Understanding and Leveraging Crowd Development in Crowdsourcing

by

Zhongzhi Liu

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2017 by the
Graduate Supervisory Committee:

Thomas Kull, Chair
Kevin Dooley
Adegoke Oke

ARIZONA STATE UNIVERSITY
August 2017

ABSTRACT

Although many examples have demonstrated the great potential of a human crowd as an alternative supplier in creative problem-solving, empirical evidence shows that the performance of a crowd varies greatly even under similar situations. This phenomenon is defined as the performance variation puzzle in crowdsourcing. Cases suggest that crowd development influences crowd performance, but little research in crowdsourcing literature has examined the issue of crowd development.

This dissertation studies how crowd development impacts crowd performance in crowdsourcing. It first develops a double-funnel framework on crowd development. Based on structural thinking and four crowd development examples, this conceptual framework elaborates different steps of crowd development in crowdsourcing. By doing so, this dissertation partitions a crowd development process into two sub-processes that map out two empirical studies.

The first study examines the relationships between elements of event design and crowd emergence and the mechanisms underlying these relationships. This study takes a strong inference approach and tests whether tournament theory is more applicable than diffusion theory in explaining the relationships between elements of event design and crowd emergence in crowdsourcing. Results show that that neither diffusion theory nor tournament theory fully explains these relationships. This dissertation proposes a *contatition* (i.e., contagious competition) perspective that incorporates both elements of these two theories to get a full understanding of crowd emergence in crowdsourcing.

The second empirical study draws from innovation search literature and tournament theory to address the performance variation puzzle through analyzing crowd attributes. Results show that neither innovation search perspective nor tournament theory fully explains the relationships between crowd attributes and crowd performance. Based on the research findings, this dissertation discovers a competition-search mechanism beneath the variation of crowd performance in crowdsourcing.

This dissertation makes a few significant contributions. It maps out an emergent process for the first time in supply chain literature, discovers the mechanisms underlying the performance implication of a crowd-development process, and answers a research call on crowd engagement and utilization. Managerial implications for crowd management are also discussed.

DEDICATION

I dedicate this dissertation to my son, Lucas Liu. He is my inspiration for being the best person I can possibly be.

ACKNOWLEDGMENTS

This dissertation would not have come into existence without the encouragement, support, and dedication of my family. My parents and my wife were always there to support me in many ways throughout my doctoral years. I sincerely appreciate their love and support.

The most prominent person that has impacted the strength and quality of this dissertation is my dissertation chair, Professor Thomas Kull. Not only did his steadfast guidance and wisdom make this research valuable, but his insightful perspectives on work, family, and life have changed me forever. In addition, the other members of my dissertation committee, Professor Keven Dooley and Professor Adegoke Oke, have each made a difference in this dissertation and in my career. They have from the beginning encouraged me to explore new topics and to focus on what really matters.

I am also grateful for many other faculty at Arizona State University who made a difference on my early academic career. Professor Thomas Choi taught me how to look critically at my own work through “the poet’s eye” so that it attains the highest standards of our discipline. Professor Eddie Davila set foundations for my teaching philosophy and style. Professor John Fowler and Professor Dale Rogers were very generous in providing financial support for my PhD study. Professor Leona Aiken and Professor Roger Millsap from the Psychology Department at Arizona State University were instrumental in developing my methodological skills and scientific paradigm.

Finally, I thank the people who were doctoral students with me in the Supply Chain Department at Arizona State University. I especially acknowledge the people in my entering year: Yousef Abdulsalam, Sangho Chae, and Zac Rogers. Each of them supported and challenged me along the way. They made the character of my doctoral studies as special and unique as they are.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure | Page

Chapter 1: Introduction

"When we harness the power of the crowd, we can innovate and iterate on products at a pace many manufacturers didn't think was possible."

Jay Rogers, Co-founder and CEO, Local Motors

Background

Statistics show that many best global brands (e.g., IBM, Cisco, GE, and Dell) are actively applying crowdsourcing to tap into external creative resources in their innovation processes (King & Lakhani, 2013; Roth, Pétavy, & Céré, 2015). Crowdsourcing is defined as a practice of outsourcing a task to a crowd rather than to a designated contract supplier in the form of an open call (Afuah & Tucci, 2012; Howe, 2006). The term "a crowd" in the crowdsourcing definition refers to a collective of suppliers who are nested within a virtual network and share a common focus to solve crowdsourced tasks (e.g., product design) (Afuah & Tucci, 2012; Forsyth, 2009). For instance, Airbus intended to develop a drone that could be used in the last-mile humanitarian logistics in 2016. Instead of relying on in-house development or contract outsourcing, Airbus teamed with Local Motors[1] and crowdsourced this task to Local Motors' community suppliers by creating the Airbus Cargo Drone Challenge. Within two months, Airbus acquired a total of 425 designs from Local Motors' community suppliers (Prassler, 2016).

[1] Local Motors is an American auto company based in Chandler, Arizona that designs and builds customized vehicles through co-creation with community members (Gerth, Burnap, & Papalambros, 2012; Randall, Ramaswamy, & Ozcan, 2013).

Cases like the Airbus Cargo Drone Challenge demonstrate the great potential of a crowd in generating solutions in innovation processes. As many companies adopt crowdsourcing to solve their innovation-related problems, the human crowd has emerged as a new type of supplier that specializes in providing knowledge in innovation processes (Boudreau & Lakhani, 2013). Empirical evidence from the pharmaceutical industry shows that the application of a crowd in the R&D domain can be more than 20 times less expensive than regular R&D paths (e.g., in-house development or contract outsourcing) (Lakhani, Jeppesen, Lohse, & Panetta, 2007; Raynor & Panetta, 2008). Statistics from TopCoder[2] show that the crowd can often provide Topcoder's clients with development work that is comparable in quality to what they would get by more traditional means but at little as 25 percent of the cost (Johns, Laubscher, & Malone, 2011). As such, some analysts and scholars anticipate that the human crowd has potential to reshape established business processes, redraw organizational boundaries, and change global labor markets, thus profoundly disrupting the supply network in the near future (Howe, 2008; Kaganer, Carmel, Hirschheim, & Olsen, 2013).

Research Phenomenon

Although the human crowd has a huge potential in creative problem-solving, not every crowd is always creative and productive (Euchner, 2010; King & Lakhani, 2013). Empirical evidence on crowdsourcing from Topcoder shows that the performance of a crowd varies even under similar situation. In this dissertation, crowd performance refers to the quantitative outcomes of a crowd in crowdsourcing (e.g., crowd productivity

[2] A company that administers crowdsourcing contests in computer programming (Archak, 2010).

defined as the numbers of solutions generated by a crowd) (Cohen & Bailey, 1997; Horwitz & Horwitz, 2007). For instance, Topcoder hosted two programing contests on data search web design challenges in August 2014. These two events had the exact same payment size (i.e., $2,250) and payment structure (first place: $1500; second place: $500; third place: $250) (Topcoder, 2014a, 2014b). The nature of the tasks and the event lengths were similar. However, the crowd performance between these two cases were significant different: One event had six submissions and the other had no submission (Topcoder, 2014a, 2014b). Scholars in the operations and supply chain management literature have also identified similar cases on performance variation in crowdsourcing (Billington & Davidson, 2013; Sloane, 2012; Tang et al., 2011). The phenomenon that the performance of a crowd in crowdsourcing varies even under similar situations is termed as the performance variation puzzle in this dissertation.

Motivations

Theories that scholars use to explain firm performance variation, such as resource-based view (Wernerfelt, 1984), knowledge-based theory of the firm (Grant, 1996), and relational view (Dyer & Singh, 1998), are at firm level or network level. These theoretical lenses are out of scope to explain the performance variation of a crowd in crowdsourcing because a crowd in crowdsourcing does not have a formal organizational structure (Forsyth, 2009; Reicher, 2001). Suppliers in a particular crowd are loosely connected and geographically distributed all over the world with limited information visibility. Moreover, because the performance of a crowd varies even when firms use the same incentives for similar tasks (Tang et al., 2011), knowledge from motivation literature (e.g., Ryan & Deci, 2000) cannot easily explain performance variation puzzle in crowdsourcing either.

Current crowdsourcing literature primarily focuses on the best practices (Guinan, Boudreau, & Lakhani, 2013), the conditions facilitating crowdsourcing (Afuah & Tucci, 2012), individuals' motivations for participation in crowdsourcing (Brabham, 2010, 2012), incentive design and its influence on individual performance (Boudreau, Lacetera, & Lakhani, 2011; Liu, Yang, Adamic, & Chen, 2014), and winners' characteristics in crowdsourcing (Bockstedt, Druehl, & Mishra, 2015; Jeppesen & Lakhani, 2010). A few studies on individual performance suggest that crowd attributes have an impact on crowd performance (Boudreau et al., 2011; Jeppesen & Lakhani, 2010). However, these studies provide contradictory findings on how crowd attributes (e.g., crowd size and crowd diversity) relate to crowd performance. For instance, Boudreau, Lacetera, and Lakhani (2011) found that an increase in the crowd size has a negative influence on solvers' effort and crowd performance due to a reduced chance of winning, but Bockstedt, Druehl, and Mishra (2015) identified a positive association between crowd size and crowd performance.

The existence of contradictory findings indicates an insufficient understanding of the crowd performance issue. Our literature review shows that little research in crowdsourcing literature examines crowd-level performance and explores the factors that can explain the performance variation puzzle in crowdsourcing. A lack of research on this puzzle creates confusion about crowdsourcing and causes scholars to question the application of a crowd in an innovation process (Euchner, 2010; Simula, 2013). Many executives and supply chain managers are thus unable to develop strategies or are hesitant to allocate resources to crowdsourcing, resulting in missed opportunities for new competitive advantages that might come from engaging crowds (Prpić, Shukla, Kietzmann, & McCarthy, 2015). As such, scholars call for research that can help organizations better manage, utilize, and organize both internal and external crowds when innovating (Felin, Lakhani, & Tushman, 2015).

Statement of the Problem

A few case studies on crowdsourcing suggest that the process through which firms develop a crowd, defined as crowd development, influences the operational processes of a crowd which, in turn, have a potential impact on crowd performance (Guinan et al., 2013; King & Lakhani, 2013). However, relatively little research describes how to develop a crowd more effectively and efficiently, despite a growing popularity and reliance on the human crowd in practice (Wooten & Ulrich, 2017). We thus do not know how crowd development works and how it explains the performance variation puzzle in crowdsourcing. Therefore, the grand research question we study in this dissertation is:

How does a crowd development impact the performance of a crowd in crowdsourcing?

Dissertation Design

To answer the above grand research question, we divide this dissertation into three closely related parts that include one conceptual framework development and two empirical tests. From structural thinking perspective (Molm, 1990; Ralston, Blackhurst, Cantor, & Crum, 2015), we first develop a double-funnel model based on four crowd-development examples to address the deficiency of no framework on crowd development in existing crowdsourcing literature. We use this model to describe the detailed process of crowd development, which includes crowd initiation, crowd formation, crowd realization, and crowd evaluation. This framework partitions a crowd-development process into crowd emergence and crowd evaluation, which maps out two empirical studies that examine the influence of event design on the emergence of a crowd and the performance implications of crowd attributes.

The first empirical study in this dissertation examines the relationships between elements of event design and crowd emergence and the mechanisms underlying these relationships. In this study, crowd emergence is defined as the arising of unexpected growth rate and crowd size in a crowd development process (Dooley & Corman, 2002; Holland, 2000). One rationale behind this study is current crowdsourcing literature lacks studies examining the influence of event design on crowd emergence. Both scholars and professionals thus have no reported knowledge on how to manage crowd emergence in crowdsourcing. Another reason is that scholars suggest two mechanisms to explain crowd emergence: competition mechanism from tournament theory (Connelly, Tihanyi, Crook, & Gangloff, 2014; Lazear & Rosen, 1981) and contagion mechanism from diffusion theory (Rogers, 2010; Strang & Soule, 1998). These two mechanisms offer different predictions on the relationships between elements of event design and crowd emergence. We thus took a strong inference approach (Davis, 2006; Platt, 1964) and developed alternative hypotheses on the relationships between elements of event design and crowd emergence. Our regression analysis based on 734 observations shows that neither competition mechanism based on tournament theory nor contagion mechanism based on diffusion theory fully explains crowd emergence in crowdsourcing. Based on our empirical findings, we propose a *contatition* (i.e., contagious competition) perspective that incorporates both elements of these two theories to get a full understanding of crowd emergence in crowdsourcing.

The second empirical study is designed to address the performance variation puzzle by analyzing the performance implications of crowd attributes (i.e., crowd size and crowd diversity). In this study, we attempt to explain some contradictory findings related to the performance implication of crowd attributes (Bockstedt et al., 2015; Boudreau et al., 2011), and resolve the confusion about the mechanisms underlying the relationships between crowd attributes and crowd performance. Some scholars argue that a competition mechanism based on tournament theory explains the relationship between crowd attributes and crowd performance (Boudreau et al., 2011), while others suggest that a search process based on innovation search literature explains the performance implications of crowd attributes (Afuah & Tucci, 2012; Jeppesen & Lakhani, 2010). These two mechanisms offer different explanations for how crowd attributes relate to crowd performance, providing us another chance to develop alternative hypotheses. We test our theory by using secondary data collected from a crowdsourcing platform company through web crawling. Results demonstrate that crowd attributes explain the crowd performance variation puzzle and that some relationships are not linear but quadratic, suggesting the complication of crowd performance. Our empirical findings also indicate that the competition mechanism plays a majority role in explaining the relationships between crowd attributes and crowd performance, but we need to consider the search mechanism due to the significant interactions between these two mechanisms. We thus propose a competition-search view on the performance implications of crowd attributes.

