0165*	0.0124***
	(0.0021)	(0.0032)	(0.0038)	(0.0083)	(0.0097)	(0.0035)
fbc_games	0.0374	-0.0234	0.0932	0.0101	0.0009	-0.0172
	(0.0273)	(0.0304)	(0.1048)	(0.0803)	(0.0074)	(0.0113)
fbc_books	-0.0078***	-0.0103***	-0.0083***	-0.0104***	-0.0071***	-0.0093**

	(0.0014)	(0.0009)	(0.0013)	(0.0009)	(0.0024)	(0.0035)
day_lvl_adop	-0.0039***	-0.0015***	-0.0027***	-0.0014***	-0.0011*	-0.0027***
	(0.0004)	(0.0005)	(0.0002)	(0.0001)	(0.0006)	(0.0007)
frnd_avg_pilot	-0.1124	-0.0116	-0.0066	-0.0092	-0.0102	-0.0076
	(0.0953)	(0.1055)	(0.0081)	(0.0058)	(0.0758)	(0.0158)
frnd_avg_passenger	-0.0804	0.0226	-0.0036	0.0015	-0.0412	-0.0221
	(0.1012)	(0.1120)	(0.0083)	(0.0055)	(0.0634)	(0.0672)
frnd_avg_photo	0.0234	0.0384	0.0145	0.0021	0.0163	0.0265
	(0.1431)	(0.1203)	(0.0157)	(0.0053)	(0.0244)	(0.0312)
frnd_avg_stories	0.1023	0.0345	-0.0186	-0.0035	0.0617	0.0214
	(0.4283)	(0.4950)	(0.0142)	(0.0088)	(0.0487)	(0.0596)
Inverse Mills Ratio	0.0458***	0.0897***	0.0354***	0.0269***	0.0627***	0.0738***
	(0.0095)	(0.0142)	(0.0084)	(0.0045)	(0.0049)	(0.0062)
_cons	-0.1701***	0.0153***	-0.8954***	-0.0803*	0.6550***	0.0340***
	(0.0122)	(0.0037)	(0.0682)	(0.0411)	(0.0656)	(0.0040)
N	8208	8208	275082	275082	787	787

Notes: This table reports parameter estimates and standard errors of the second stage of the Heckman model as specified in Section 2.4 for App diffusion outcomes (a) Log(Outdegree) and (b) Log(Diameter) for all three samples (1) FanPage ,(2)Top10CC, and (3) Intersection.
*p< 0.10, **p< 0.05, ***p< 0.005. Standard errors shown in parentheses.

Table 2.7 2SLS App Diffusion Regression Models

Parameter	Variable	(1) Log(Diameter)	(2) Log(Outdegree)
AppUse	ln_total_AppUse	0.1302***	0.1739***
		(0.0405)	(0.0301)
Adoption Source	source_ad	0.0376	0.0275
		(0.1376)	(0.0902)
	source_email	0.1194	0.0986
		(0.0806)	(0.0613)
	source_invited	0.1332***	0.1015***
		(0.0075)	(0.0118)
	source_nf	0.3881***	0.0990***
		(0.0295)	(0.0124)
	source_socialad	0.2704***	0.0981***
		(0.0203)	(0.0090)
Demographic Controls	Female	0.0014	0.0056
		(0.0076)	(0.0092)
	location_GB	-0.0196	-0.0378**
		(0.0125)	(0.0138)
	location_US	-0.0193	-0.0698***
		(0.0165)	(0.0199)
Facebook Activities Controls	fbc_interests	0.0037**	-0.0019
		(0.0015)	(0.0019)
	fbc_activities	-0.0052**	-0.0006
		(0.0023)	(0.0021)
	fbc_television	0.0007	0.0016
		(0.0011)	(0.0013)
	fbc_sports	0.0030	-0.0067
		(0.0036)	(0.0045)
	fbc_movies	-0.0003	-0.0018
		(0.0011)	(0.0013)
	fbc_other	0.0009**	0.0008
		(0.0004)	(0.0005)
	fbc_music	0.0001	-0.0004
		(0.0010)	(0.0012)
	fbc_favoriteTeams	-0.0027	0.0015
		(0.0017)	(0.0019)
	fbc_favoriteAthletes	-0.0010	0.0038**
		(0.0015)	(0.0018)
	fbc_games	0.0014	-0.0046
		(0.0020)	(0.0035)
	fbc_books	0.0010	-0.0012

		(0.0012)	(0.0015)
	num_of_FBfriends	0.0000	0.0000*
		(0.0000)	(0.0000)
Privacy Controls	ttl_ln_privacy	-0.0007	-0.0002
		(0.0005)	(0.0005)
Adoption Rate Controls	day_lvl_adop	0.0063***	0.0016***
		(0.0008)	(0.0002)
Network App Controls	thresholddaily_percOfFriendsLike	0.0163*	-0.0334***
		(0.0092)	(0.0070)
	FirstDegFrndAdop	0.2711***	0.3057***
		(0.0451)	(0.1035)
	SecondDegFrndAdop	0.0201**	0.0136***
		(0.0077)	(0.0019)
Friend App activity Controls	frnd_avg_pilot	-0.0054	-0.0084
		(0.0079)	(0.0053)
	frnd_avg_passenger	-0.0025	-0.0005
		(0.0078)	(0.0051)
	frnd_avg_photo	0.0138	0.0014
		(0.0107)	(0.0042)
	frnd_avg_stories	-0.0158	-0.0027
		(0.0120)	(0.0072)
	_cons	-0.1192	-0.4958***
		(0.1081)	(0.1310)
	N	173050	173050
	r2	0.3976	0.3318

Notes: This table reports parameter estimates and standard errors of a two stage least squares model as specified in Section 2.5.5. I use Facebook Likes_count as an IV in the first stage to instrument for App Use (ln_total_activity). App use is measured at time t whereas diffusion metrics are measured at time t+1, such that the total App activity of a user came before the friend's acceptance of the invitation.
*p< 0.10, **p< 0.05, ***p< 0.005. Standard errors shown in parentheses.

Table 2.8 Tradition ads versus Social ads With Friends Liked

Count	Friends who Liked	No Friends Liked	Total
Before Social Ads	4446	13290	17736
After Social Ads	6568 (SOCIAL ADS)	25050	31618
Total	11014	38340	49354

Table 2.9 T-Tests and Averages of Before and After Groups of Ads

Group	Sum of outdegree	Sum of url_dum	Sum of likes_count	Sum of friends_dum	Sum of send_message	Sum of add_as_friend	Sum of num_of_friends	Sum of size_of_cc	Sum of thresholddaily_perfrnd like	Sum of diameter
Before Friends Liked	4490	5293	89542	4098	4893	5079	3137604	15397	40.3865	4935
After Friends Liked	6165	8282	132160	6385	7735	7986	4741782	21418	75.4139	6434
Before NO Friends Liked	2946	16730	287777	13355	15585	16150	5507843	30371	0	3285
After NO Friends Liked	4945	33459	567547	27387	31224	32374	10835949	56075	0	5250
t-test on Before and After P(T<=t) two-tail	0.3101	0.2944	0.4648	0.3274	0.4291	0.4471	0.5283	0.5873	0.2847	0.2463

Table 2.10a Social ads versus Traditional ads in Natural Experiment – App Use

Parameter	Variable	(1) Log(Passenger)	(2) Log(Stories)	(3) Log(Photos)
Adoption Source	after_friendliked- Social Ad	0.0314***	0.1275***	0.0237***
		(0.0037)	(0.0065)	(0.0035)
	after_nofriendliked- Traditional Ad	-0.0100	-0.1251	-0.0221
		(0.0931)	(0.1183)	(0.0144)
	before_nofriendsliked- Traditional Ad	-0.0217**	-0.0477*	-0.0298
		(0.0081)	(0.0262)	(0.0239)
Demographics Controls	Female	-0.0093***	0.0412***	0.0054**
		(0.0023)	(0.0040)	(0.0021)
	location_GB	-0.0076	0.0011	-0.0012
		(0.0078)	(0.0135)	(0.0072)
	location_US	-0.0031	0.0321**	-0.0060
		(0.0066)	(0.0115)	(0.0062)
Network App Controls	FirstDeg_Frnd	0.8433***	0.1143***	0.0734***
		(0.0383)	(0.0168)	(0.0227)
	SecDeg_Frnd	0.0237***	0.0087***	0.0069***
		(0.0025)	(0.0009)	(0.0005)
	thresholddaily_percOfFriendsLike	0.2474***	0.1173***	-0.0569***
		(0.0209)	(0.0365)	(0.0195)
Facebook Activities Controls	fbc_interests	0.0004	-0.0002	0.0016*
		(0.0008)	(0.0015)	(0.0009)
	fbc_activities	-0.0025	0.0128***	-0.0003
		(0.0016)	(0.0028)	(0.0015)
	fbc_television	0.0001	-0.0040*	-0.0010
		(0.0009)	(0.0022)	(0.0009)
	fbc_sports	0.0059	-0.0120*	0.0084**
		(0.0043)	(0.0067)	(0.0031)
	fbc_movies	-0.0008	-0.0014	0.0013
		(0.0010)	(0.0017)	(0.0009)
	fbc_other	0.0007	-0.0027***	0.0003
		(0.0005)	(0.0007)	(0.0004)
	fbc_music	0.0020**	-0.0021	-0.0005
		(0.0009)	(0.0016)	(0.0008)
	fbc_favoriteTeams	0.0014	-0.0025	0.0032**
		(0.0013)	(0.0023)	(0.0012)
	fbc_favoriteAthletes	-0.0012	-0.0026	-0.0014
		(0.0012)	(0.0022)	(0.0012)

	fbc_games	-0.0004	0.0055*	0.0025*
		(0.0013)	(0.0032)	(0.0014)
	fbc_books	0.0021*	-0.0028	0.0006
		(0.0011)	(0.0019)	(0.0009)
	likes_count	-0.0007	0.0021**	-0.0004
		(0.0006)	(0.0009)	(0.0004)
	num_of_FBfriends	-0.0000***	-0.0000***	0.0000
		(0.0000)	(0.0000)	(0.0000)
Facebook Privacy Controls	ln_ttl_publicity	0.3169	-0.4957*	0.0213
		(0.2212)	(0.2641)	(0.0989)
Adoption Rate Controls	day_lvl_adop	-0.0004**	-0.0071***	-0.0007***
		(0.0002)	(0.0003)	(0.0002)
Friend App Activity	frnd_avg_pilot	0.0098	-0.0068	-0.0035
		(0.0095)	(0.0100)	(0.0042)
	frnd_avg_passenger	0.0033	-0.0031	-0.0003
		(0.0025)	(0.0027)	(0.0011)
	frnd_avg_photo	-0.0494	-0.0827**	0.0126
		(0.0366)	(0.0380)	(0.0119)
	frnd_avg_stories	0.0308	0.0421*	-0.0038
		(0.0198)	(0.0216)	(0.0071)
	_cons	0.0221**	0.2364***	0.0376***
		(0.0088)	(0.0154)	(0.0082)
	N	49354	49354	49354
	r2	0.5340	0.1939	0.0407

Notes: This table reports parameter estimates and standard errors of an ordinary least squares. Dependent variables are in logarithmic form.
*p< 0.10, **p< 0.05, ***p< 0.005. Standard errors shown in parentheses.