Contributions

This dissertation contributes to the current crowdsourcing literature and supply chain field in several significant ways. First, it maps out an emergent process in supply chain literature by proposing a double-funnel framework on crowd development. This is the first time in crowdsourcing and supply chain literature to describe this new process and to explore the performance implications of this process. This dissertation thus fills a void in crowdsourcing and supply chain literature. The proposed double-funnel model has significant implications for scholars and supply chain managers. For scholars, this framework advances academic understanding of supplier development from a controlled, deliberate perspective in outsourcing literature to an emergent, unsystematic perspective in crowdsourcing. It also provides a framework for scholars to explore crowd development from many other lenses like system dynamics. For managers, this model offers a holistic view on engaging with a crowd in the innovation processes through adjusting elements of event design.

Second, this dissertation uncovers the contatition mechanism that underlies the relationships between elements of event design and crowd emergence. This contatition mechanism indicates that the crowd emergence based on suppliers' interactive participation follows neither a full competition process as suggested by tournament theory (Connelly et al., 2014; Lazear & Rosen, 1981) nor a full contagion process as implied by diffusion theory (Rogers, 2010; Van den Bulte & Stremersch, 2004). Instead, it demonstrates both competition and contagion elements. This finding suggests that some suppliers (i.e., participants) influence crowd emergence through competition while others, especially senior ones with winning records, exert their influence on crowd emergence by triggering imitation within a crowd. This dissertation thus deepens our understanding on suppliers' participation behaviors in the crowd-development process. By discovering this contatition mechanism underlying crowd emergence, this dissertation also answers a research call on crowd management (Felin et al., 2015). Meanwhile, managers can better manage and engage with a human crowd in innovation processes by leveraging this contatition mechanism underlying crowd emergence. For instance, managers can take a less homogenous view towards the crowd members and keep a close eye on the influential suppliers in the crowd formation process.

Finally, our dissertation explains the performance variation puzzle by revealing the complicated relationships between crowd attributes and crowd performance and by discovering the competition-search mechanism underneath the complicated relationships between crowd attributes and crowd performance. The competition-search mechanism means that the logic linkage between crowd attributes and crowd performance includes not only the competition process driven by solvers' utility maximization but also a search process over a solution landscape. This dissertation shows that these two forces are not necessarily exclusive in explaining performance. Instead, they are complementary to each other. This finding is different from the predominant thinking in crowdsourcing literature (Afuah & Tucci, 2012) and tournament literature (Boudreau et al., 2011; Fullerton & McAfee, 1999; Lazear & Rosen, 1981). The competition-search mechanism also suggests that the crowd-level attributes have direct influence in causing the variations on crowd performance. Although many of the event design elements such as payment size and payment structure are similar, crowd performance can vary if the emergent crowd-level attributes are different.

Organization

This dissertation organizes as follows. Chapter 2 is a literature review that summarizes the current state of crowdsourcing, identifies the main deficiencies of existing literature, and explores the theoretical lenses that we can use to understand crowd development and crowd performance variation. Chapter 3 proposes a process model to elaborate the stages of crowd development and identify the constructs that are relevant to different stages of crowd development. Chapter 4 is our theory development section that includes two theory developments: One is for understanding the relationships between elements of event design and crowd emergence; the other for the performance implications of crowd attributes. Chapter 5 describes the methodology design and data collection process. Chapter 6 covers the detailed data analysis and empirical findings from our two studies. Chapter 7 is our discussion chapter that addresses the theoretical contributions and managerial implications. Chapter 8 is the conclusion section that summarizes the whole dissertation, addresses methodology-related limitations, and proposes future research directions. A publication plan is also discussed in this chapter.

Chapter 2: Background Literature

Overview of Crowdsourcing Practice

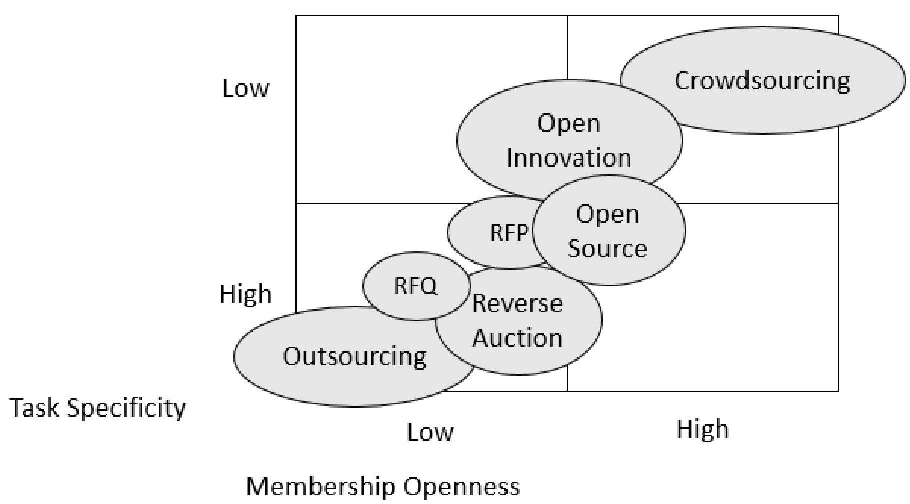
This chapter addresses the literature background of crowdsourcing. The practice of outsourcing a task to a crowd in an open call can be traced back to the Longitude Prize organized by the British government to determine the position of ships in the sea in 1714 (Economist, 2008). History is filled with examples similar to the Longitude Prize, especially in the architecture design industry. The use of architecture design contests have led to some of the most notable buildings in the world, including the Sydney Opera House, the White House, the British Houses of Parliament, and the Berlin Central Station (Afuah & Tucci, 2012). However, the research stream on crowdsourcing did not occur until the notion of crowdsourcing was introduced one decade ago (Howe, 2006). We thus review the research on crowdsourcing in this chapter after the notion was created. Through this review, we intend to identify deficiencies in crowdsourcing literature and the theoretical gaps that this dissertation can fill. We also review other relevant literature (e.g., supplier development) and the theoretical lenses (e.g., diffusion theory, tournament theory, and structural thinking) that can help us understand crowd development and develop our research framework.

Definition. The earliest references to the term “crowdsourcing” can be traced to Jeff Howe in a 2006 *Wired* magazine article to describe a web-based business practice that companies use to harness the creative solutions of a distributed network through an open-call process (Howe, 2006). According to Howe (2006), crowdsourcing represents an act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of suppliers in the form of an open call. In Howe's (2006) definition of crowdsourcing, the crucial prerequisites of crowdsourcing are outsourcing an internally performed function, the use of the open-call format, and a large network of potential suppliers, i.e., a crowd. As crowdsourcing gets more popular, scholars observe that companies or institutions crowdsource many activities that never have been performed by their employees (Billington & Davidson, 2013; Boudreau & Lakhani, 2013). Google's Lunar X Prize, a crowdsourcing competition that called for privately funded spaceflight teams to land robotic spacecraft on the moon, is a case in point (Kay, 2012). Afuah and Tucci (2012) thus redefine crowdsourcing as "the act of outsourcing a task to a 'crowd', rather than to a designed 'agent' (an organization, informal or formal team, or individual), such as a contractor, in the form of an open call" (p.355). This definition has been commonly cited in crowdsourcing literature.

Crowdsourcing as a New Outsourcing Practice. By definition, crowdsourcing falls within the domain of outsourcing. The open-call process involved in crowdsourcing makes crowdsourcing seem like other common business practices such as reverse auction, request for quotes (RFQ), or request for bidding (RFB). Because of the open call process, some scholars argue that crowdsourcing overlaps with open innovation in innovation literature (Chesbrough, 2006) and open source in computer science literature (Daniel, Agarwal, & Stewart, 2013; Roberts, Hann, & Slaughter, 2006). However, crowdsourcing differs significantly from these traditional practices in terms of task specificity and membership openness (Figure 1). Task specificity refers to the extent to which the inputs for a task are specified (Piller & Walcher, 2006). In general, tasks in crowdsourcing have a low level of specificity. For instance, Harvard Catalyst organized a crowdsourcing challenge titled "What do we not know to cure Type 1 diabetes?" (Guinan et al., 2013). Harvard Catalyst did not specify specific requirements for this challenge. Instead, participants had to formulate their own well-defined problems and/or hypotheses to advance knowledge about Type 1 diabetes research in new and promising directions. Membership openness refers to the extent of filtering in the selection process of external participants (i.e., suppliers) for a particular task (Chesbrough, 2006; Lakhani et al., 2007). In reality, crowdsourcing has a high level of membership openness because each agent (i.e., individuals, teams, and/or organizations) can self-select to participate for a particular task.

Figure 1

Uniqueness of Crowdsourcing



These differences contribute to the operation of focal buying firms in several significant ways. First, high levels of membership openness allows focal buying firms to expand organizational boundaries. Unlike traditional outsourcing, crowdsourcing does not establish an *ex ante* contract relationship between a focal buying firm and its potential suppliers. Suppliers in a crowd self-select to compete and cooperate with each other for a specific crowdsourced event. Buying firms thus can avoid the classic principal-agency and moral hazard issues associated with suppliers in outsourcing if they use a contract supplier to solve their innovation-related problems (Afuah & Tucci, 2012). Due to the high level of membership openness, crowdsourcing proves to be a cost-effective solution to innovation related problems (Johns et al., 2011; Lakhani et al., 2007). Second, low levels of task specificity in crowdsourcing facilitate focal buying firms to tap into creative resources outside their organizations in a large scale. Crowdsourcing provides a solution for firms to conduct distant search in their innovation process (Afuah & Tucci, 2012), which can help firms find the optimal solution and increase their innovation performances.

Classification of Crowdsourcing. Crowdsourcing can take the form of peer production in which self-selected suppliers work together on a particular problem, while the result is one solution or multiple solutions generated from the crowd (Afuah & Tucci, 2012). This type of crowdsourcing is termed as collaboration-based crowdsourcing (Afuah & Tucci, 2012), also called community-based crowdsourcing (Bayus, 2013), or online open collaboration (Ren, Chen, & Riedl, 2015). Wikipedia is a classic example of collaboration-based crowdsourcing in which a group of editors collaborate with each other through the internet to perform encyclopedic work (Ren et al., 2015). Another example is Dell's IdeaStorm through which Dell collects product/process improvement ideas from its cooperative online community (Bayus, 2013). This cooperative type of crowdsourcing shares similarities with phenomena such as open source in computer science literature (Daniel et al., 2013; Roberts et al., 2006) and information system literature (Dissanayake, Zhang, & Gu, 2015). Many scholars from these two research streams examine the phenomenon of crowdsourcing from different perspectives (Dissanayake et al., 2015).

Crowdsourcing can also take the form of peer competition in which each supplier self-selects to work on its own solution(s) and compete with others to provide the best solution. Only the winner(s) chosen by the focal buying firms can receive financial payment, which is always publicly announced at the beginning of a crowdsourcing event. Scholars call this competitive type of crowdsourcing competition-based crowdsourcing (Afuah & Tucci, 2012), broadcast search (Jeppesen & Lakhani, 2010), or innovation contest (Bockstedt et al., 2015). For example, Netflix crowdsourced a task of developing an algorithm to improve its movie recommendation system in 2007 in the form of an open call to the world. Anyone who could come up with an algorithm that improved Netflix's existing recommendation system by at least 10 percent could win $1 million (King & Lakhani, 2013). This competitive type of crowdsourcing shares similarities with the tournament in economics literature, such as rewarding policy and self-selected participation (Connelly et al., 2014).

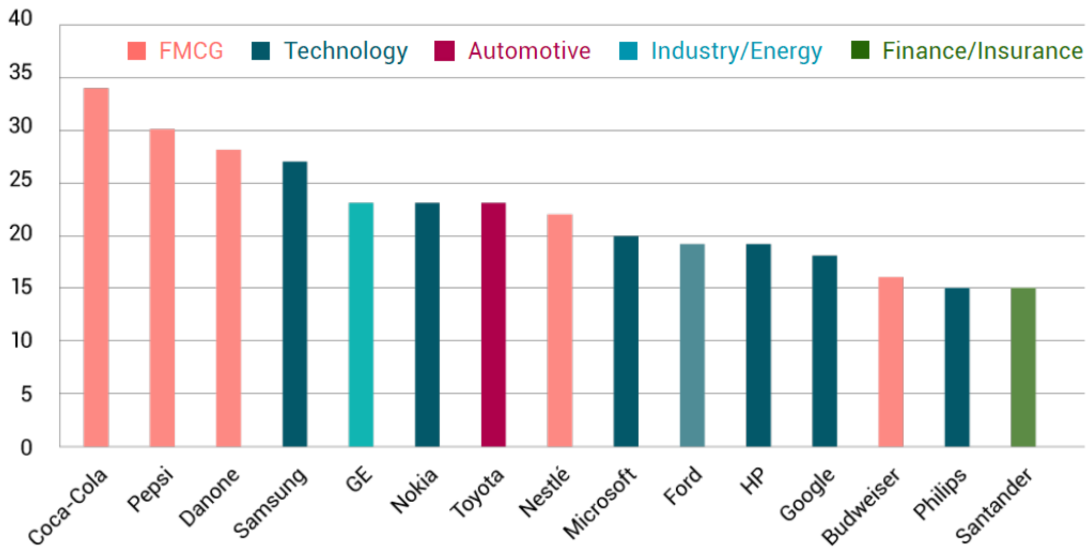
In the current crowdsourcing industry, competition-based crowdsourcing is more popular than cooperation-based crowdsourcing for a few practical reasons. First, because of the low level of information visibility and loose connection among crowd members, it is difficult for self-selected crowd members to develop high-level of interpersonal trust to function efficiently and effectively as a team in a cooperation-based crowdsourcing (Dirks, 1999; Nirwan, 2014). Another obstacle for cooperation in crowdsourcing is the potential leakage of intellectual property (King & Lakhani, 2013). Cooperation in crowdsourcing thus becomes challenging. On the other site, the emergence of many platforms (e.g., Topcoder, InnoCentive, Eyeka, and Kaggle) that specialize in organizing contests makes competition-based crowdsourcing more attractive to managers (Billington & Davidson, 2013). Also, competition is a different from cooperation, which means that competition-based crowdsourcing is a different phenomenon from cooperation-based crowdsourcing. We thus mainly focus on competition-based crowdsourcing in this dissertation and use crowdsourcing to represent competition-based crowdsourcing.