Table 2.10b Social ads versus Traditional ads in Natural Experiment – App Diffusion

Parameter	**Variable**	**(1) Log(Diameter)**	**(2) Log(Outdegree)**
AppUse	ln_total_activity	0.0211***	0.2074***
		(0.0056)	(0.0569)
AdoptionSource	after_friendliked- Social Ad	0.0721***	0.0217***
		(0.0205)	(0.0066)
	after_nofriendliked- Traditional Ad	-0.2699	-0.1570
		(0.2864)	(0.1984)
	before_nofriendsliked- Traditional Ad	-0.0302*	-0.0324**
		(0.0164)	(0.0120)
Demographic Controls	Female	0.0036	0.0027
		(0.0149)	(0.0143)
	location_GB	-0.0088	-0.0085
		(0.0136)	(0.0126)
	location_US	-0.0095	-0.0244
		(0.0193)	(0.0184)
Facebook Activity Controls	fbc_interests	0.0044**	0.0037*
		(0.0016)	(0.0021)
	fbc_activities	-0.0040	-0.0033
		(0.0030)	(0.0029)
	fbc_television	-0.0005	0.0020
		(0.0016)	(0.0015)
	fbc_sports	0.0021	-0.0040
		(0.0048)	(0.0045)
	fbc_movies	0.0005	-0.0003
		(0.0014)	(0.0013)
	fbc_other	0.0014**	0.0021***
		(0.0005)	(0.0005)
	fbc_music	0.0000	0.0011
		(0.0013)	(0.0013)
	fbc_favoriteTeams	-0.0041*	-0.0031
		(0.0023)	(0.0020)
	fbc_favoriteAthletes	0.0024	0.0036*
		(0.0018)	(0.0019)
	fbc_games	0.0008	-0.0009
		(0.0028)	(0.0026)
	fbc_books	0.0022	0.0025
		(0.0015)	(0.0015)
	num_of_FBfriends	-0.0000**	-0.0000
		(0.0000)	(0.0000)

Facebook Privacy Controls	ln_ttl_privacy	-0.0011	-0.0016
		(0.0015)	(0.0015)
Adoption Rate Controls	day_lvl_adop	0.0010	0.0013
		(0.0007)	(0.0009)
Network App Controls	thresholddaily_percOfFriendsLike	0.2472**	-0.1193*
		(0.1202)	(0.0638)
	FirstDeg_Frnd	0.0648***	0.0421***
		(0.0212)	(0.0116)
	SecDeg_Frnd	0.0355**	0.0286***
		(0.0127)	(0.0098)
Friend App Activity Controls	frnd_avg_pilot	-0.0139	-0.0130
		(0.0146)	(0.0117)
	frnd_avg_passenger	-0.0015	-0.0020
		(0.0071)	(0.0055)
	frnd_avg_photo	0.0424	0.0090
		(0.1380)	(0.0929)
	frnd_avg_stories	-0.0288	-0.0088
		(0.0635)	(0.0427)
	_cons	0.1100	-0.0606
		(0.1374)	(0.1321)
	N	49354	49354
	r2	0.1329	0.1811

Notes: This table reports parameter estimates and standard errors of a two stage least squares model as specified in Section 2.5.5. I use Facebook Likes_count as an IV in the first stage to instrument for App Use (ln_total_activity). App use is measured at time t whereas diffusion metrics are measured at time t+1, such that the total App activity of a user came before the friend's acceptance of the invitation.

*p< 0.10, **p< 0.05, ***p< 0.005. Standard errors shown in parentheses.

Appendices for Essay 2

Figure 3.1 Contest Process

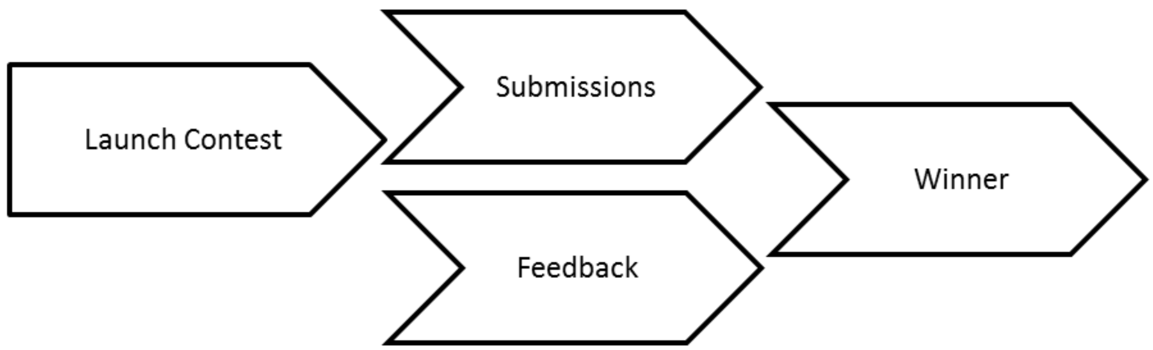


Figure 3.2 Contests in the Crowdsourcing Community

Contest Type	Contest Title	Contest Holder	Ends	Entries	Prize
Open	Help Community Care with a new logo	***	14mins, 28 secs	116	295
Blind	Help Chils Play Qid with a new logo	***	18mins, 48 secs	92	$495
Open	Huniu Photography needs a new logo	***	35mins, 16 secs	124	$295
Open	New logo wanted for Football Die hards	***	42 mins, 43 secs	73	$395
Blind	Logo redesign for 76 year old company	***	47 mins, 23 secs	128	$495

Notes: This figure is adapted from an open innovation marketplace. Contest Holder name is masked.

Figure 3.3 Contest Holder Profile

Contest Holder X		
PROFILE	**Open Contests**	
Activity		LOGO needed for new web application!
Contests Held	2	
Contests Active	1	
Contests Awarded	2	
Prizes Awarded	$595	
Average Feedback	100%	

Notes: This figure is adapted from an open innovation marketplace.

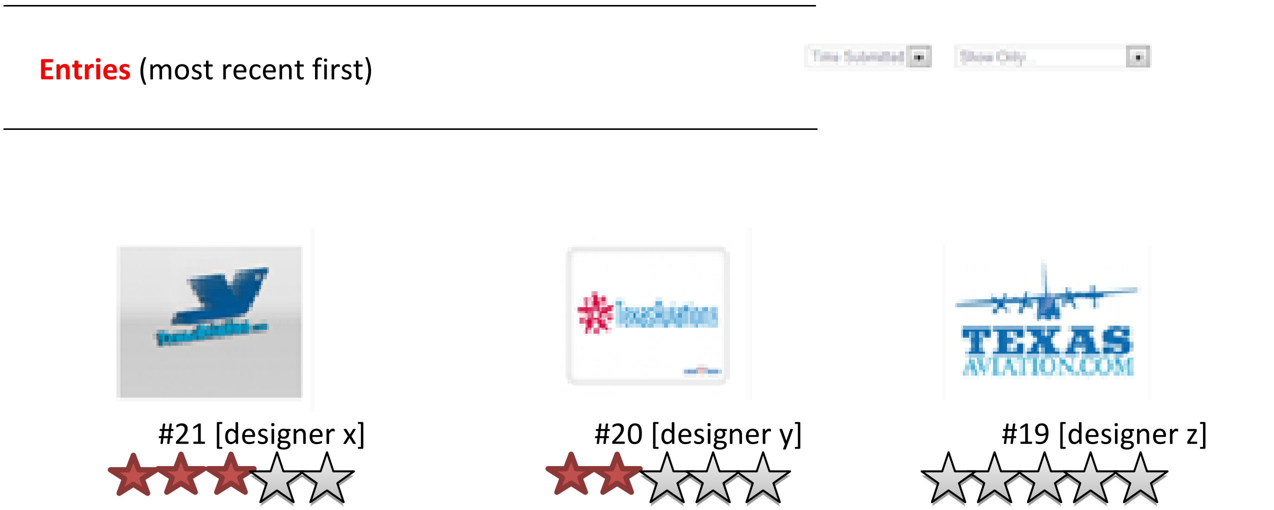
Figure 3.4 Contest Description

Design Brief		
For Contest : New Logo for Avaiation Company , Held by : *** , in the logo design community		
Open	**Entries**	**Prize**
Contest accepting entries 6 days, 8 hours remaining	20	$600

Brief Overview	We are a new aviation company. Our goal is to provide the premier club for aviation enthusiasts in the stage of Texas. We will be hosting events for owners of aircraft, and also provide a complete resource for all kinds of aviation information for the state of Texas.
Brand Name	***
Target Audience	We are targeting Aviation enthusiasts for the state of Texas. Pretty much anyone that is interested in flying.
Requirements	Please include .com in the logo. I am looking for something professional and also creative. You are free to be creative on this logo.

Notes: This figure is adapted from an open innovation marketplace. Brand name is masked.

Figure 3.5 Contestants and their Uploads to a Contest



Notes: This figure is adapted from an open innovation marketplace. It shows a few submission entries for a contest along with the star rating for each submission. A star is given by the contest holder to a submission. Star ratings are out of 5. For example, submission #21 has a star feedback of 3 out of 5. Submission #20 has a star feedback of 2 out of 5, and submission #19 has no star feedback. Designer names are masked.

Figure 3.6 Contestant Profile

Profile for Designer X

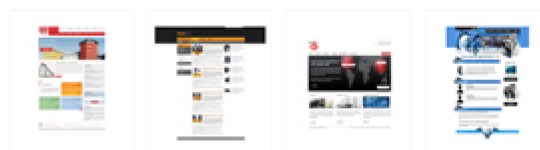
Activity	
Contests Entered	230
Contests Won	16
Contests Participating	2

Bio

Portfolio by request.

Folio Highlights all folio >

Recent Wins all wins >



Notes: This figure is adapted from an open innovation marketplace. Designer name is masked.

Figure 3.7 Submission Order Frequency

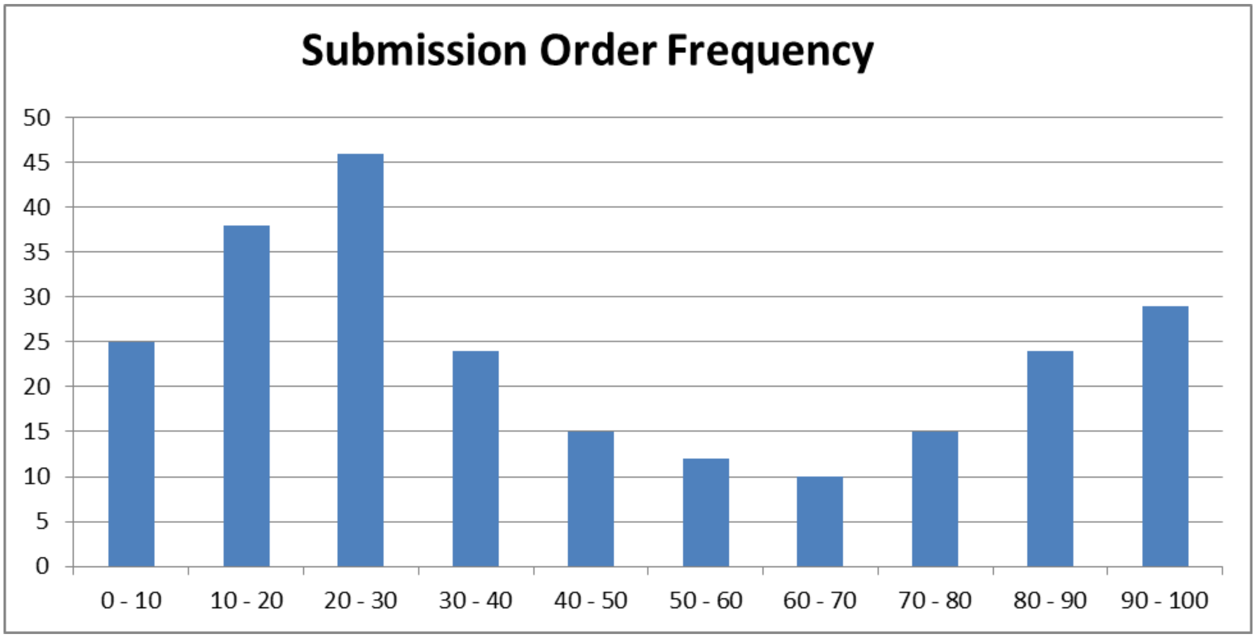


Figure 3.8 Text Feedback Frequency

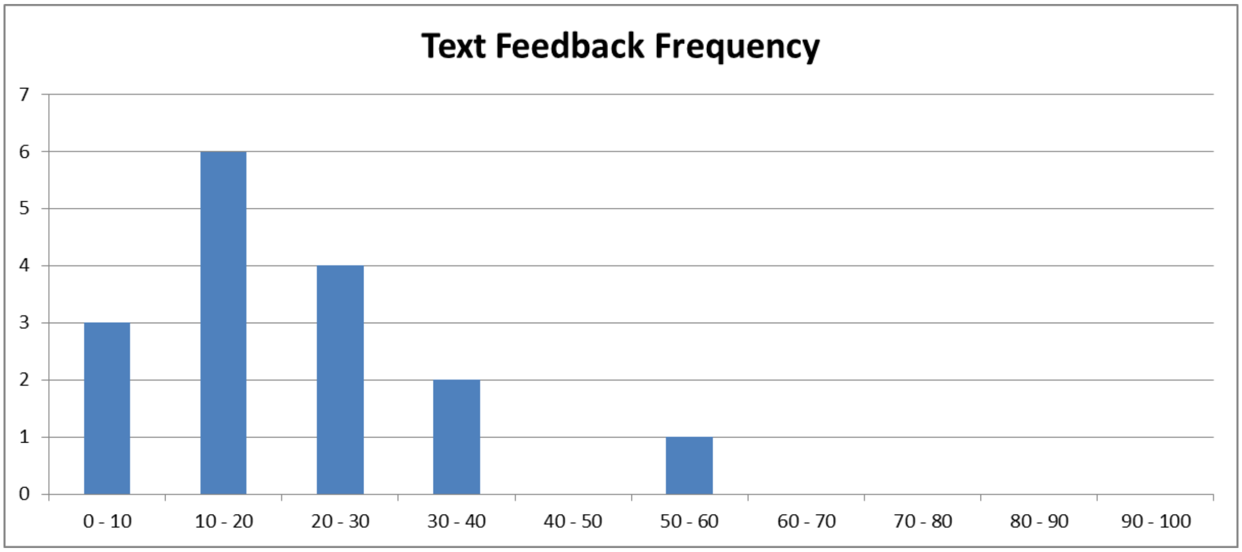


Figure 3.9 Star Feedback Frequency

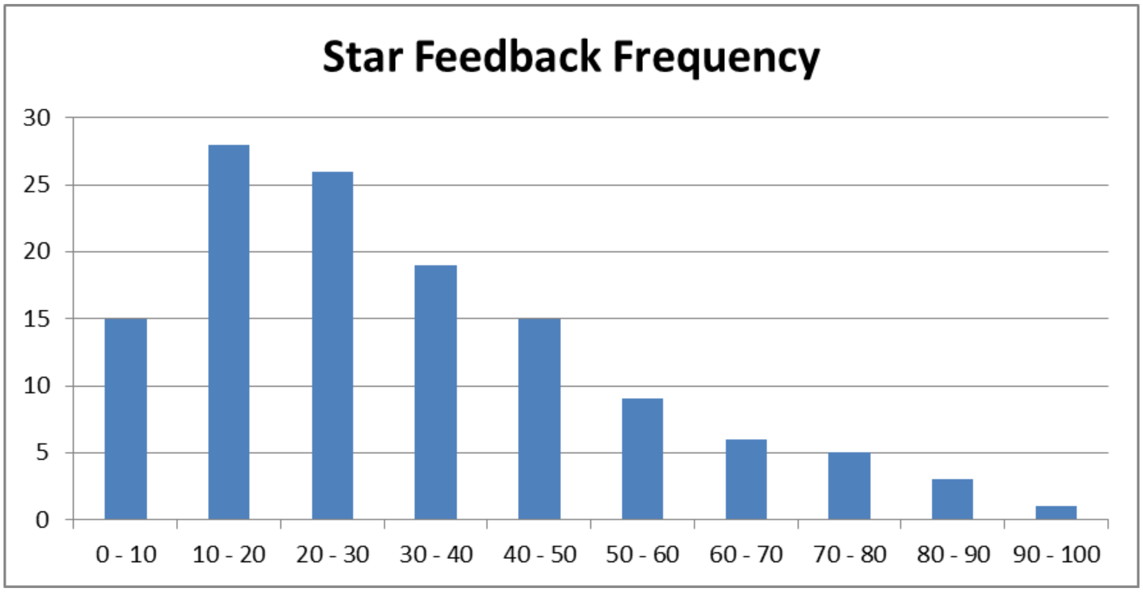


Figure 3.10a Benefits and Drawbacks of Open and Blind Contests for a Contest Holder

Contest Type	**Benefits**	**Drawbacks**
Open	- On average more participation. - Visible in search engines like Google, and thus draws on more designers not necessarily limited to the community.	- More derivative work. - Everyone can see final logo design; less confidential.
Blind	- No derivative work – contestants cannot see each other's submissions - No one can see the final logo design ; more confidential	- Less designer participation - Not visible in search engines

Figure 3.10b Benefits and Drawbacks of Early Versus Late Submissions for a Contestant

Open Contests With Feedback	**Positive**	**Negative**
Early	- More likely to receive feedback and improve upon feedback. - Might be able to deter entry of others by submitting high quality solutions early.	- More likely to be copied by others
Late	- More time to work on solutions - More information in a contest- more chances to learn from competing submissions. - Less likely to be copied by others	- Less likely to receive feedback - More likely to have similar competing submissions

Figure 3.11 Feedback Categorization

(1) Generic Feedback	**(2) Specific Feedback**
Overall Feedback for all, with no specific callouts.	Specific callout to a submission (Use of submission number or username).
Examples:	*Examples:*
Hi again everyone, thanks for all the designs. The designs are great however many are either more masculine or more feminine. The primary target is female (they will likely be the purchaser of the gifts) however they will buy for both males and females. This means the logo on the packaging will need to be more gender neutral and not too feminine. Female buyers might be turned off if they think the product looks too girly and not suitable for a male. I hope this makes sense and helps with your designs.	*Did we lose BigBaldBeard? We liked number #305...*
Hi everyone thanks for the designs so far. We're looking for something pretty simple with more emphasis and refinement on the typography. It looks like the device we asked for might be a bit difficult to crack so we would prefer you concentrate on making the typography look simple and premium. If you have a great idea for a device/mark by all means submit it. Thanks again everyone.	*#82, #68 and #59 are the leading designs. #59 is the only contender for the logo but it is still not quite perfect. Some additional shaping to the G to make it look more smiley is desired. I would also like to see the text in different colors for $59. For #68 - I love this but it needs a body. #82is the most fun body because of the attire. His shirt is open and the collar is outside of his lapel. He's a bit "cooler" and less formal. He is having fun, which is important for my character to portray. I don't know what the rules are with merging designs, but I feel like the purple head of #68 is much more elaborate and higher quality, so if anything, I would like to see that design adopt a better body / outfit. #82 really only has the attire correct. Everything else is not something I care for too much.*
Try put signal strength, globalization, and tower within the text of the logo.	*We like #71 #52. But we would like to some more bright colors in both.*

Figure 3.12a Contest Holders Survey - Email

Dear

My name is Abrar Al-Hasan, a PhD student at the University of Maryland College Park. As part of my dissertation I am studying the dynamics of online crowdsourcing markets. This survey seeks to understand how contest holders make choices in crowdsourcing markets and the factors that drive these choices.