Literature Review on Crowdsourcing

A recent report on the state of crowdsourcing published by eYeka (one of the largest crowdsourcing platforms) in 2015 indicates that 85 percent of the best global brands (e.g., Coca-Cola, Pepsi, Samsung, and GE) has used crowdsourcing in the last ten years (Figure 2) (Roth et al., 2015). According to this industry report, Toyota used crowdsourcing 23 times in the last ten years to tap into the creativity of the crowd. The fact that crowdsourcing was successfully utilized in a company like Toyota, whose innovation was traditionally assumed to be fully driven by its internal employees and external tier-structured suppliers (Girotra & Netessine, 2013), demonstrates the great potential of crowdsourcing for contemporary business operations. As crowdsourcing gets more popular, scholars pay increasing attention to the issues related to the application of this new practice. Current literature on crowdsourcing can be segregated into three streams of research based on methodological approaches: qualitative stream (including conceptual thinking), empirical stream, and analytical stream.

Figure 2

The 15 Best Global Brands that Most Use Crowdsourcing Since 2004



Source: Roth, Petavy, & Cere (2015, p.8)

Qualitative Research Stream. This research stream mainly focuses on identifying the best practices through case study and conceptually understanding crowdsourcing as a solution of distant search. Jeff Howe's (2006) qualitative article on crowdsourcing in *Wired* magazine represents the start of academic research on crowdsourcing. In this article, Howe (2006) provided successful crowdsourcing applications and predicted the rise of crowdsourcing. Howe's (2006) prediction was so insightful that Howe's (2006) article has been cited 2,779 times since its publication, according to Google Scholar as of March 7, 2016. Inspired by Howe's (2006) work, scholars from strategy literature, innovation literature, and other disciplines start to qualitatively examine successful crowdsourcing cases from Fortune 500 companies, such as IBM, Cisco, and GE. A series of cases studies were published in the following years in managerial journals such as *Harvard Business Review* (HBR), *MIT Sloan Management Review*, and *California Management Review* (Bjelland & Wood, 2008; Chesbrough, 2012; Jouret, 2009). These qualitative case studies demonstrate the value of crowdsourcing as another successful mechanism for creative problem-solving beyond two traditional mechanisms (i.e., internal development and contract outsourcing) (Brabham, 2008).

As the practice of crowdsourcing becomes popular, scholars start to question when crowdsourcing might be a better mechanism for solving problems than the other two traditional mechanisms. Afuah and Tucci (2012) addressed this question in their conceptual paper that was published in *Academy of Management Review* (AMR). Afuah and Tucci (2012) conceptualized crowdsourcing as a solution for distant search. This conceptualization is consistent with the "search thinking" in the innovation literature, which argues that a problem-solving or innovation process falls a recombinant search over a solution landscape (Fleming & Sorenson, 2004; Katila & Ahuja, 2002; Laursen & Salter, 2006). In their conceptual paper, Afuah and Tucci (2012) explored conditions that could increase the likelihood of crowdsourcing by considering the characteristics of the focal problem, the knowledge required for solution, the crowd, and the solutions to be evaluated, as well as the pervasiveness of information technology. Afuah and Tucci's (2012) conceptual thinking was so thought-provoking that it was awarded the 2012 AMR Best Paper Award. This paper represents the most advanced conceptual thinking in crowdsourcing literature.

Another interesting phenomenon in the qualitative research stream is the emergence of many intermedia (e.g., InnoCentive, Topcoder, and eYeka) that specialize in organizing crowdsourcing events for focal buying companies and institutions (Billington & Davidson, 2013). These intermedia act as bridges between focal buying firms and potential suppliers from a crowd (i.e., individuals, teams, or organizations). Scholars thus start to qualitatively examine how managers can better leverage crowdsourcing and manage the human crowd by using these crowdsourcing intermedia. Many qualitative case studies on crowdsourcing intermedia appear in *HBR and MIT Sloan Review* (Boudreau & Lakhani, 2013; Guinan et al., 2013; Kaganer et al., 2013). Although these case studies are quite useful to demonstrate the power of a crowd for decision-makers, no specific framework on crowd management has been developed yet. Besides, current crowdsourcing literature still witnesses quite a few crowdsourcing failures and many unproductive or even destructive crowds (Harris, 2015; Rosenfeld, 2012). Scholars are thus calling for research that can help organizations better manage, utilize, and organize both internal and external crowds when innovating (Felin et al., 2015).

Empirical Research Stream. In the empirical research stream, the level of analysis is primarily at an individual level instead of the crowd level. Scholars in this stream use different methods (e.g., filed study, secondary data, and experiment) to examine issues related to individual participants in a crowd. These issues include but are not limited to participation motivations (Brabham, 2010, 2012), factors that influence individuals' performance (Bayus, 2013; Bockstedt et al., 2015; Jeppesen & Lakhani, 2010), individuals' expected fairness as well as its effects in crowdsourcing (Franke, Keinz, & Klausberger, 2013), individuals' attention allocation in crowdsourcing contests (Haas, Criscuolo, & George, 2015), and incentive design as well as its effect on individuals' participation (Boudreau et al., 2011; Liu et al., 2014).

Findings related to individual motivations in crowdsourcing are helpful to understand suppliers' self-selected participation in a crowd-development process. Through these empirical findings, we know that supplier's participation behaviors in crowdsourcing are driven by economic and social reasons. For instance, Brabham (2010, 2012) identified that individuals are motivated by both extrinsic motivations (e.g., financial return, reputation, and status) and intrinsic motivations (e.g., fun, learning, a sense of satisfaction and accomplishment) to participate in crowdsourcing events. In their study on how individuals allocate attention for crowdsourced problems on-line, Haas, Criscuolo, and George (2015) identified that individuals are more likely to participate in solving problems that closely match their expertise, but that their participation decisions are influenced by problem characteristics (e.g., length, breadth, and novelty). Liu and her colleagues (2014) found that a higher reward could induce significantly more submissions and higher quality submissions. They also found that high-quality participants were less likely to participate in crowdsourcing tasks where a high-quality solution had been posted as a benchmarking, suggesting that competition within a crowd could have a negative influence on supplier's participation behaviors (Liu et al., 2014). All the findings related to suppliers' participation in crowdsourcing suggest that managers can exert their impact on the crowd-development process indirectly through adjusting the precedents of suppliers' self-selected participation (e.g., payment size). However, these findings do not provide specific implications for managers to exert their influence because no research in the empirical stream has examined the mechanisms underlying crowd development.

Meanwhile, the empirical findings related to individual performance are inconsistent and have conflicting implications for crowd-level performance. The inconsistent findings are mainly related to the performance implication of crowd size and crowd diversity. For instance, some scholars conceptualize crowdsourcing as a distant search for solutions over a rugged landscape (Afuah & Tucci, 2012; Boudreau, Guinan, Lakhani, & Riedl, 2016). This conceptualization runs parallel with the search view in the innovation literature, which claims that the progress of science follows a recombinant search process through either recombining existing elements or combining new elements (Fleming, 2001; Fleming & Sorenson, 2004). Following these two lines of thinking, scholars argue that an increase in crowd size (i.e., the number of participants) allows firms to search in a wide landscape and thus acquire more solutions (Laursen & Salter, 2006). Through an empirical analysis based on secondary data from Logomyway.com[3], Bockstedt, Druehl, and Mirsha (2015) found a positive association between the number of participants (i.e., crowd size) and the number of submissions per task (i.e., crowd productivity). From a tournament theory perspective, however, Boudreau, Lacetera, and Lakhani (2011) found that an increase in the number of constants leads to poor performance outcomes. This is because increasing the number of competitors a contest reduces the likelihood of winning for any one competitor, thereby reducing contestants’

[3] A popular competition-based crowdsourcing platform that matches graphic designers with organizations in need of new logos (Bockstedt et al., 2015)

motivation to invest or exert effort and then lowering overall performance (Che & Gale, 2003; Fullerton & McAfee, 1999).

As for the performance implication of diversity, Jeppesen and Lakhani (2010) identified a "marginality effect" in tournament-based crowdsourcing which means individuals who are technically and socially far away from the focal buying firms are more likely to offer creative solutions and thus become the winners in competitive crowdsourcing events. This finding provides strong support for firms to conduct distant search for their innovation-related problems (Afuah & Tucci, 2012). Franke, Poetz, and Schreier (2014) also found that integrating problem solvers in ideation crowdsourcing could increase the chance to generate more novel solutions. However, in their study on individuals' problem-solving effort and success in innovation contests, Bockstedt, Druehl, and Mishra (2015) found that individuals with greater similarity to focal buying firms in terms of cultural background and economic wealth are more likely to be winners. Scholars call this phenomenon as the "homophily effect" which refers to the propensity of individuals to associate with other individuals who have similar social, cultural, economic, and/or demographic characteristics (McPherson, Smith-Lovin, & Cook, 2001; Milliken & Martins, 1996).

These above contradictory findings provide conflicting implications for managers on how to develop a crowd for a particular crowdsourcing contest, which demonstrates the necessity for further exploration on the issue of winners' characteristics and winner selection in tournament-based crowdsourcing. The search view in the innovation literature and the competition view in tournament theory seems to have conflicting implications in crowdsourcing. We found no research in the current crowdsourcing literature that has tested the relative power of these two views in explaining the relationships between crowd attributes and crowd performance. This literature gap offers a great opportunity for this dissertation to make a contribution in crowdsourcing literature.

Analytical Research Stream. Based on tournament theory in economics literature (Che & Gale, 2003; Lazear & Rosen, 1981), many scholars from operations management (OM) and operation research (OR) apply an analytical modeling approach to study optimal design of innovation contests (i.e., competition-based crowdsourcing) (Ales, Cho, & Körpeoğlu, 2017; Terwiesch & Xu, 2008) and the behavior of contestants (Boudreau, Lakhani, & Menietti, 2016; Chen, Ham, & Lim, 2011). This research stream assumes agents (i.e., suppliers) are rational. Agents' participation behavior and investment (i.e., effort) are driven by their utility maximization functions. Findings from this research stream provide a full economic view on suppliers' behaviors in crowdsourcing and offers some support for us to understand the performance implications of crowd size.

The dominant view on the contest design is that having many people work on an innovation problem simultaneously will lead to a lower equilibrium effort for each participant (Che & Gale, 2003; Fullerton & McAfee, 1999). This result is undesirable from the perspective of focal buying firms and suggests that firms should limit the number of participants. However, Terwiesch and Xu (2008) found that buying firms can benefit from a large crowd because they obtain a more diverse set of solutions, which mitigates and outweighs the effect of participants' underinvestment in effort. This conclusion supports the distant search view in the empirical steam which argues that an increase in the crowd size is positively associated with crowd productivity. Terwiesch and Xu (2008) also found that the inefficiency of the innovation contest caused by participants' underinvestment can be reduced by changing the incentive structure from a fixed-price to a performance-contingent award.

From a behavioral perspective, Boudreau, Lakhani, and Menietti (2016) found that the performance response to added contestants varies non-monotonically across contestants of different abilities: Most participants respond negatively, whereas the highest skilled contestants respond positively. Chen, Ham, and Lim (2011) examined how a change in the prize structure affects the effort of contestants in a multi-person tournament where contestants have different initial endowments. In particular, Chen, Ham, and Lim (2011) found that when the number of prizes increases from one to two, both high-level initial endowments and low-level initial endowment participants increase their efforts. This is because high-level initial endowments might perceive psychological losses from losing while low-level initial endowments think about psychological gains from winning (Chen et al., 2011). This finding shows the importance of payment structure on suppliers' behaviors in a contest. We need to control for the number of payments in our empirical test analysis.

Summary of Literature Review. The above literature review on three streams of research on crowdsourcing identifies a few gaps that hinder academic development of crowdsourcing in the supply chain field. First, current crowdsourcing literature improperly assumes the preexistence of a crowd and thus lacks a developmental view to look at the crowd-development issue. Second, there exists no framework that can be useful for studying crowd development, which provides an opportunity for this dissertation to make a meaningful contribution. Third, there exist some contradictory findings related to performance implication of crowd attributes, suggesting a paradox of applying knowledge from different literature (e.g., tournament theory and innovation search) to understand crowdsourcing. Current crowdsourcing literature lacks studies that compare the relative power of these different views in explaining the relationships between crowd attributes and crowd performance, thus motiving us to develop empirical tests to fill this gap.