This survey can be easily accessed online by clicking on the link below, and it will only take 5 minutes of your time. To show our appreciation, we will also send you a summary of the survey results once it is completed.

[Link]

All responses to our survey will be kept strictly confidential. Thank you very much for your help!

Abrar Al-Hasan
University of Maryland, College Park
Email: aalhasan@rhsmith.umd.edu

Figure 3.12b Survey Questions

Welcome to the Contest Holder Assessment Survey

Thank you for participating in the survey. Your response to fill out each question in this survey is extremely important. This survey will take approximately 5 minutes to complete.

Please fill out the following questions by choosing the appropriate answer(s):

Questions:

1- How many contests have you launched at [the online design community][9]?

- 1-2
- 3-4
- 5-6
- More than 6 : ____

2- When you launch the contest(s), do you usually have a design in mind or are you looking for new ideas from designers? You can choose more than one answer.

- I have a rough idea of the design I want
- I usually have a fair idea of the design I want
- I don't have any prior ideas
- I am looking for new designs
- Others (please specify): ______________

Types of Contests Launched

3a- What types of contests have you launched?

- Open
- Blind
- Both

3b- In your opinion, what advantages do open contests have over blind contests?

3c- In your opinion, what advantages do blind contests have over open contests?

Submissions Received

4a- If you launched an Open contest(s), did you find that the later submissions are sometimes similar to the earlier submissions?

- Yes
- No

[9] The name of the online community has been masked in this dissertation for privacy concerns

4b- If you launched a Closed contest(s), did you find that the later submissions are sometimes similar to the earlier submissions??

- Yes
- No

Feedback

5a- Have you given feedback to participants (designers) in your contests?

- Yes
- No

5b- How long do you wait before providing feedback to participants (designers) in your contests?

- I don’t provide feedback
- I provide feedback as soon as I see a (potentially) good design
- I wait for a few days before providing feedback
- I wait for a certain number of submissions before providing feedback

5c- Did you give feedback to multiple submissions? If yes, did you provide feedback to

- I provide feedback only to the best submission.
- I provide feedback to the top 2 or 3 submissions.
- I provide feedback to the top 4 or 5 submissions.

5d- The primary purpose of providing feedback to contestants is

- For better reputation in the [the online design community] community.
- To suggest improvements to an existing design.
- To attract more submissions.
- Others (please specify): ________________

5e- Do you typically see more submissions from other contestants upon providing feedback?

- Yes
- No

5f- After providing feedback, did you get a better design from the contestant(s) to whom you provided feedback?

- Yes
- No

5g- Do you find that providing feedback also benefits other contestants who submit later?

- Yes
- No

Figure 3.13 Informal Interview

Contests Entered	**Contests Won**	**Contests Currently Participating**	**Geographical Location**	**Profession. Do you design logos as a hobby or a main source of income?**	**How do you choose what contests to join?**	**Do you have a preference as to when you post up your design? (E.g. early on the competition, sometime in the middle, or right before the competition ends).**
318	47	0	Serbia	Mother by day, designer when night falls.Working as a freelance on [design community] and huge number of clients out of [design community].Self-learned designer.	I just like it or don't. - When I see a contest I may have a vision of what I can do in it, or I just skip it if my brain can't think of something. Just feel what is good for me.	Sometimes I am the first one. Sometimes in the middle, but I don't like to enter at the end. I did enter few at the end, but prefer beginning and middle.
145	21	1	Indonesia	I'm still studying in Gadjah Mada University. Designing is my hobby and secondary profession. It's main source of income. When I have more time, I sometimes spend it to join [design community] contest.	I usually view more open contests first. If I like to join, I join it. I don't like if too much participant in a contest. I even don't like a logo contests. I prefer to join another logo contests.	When I finished doing my design, I submitted it soon. I give the contest holder more time to revise my design.
301	22	3	Philippines	Graphic Artist. Hobby just to augment income.	I choose a contest that the contest holder is actively participating. Meaning that he/she gives regular feedback and the way the entries are rated truly reflects what he/she wants. Most importantly, if the contest is also interesting.	No actual preference. It depends on the contest and inspiration. I usually join early if I have an inspiration for an entry or if the contest is very interesting. I usually wait for participation from others. And then I would wait for a rating to happen to see if the CH is giving feedback and also to get a better idea of what he is after.
1471	26	19	Arizona, USA	[design community] is my sole source of self-generated income.	I look mainly at logo contests of $300 or more. From there, I see if the contest title and brief are clear and inspire an idea that I believe will BENEFIT the CH and help him become successful.	It's rare that I enter a contest the first or second day. That usually is a waste of time unless the brief is very clear and specific. I watch for the amount of feedback and participation by the CH to understand what he wants better and then decide if I will enter.

Table 3.1 Variables and Descriptions

	Variable name	Description
Dependent Variable	Winner_dum	Binary variable that indicates whether the contestant won the contest.
Independent Variables		
Contestant Skill & Expertise	c_skill	Contestant skill = (Total number of previous wins) / (Total number of previous contest participation)
	c_expertise_dum	Dummy of whether contestant has more than one design Expertise e.g. logo , graphics, website , etc.
	c_experience	Contestant total contests entered.
Strategic Behavior	c_subOrder	Time of contestant's first submission in contest X ,as a percentage of total entries in contest X.
Direct Feedback to Contestant	c_maxstar	For contests with Feedback : Maximum star rating of the contestant received by the contest holder (out of 5) in contest X.
	c_feedback_dum	Dummy whether the contestant was given feedback through stars or through text.
Feedback to Others		
Text- Type of Feedback	contest_SpecificFdbk	For contests with Feedback: Total count of the specific feedback given in a contest at time of contestant entry (see Figure 3.11).
	contest_GenericFdbk	For contests with Feedback: Total count of the generic feedback given in a contest at time of contestant entry (see Figure 3.11).
Star	c_maxstar_prior	Max star rating given in a contest at time of contestant entry to contest.
	c_avgstar_prior	Average star rating given in a contest at time of contestant entry to contest.
	c_ttlnumuserswithstars_prior	Total number of users given a star rating at time of contestant entry to contest.
Resubmission	c_ttl_resubs	Total number of resubmissions of contestant in contest X.
	c_timeToResubmit	The average lag in number of submissions from one submission to next for a contestant in contest X.
Control Variables		
Contestant Controls	c_participating_in	Total number of contests user is participating in at time of submission in contest X.
	c_membershipAge	Total number of days the contestant has been in the design community.
Contest Controls	contest_entries	Total number of entries in contest X.
	contest_prize	Prize awarded ($) in contest X.
	contest_ttldescription_length	Count of the total number of words in the design description.
Contest Holder Controls	ch_matchesHeld	Contest Holder total number of matches held.
	ch_matchesPrizes	Contest Holder total number of prizes awarded.
	ch_avgfdbk	Contest Holder average feedback.

Table 3.2a Descriptive Statistics

Unit	Variable	Description	Mean	Std.Dev	Min	Max
Contestant *i*						
	c_subOrder	First submission time of contestant *i*.	0.5698	0.3888	0	1
	c_maxstar	Max star rating of contestant *i*.	1.9656	1.8223	0	5
	c_won	Total number of contests won for contestant *i*.	13.3417	21.1542	0	352
	c_experience	Total number of contests entered for contestant *i*.	287.0228	319.4493	2	6043
	c_skill	Contestant *i*'s skill = Won/Entered.	0.0405	0.0545	0	1
	c_participating_in	Contestant *i* total number of contests currently participating in.	1.2349	2.3983	0	72
	c_ttl_resubs	Contestant *i*'s total number of resubmissions.	1.8375	3.2347	0	14
	c_membershipAge	Contestant *i*'s total number of days in community.	422.0544	612.7567	102	3102
Contest Details						
	contest_Entries	Total number of Entries in a contest.	211.1656	354.0386	20	4531
	contest_Designers	Total number of Designers in a contest.	69.2919	76.9834	5	1523
	contest_prize	Total USD Prize of Contest.	517.4523	201.9317	100	1945
	contest_tt_comments	Total number of comments in a contest.	10.5048	12.9505	0	145
	contest_ttldescription_length	Count of the total number of words in the design description	223.2715	162.1385	38	1793
Contest Holder						
	ch_matchesheld	CH -Total number of matches held.	2.8316	3.36956	1	156
	ch_matchesPrizes	CH -Total USD awarded as Prizes.	718.6134	1275.0500	0	32412
	ch_AvgFdbk	Average Feedback of CH.	0.7013	0.3311	0	1
	ch_Lastseen	CH - Number of days last seen.	286.2236	485.7173	0	2103
	ch_DaysSinceLastFeedback	CH - Number of Days since last feedback.	689.4185	478.0628	0	2413

Note: Whole sample statistics. CH denotes Contest Holder.