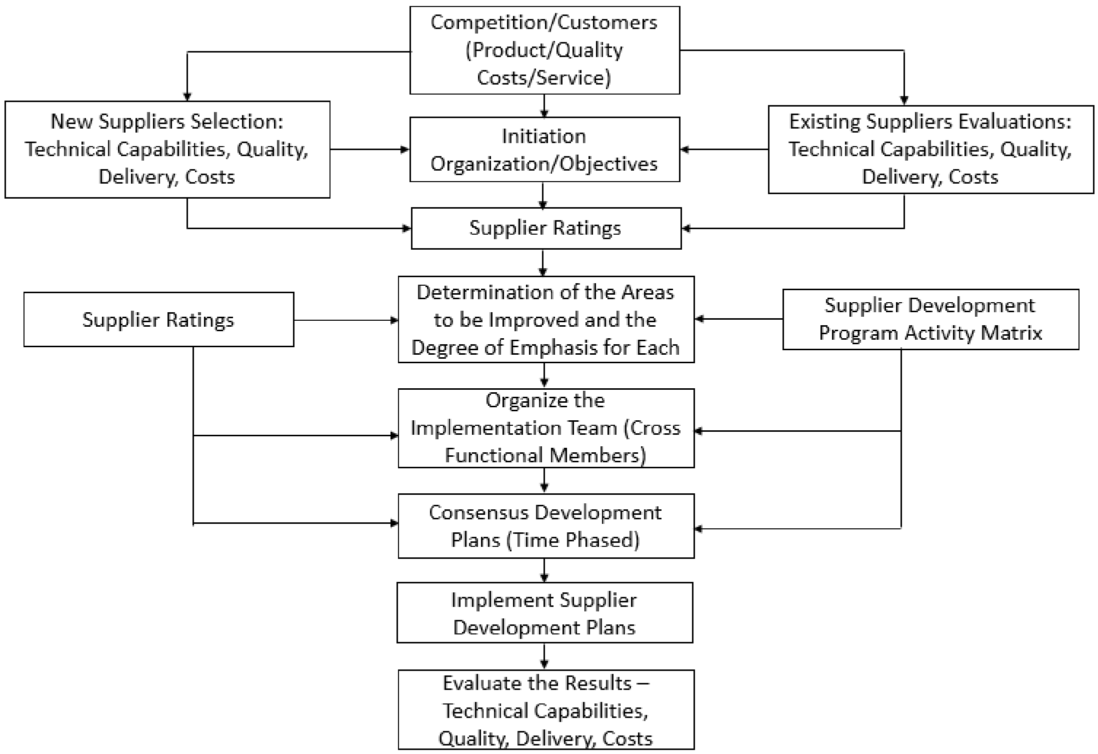
Theoretical Background of Crowd Development

The above literature review shows that crowdsourcing literature lacks a theoretical framework that addresses crowd development. We thus review multiple research streams from different disciplines to increase our understanding on crowd development. In its essence, crowd development is a process through which firms identify a collective of suppliers for a particular task. This process is similar to supplier development in traditional sourcing literature. We first briefly summarize supplier development in sourcing literature in the following section. Because a crowd in crowdsourcing shares similarities with a crowd in sociology (e.g., fuzzy boundary, no specific structure, and transience), we also review the contagion thinking and diffusion theory in sociology literature that are related to crowd formation. As we state in the beginning of this chapter, we mainly focus on tournament-based crowdsourcing in this dissertation. We then review tournament theory and structural thinking to help us better understand the influence of crowdsourcing event design on crowd development.

Supplier Development. In sourcing literature, supplier development refers to any effort or attempts of a buying company to increase performance and/or capabilities of its suppliers to meet its short- and/or long-term needs (Krause, 1997; Krause & Ellram, 1997a). The research stream on supplier development occurred in the early 1990s. Because of global sourcing at that time, suppliers played an important role in determining buying firms' competitive advantages (Krause, 1997; Krause & Ellram, 1997a). As such, buying firms increasingly relied on their suppliers to deliver technologically advanced, defect-free products in a timely and cost-effective manner and thus developed many managerial practices to develop their suppliers' capabilities and skills (Hahn, Watts, & Kim, 1990; Krause & Ellram, 1997a). Based on many supplier development practices, Hahn, Watts, and Kim (1990) proposed the first conceptual model for supplier development (Figure 3), which has been widely cited in the supplier development literature. This process model proposed by Hahn, Watts, and Kim (1990) demonstrates the detailed and sequential steps of a supplier development. This framework also suggests that s supplier development is a systematic, deliberate, and controlled process. This is because buying firms have full decision power over which supplier needs to improve, what needs to be done, and what the expected results would be (Hahn, Watts, & Kim, 1989; C. K. Hahn et al., 1990).

Figure 3

Supplier Development Framework



Source: Hahn, Watts, and Kim (1990, p.4)

Krause and Ellram (1997) systematically reviewed the critical elements of supplier development from a buying-firm perspective. This review indicates that main identified supplier development practices include effective two-way communication, top management involvement, cross-functional buying firm teams, and large percentage of supplier’s annual sales (Krause & Ellram, 1997a; Watts & Hahn, 1993). The identified facilitators for supplier development are buying firms’ communication efforts with suppliers (Krause & Ellram, 1997b), and buying firms’ proactive attitude toward supply-base performance (Krause, Handfield, & Scannell, 1998; Monczka, Trent, & Callahan, 1993). Scholars in the supplier development literature also identified many barriers for supplier development, which include lack of buying firm power measured in terms of the percentage of a supplier's output purchased by the buying firm (Lascelles & Dale, 1989), lack of effective communication, and lack of buying firms' credibility (Galt & Dale, 1991; Lascelles & Dale, 1989). Empirical studies shows that buying firms' supplier development programs have significant performance implications (e.g., suppliers' performance improvement, buyer's competitive advantage, and buyer-supplier relationship improvement) (Humphreys, Li, & Chan, 2004; Modi & Mabert, 2007). The whole supplier development involves into early supplier involvement (Dowlatshahi, 1998; Neal, 1993; Zsidisin & Smith, 2005).

A crowd development by definition is a process of identifying a collective of suppliers for a particular crowdsourced task. The supplier development literature thus offers some insights for us to understand crowd development. According to the supplier development framework proposed by Hahn, Watts, and Kim (1990), a crowd development process involves multiple steps such as initiation, development, and evaluation. As indicated by the empirical studies on supplier development, the developmental practices taken by the focal buying firms such as communication and information sharing could impact the operational process of a crowd development as well as its performance implication.

However, we believe that the application of supplier development in crowd development is limited because of several significant differences between a crowd development and a supplier development. First, crowd development involves an open-call process, while supplier development is a closed-call process. In crowdsourcing, it is the suppliers that make their own decisions (i.e., self-selection) to participate in solving a particular crowdsourced task. Buying firms thus have very limited decision power over which supplier gets involved in the crowd development process. Second, the information visibility is very low in crowdsourcing. Suppliers are nested in a virtue network. Low information means high uncertainty since buying firms have no information or very limited visibility who might self-selection to participate. Third, the task is less specified in crowdsourcing, which means that buying firms cannot apply specific criteria to evaluate, select, and engage with suppliers as they do in supplier development. Because of these significant differences, we believe that knowledge from the supplier development cannot fully explain crowd development. We need other theoretical lenses that is reviewed in the remaining sections.

Contagion Thinking. The crowd in crowdsourcing shares three similarities with the crowd in sociology. First, the boundary of a crowd is not clearly defined. Second, the existence of the crowd is temporary. Once the task is completed in crowdsourcing or the common focus disappears in a social setting, the crowd dissolves. Third, the relationships among crowd members are loosely coupled. Members self-select to form a crowd. Scholars in sociology literature argue that the formation of a crowd is due to the contagion influence existing within a crowd (Christakis & Fowler, 2013; Freedman & Perlick, 1979; Wheeler, 1966). This research stream is referred to contagion thinking in sociology literature.

Contagion thinking was developed to describe the phenomenon of contagion in crowd formation process, which refers to the spreading of behaviors, attitudes, affect, and emotions through a crowd and other types of social aggregations from one member to another (Forsyth, 2009; Le Bon, 1897, 1960). In this research stream, a crowd refers to a gathering of individuals sharing a common focus and concentrated in a single location (Forsyth, 2009). Gustave Le Bon (1897, 1960) was the first scholar who observed the phenomenon of contagion in social psychology. According to Le Bon, emotions and behaviors could be transmitted from one person to another just as germs could be passed along, and he believed that contagion accounted for the tendency of crowd members to behave in very similar ways (Le Bon, 1897, 1960; Wheeler, 1966). In Le Bon's own words, "In a crowd every sentiment and act is contagious" (Le Bon, 1960, p. 50).

Le Bon (1897, 1960) recognized the contagion issue in a crowd but did not offer explanations for the mechanisms underlying this issue. Scholars in social psychology have pondered and debated crowd behavior for centuries, seeking to specify the factors that transform individuals so thoroughly and so unexpectedly (Forsyth, 2009). Various explanations have been offered for the occurrence of contagion in society, including imitation, social facilitation, normative pressure, herding, and/or conformity (Baddeley, 2010; Chapman, 1973; Freedman & Perlick, 1979; Raafat, Chater, & Frith, 2009). Factors that can contribute to the contagion in a crowd include physical density of a crowd (Freedman & Perlick, 1979), similarity of crowd members in terms of needs, values, goals (Hoffer, 1951; Van Zomeren, Postmes, & Spears, 2008), and crowd size (Gladwell, 2006; Newton & Mann, 1980).

Following this contagion thinking, we believe that suppliers' participation behavior is contagious and can spread to other members within the same nested network. Thus, the crowd formation/development is an automatic and spontaneous process. As indicated by the empirical studies on contagion in society (e.g., Gladwell, 2006; Hoffer, 1951), factors that are beneficial to crowd development in crowdsourcing include the closeness of crowd members, the similarity of crowd members, and the number of participants.

Although the contagion thinking seems promising in explaining crowd development, the application of this literature should be tested due to its own limitations. First, it does not tell us how the process starts, that is, the contagion thinking does not address the initiation of a crowd development process. Second, the contagion thinking mainly addresses the contagious phenomena in a physical crowd (e.g., street crowds, mobs, and riots) (Forsyth, 2009). The crowd in crowdsourcing is virtual. It remains unclear whether the contagion thinking still holds in explaining the spreading of suppliers' participation behavior in a virtual setting like crowdsourcing. Due to these obvious limitations of contagion think, this dissertation further reviews other theoretical lenses that are related to crowd development and crowdsourcing in the following sections.

Diffusion Theory. From an emergence perspective (Dooley & Corman, 2002; Holland, 2000), a crowd for a particular crowdsourcing event arises from suppliers' interactive participation behavior in the crowd-development process. The spreading of suppliers' participation within a social network then forms the foundation of crowd development. Diffusion theory is thus an appropriate theoretical lens to look at crowd development. This is because diffusion theory seeks to explain the spreading of behavior, new ideas, products, and technologies (i.e., innovations) through certain channels over time among members of a social system (Rogers, 1962, 2010).

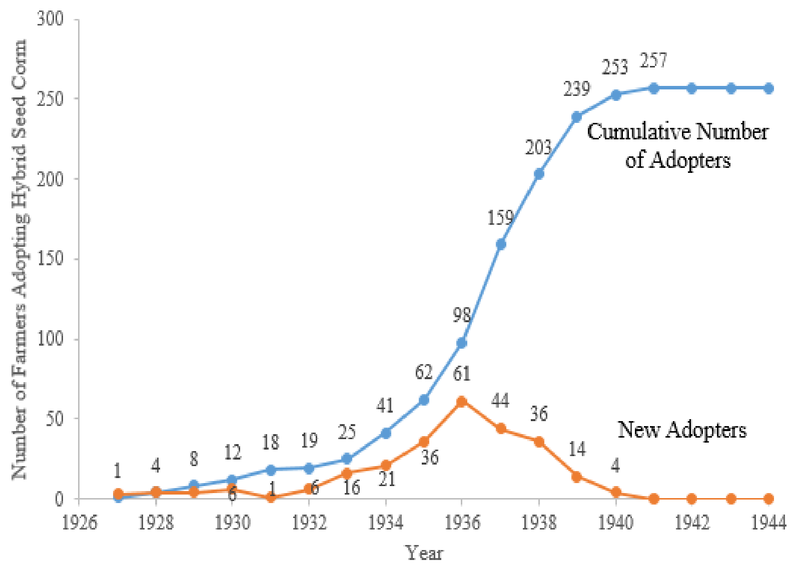
The diffusion paradigm was developed by Ryan and Gross (1943), two rural sociologist who studied the diffusion of hybrid corn seed in two Iowa communities. By surveying more than 300 farmers in two communities, Ryan and Gross found that diffusion is a social process that spreads adoption in the community through subjective evaluation and social imitation, rather than individual rational decision-making (Ryan & Gross, 1943). After Ryan and Gross (1943), the issue of diffusion has been studied in many disciplines such as anthropology, marketing, rural sociology, economics, agriculture, and communications science (Chandrasekaran & Tellis, 2007; Valente & Rogers, 1995). Everett Rogers, a professor in communications studies, synthesized the work of many studies on diffusion and developed the theory of diffusion in his book *Diffusion of Innovations* (Rogers, 1962, 2010).