Table 3.2b Descriptive Statistics of Different Contests

Parameter	**Open With Feedback** Mean (stdev)	**Open With No Feedback** Mean (stdev)	**Blind With Feedback** Mean (stdev)	**Blind With No Feedback** Mean (stdev)
Resubmissions				
Early	6.2511 (2.9281)	2.5012(0.5101)	4.1231(0.8165)	0.8283(0.3102)
Middle	4.3221(3.4010)	0.5129(0.9371)	2.1232(0.5774)	0.1023(0.2938)
Late	0.5281(0.8210)	0.2574(0.5023)	0.3984(0.5012)	0.0023(0.0109)
Skill				
Early	0.0425(0.04102)	0.0368(0.03483)	0.0432(0.03979)	0.0423(0.04321)
Middle	0.0382(0.02307)	0.0375(0.03401)	0.0388(0.03521)	0.0402(0.04117)
Late	0.0432(0.03918)	0.0401(0.04032)	0.0415(0.04019)	0.0416(0.04123)
Experience				
Early	280.1923(340.3910)	276.3918(376.1920)	289.1920(411.3201)	293.1029(417.1927)
Middle	275.1263(382.4263)	273.2837 (401.2371)	295.2736(421.2929)	290.1139(411.2039)
Late	293.1249(312.3010)	285.1029(412.1029)	301.2394(370.1820)	292.1298(428.3098)
Membership Age				
Early	417.2707(687.1982)	415.2380(632.1028)	434.1039(589.2981)	431.2983(597.4126)
Middle	414.1923(662.9247)	409.2938(640.2038)	410.2981(620.1092)	411.2931(610.2986)
Late	426.9810(680.1725)	421.2981(664.2081)	438.1029(601.2091)	435.2827(593.1092)
Contest Details				
Contest Entries	267.4165 (363.9745)	228.5493 (231.5256)	197.3981(221.3382)	151.2983 (215.0192)
Designers	94.7061 (73.2588)	71.7773 (49.2803)	63.4918 (50.2983)	47.1923 (52.3948)
Prize	520.2096 (290.8650)	473.3077 (148.0056)	584.1201 (262.9384)	512.1717 (200.8537)
Contest Description Length	215.2093 (166.2981)	220.1837 (170.2721)	222.3948 (171.2481)	235.2985 (150.2938)
Contest total comments/Feedback	8.6181 (18.6417)	0	12.39155 (22.2726)	0

Notes: Early, Middle and Late are defined as, the time of submission as a percentage of total entries; Early: 0-20%, Middle: 20-80%, and Late: 80-100%.

Table 3.3 Main Findings

Category	Parameter	Open With Feedback	Open With No Feedback	Blind with Feedback	Blind With No Feedback	Outcome = Winning
Submission Strategy	subOrder & subOrder^2	U-shaped	+	-	NS	Winners capitalize on design visibility and feedback visibility and submit late in open contests, whereas it is better to submit early when there is no design visibility. When there is neither design nor feedback, there is no significant strategic behavior. **Supports Proposition 1.**
	Skill*suborder	+	+	-	NS	Higher skill and later submission, lead to higher probability of winning in open contests.
Feedback to Others	star	-		-		Competition.
	star*subOrder	+		-		In open contests, winners capitalize on design visibility and feedback given to others and submit later. Whereas in blind contests, users cannot capitalize on information spillover as such, thus early submission is best. **Supports Proposition 2a.**
Feedback to Others – Text	Specific	+		-		Open contests: Winners capitalize on specific feedback (information spillover). Whereas in blind contests, specific feedback translates to competition. **Supports Proposition 2b.**
	Generic	NS		+		Generic feedback is more information to everyone and is more beneficial in blind contests.
	subOrder*Specific	+		-		Open contests: later submission & more specific leads to more information spillover, resulting in higher probability of winning. Blind contests: specific feedback is competition. **Supports Proposition 2b.**
	subOrder*Generic	NS		-		Blind contests, early submission and generic feedback leads to higher probability of winning.
Feedback to User	Feedback to User	+		+		Higher feedback leads to higher probability of winning.
Expertise	All measures	+	+	+	+	Higher skill leads to higher probability of winning
Resubs	ttlResubs	+	+	+	NS	More resubmissions lead to higher probability of winning.
	TimetoResubmit	-	NS	-	NS	Shorter time to resubmit leads to higher probability of winning.
	Feedbk*TimetoResubmit	-		-		Higher feedback and earlier resubmission lead to higher probability of winning.
	ttlResubs*Feedback	+		+		Higher total resubmission of users with feedback as opposed to users no feedback is positively associated with probability of winning.

Note: (+) indicates a positive significant coefficient, (-) indicates a negative significant coefficient, NS indicates non-significant coefficient.

Table 3.4 Regression Analysis

	Open With Feedback	Open With Feedback	Open With No Feedback	Open With No Feedback	Blind With Feedback	Blind With Feedback	Blind With No Feedback	Blind With No Feedback
Variable	**(1) FE-IV**	**(2) RE-IV**	**(3) FE-IV**	**(4) RE-IV**	**(5) FE-IV**	**(6) RE-IV**	**(7) FE-IV**	**(8) RE-IV**
c_subOrder	-0.1223***	-0.1455***	0.1634***	0.2176***	-0.2544***	-0.3246***	-0.2321	-0.2551
	(0.0187)	(0.0173)	(0.0166)	(0.0146)	(0.0432)	(0.0328)	(0.7241)	(0.6394)
c_subOrder^2	0.1667***	0.1976***	0.0076***	0.0084***	-0.0302***	-0.0306***	-0.0864	-0.0869
	(0.0419)	(0.0402)	(0.0012)	(0.0013)	(0.0077)	(0.0080)	(0.6741)	(0.6743)
contest_SpecificFdbk	0.2421***	0.2822***			-0.1052***	-0.1053***		
	(0.0543)	(0.0512)			(0.0282)	(0.0283)		
contest_GenericFdbk	-0.0052	-0.0053			0.0344***	0.0352***		
	(0.3011)	(0.3015)			(0.0093)	(0.0091)		
c_subOrder*GenericFdbk	0.0283	0.0285			-0.2152***	-0.2777***		
	(0.1071)	(0.1062)			(0.0343)	(0.0317)		
c_subOrder*SpecificFdbk	0.2841***	0.4647***			-0.1931***	-0.2233***		
	(0.1001)	(0.1020)			(0.0465)	(0.0455)		
c_skill	0.4666***	0.4959***	0.4841***	0.5269***	0.5066***	0.5787***	0.5684***	0.6188***
	(0.1264)	(0.1213)	(0.1034)	(0.1029)	(0.0964)	(0.0952)	(0.0634)	(0.0626)
c_experience	0.2583***	0.2921***	0.2733***	0.3337***	0.3361***	0.3462***	0.3428***	0.3530***
	(0.0693)	(0.0667)	(0.0733)	(0.0724)	(0.0831)	(0.0822)	(0.0808)	(0.0789)
c_expertise_dum	0.1159***	0.1467***	0.0718***	0.0727***	0.1513***	0.1567***	0.2113***	0.2417***
	(0.0267)	(0.0214)	(0.0177)	(0.0179)	(0.0428)	(0.0429)	(0.0310)	(0.0304)
c_skill*c_subOrder	0.4953***	0.5062***	0.5236***	0.5429***	-0.2563***	-0.2711***	-0.1650	-0.1652
	(0.0829)	(0.0816)	(0.0664)	(0.0641)	(0.0429)	(0.0426)	(0.4323)	(0.4325)
c_experience*c_subOrder	0.3123***	0.3430***	0.3239***	0.3347***	-0.1601***	-0.1604***	-0.1287	-0.1289

	(0.0624)	(0.0617)	(0.0834)	(0.0836)	(0.0341)	(0.0344)	(0.6472)	(0.6457)
c_expertise_dum*c_subOrder	0.2043***	0.2059***	0.1302***	0.1310***	-0.1087***	-0.1092***	-0.0247	-0.0251
	(0.0441)	(0.0425)	(0.0165)	(0.0158)	(0.0212)	(0.0205)	(0.7566)	(0.7568)
c_maxstars	0.0751***	0.0973***			0.2011***	0.2034***		
	(0.0245)	(0.0246)			(0.0380)	(0.0289)		
c_timeToResubmit	-0.0134***	-0.0157***	-0.0015	-0.0016	-0.0341***	-0.0392***	-0.0342	-0.0347
	(0.0024)	(0.0034)	(0.0532)	(0.0543)	(0.0074)	(0.0071)	(0.7213)	(0.7215)
c_maxstars*c_timeToResubmit	-0.0976***	-0.0991***			-0.0339***	-0.0361***		
	(0.0201)	(0.0197)			(0.0079)	(0.0071)		
c_ttl_resubs	0.0557***	0.0533***	0.0089***	0.0115***	0.1043***	0.1144***	0.0438	0.0439
	(0.0128)	(0.0121)	(0.0012)	(0.0025)	(0.0113)	(0.0107)	(0.0322)	(0.0324)
c_ttl_resubs*feedback_dum	0.0835***	0.0846***			0.1721***	0.1922***		
	(0.0209)	(0.0214)			(0.0135)	(0.0128)		
feedback_dum	0.0624***	0.0628***			0.1319***	0.1328***		
	(0.0151)	(0.0146)			(0.0143)	(0.0144)		
c_maxstar_prior	-0.0007***	-0.0009***			-0.0896***	-0.0902***		
	(0.0002)	(0.0003)			(0.0221)	(0.0227)		
c_maxstar_prior*suborder	0.1284***	0.1321***			-0.0054***	-0.0067***		
	(0.0322)	(0.0314)			(0.0014)	(0.0013)		
c_avgstar_prior	-0.0035***	-0.0038***			-0.0507***	-0.0518***		
	(0.0009)	(0.0008)			(0.0123)	(0.0117)		
c_avgstar_prior*suborder	0.1174***	0.1181***			-0.0139***	-0.0144***		
	(0.0320)	(0.0327)			(0.0031)	(0.0029)		
c_ttlnumuserswithstars_prior	-0.0050***	-0.0067***			-0.1126***	-0.1148***		
	(0.0008)	(0.0007)			(0.0135)	(0.0137)		
c_ttlnumuserswithstars_prior*subOrder	0.1055***	0.1073***			-0.0729***	-0.0732***		
	(0.0155)	(0.0154)			(0.0108)	(0.0113)		

c_participating_in	0.0000	0.0000	0.0000	0.0000	-0.0008	-0.0012	-0.0001	-0.0009
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0212)	(0.0203)	(0.0012)	(0.0011)
contest_entries	-0.0005**	-0.0008**	-0.0006***	-0.0009***	-0.0008***	-0.0009***	-0.0001**	-0.0001**
	(0.0002)	(0.0004)	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0000)	(0.0000)
contest_prize	-0.0026***	-0.0027**	-0.0011***	-0.0012***	-0.0019**	-0.0019**	-0.0010***	-0.0011***
	(0.0009)	(0.0012)	(0.0003)	(0.0004)	(0.0007)	(0.0007)	(0.0002)	(0.0003)
ch_matchesHeld	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0008	-0.0009
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0020)	(0.0021)
ch_matchesPrizes	-0.0401***	-0.0408**	-0.0040***	-0.0040***	-0.0000***	-0.0003***	-0.0000***	-0.0002***
	(0.0064)	(0.0186)	(0.0003)	(0.0005)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
ch_avgfdbk	-0.0000***	-0.0002***	-0.0176***	-0.0172***	-0.0007***	-0.0007***	-0.0104***	-0.0107***
	(0.0000)	(0.0000)	(0.0018)	(0.0019)	(0.0001)	(0.0001)	(0.0013)	(0.0015)
c_membershipAge	0.0218***	0.0241***	0.0421***	0.0448***	0.0545	0.0579	0.0240	0.0253
	(0.0057)	(0.0069)	(0.0076)	(0.0112)	(0.1152)	(0.1143)	(0.0845)	(0.1153)
contest_ttldescription_length	-0.0112***	-0.0148***	-0.0094***	-0.0109***	-0.0124***	-0.0132***	-0.0127***	-0.0138***
	(0.0011)	(0.0014)	(0.0021)	(0.0024)	(0.0031)	(0.0036)	(0.0032)	(0.0033)
_cons	0.1551***	0.2682***	0.3151***	0.3353***	0.6422***	0.6249***	0.4285***	0.4376***
	(0.0095)	(0.0588)	(0.0824)	(0.0775)	(0.1467)	(0.1143)	(0.1185)	(0.1041)
R-sq between	0.1063	0.1085	0.0141	0.0187	0.0061	0.0083	0.1047	0.1145
R-sq within	0.1116	0.2101	0.0261	0.0298	0.0115	0.1108	0.1246	0.1358
R-sq overall	0.1389	0.1969	0.0871	0.0945	0.0376	0.0964	0.1189	0.2104
F-Stat	22.64***		25.53***		24.48***		27.43***	
N	1630457	1630457	1632492	1632492	477427	477427	152845	152845

Notes: This table reports parameter estimates and standard errors from the two stage fixed effects and random effects model specified in Section 3.4. The IV used for a contestants' submission order (c_subOrder) is the average submission order in previous periods. The IV used for a contestants feedback (c_maxstars) is the average star rating in previous periods. The sample consists of 6091 open contests with feedback, 7134 open contests without feedback, 2417 blind contests with feedback, and 1003 blind contests with no feedback.
*p< 0.10, **p< 0.05, ***p< 0.005. Standard errors shown in parentheses.

Bibliography

Aaker, D. A., & Brown, P. K. (1972). Evaluating vehicle source effects. *Journal of advertising Research*, *12*(4), 11-16.

Agarwal, R., & Lucas Jr, H. C. (2005). The information systems identity crisis: Focusing on high-visibility and high-impact research. *Mis Quarterly*, *29*(3), 381-398.

Agarwal, R., Animesh, A., & Prasad, K. (2009). Research Note—Social Interactions and the "Digital Divide": Explaining Variations in Internet Use. *Information Systems Research*, *20*(2), 277-294.

Aral, S., & Walker, D. (2011). Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Management Science*,*57*(9), 1623-1639.

Aral, S., Muchnik, L., & 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.

Archak, N. (2010). Money, glory and cheap talk: analyzing strategic behavior of contestants in simultaneous crowdsourcing contests on TopCoder. com. *Proceedings of the 19th international conference on World wide web* (pp. 21-30). ACM.

Atkinson, J. W. (1964). An introduction to motivation. Princeton, NJ: Van Nostrand.

Baik, K. H., & Lee, S. (2000). Two-stage rent-seeking contests with carryovers. *Public Choice*, *103*(3), 285-296.

Bain, J. S. (1956). *Barriers to new competition: their character and consequences in manufacturing industries.* Cambridge, MA: Harvard University Press.

Bajari, P., & Hortacsu, A. (2003). The winner's curse, reserve prices and endogenous entry: Empirical insights from eBay auctions. *RAND Journal of Economics, 34*(2), 329–355.

Bajari, P., & Hortacsu, A. (2004). Economic Insights from Internet auctions. *Journal of Economic Literature*, *42*(6), 457–486

Bala, V., & Goyal, S. (1998). Learning from neighbours. *The Review of Economic Studies*, *65*(3), 595-621.

Bandura, A. (1991). Self-regulation of motivation through anticipatory and self-regulatory mechanisms. In R. A. Dienstbier (Ed.), Perspectives on motivation: Nebraska Symposium on Motivation (Vol. 38, pp. 69–164). Lincoln: University of Nebraska Press

Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of personality and social psychology*, *45*(5), 1017- 1028.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, *107*(3), 797-817.

Batra, R., & Ray, M. L. (1983). Operationalizing involvement as depth and quality of cognitive response. *Advances in consumer research*, *10*(1), 309-313.

Becker, G. S. (1991). A note on restaurant pricing and other examples of social influences on price. *Journal of Political Economy*, *99*(5), 1109-1116.

Bensen, C. (2013). When Brand Advocates Become Employees: IdeaStorm PM Cy Jervis Shares Tips for Businesses. *Dell*, Jan 7. Retrieved from http://en.community.dell.com/dell-groups/sbc/b/weblog/archive/2013/01/07/when-brand-advocates-become-employees-ideastorm-pm-cy-jervis-shares-tips-for-businesses.aspx

Benveniste, L. M., Busaba, W. Y., & Wilhelm, W. J. (2002). Information externalities and the role of underwriters in primary equity markets. *Journal of Financial Intermediation*, *11*(1), 61-86.

Biyalogorsky, E., Gerstner, E., & Libai, B. (2001). Customer referral management: Optimal reward programs. *Marketing Science*, *20*(1), 82-95.

Bramoullé, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of econometrics*, *150*(1), 41-55.

Bronner, F., & Neijens, P. (2006). Audience experiences of media context and embedded advertising: a comparison of eight media. *International Journal of Market Research*, *48*(1), 81.

Brown, C. L., & Lattin, J. M. (1994). Investigating the relationship between time in market and pioneering advantage. *Management science*, *40*(10), 1361-1369.

Brown, J. J., & Reingen, P. H. (1987). Social ties and word-of-mouth referral behavior. *Journal of Consumer Research*, *14*(3), 350-362.

Brown, S. A., & Venkatesh, V. (2005). Model of adoption of technology in households: A baseline model test and extension incorporating household life cycle. *MIS quarterly*, *29*(3), 399-426.

Brynjolffson, E., & Hitt, L. M. (2000). Beyond computation: Information technology, organizational transformation and business performance. *The Journal of Economic Perspectives*, *14*(4), 23-48.

Brynjolfsson, E., & Hitt, L. M. (2003). Computing productivity: Firm-level evidence. *Review of economics and statistics*, *85*(4), 793-808.

Calder, B. J., Malthouse, E. C., & Schädel, U. (2009). An experimental study of the relationship between online engagement and advertising effectiveness. *Journal of Interactive Marketing*, *23*(4), 321-331.

Carpenter, G. S., & Nakamoto, K. (1989). Consumer preference formation and pioneering advantage. *Journal of Marketing Research*, *26*(3), 285-298.

Chang, Y. B., & Gurbaxani, V. (2012). The Impact of IT-Related Spillovers on Long-Run Productivity: An Empirical Analysis. *Information Systems Research*, *23*(3-Part-2), 868-886.

Chatterjee, P., Hoffman, D. L., & Novak, T. P. (2003). Modeling the clickstream: Implications for web-based advertising efforts. *Marketing Science*, *22*(4), 520-541.

Chatterjee, Patrali. (1998). *Modeling consumer network navigation at World Wide Web sites—Implications for advertising.* (Unpublished doctoral dissertation). Vanderbilt University, Nashville, TN.

Chen, Y., Wang, Q., & Xie, J. (2010). Online social interactions: A natural experiment on word of mouth versus observational learning. *Available at SSRN 1501843*.

Cheng, J., & Nault, B. (2007). Industry level supplier-driven IT spillovers. *Management Science*. *53*(8), 1199–1216.

Chevalier, J., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, *43*(8), 345–354.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine*, *357*(4), 370-379.

Coe, D. T., & Helpman, E. (1995). International r&d spillovers. *European Economic Review*, *39*(5), 859-887.

Conrad, C. A. (1983). The advantage of being first and competition between firms. *International Journal of Industrial Organization*, *1*(4), 353-364.

Cooper, R. G. (1979). The dimensions of industrial new product success and failure. *The Journal of Marketing*, *43*(3), 93-103.

Coulter, K. S. (1998). The effects of affective responses to media context on advertising evaluations. *Journal of Advertising*, *27*(4), 41-51.

Cunningham, T., Hall, A. S., & Young, C. (2006). The advertising magnifier effect: An MTV study. *Journal of Advertising Research-New York*, *46*(4), 369.

Dahlén, M. (2005). The medium as a contextual cue: Effects of creative media choice. *Journal of Advertising*, *34*(3), 89-98.

Dahlén, M., & Edenius, M. (2007). When is advertising advertising? Comparing responses to non-traditional and traditional advertising media. *Journal of Current Issues & Research in Advertising*, *29*(1), 33-42.

D'Aspremont, C., & Jacquemin, A. (1988). Cooperative and noncooperative R & D in duopoly with spillovers. *The American Economic Review*, *78*(5), 1133-1137.

De Pelsmacker, P., Geuens, M., & Anckaert, P. (2002). Media context and advertising effectiveness: The role of context appreciation and context/ad similarity. *Journal of Advertising*, *31*(2), 49-61.

Dean Jr., J. W., & Brass, D. J. (1985). Social interaction and the perception of job characteristics in an organization. *Human Relations*, *38*(6), 571-582.

Dewan, R., Jing, B., & Seidmann, A. (2003). Product customization and price competition on the Internet. *Management Science*, *49*(8), 1055-1070.

Dietrich, J. R., Kachelmeier, S. J., Kleinmuntz, D. N., & Linsmeier, T. J. (2001). Market efficiency, bounded rationality, and supplemental business reporting disclosures. *Journal of Accounting Research*, *39*(2), 243-268.

Dutta, P. K., Lach, S., & Rustichini, A. (2005). Better late than early: Vertical differentiation in the adoption of a new technology. *Journal of Economics & Management Strategy*, *4*(4), 563-589.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological review*, *95*(2), 256-273.

Elliott, M. T., & Speck, P. S. (1998). Consumer perceptions of advertising clutter and its impact across various media. *Journal of Advertising Research*, *38*(1), 29-42.

Ellison, G., & Fudenberg, D. (1993). Rules of thumb for social learning. *Journal of Political Economy*,*101*(4), 612-643.