According to Rogers (1962, 2010), diffusion is the process by which an innovation is communicated over time among the participants in a social system. In this line of thinking, the concept of "innovation" is a generic term that includes not only new ideas, products, and technologies (Rogers, 1962, 2010), but also human behaviors such as communication of information, policy decision-making, and adoption of technology in a network (Rogers & Shoemaker, 1971; Valente, 1993, 1995). Two main indicators used to capture a diffusion process are diffusion rate and diffusion scale (Rogers, 2010; Van den

Bulte, 2000). Diffusion rate is defined as “the relative speed with which an innovation is adopted by members of social system” (Rogers, 2010, p.221). Diffusion scale, sometimes called market size, captures the aggregate number of people who adopt an innovation over a certain period of time (Rogers, 1962; Ryan & Gross, 1943). By plotting the diffusion rate or aggregate number of adopters over time in a curve, scholars in the diffusion literature found that the innovation diffusion follows a so-called S-curve (Figure4). This curve is also termed as the growth curve (Mahajan & Muller, 1979; Peres, Muller, & Mahajan, 2010). The diffusion rate is a numerical indicator of the steepness of the diffusion curve for an innovation (Rogers, 2010).

Figure 4

Innovation Diffusion Process

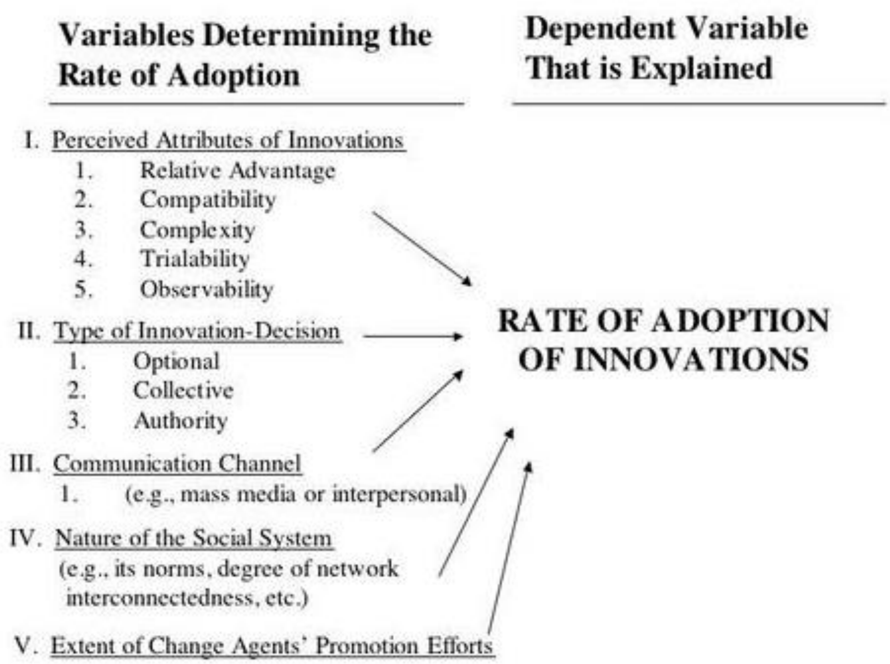


Source: Rogers (2010, p.273)

According to the conceptual diffusion framework proposed by Rogers (1962, 2010), the main structural elements of social systems that influence the diffusion rate of an innovation include perceived attributes of an innovation (e.g., relative advantage and complexity), communication channels (e.g., mass media or interpersonal), and characteristics of a network (e.g., degree of network interconnectedness) (Figure 5). Researchers in the marketing literature extended

Figure 5

Innovation Diffusion Framework



Source: Rogers (2010, p.222)

Roger's diffusion framework and considered the influence of opinion leaders (i.e., influentials) to capture the social imitation underlying a diffusion process (Goldenberg, Han, Lehmann, & Hong, 2009; Keller & Berry, 2003; Van den Bulte & Joshi, 2007). Inspired by Rogers' (1962) framework and its extended versions, scholars from different disciplines have extensively examined the issue of diffusion and tested the conceptual diffusion framework. For a detailed understanding, refer to the systematic literature review in the innovation diffusion domain (e.g., Chandrasekaran & Tellis, 2007; Greenhalgh, Robert, Macfarlane, Bate, & Kyriakidou, 2004; Valente & Rogers, 1995).

The diffusion framework proposed by Rogers (1962, 2010) offers a structural guideline for us to conceptualize factors that might influence crowd development (e.g., crowd growth rate and crowd size). The detailed application of this theoretical lens will be addressed in chapter 3 and chapter 4 of this dissertation. In the empirical stream of diffusion literature, the diffusion rate is generally measured as the number of individuals who adopt a new idea in a specified period, such as a year (Roger, 2010). This measurement makes the diffusion rate contingent on the time specified for the diffusion of an innovation. If the time specified for innovations is different, the calculated rates might not be fully comparable for a large-scale empirical study like this dissertation that involves thousands of observations each with unique event length. Moreover, the role of social imitation that is similar to the contagion thinking in a previous section is underplayed in Roger’s (1962, 2010) diffusion framework. The Bass Model developed by Frank M. Bass (1969) in the analytical stream of diffusion literature resolved these two issues. Because we adopt the Bass Model to operationalize the crowd growth rate in this dissertation, a brief summary of this model is provided in the following paragraphs.

The Bass model emphasizes the role of communication, namely external influence via advertising and mass media, and social imitation (i.e., contagion) (Bass, 1969; Van den Bulte & Stremersch, 2004). The basic assumption of this model is that the timing of consumers' initial purchase (i.e., adoption) is related to the number of previous buyers (Bass, 1969). In a mathematical term, this assumption suggests that "the probability that an initial purchase [$P(T)$] will be model at T given that no purchase has yet been made is a linear function of the number of previous buyers" (Bass, 1969, p. 1826), that is,

$$\frac{f(T)}{1-F(T)} = P(T) = p + \frac{q}{m}Y(T),$$

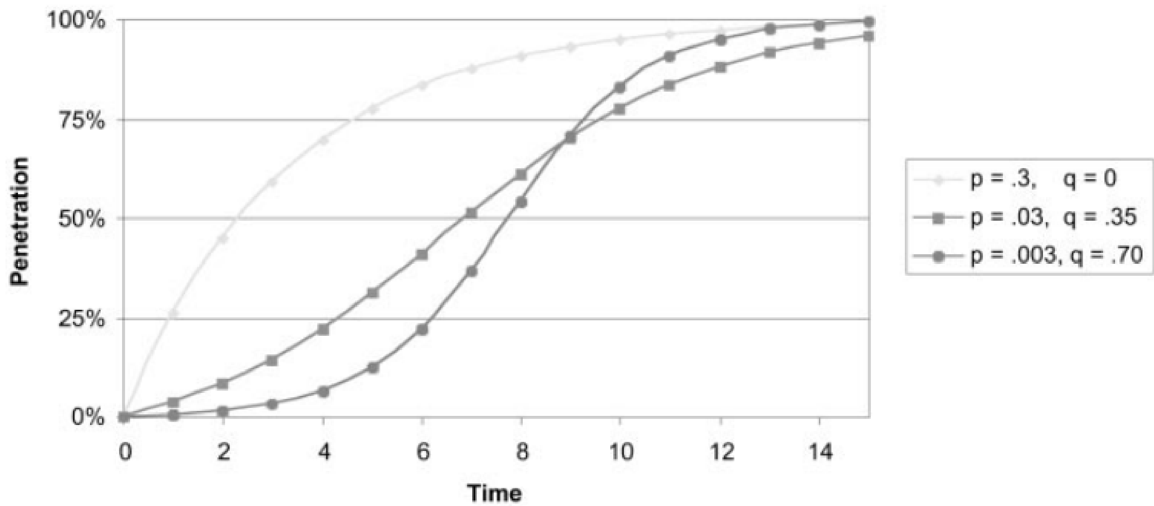
where $f(T)$ is the probability of purchase at time T, $F(T)$ is accumulated probability of purchase at time T, and $Y(T)$ is the total number of previous buyers. The three key parameters in the Bass Model are the coefficient of innovation (p), which captures the intrinsic tendency to make an initial purchase (i.e., adopt an innovation), the coefficient of imitation (q), which captures social influence on making initial purchase, and the potential market size(m). Through algebra and calculus transformation, Bass (1969) identified cumulative sales $S(T)$ at time T (i.e., the growth curve of an innovation diffusion) as a function of $p, q,$ and m, that is,

$$S(T) = m(p+q)^2/p[e^{-(p+q)T}/(\frac{q}{pe^{-(p+q)T}}+1)^2]$$

When $q > p$, the diffusion curve has a S-shape and a differentiation of the $S(T)$ function can get the maximum of diffusion rate at the reflection point $T^* = 1/(p+q)\ln(\frac{q}{p})$. When $q < p$, the diffusion curve is concave and has no reflection point (Figure 6).

Figure 6

Bass Diffusion Curve



Source: Van de Bulte (2002, p.13)

The Bass Model provides a very scientific approach to operationalize the diffusion rate of an innovation curve. The sum of innovation coefficient (p) and imitation coefficient (q) offers an estimate of total diffusion rate (Lawrence & Lawton, 1981; Sultan, Farley, & Lehmann, 1990). Lawrence and Lawton (1981) found that $p + q$ ranged from 0.3 to 0.7 over several innovations. The analytical stream of the diffusion literature offers many approaches to estimate the three parameters in the Bass Model, which include ordinary least squares (OLS) regression (Bass, 1969), nonlinear least squares regression (Jain & Rao, 1990; Srinivasan & Mason, 1986), genetic algorithm (Venkatesan, Krishnan, & Kumar, 2004), and agent-based simulation (Kiesling, Günther, Stummer, & Wakolbinger, 2012; Rand, Herrmann, Schein, & Vodopivec, 2015). Each method has its advantages and disadvantages. This issue will be discussed further when we use the Bass Model to operationalize the crowd growth rate for each crowdsourcing event in our sample.

Tournament Theory. Because the crowdsourcing events that this dissertation mainly focuses on are competition-based, tournament theory is an appropriate framework for explaining the structure, design, and outcomes of a competition-based crowdsourcing event (Terwiesch & Xu, 2008). That is the main reason that we extensively review the structure of a tournament in tournament literature in this section, which will be useful in chapter 3 when we describe the event design of a crowdsourcing event. In this section, we also review the main constructs in the tournament theory and their associated analytic and empirical findings. We will use these findings to look at crowd development and its performance implication in crowdsourcing, which will be addressed in the theory development in chapter 4.

Tournament theory mainly focuses on designing contests (i.e., tournaments) that promote effective competition among agents (i.e., participants), which, in turn, leads to more positive final performance outcomes (Lazear & Rosen, 1981; Lin, Yeh, & Shih, 2013). It is originally developed in personnel economics to study human behavior when reward structures are based on relative rank rather than absolute levels of outcomes (Connelly et al., 2014). This research stream was first proposed by economists Edward Lazear and Sherwin Rosen in the early 1980s when they examined the optimal labor contract design based on relative ranking instead of absolute levels of output (Lazear & Rosen, 1981). Since then, this theoretical lens has expanded to a wider range of other disciplines, such as law (Anabtawi, 2005), ecology (Zabel & Roe, 2009), psychology (Nieken & Sliwka, 2010), finance (Kale, Reis, & Venkateswaran, 2009), sports (Bothner, Kang, & Stuart, 2007; Frick, 2003), management (Henderson & Fredrickson, 2001; Lin et al., 2013), and supply chain management (Bockstedt, Druehl, & Mishra, 2016; Wowak, Craighead, Ketchen, & Hult, 2016).

Scholars in the tournament research stream conceptualize tournaments as contests in which agents compete for a prize that is awarded based on relative rank and are designed to incentivize an optimal level of effort (Becker & Huselid, 1992; Lazaer, 1999). Within this line of thinking, a tournament has four main structural elements (Connelly et al., 2014; Lin et al., 2013): (1) a specific task for agents to participate and compete, such as a promotion contest in management setting, a tennis tournament in sport setting, and logo design in a product development context; (2) an effective time frame associated with this task (e.g., annual golf tournament, monthly sales contest, and weekly logo design); (3) a disclosed reward policy based on agents' relative performance ranking (i.e., prize); and (4) a participation policy (e.g., qualifications in sports tournaments and open call in innovation contests). In the tournament research stream, agents are assumed to be rational, and agents' decisions (e.g., participation and effort investment) are based on utility maximization (Lazear & Rosen, 1981). Scholars in the tournament literature thus pay much attention to issues that influence agents' expected utilities when they study tournament design (e.g., payment size, pay gap, and tournament size) (Connelly et al., 2014).

Connelly and his coauthors (2014) extensively reviewed the development of tournament theory in the past thirty years. According to this review, key constructs addressed in the tournament theory include payment size, pay gap, and tournament size. Payment size refers the financial reward for tournament winner(s) designed to incent the effort of all participants (Knoeber & Thurman, 1994). A payment size is considered "optimal" when it maximizes the productive outcome of the tournament, including all participants (Knoeber, 1989; Knoeber & Thurman, 1994). Empirical evidence shows that payment size is positively associated with the number of participants (Liu et al., 2014; Morgan, Orzen, & Sefton, 2012), suggesting that tournaments with a large payment size are more attractive to participants. As for the performance implication of payment size, scholars suggest that what matters is not the payment size but the pay gap defined as the difference between winning and losing or between relative ranks (Connelly et al., 2014).

When the pay gap is small, agents are not motivated to compete (Knoeber & Thurman, 1994; Lazear & Rosen, 1981). Under this situation, the total productive output of the tournament drops. However, a very high pay gap can have a detrimental effect on tournament efficiency because it induces too much effort that agents must be broadly compensated (Connelly et al., 2014; Wowak et al., 2016). These findings suggest the complexity of payment design and imply the existence of a quadratic relationship between payment and tournament performance. Empirical studies among executives in corporate tournaments demonstrate that the executives' pay gap has positive implications for performance (e.g., ROA) in general, but large pay gaps do not necessarily lead to high firm performance (Henderson & Fredrickson, 2001; Lin et al., 2013). In addition to studying the consequences of pay gap, scholars in the corporate tournament stream examined the antecedents of pay gap and found that job-related risks and uncertainties are positively associated with pay gap (Bloom & Michel, 2002; Gupta, Conroy, & Delery, 2012).

Another key issue in a tournament design is the tournament size that is defined as the number of participants in a tournament (Connelly et al., 2014). Based on the utility maximization assumption, analytic scholars in the tournament literature have studied the issue of tournament size and its performance implication for a while (Fullerton & McAfee, 1999; Körpeoğlu & Cho, 2017; Terwiesch & Xu, 2008). The main notion is that a small increase in tournament size can motivate participants to exert effort to improve performance, but too many participants actually reduces the winning chance for each participants, thereby reducing incentives to invest or exert effort and lowering overall performance outcomes (Che & Gale, 2003; Fullerton & McAfee, 1999; Terwiesch & Xu, 2008). Similar predictions and findings associated with the negative performance implication of a large increase in tournament size have been found in competition situations in sociology, a phenomenon called the "N-effect" which means that more competitors lead to less competition and worse performance outcome (Garcia & Tor, 2009; Mukherjee & Hogarth, 2010). Based on an empirical analysis on innovation contests, Boudreau, Lacetera, and Lakhani (2011) found a negative association between the number of competitors (i.e., tournament size) and the overall performance outcomes.

However, the above notion that a large increase on tournament size leads to lower performance outcomes due to reduced winning chance and less effort investment is challenged by the latest findings in the tournament literature. By taking participants' heterogeneity into consideration, Körpeoğlu and Cho (2017) found that participants with high-expertise actually raise their effort and improve their performance in response to increased competition. This is because an increase in the tournament size raises the expected best performance among other participants, creating positive incentives for participants to exert higher effort to win the contest, thereby increasing the overall performance (Körpeoğlu & Cho, 2017). This latest finding not only justifies the increase popularity of tournament-based crowdsourcing to attract a large number of participants, but also indicates the importance of heterogeneity (i.e., diversity) related to tournament size. This finding is consistent with the concept of “relative deprivation” in social comparison situation, which means that participants who suspect that they might be left behind by their peers (i.e., structural equivalent participant) are motivated to exert and improve their performances (Bothner et al., 2007; Burt, 1982; Garcia, Tor, & Schiff, 2013). These latest findings in the tournament literature call for further research on the performance implication of crowd attributes, which is one of the main objectives of this dissertation.

Structural Thinking. The grand research question of this dissertation is to answer how crowd development impacts the performance of a crowd in crowdsourcing. The theoretical lenses that this dissertation reviews so far (e.g., contagion thinking, diffusion theory, and tournament theory) mainly address crowd formation that is based on suppliers’ interactive participation behavior in crowdsourcing (e.g., Rogers, 2010; Connelly et al., 2014). Tournament theory does examine participants’ efforts and final performances in tournaments (e.g., Garcia & Tor, 2009; Terwiesch & Xu, 2008), but the level of analysis is not at a crowd level but mainly at individual level. We thus review another research stream called structural thinking which links crowdsourcing event design not only with suppliers’ interactive participation behavior but also crowd-level performance.

Structural thinking evolves from multiple research streams such as the structure-conduct-performance in industrial organization economics (Bain, 1956; Caves, 1987) and the structure-process-outcome in service quality literature (Donabedian, 1966, 1988). The basic tenet of structural thinking is that the performance of a social system (e.g., industry, firm, or team) is a function of the conduct of the agents in the system and the process underlying agents' conduct which, in turn, are a function the system's structure (Caves, 1987; Donabedian, 1988; Harper, 2015). For instance, scholars argue that firms derive competitive advantages by responding to the characteristics of the industry in which they compete (e.g., R&D, merge, and acquisition) (Bain, 1956; Caves, 1987). The attributes of the service settings (e.g., facilities, equipment, and human resources etc.) denotes what is actually done in giving and receiving services, which directly influences the service quality (e.g., customer satisfaction) (Campbell, Roland, & Buetow, 2000; Donabedian, 1988).

In structure thinking literature, a structure is generally defined as a system (such as an organization) made up of individual elements or parts (such as people, resources, aspirations, values, market trends, levels of competence, reward systems, departmental mandates, capital, workload/capacity relationship, and so on) that impact each other by the relationships they form (Fritz, 1996; Harper, 2015; Molm, 1990). A structure includes tangible elements (e.g., hierarchy, policy statement, procedures, rules, regulation, and reward systems) and intangible ones such as norms, values, beliefs, and roles (Fritz, 1989, 1996, 1999). Conduct refers to the activities of the agents in the system. Depending on the situations, the conduct in structural think literature can refer to installation and utilization of capacity in management (McWilliams & Smart, 1993), strategic supply chain integration in supply chain management (Ralston et al., 2015), and adoption of IT-assisted communication technology in healthcare (Angst, Devaraj, & D'Arcy, 2012). The process denotes how the agent in the system interact with different elements within the system to provide different activities, while the outcomes refer to the final performance.

The structural thinking framework makes statements about how elements of a social system (e.g., an organization) can be configured and how they causally relate to each other (Größler, Thun, & Milling, 2008). This framework is a theoretical lens that examines social phenomena for scholars in disciplines such as marketing (Geyskens, Steenkamp, & Kumar, 1999), sociology (Molm, 1990), E-commerce (Devaraj, Fan, & Kohli, 2006), management (Mathieu, Maynard, Rapp, & Gilson, 2008), and supply chain management (Ashenbaum, Salzarulo, & Newman, 2012; Samaddar, Nargundkar, & Daley, 2006). Through these many studies, the level of analysis for the structural thinking has been expanded from organizational level (Angst et al., 2012; Donabedian, 1988) to network level (Molm, 1990; Samaddar et al., 2006) and group or team level (Mathieu et al., 2008). The structural thinking has also evolved to study the influence of structure on outcomes through not only different processes (e.g., establishing technical protocols of care) (Angst et al., 2012), but also unique organizational or individual actions (e.g., information sharing and usage of power) (Geyskens et al., 1999; Molm, 1990) and some emergent states (e.g., team efficacy and group cohesion) (Mathieu et al., 2008).

According to these studies, the causal mechanisms through which a structure influences the processes and outcomes include but are not limited to resource allocation, levels of dependence or interdependency, and generating incentives (Molm, 1990; Yin &

Zajac, 2004). As indicated by structural thinking, the structure of a system (e.g., team, organization, or network) affects the behaviors and processes of the system which, in turn, determines the outcomes of this system (DeCanio, Dibble, & Amir-Atefi, 2000; Hansen & Wernerfelt, 1989; Reagans & McEvily, 2003). If we apply this theoretical lens to look at the crowd development, we argue that different elements of event design (e.g., specific crowdsourcing task, evaluation criteria, participation rule, payment size, and event length) constitute a unique structure in crowdsourcing. This particular structure influences how solvers and sponsors (i.e., focal buying companies) interact with each other, thus determining the outcomes of a crowd in crowdsourcing (e.g., crowd productivity, efficiency, and effectiveness). The detailed application of structural thinking in this dissertation is discussed in the following chapter.

Summary

In this chapter, we conducted a comprehensive review on the recent development of crowdsourcing literature. Because of the increasing popularity of crowdsourcing in innovation processes, the human crowd has emerged as an alternative collective supplier. As such, academic research on crowdsourcing has develops quickly in recent years. However, this research stream is at an early stage. Our extensive literature review indicates the existence of several significant research gaps that further motives us to develop this dissertation to enrich crowdsourcing literature and supply chain literature.

First, issues such as the crowd and crowd development are under-examined. Scholars implicitly assumed that a crowd exists before a crowdsourcing event initiates (e.g., Afuah & Tucci, 2012). Crowd characteristics (i.e., pervasiveness of problem-solving know-how in a crowd and motivation of potential solvers with solution knowledge) are also assumed to be known to decision-makers when they consider the possibility of crowdsourcing. Following these two assumptions, managers do not need to think of the issue of crowd development. However, these assumptions are not justified since a crowd emerges only after the crowdsourcing decision has been made. Since these participants are nested online and are located all over the world (Howe, 2006), it is impossible for decision makers to know the characteristics of a crowd in advance.

Besides, individuals' decisions to participate in or withdraw from a particular crowdsourcing event are greatly influenced by the conditions specified by focal buying firms for the crowdsourced tasks (Haas et al., 2015; Liu et al., 2014). The characteristics of a crowd for a particular crowdsourcing event thus remain unknown to decision-makers. Therefore, it is theoretically possible but unrealistic to discuss how the characteristics of a crowd in crowdsourcing influence the manager's decision on crowdsourcing. It is no wonder that current crowdsourcing literature lacks empirical studies that test the crowdsourcing theory proposed by Afuah and Tucci (2012). There also exists no framework that explains crowd development in current crowdsourcing literature. This research void will be addressed in chapter 3.

Second, our literature review identifies multiple theoretical lenses (e.g., contagion thinking, diffusion theory, and tournament theory) that are useful in explaining crowd development, but they suggest different mechanisms underlying crowd development. For instance, both contagion thinking and diffusion theory propose a contagion mechanism through which managers can facilitate the spreading of suppliers' participation behavior to form a crowd for a particulate crowdsourcing event (e.g., Le Bon, 1960; Rogers, 2010). However, tournament theory advocates a competition mechanism through which managers can administrate the growth of a crowd for an event by creating beneficial conditions for suppliers to compete with each other (e.g., Connelly et al., 2014; Terwiesch & Xu, 2008). It remains unclear which mechanism is more applicable in explaining crowd development in crowdsourcing since no research in current crowdsourcing literature has examined this issue. This deficiency will be addressed in the first empirical study in first section of chapter 4.

Third, our literature review discovers some conflicting findings on the performance implication of crowd attributes. For instance, the "homophily effect" (e.g., Bockstedt et al., 2015) and "marginality effect" (e.g., Jeppesen & Lakhani, 2010) indicate that crowd diversity could have opposite implications for performance: the former would suggest a negative influence while the later would indicate a positive effect. These conflicting findings offer opposite implications for managers to organize a crowd in crowdsourcing. The existence of these contradictory findings suggests an insufficient understanding of the associations between crowd attributes and crowd performance, which motivates us to develop the second empirical study in this dissertation. The theory development of this study will be addressed in the second section of chapter 4.

Finally, our literature review indicates that some scholars with the distant search view in the empirical research stream argue that crowdsourcing can allow firms to find "novel" solutions (Afuah & Tucci, 2012; Jeppesen & Lakhani, 2010). This argument suggests that crowd solution quality is a crowd-level performance indicator, in addition to crowd productivity and crowd efficiency introduced in the introduction section. Through a few empirical studies, we found that managers have a strong selection bias caused by limited cognitive attention and familiarity bias when they evaluate the solution quality (Boudreau, Guinan, et al., 2016; Piezunka & Dahlander, 2015). To avoid this section bias, we focus on quantitative crowd-level performance (i.e., crowd productivity and crowd efficiency) in our empirical tests.

Chapter 3: Crowd Development Framework

This chapter conceptually addresses the issue of crowd development from a structural thinking perspective. Based on four descriptive cases on crowd development, this chapter proposes a crowd development framework termed as the double-funnel model. Through this conceptual development, this dissertation lays down the theoretical foundations for the two empirical examinations in the following chapter.

Introduction

A crowd in crowdsourcing refers to a collective of agents (e.g., individuals, teams, and/or firms) who are nested within a network and share a common focus such as a scientific problem-solving, a product design, or a logo design (Afuah & Tucci, 2012; Howe, 2006). As the application of crowdsourcing becomes popular (Roth et al., 2015), using a human crowd for solving innovation-related tasks is growing rapidly. Statistics show year-over-year growth in the global revenue of human crowd platforms was 53 percent in 2010 and 74 percent in 2011 (Kaganer et al., 2013). Some scholars and analysts say that the application of human crowd is potentially more disruptive than the previous outsourcing or global sourcing (DeViney, Sturtevant, Zadeh, Peluso, & Tambor, 2012; Kaganer et al., 2013). They claim that the application of human crowd will "reshape established business processes, redraw organizational boundaries, and – most importantly – profoundly change global labor markets" (Kaganer et al., 2013, p. 24). Due to the increasing application of crowdsourcing in innovation processes, crowd management becomes an important and arising issue for supply chain managers (Kaganer et al., 2013; Malhotra & Majchrzak, 2014).

Despite a growing popularity and reliance on a crowd in practice, relatively little research prescribes how to manage a crowd more effectively and efficiently (Wooten & Ulrich, 2017). Specifically, current crowdsourcing literature lacks research framework to describe the crowd development process. Unlike a supplier development in outsourcing, crowd development in crowdsourcing involves an open-call process, a loosely-coupled buyer-supplier relationship, very limited suppliers' information visibility, and suppliers' self-selection. These differences make it impossible for supply chain managers to apply knowledge from supplier development to manage a crowd development in crowdsourcing. Due to a lack of understanding, many professionals and academic scholars express their concerns and doubts on applying crowds in their innovation processes (Clough, Sanderson, Tang, Gollins, & Warner, 2013; Euchner, 2010). Scholars thus call for research that can help managers better engage, utilize, and organize crowds in innovation processes (Felin et al., 2015).