Feltham, T. S., & Arnold, S. J. (1994). Program involvement and ad/program consistency as moderators of program context effects. *Journal of Consumer Psychology*, *3*(1), 51-77.

Festinger, L. (1954). A theory of social comparison processes. *Human relations*, *7*(2), 117-140.

Friestad, M., & Wright, P. (1995). Persuasion knowledge: Lay people's and researchers' beliefs about the psychology of advertising. *Journal of consumer research*, *22*(6), 62-74.

Fu, Q., & Lu, J. (2006). The optimal multi-stage contest. *Economic Theory*, *51*(2), 1-32.

Fudenberg, D., & Tirole, J. (1985). Preemption and rent equalization in the adoption of new technology. *The Review of Economic Studies*, *52*(3), 383-401.

Gallagher, K., Foster, K. D., & Parsons, J. (2001). The medium is not the message: advertising effectiveness and content evaluation in print and on the web. *Journal of Advertising Research, 41*(4), 57-70.

Ghose, A., & Han, S. P. (2010). An empirical analysis of user content generation and usage behavior in mobile media. *Working paper, Stern School of Business, New York University,* New York, NY.

Glazer, A., & Hassin, R. (1988). Optimal contests. *Economic Inquiry*, *26*(1), 133-143.

Godes, D., & Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. *Marketing Science*, *23*(4), 545-560.

Goldenberg, J., Han, S., Lehmann, D., & Hong, J. (2009). The role of hubs in the adoption processes. *Journal of Marketing*, *73*(2), 1-13.

Goodstein, R. C. (1993). Category-based applications and extensions in advertising: motivating more extensive ad processing. *Journal of Consumer Research*, *20*(1), 87-99.

Gradstein, M., & Konrad, K. A. (1999). Orchestrating rent seeking contests. *The Economic Journal*, *109*(458), 536-545.

Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, *78*(6), 1360-1380.

Greene, W. H. (1993). *Econometric Analysis (2nd ed.)*. New York: Macmillan.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of Economics*, *10*(1) 92–116.

Griliches, Z. (1998). Issues in assessing the contribution of research and development to productivity growth. In *R&D and Productivity: The Econometric Evidence* (pp. 17-45). University of Chicago Press.

Guillard, H. (2005): "Qu'est-ce que le web 2.0?" Retrieved Online. http://www.internetactu.net/

Ha, L., & Litman, B. R. (1997). Does advertising clutter have diminishing and negative returns? *Journal of Advertising*, *26*(1),31-42.

Hand, J. R. (1990). A test of the extended functional fixation hypothesis. *Accounting Review*, *65*(4), 740-763.

Hartman, R. L., & Johnson, J. D. (2006). Social contagion and multiplexity communication networks as predictors of commitment and role ambiguity. *Human Communication Research*, *15*(4), 523-548.

Hasker, K., Gonzalez, R., & Sickles, R. C. (2004). An analysis of strategic behavior and consumer surplus in eBay auctions. *Working paper, Dept. Economics, Rice Univ, Texas.*

Hauser, J. R., & Shugan, S. M. (1983). Defensive marketing strategies. *Marketing Science*, *2*(4), 319-360.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the Econometric Society*, *46*(6), 1251-1271.

Hausman, J. A., & Taylor, W. E. (1981). Panel data and unobservable individual effects. *Econometrica: Journal of the Econometric Society*, *49*(6), 1377-1398.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, *47*(1), 153-161.

Highsmith, J. A. (1978). Solving design problems more effectively. *MIS Quarterly*, *2*(4), 23-30.

Hirst, D. E., & Hopkins, P. E. (1998). Comprehensive income reporting and analysts' valuation judgments. *Journal of Accounting Research*, *36*, 47-75.

Hoffmann-Burchardi, U. (2001). Clustering of initial public offerings, information revelation and underpricing. *European Economic Review*, *45*(2), 353-383.

Hopkins, P. E., Houston, R. W., & Peters, M. F. (2000). Purchase, pooling, and equity analysts' valuation judgments. *The Accounting Review*, *75*(3), 257-281.

Hoppe, H. C. (2000). Second-mover advantages in the strategic adoption of new technology under uncertainty. *International journal of industrial organization*,*18*(2), 315-338.

Hoppe, H. C., & Lehmann-Grube, U. (2004). Second-Mover Advantages in Dynamic Quality Competition. *Journal of Economics & Management Strategy*,*10*(3), 419-433.

Huang, Y.,Singh, P., & Mukhopadhyay, T.(2012). *How to Design Crowdsourcing Contest: A Structural Empirical Analysis*. *Workshop of Information Systems and Economics (WISE) 2012*. Retrieved from http://www.krannert.purdue.edu/faculty/kkarthik/wise12/papers%5Cwise12_submission_149.pdf

Huff, L. C., & Robinson, W. T. (1994). Note: the impact of leadtime and years of competitive rivalry on pioneer market share advantages. *Management Science*,*40*(10), 1370-1377.

Ibarra, H., & Andrews, S. B. (1993). Power, social influence, and sense making: Effects of network centrality and proximity on employee perceptions. *Administrative science quarterly*, *38*(2), 277-303.

Iyengar, R., Van den Bulte, C., & Valente, T. W. (2011). Opinion leadership and social contagion in new product diffusion. *Marketing Science*, *30*(2), 195-212.

Kalra, A., & Shi, M. (2001). Designing optimal sales contests: A theoretical perspective. *Marketing Science*, *20*(2), 170-193.

Kalyanaram, G., & Wittink, D. R. (1994). Heterogeneity in entry effects between nondurable consumer product categories. *International Journal of Research in Marketing*, *11*(3), 219-231.

Kalyanaram, G., Robinson, W. T., & Urban, G. L. (1995). Order of market entry: Established empirical generalizations, emerging empirical generalizations, and future research. *Marketing Science*, *14*(3 supplement), G212-G221.

Kamien, M. I., Muller, E., & Zang, I. (1992). Research joint ventures and R&D cartels. *The American Economic Review*, *82*(5), 1293-1306.

Katz, E., & Lazarsfeld, P. F. (1955). *Personal influence: The part played by people in the flow of communications*. Glencoe, III:The Free Press.

Katz, R. (1982). The effects of group longevity on project communication and performance. *Administrative Science Quarterly*, *27*(1), 81-104.

Katz, R. (1982). The effects of group longevity on project communication and performance. *Administrative Science Quarterly*, *27*(1), 81-104.

Kelejian, H.H., & Prucha. I.R. (1998) . The method of elimination and substitution in the GMM estimation of mixed regressive, spatial autoregressive models. *Journal of Real Estate Fin. & Econ*. *17*(1), 99-121.

Keller, R. T. (1986). Predictors of the Performance of Project Groups in R&D Organizations. *Academy of Management Journal*, *29*(4), 715-726.

Kelman, H. C. (1958). Compliance, identification, and internalization: Three processes of attitude change. *The Journal of Conflict Resolution*, *2*(1), 51-60.

Kerin, R. A., Varadarajan, P. R., & Peterson, R. A. (1992). First-mover advantage: A synthesis, conceptual framework, and research propositions. *The Journal of Marketing*, *56*(4) 33-52.

Kleemann, F., Voß, G. G., & Rieder, K. (2008). Un (der) paid Innovators: The Commercial Utiliza-tion of Consumer Work through Crowdsourcing. *Science*, *Technology & Innovation Studies*, *4*(1), PP-5.

Klemperer, P. (1987). The competitiveness of markets with switching costs.*The RAND Journal of Economics*, *18*(1), 138-150.

Krouse, C. G. (1984). Brand Name as a Barrier to Entry: The Real Lemon Case. *Southern Economic Journal*, *51*(2), 495-502.

Lee, R. P., & Grewal, R. (2004). Strategic responses to new technologies and their impact on firm performance. *Journal of Marketing*, *68*(4),157-171.

Levin, R. C., Klevorick, A. K., Nelson, R. R., Winter, S. G., Gilbert, R., & Griliches, Z. (1987). Appropriating the returns from industrial research and development. *Brookings papers on economic activity*, *1987*(3), 783-831.

Lewin, K. (1935). A dynamic theory of personality. New York: McGraw-Hill.

Lieberman, M. B., & Montgomery, D. B. (1988). First-mover advantages. *Strategic management journal*, *9*(S1), 41-58.

Lieberman, M. B., & Montgomery, D. B. (1998). First-mover (dis) advantages: Retrospective and link with the resource-based view. *Strategic management journal*, *19*(12), 1111-1125.

Lilien, G. L., & Yoon, E. (1990). The timing of competitive market entry: An exploratory study of new industrial products. *Management Science, 36*(5), 568-585.

Lipe, M. G. (1998). Individual investors' risk judgments and investment decisions: The impact of accounting and market data. *Accounting, Organizations and Society*, *23*(7), 625-640.

Liu, D., Geng, X., & Whinston, A. B. (2007). Optimal design of consumer contests. *Journal of Marketing*, *71*(4), 140-155.

Locke, E. A., & Latham, G. P. (1990). A theory of goal setting and task performance. Englewood Cliffs, NJ: Prentice-Hall.

Lynch, J. G., & Zax, J. S. (2000). The Rewards to Running Prize Structure and Performance in Professional Road Racing. *Journal of Sports Economics*, *1*(4), 323-340.

Lynley, M. (2013). What Facebook’s Graph Search Means for Google. *Wall Street Journal*, Jan 15. Retrieved from http://blogs.wsj.com/digits/2013/01/15/what-facebooks-graph-search-means-for-google/

MacInnis, D. J., Moorman, C., & Jaworski, B. J. (1991). Enhancing and measuring consumers' motivation, opportunity, and ability to process brand information from ads. *The Journal of Marketing*, *55*(4), 32-53.

Maloney, M. T., & McCormick, R. E. (2000). The Response of Workers to Wages in Tournaments Evidence From Foot Races. *Journal of Sports Economics*, *1*(2), 99-123.

Mansfield, E., Schwartz, M., & Wagner, S. (1981). Imitation costs and patents: an empirical study. *The Economic Journal*, *91*(364), 907-918.

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

Mascarenhas, B. (2006). Order of entry and performance in international markets. *Strategic Management Journal*, *13*(7), 499-510.

McKendrick, J. (2012). Crowdsourcing is so 2010: here's what's next. *Smartplanet*, Aug 8. Retrieved from http://www.smartplanet.com/blog/business-brains/crowdsourcing-is-so-2010-heres-whats-next/25653

Miniard, P. W., & Cohen, J. B. (1983). Modeling personal and normative influences on behavior. *Journal of Consumer Research*, *10*(3), 169-180.

Moldovanu, B., & Sela, A. (2001). The optimal allocation of prizes in contests. *American Economic Review*, *91*(3), 542-558.

Mounier, P. (2005). SHS 2.0. *Homo-Numericus*, Sept 13. Retrieved from http://blog.homo-numericus.net/article32.html

Mueller, D. C. (1997). First-mover advantages and path dependence. *International Journal of Industrial Organization*, *15*(6), 827-850.

Mun, S. B., & Nadiri, M. I. (2002). *Information technology externalities: Empirical evidence from 42 US industries* (No. w9272). National Bureau of Economic Research, New York.

Muncy, J. A., & Hunt, S. D. (1984). Consumer involvement: definitional issues and research directions. *Advances in consumer research*, *11*(1), 193-196.

Nam, S., Manchanda, P., & Chintagunta, P. K. (2010). The effect of signal quality and contiguous word of mouth on customer acquisition for a video-on-demand service. *Marketing Science, 29*(4), 690-700.

Nekipelov, D. (2007). Entry deterrence and learning prevention on eBay. *Working Paper, Duke University, Durham, NC*.

Newman, J. L., & McCulloch, C. E. (1984). A hazard rate approach to the timing of births. *Econometrica: journal of the Econometric Society*,*52*(4) 939-961.

Nicovich, S. G. (2005). The effect of involvement on ad judgment in a video game environment: The mediating role of presence. *Journal of Interactive Advertising*, *6*(1), 38-51.

Ockenfels, A., & Roth, A. E. (2002). The timing of bids in internet auctions: Market design, bidder behavior, and artificial agents. *AI magazine*, *23*(3), 79.

Ockenfels, A., & Roth, A. E. (2006). Late and multiple bidding in second price Internet auctions: Theory and evidence concerning different rules for ending an auction. *Games and Economic Behavior*, *55*(2), 297-320.

Oestreicher-Singer, G., & Sundararajan, A. (2012). The Visible Hand? Demand Effects of Recommendation Networks in Electronic Markets. *Management Science*, *58*(11), 1963-1981.

Parry, M., & Bass, F. M. (1990). When to lead or follow? It depends. *Marketing Letters*, *1*(3), 187-198.

Rao, R. C., & Rutenberg, D. P. (1979). Preempting an alert rival: Strategic timing of the first plant by analysis of sophisticated rivalry. *The Bell Journal of Economics*, *10*(2), 412-428.

Reinganum, J. F. (1981). Dynamic Games of Innovation. *Journal of Economic Theory*, *25*(1), 21-41.

Rice, R. E., & Aydin, C. (1991). Attitudes toward new organizational technology: Network proximity as a mechanism for social information processing. *Administrative science quarterly*, *36*(2), 219-244.

Robinson, W. (1988). Sources of Market Pioneer Advantages: The Case of Industrial Goods Industries. *Journal of Marketing Research*, *25*(2), 87-94.

Robinson, W. T., & Fornell, C. (1985). Sources of market pioneer advantages in consumer goods industries. *Journal of Marketing Research*,*22*(3),305-317.

Robinson, W. T., Kalyanaram, G., & Urban, G. L. (1994). First-mover advantages from pioneering new markets: A survey of empirical evidence. *Review of Industrial Organization*, *9*(1), 1-23.

Rodgers,Z. (2011). Lego Takes Crowdsourcing up a Notch. *ClickZ*, Jun 20. Retrieved from http://www.clickz.com/clickz/news/2080033/lego-takes-crowdsourcing-notch

Rogers, E. M. (1995). *Diffusion of innovations*. New York, NY: Simon and Schuster.

Romer, P. M. (1986). Increasing returns and long-run growth. *The Journal of Political Economy*, *94*(5),1002-1037.

Romer, P. M. (1994). The origins of endogenous growth. *The Journal of Economic Perspectives*, *8*(1), 3-22.

Roth, A. E., & Ockenfels, A. (2002). Last-minute bidding and the rules for ending second-price auctions: Evidence from eBay and Amazon auctions on the Internet. *The American Economic Review*, *92*(4), 1093-1103.

Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *Quarterly Journal of Economics,116*(2), 681-704.

Salancik, G. R., & Pfeffer, J. (1978). A social information processing approach to job attitudes and task design. *Administrative science quarterly*, *23*(2), 224-253.

Schindler, J. (2003). Late bidding on the internet. *Working Paper, Vienna University of Economics and Business Administration, Austria*.

Seshadri, S., & Shapira, Z. (2003). The flow of ideas and timing of evaluation as determinants of knowledge creation. *Industrial and Corporate Change*, *12*(5), 1099-1124.

Shankar, V., Carpenter, G. S., & Krishnamurthi, L. (1998). Late mover advantage: How innovative late entrants outsell pioneers. *Journal of Marketing research*, *35*(1), 54-70.

Sheremeta, R. M. (2011). Contest design: An experimental investigation. *Economic Inquiry*, *49*(2), 573-590.

Silverman, R. (2012). Big Firms Try Crowdsourcing. *Wall Street Journal*, Jan 17. Retrieved from http://online.wsj.com/article/SB10001424052970204409004577157493201863200.html

Smiley, R. H., & Ravid, S. A. (1983). The importance of being first: learning price and strategy. *The Quarterly Journal of Economics*, *98*(2), 353-362.

Snijders, T.A.B., Steglich, C.E.G., Schweinberger, M. (2007). Modeling the co-evolution of networks and behavior. In: van Montfort, K., Oud, H., Satorra, A. (Eds.), *Longitudinal models in the behavioral and related sciences*. Lawrence Erlbaum, Mahwah, NJ, pp. 41–71

Speck, P. S., & Elliott, M. T. (1997). Predictors of advertising avoidance in print and broadcast media. *Journal of Advertising*, *26*(3), 61-76.

Spors, K. (2006). How are We Doing?. *Wall Street Journal*, Nov 13, R9. Retrieved from http://online.wsj.com/article/SB116309089183218694.html

Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica. 65*(3), 557–586.

Stern, B. B. (1994). A revised communication model for advertising: Multiple dimensions of the source, the message, and the recipient. *Journal of Advertising*, *23*(2), 5-15.

Stern, B. B. (1997). Advertising intimacy: relationship marketing and the services consumer. *Journal of Advertising*, *26*(4), 7-19.

Sullivan. M. W. (1992).Brand Extensions: When to Use Them. *Management Science. 38* (6). 793-806.

Terwiesch, C., & Ulrich, K. (2009). *Innovation tournaments: Creating and selecting exceptional opportunities*. Harvard Business Press.

Terwiesch, C., & Xu, Y. (2008). Innovation contests, open innovation, and multiagent problem solving. *Management Science, 54*(9), 1529-1543.

Tjosvold, D., & McNeely, L. T. (1988). Innovation through communication in an educational bureaucracy. *Communication Research*, *15*(5), 568-581.

Trusov, M., Bucklin, R. E., & Pauwels, K. (2009). Effects of Word-of-Mouth Versus Traditional Marketing: Findings from an Internet Social Networking Site. *Journal of Marketing*, *73*(9), 90-102.

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

Urban, G. L., & Star, S. H. (1991). Advanced marketing strategy: Phenomena. *Analysis, Decisions*. Prentice Hall.

Uzzi, B. (1999). Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing. *American sociological review*, *64*(8), 481-505.

Vadovic,R. (2009). Early, Late, and Multiple Bidding in Internet Auctions. *Working Paper, Centro de Investigacion Economica, ITAM.*

Van den Bulte, C., & Lilien, G. L. (2001). Medical Innovation Revisited: Social Contagion versus Marketing Effort1. *American Journal of Sociology*, *106*(5), 1409-1435.

Van der Lans, R., Van Bruggen, G., Eliashberg, J., & Wierenga, B. (2010). A viral branching model for predicting the spread of electronic word of mouth.*Marketing Science*, *29*(2), 348-365.

Vargo, S. L., & Lusch, R. F. (2004). The four service marketing myths. *Journal of Service Research*, *6*(4), 324–335.

Villanueva, J., Yoo, S., & Hanssens, D. M. (2008). The impact of marketing-induced versus word-of-mouth customer acquisition on customer equity growth. *Journal of Marketing Research*, *45*(1), 48-59.

Villas-Boas, J. M., & Winer, R. S. (1999). Endogeneity in brand choice models. *Management Science*, *45*(10), 1324-1338.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and instruction*, *4*(1), 45-69.

Wang, A. (2006). Advertising engagement: A driver of message involvement on message effects. *Journal of Advertising Research-New York*, *46*(4), 355.

Watts, D. J., & Dodds, P. S. (2007). Influentials, networks, and public opinion formation. *Journal of consumer research*, *34*(4), 441-458.

Weiner, B. (Ed.). (1974).*Achievement motivation and attribution theory*. Morristown,N.J.: General Learning Press.

Wellman, B. (1983). Network analysis: Some basic principles. *Sociological theory*, *1(*1), 155-200.

Wooldridge, J. M. (2001). Econometric analysis of cross section and panel data. MIT press.

Wooten, J., & Ulrich, K. (2011). Idea Generation and the Role of Feedback: Evidence from Field Experiments with Innovation Tournaments. *Available at SSRN 1838733*.

Yang, Y., Chen, P. Y., & Banker, R. (2011). *Winner Determination of Open Innovation Contests in Online Markets*. *ICIS 2011 Proceedings* (pp. 1-16). Retrieved from http://aisel.aisnet.org/icis2011/proceedings/ebusiness/16

Zahra, S. A., Nash, S., & Bickford, D. J. (1995). Transforming technological pioneering into competitive advantage. *The Academy of Management Executive*,*9*(1), 17-31.

Zajonc, R. B., & Brickman, P. (1969). Expectancy and feedback as independent factors in task performance. *Journal of Personality and Social Psychology*, *11*(2), 148-156.

Zhang, J., & Huang, L. (2011). New Product Diffusion: A Dual Word-Of-Mouth Perspective. *PACIS 2011 Proceedings*. Paper 222. Retrieved from http://aisel.aisnet.org/pacis2011/222