In this chapter, we intend to develop a crowd development framework that can help both managers and scholars in the supply chain field better understand and utilize a crowd in crowdsourcing. Specifically, we first identify stages of crowd development by reviewing four illustrative crowd development cases. Based on these illustrative cases, we compare the differences between a crowd development and supplier development in traditional sourcing literature. We then draw from structural thinking perspective to discuss how each stage of crowd development process relates to crowd development and crowd performance. In this process, we also identify theoretical constructs related to each stage of crowd development so as to facilitate academic research on crowd management (Felin et al., 2015).

Illustrative Examples of Crowd Development

As indicated by our literature review in the previous chapter, the issue of crowd development is under-developed in current crowdsourcing literature and there exist no research frameworks for crowd development in this research stream. We thus illustrate four crowd development cases to increase academic understanding of this emergent crowd-development process. The four cases come from four industries (e.g., aviation, recreation, services, and medical) and three categories of business: corporations (e.g., Airbus and Netflix), a crowdsourcing platform (e.g., Topcoder), and a non-government organization (e.g., Harvard Catalyst). These cases represent crowd development under different situations, i.e., outsourcing crowd development (e.g., Airbus Cargo Challenge, Harvard Catalyst Experiment, and Topcoder Programming Contest) and making crowd development (e.g., Netflix Prize Challenge). We believe that these cases are representative for crowd development in tournament-based crowdsourcing. We pull information from multiple sources (e.g., Web of Science, Google Scholar, and Website) to recapture the crowd-development process for these four cases.

Airbus Cargo Drone Challenge. In early 2016, Airbus intended to identify the next generation of multi-purpose drones and to seek a safe, easy-to-operate and affordable drone solution that could be used for many civil applications including last-mile humanitarian logistics (Local Motors, 2016). Instead of relying on internal development or contract outsourcing, Airbus created the Airbus Cargo Drone Challenge in partnership with Local Motors, a US-based vehicle innovation company focused on low-volume manufacturing of pen-sourced motor vehicle designs using co-creation (Moritz, Redlich, & Wulfsberg, 2016).

Airbus first created the design specifications regarding size, weight, and operation mode (Moritz et al., 2016). For instance, the design should be capable of vertical takeoff and landing, and vehicle weight when fully loaded should be less than 25 kg (i.e., around 55 pounds) (Local Motors, 2016). Total payment size was set as $117,500, which would be awarded in three categories (main award voted by Airbus executives: 1st: $50,000, 2nd: $20,000, 3rd: $10,000; cargo prize voted by cargo industry experts: 1st: $15,000, 2nd: $5,000, 3rd: $2,500; community prize voted by Local Motors' community designers: 1st: $10,000; 2nd: $3,000, 3rd: $2,000). After Airbus identified the design requirement and specified the payment policy, Local Motors broadcasted this challenge to its online design community, which has around 300,000 members including engineers, fans, investors, and enthusiasts from all over the world (Warwick, 2016).

This event started on April 12, 2016. After the initiation date, thousands of designers from Local Motors' design community started to participate in this competition and submitted their designs. Each submission has its own webpage where all information (e.g., text, design, drawings etc.) on the design is posted online and other community members can make comments. All submissions were publicly available and licensed under Creative Commons (CC-BY-NC-SA) (Warwick, 2016). By May 22, 2016 (i.e., six weeks later after the initiation day), Airbus and Local Motors received 425 solutions. After the submission deadline, all solutions were checked for validity according to the specified requirements. Then, the voting was conducted by a panel including Airbus executives, cargo industry experts, and Local Motors' community designers. Winners were announced June 15, 2016. This Airbus Cargo Drone Challenge demonstrates not only the great potential of a crowd in generating designs but also the different stages of crowd development (e.g., initiation, formation, realization, and evaluation).

Harvard Catalyst's Experiment. Harvard Catalyst, a pan-university clinical translational science center situated at Harvard Medical School, intended to generate research topics to cure Type 1 diabetes in early 2010 (Guinan et al., 2013). Rather than working with the 17 health centers and more than 20,000 faculty, research staff, and graduate students affiliated with Harvard Medical School, Harvard Catalyst partnered with InnoCentive, an online crowdsourcing platform, and organized a challenge titled "What do we not know to cure Type 1 diabetes?"

This crowdsourcing event, which was open for six weeks in 2010, was advertised throughout the Harvard and InnoCentive communities, and in the journal *Nature* as well. Harvard Catalyst offered $30,000 in awards. Within six weeks, 779 individual agents self-selected to compete in this contest. In the end, 163 agents submitted 195 solutions. These participants represented 17 counties and every continent except Antarctica. Their solutions encompassed a broad range of therapeutic areas including immunology, nutrition, stem cell/tissue engineering, biological mechanisms, prevention, and patient self-management (Guinan et al., 2013). A total of 150 submissions was identified as ready for evaluation after duplicates and incomplete submissions were filtered out. Harvard Catalyst opened the process of evaluation by inviting experts with widely disparate knowledge bases to select noteworthy solution submissions.

In the end, 142 Harvard Medical School faculty reviewed and evaluated 150 submissions. After aggregating the anonymous evaluations form all reviewers, Harvard Catalyst provided awards to the 12 best submissions based on the average score. Winners included a human resources professional with Type 1 diabetes, a college senior, an associate professor of biostatistics, a retired dentist with a family member with diabetes, faculty biomedical researchers, and an endocrinologist (Guinan et al., 2013). The background of the winners indicates the importance of diversity in creative problem-solving. The Harvard Catalyst's experiment shows the power of a crowd in generating solutions and the sequential process of crowd development such as initiation, formation, and evaluation.

Netflix Prize Challenge. Netflix desired to develop a software that would achieve a 10 percent improvement in the DVD rental firm's algorithm-based movie recommendation system in 2006 (Bennett & Lanning, 2007; Zhou, Wilkinson, Schreiber, & Pan, 2008). Netflix provided over 100 million ratings from over 480,000 randomly chosen, anonymous subscribers on nearly 18 thousand movie tiles. Netflix made this data publically available on its website and created the 2006 Netflix Prize Challenge. A grand prize of $1 million would be awarded to the first person or team that reached the goal of 10 percent improvement (Bell & Koren, 2007).

This competition began on October 2, 2006. Hundreds of thousands of people competed in this challenge. By Jun 2007, over 20,000 teams had registered for this competition from over 150 countries, and 2,000 teams had submitted over 13,000 solutions (Bennett & Lanning, 2007). Due to the complexity of this challenge, no team had ever achieved the goal of 10 percent improvement before June 2009. To maintain and stimulate the crowd growth, Netflix offered two progress prizes with $50,000 in 2007 and 2008 to the team that achieved the best performance among all participants. Netflix stopped gathering submissions for the Netflix Prize Challenge on July 26, 2009 and announced the $1 million grand prize to a team who achieved a 10.05 percent improvement (Netflix, 2009). This Netflix Prize Challenge denotes the productivity of a crowd and also the complexity of a crowd development under a situation for a very complex task.

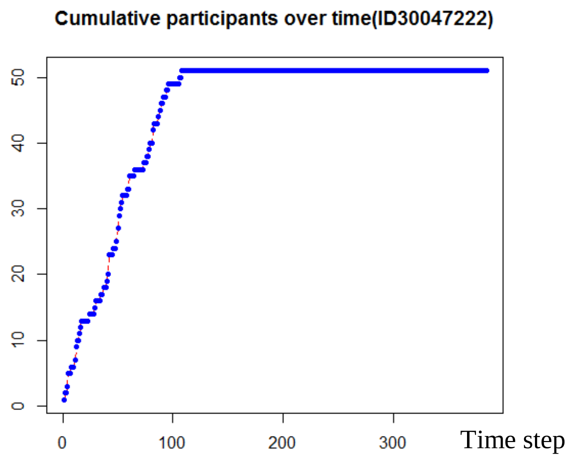
Topcoder – IBM Discount Mobile Apps Design Challenge. Established in 2001, Topcoder is a leading platform for delivering crowdsourced software solutions for IT-intensive organizations by soliciting independent programmers from around the world to compete in a regular stream of software contests (Boudreau et al., 2011). Over the years, Topcoder has served companies such as Best Buy, Eli Lilly, IBM, and GEICO (Lakhani, Garvin, & Lonstein, 2010). Latest statistics from Topcoder community shows that Topcoder has more than one million active online programmers from all over the world, and that there are hundreds of them competing in programming contests every day (Topcoder, 2017b). Our observations collected from Topcoder's website show that Topcoder organized 6,825 programing contests on behalf of its clients between July 4, 2014 and October 18, 2016. On average, there were eight programming events every day during our data collection time frame. The following paragraph describes the crowd development for a discount mobile application design challenge organized by Topcoder for IBM in November 2014 (ID: 30047222).

Topcoder first worked with IBM to identify software needs. The goal of this "IBM – Discount Mobile Apps Design Challenge" was to create a new mobile application that would help IBM employees find places where they could use their IBM company discount. According to the identified needs, Topcoder transferred specific needs to programming design requirements (e.g., screen features, navigation, and dashboard), guidelines (e.g., font size, colors, platform), and judging criteria (e.g., visual effect, cleanliness of design, and compatibility with smartphone). After the requirements were clarified, Topcoder posted this event on its website. The total payment size was decided at $2,450 (i.e., 1st: $1,500; 2nd: $650; 3rd: $300). The event start data and end data were specified at 8:00 EST, Nov 15, 2014 and 8:01 EST, Dec 1, 2014, respectively (i.e., event length is 16 days/384 hours). After this event went alive on its website, designers from Topcoder's community registered online for participating in this design challenge and started to submit solutions. During the crowd formation process, IBM offered feedbacks to suppliers who submitted solutions at 8:14 EST, Nov 20, 2014.

In total 51 programmers participated in this competition. The first submission was made on 11:46 EST, Nov 15, 2014 (i.e., 3.16 days after the starting date), and the last submission was on 7:53 EST, Dec 1, 2014 (i.e., 16 days after the starting date). Submissions were closed at the announced end time. For this particular design challenge, Topcoder received 19 submissions. Each contest submission was evaluated by a peer-review panel of three expert members according to the judging criteria. Three winners were announced at 18:54 EST, Dec 3, 2014. According to each participant's unique participation time, we plotted the accumulated number of registrants over the event length and got the following crowd emergence trajectory for this event (Figure 1). These participants represented 11 countries including the United States, China, and India. 14 out of 51 participants had more than one year membership with Topcoder, and four had more than two years' membership experience.

Figure 7

Crowd Emergence Trajectory for Event 30047222



Data source: Topcoder[4]

Summary of Crowd Development Cases

Through the above four illustrative crowd development cases, we know that a crowd development process involves multiple parties such as focal buying firms and multiple suppliers who are normally outside the buying firms' network and might have no prior business relationships. If buying firms outsource the crowd development process like the Airbus case, a crowd development can also involve a crowdsourcing platform

[4] https://www.topcoder.com/challenge-details/30047222/?type=design

(e.g., InnoCentive). Each party has different objectives. Focal buying firms attempt to solve a particular task or problem (e.g., drone design or improving the prediction accuracy) by leveraging the distributed creativity outside the organizational boundaries. Suppliers self-select to participate and compete for winning. The crowdsourcing platform that is involved in crowd development process facilitates the financial transactions and IP transfer between buying companies and selected winner(s).

The crowd development process involves an open call through which suppliers make their own participation decisions. Suppliers' self-selection makes the crowd development process filled with uncertainties. For instance, buying firms do not know which supplier might participate in the crowdsourced event, how many suppliers will participate, how many solutions suppliers will generate, and what the quality of solutions will be. All these puzzles will not be resolved until the crowd emerges at the end of a crowdsourced event. The crowd for a crowdsourced event dissolves when the event reaches its deadline. In a sense, a crowd in crowdsourcing is not only transient but also an outcome of crowd-development process.

As our literature review indicated in the previous section, crowd development is a process of identifying a collective of suppliers for a particular crowdsourcing event. This process shares some similarities to supplier development in traditional source literature. According to the above descriptions, we believe that significant differences exist between a crowd development in crowdsourcing and a supplier development in traditional outsource situation. Table 1 summarizes the main differences between these two processes. These significant differences indicate that crowd development in crowdsourcing is not a controlled, deliberate, systematic process but an emergent, not calculated, unsystematic process. Due to these significant differences, we cannot use the knowledge about supplier development to understand and manage crowd development in crowdsourcing. Thus, a detailed description on crowd development becomes meaningful for both scholars and managers. We address this task in the following section.

Table 1

Difference between Crowd Development and Supplier Development

	Crowd Development	Supplier Development
Process openness	An open call	A closed, systematic call
Information visibility	Low	High
Outcome uncertainty	High	Relative low
Supplier autonomy	High	Low
Relationship proximity	Loosely coupled	Closely connected
Time horizon	Short	Relative long

Double-Funnel Crowd Development Framework

The above four illustrative cases suggest that a crowd development starts with the design of a crowdsourcing event (e.g., task specifications, payment, and event length etc.) and ends with winner announcement. Although the specific operations of crowd development might vary under different situations, the whole process generally goes through four stages: crowd initiation, crowd formation, crowd realization, and crowd evaluation. We describe each stage of a crowd development in the following section and discuss the application of each stage from the structural thinking perspective (Fritz, 1996; Molm, 1990).

Crowd Initiation. This is the starting point of a crowd development process. The main parties involved in this stage include focal buying firms (i.e., sponsors). This stage also involves a crowdsourcing platform (i.e., organizer) if buying firms outsource their crowd development like in the Airbus Cargo Drone Challenge. The purpose of this stage is to design the crowdsourcing event, which is called tournament design in tournament literature (Che & Gale, 2003; Chen et al., 2011). At this stage, buying firms first identify the task that will be crowdsourced. The task can be very specific in the Netflix case and also very abstract in Harvard Catalyst's case. They then need to clarify the requirements for the task and the criteria for evaluating solutions and selecting winners. Once the scope and requirements of a task are identified, the complexity of a task is determined from a task design perspective (Campbell, 1988; Wood, 1986; Zheng, Li, & Hou, 2011).

The event design at this stage also includes specifying reward policy (i.e., payment size and payment structure), participation policy (e.g., individual based or team based), and determining the event length (i.e., event starting time and ending time). Specifications related to payment size and event length are objective and can be used to differentiate crowdsourcing tasks. One last element of the event design is to identify the target audience. This identification can be broad, as in the Netflix case (i.e., any online users who are interested in data analytics and algorithm design), or specific, as in the Airbus case (i.e., Local Motors' community members). The target audience forms an "intended" crowd for a particular crowdsourcing event, which includes influential suppliers who are deemed to be most qualified based on their profiles (e.g., winning records, prior participation history, and skills) and non-influential ones.

From a structural thinking perspective (Fritz, 1996; Molm, 1990), we believe that all these specifications and requirements made by the buying firms or crowdsourcing platform constitute a structure for crowd development. This structure is made of very tangible elements such as crowdsourcing task, payment size, rewarding policy, and event length and some intangible elements like the target audience. Different combinations of these structural elements form unique structures for each crowdsourcing event, which influence how the targeted suppliers (e.g., agents) might interact with each other. Based on structural thinking (Caves, 1987; Harper, 2015; Molm, 1990), we argue that the structure of a crowdsourcing event will exert influence on suppliers' subsequent behaviors (e.g., participation, effort-investment, risk-taking) in the crowd formation process that, in turn, will impact the final crowd performance outcomes.

Crowd Formation. After its initiation, a crowd-development process goes into the formation stage which is the second and a very important stage of crowd development. Our selected cases indicate that this stage directly determines the outcomes of a crowdsourcing event. The main party involved in this stage is the suppliers (i.e., solvers, participants, or agents) who are nested in a virtual network. These suppliers can be individuals in most situations, but they can also be teams, as in the Netflix case. Suppliers make many decisions at this stage. For instance, they decide whether and when they participate in a particular contest, whether they withdraw or sustain their participation, and whether and when they submit their solutions. As the crowd emergence trajectory for the Topcoder – IBM case indicates, suppliers' participation decisions are not made simultaneously but gradually. This trajectory suggests that crowd formation is mainly based on suppliers' participation decisions. Depending on the situations, buying firms or crowdsourcing platform might interact with suppliers by providing feedback to suppliers who participate in a contest at this stage (e.g., Topcoder – IBM case).

The information available in our four cases provides a brief description on what's going on supplier' behaviors (i.e., participation, withdrawal, and submission) in a crowdsourcing contest. The knowledge from tournament literature suggest that suppliers' behaviors depend on many factors such as payment size, skills, and expected chances of winning (Connelly et al., 2014). Current academic interests on this stage focuses on suppliers' motivations for participation (Brabham, 2008, 2010, 2010) and interaction mechanisms (Bothner et al., 2007). The identified motivations include extrinsic motivations (e.g., money, reputation, and skills development) and intrinsic motivations (e.g., fun, a sense of belonging, and achievement) (Brabham, 2010, 2012). Potential interaction mechanisms among suppliers include imitation (i.e., social contagion) (Brabham, 2010; Le Bon, 1897) and competition (Bognanno, 2001; Boudreau, Lakhani, et al., 2016; Morgan & Wang, 2010). There also exists many other potential issues at the crowd formation stage that are worthy of further exploration. For instance, the growth speed of a crowd and its antecedents.

From a structural thinking perspective, crowd formation is an intermediate process in which suppliers interact to form a crowd for a particular crowdsourcing event. Specifically, suppliers evaluate the attractiveness of a crowdsourcing event according to not only the tangible structural elements of an event (e.g., task complexity, payment size, and event length) but also the emergent structural elements like competition intensity (i.e., the number of participants) (Connelly et al., 2014; Terwiesch & Xu, 2008). Since suppliers make their participation decision gradually, the competition intensity is dynamic in the crowd formation process. As indicated by the tournament literature, competition intensity determines the suppliers' winning chances for a contest that, in turn, impacts suppliers' effort investments and potential final outcomes (Körpeoğlu & Cho, 2017; Terwiesch & Xu, 2008). We thus believe that crowd formation is a dynamic and complicated process and the competition mechanism underlying suppliers' multiple behaviors can explain the relationship between the structure of a crowdsourcing event and final performance outcomes.

Crowd Realization. From a structural thinking perspective, the crowd realization is a transient stage between crowd formation and crowd evaluation. The reason that we argue this stage is transient is because a crowd for a particular crowdsourcing contest dissolves immediately after the contest ends. The main actors at this stage are suppliers whose main decisions are to submit their solution(s) before the ending time. After the event ending time, buying firms and/or crowdsourcing platforms close the solution submission link and automatically announce the completion of a crowdsourcing event. This stage is the time that a crowd for an event finally emerges. We refer the crowd that emerges in the end as the realized crowd.

Through this realized crowd, buying companies can get crowd-level attributes such as crowd size (i.e., the number of participants) and crowd diversity (i.e., the extent of differences in terms of demographic background and skills). For instance, the crowd size for the IBM Discount Mobile Apps Contest was 51. The crowd members came from 11 different countries including the United States, China, and India. Some crowd members were senior programmers with many years of programming contest experience. Some quantitative crowd outcomes are available at the realization stage. For instance, once suppliers submit their solutions by the event expiration day, buying firms and crowdsourcing platforms know the number of solutions generated by a crowd (i.e., crowd productivity). Another quantitative crowd-level performance indicator that can be operationalized at the realization stage is crowd efficiency defined as the relative speed with which a crowd finishes a crowdsourcing task. For instance, the shortest task time for the IBM Discount Mobile Apps Contest was 3.16 days and the average task time for this contest was 7.53 days.

Crowd Evaluation. After realization, the crowd development process moves into crowd evaluation (i.e., solution evaluation), the last stage of a crowd development. Main parties that are involved in this stage are buying firms and/or crowdsourcing platforms, whose main actions are to organize a panel and evaluate solutions generated by a crowd according to the specified criteria. Based on the evaluation outcomes and relative ranking of solutions, buying firms or the crowdsourcing platform announce the winners, reward the winners according the announced payment size and structure, and manage the transfer of intellectual property.

In the evaluation process, buying firms and crowdsourcing platform acquire quality- related crowd outcomes, such as the percentage of viable solutions. In the Harvard Catalyst case, 12 out of the 150 submissions, i.e., 8 percent, is deemed to be winning solutions. According to structural thinking, the qualitative crowd performance is a function of the conducts of suppliers in the crowd formation (e.g., competition, risk-taking, and effort investment) which, in turn, is a function of the structure of crowd development (e.g., task complexity, payment size, and event length) (Bain, 1956; Harper, 2015).

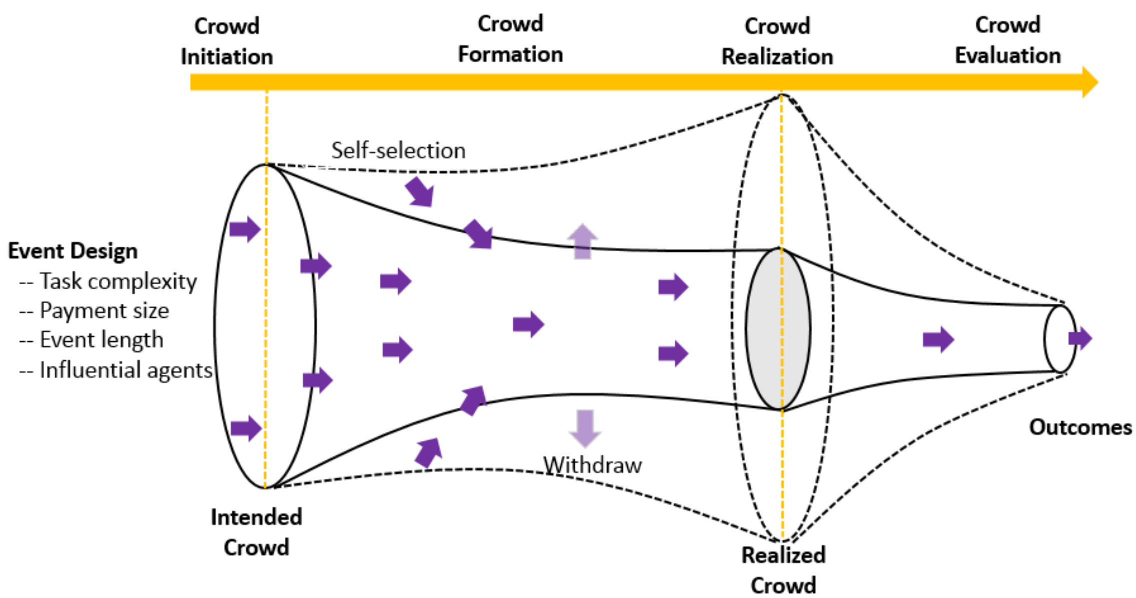
Although one of the main purposes of using a crowd in an innovation process is to acquire novel solutions (i.e., distant search), empirical studies show that managers unintentionally give low scores to the solutions that they are not familiar with when they evaluate the quality of solutions due to their narrow attention and familiarity mentality (Boudreau, Guinan, et al., 2016; Piezunka & Dahlander, 2015). Because of the existence of evaluation bias, we believe that the quantitative outcomes (e.g., crowd productivity and crowd efficiency) are more objective than quality outcomes in crowdsourcing. We thus mainly focuses on qualitative crowd performance in this dissertation.

Summary

In this chapter, we illustrate crowd development through four cases and identify different stages of crowd-development process. By putting the four stages of crowd development on a temporal scale and considering relevant variables involved at each stage, we develop the following process framework for crowd development (Figure 8). As shown in this framework, a crowd development starts from an initiation stage in which firms specify elements of a crowdsourcing event design such as the specific task, reward policy, and targeted suppliers. Through these specifications, we can operationalize and differentiate crowdsourcing events via a series of constructs such as task complexity, payment size, and event length. After the initiation stage, the crowd development process goes into formation stage. In this stage, suppliers from the "intended crowd" (i.e., targeted suppliers) and outside of the specification boundary self-select to participate and compete with each other to provide the best solutions. During this stage, some supplier might sustain their participation and submit their solutions, while others might withdraw their participation. A crowd for a particular crowdsourcing event finally emerges at the specified ending time of this event. After that, the whole process goes into evaluation stage. This is the time that managers evaluate the productivity and efficiency of a crowd as well as the usefulness and novelty of the solutions generated by a crowd.

Figure 8

Double-Funnel Crowd Development Framework



This framework indicates that a crowd-development process incorporates two filtering funnels. We thus call our proposed crowd development framework a “Double-Funnel Model”. This framework not only maps out the crowd development process but also partitions the whole process into two testable potions. From a structural thinking perspective, the first portion addresses the influence of event design on the crowd formation (i.e., the structure – conduct link), which will be the first empirical study in this dissertation. The second portion deals with the performance implication of crowd attributes (i.e., the conduct – performance link), which the second empirical test in this dissertation. The following chapter 4 will cover the detailed theory development of these two tests.

Beyond the above support for this dissertation, the proposed double-funnel framework offers other implications for scholars and managers in the supply chain management field. For instance, scholars can compare the differences between intended crowd and realized crowd in a crowd-development process and further explore the conditions that are related to the differences. The double-funnel process framework is also promising in applying system dynamics lens (Choi, Dooley, & Rungtusanatham, 2001; Dooley, 1997; Größler et al., 2008) to examine suppliers' interactive behavior in crowd development and to study the influence of event design on the whole process. As for supply chain managers, this framework is the first time in current supply chain literature to map out an emergent process (i.e., crowd development) in innovation processes. Specific managerial implications will be addressed in the discussion section of this dissertation.

Chapter 4: Theory Development

Overview of Theory Development

This chapter addresses the theory development for the two empirical tests in this dissertation. Our proposed “Double-Funnel Crowd Development Framework” partitions the whole crowd-development process into two parts: formation and evaluation. Our theory development thus includes two sections: the first one covers the influence of event design on crowd formation, and the second is related to the performance implication of crowd attributes.

The first theory development addresses the issue of crowd emergence, which refers to the arising of unexpected growth rate and crowd size in the crowd development process for a particular crowdsourcing event (Dooley & Corman, 2002; Holland, 2000). One rationale behind this study is current crowdsourcing literature lacks studies examining the influence of event design on crowd emergence. Both scholars and professionals thus have no reported knowledge on how to manage crowd emergence in crowdsourcing. Another reason is that tournament theory (Connelly et al., 2014; Lazear & Rosen, 1981) and diffusion theory (Rogers, 2010; Strang & Soule, 1998) suggest two mechanisms underlying crowd emergence (i.e., competition and contagion). These two mechanisms offer different predictions, sometimes even the opposite predictions, on the relationships between elements of event design and crowd emergence. We thus have two specific purposes for the first theory development. The first purpose is to address the relationships between elements of event design and crowd emergence. The second is to examine whether tournament theory is more applicable than diffusion theory in explaining the relationships between elements of event design and crowd emergence.

Our second theory development addresses the performance variation puzzle by analyzing the performance implication of crowd attributes. We examine two main crowd attributes suggested by the crowdsourcing and management literature: crowd size and crowd diversity. Crowd size is defined as the number of solvers (i.e., agents) participating in a crowdsourcing event (Liu et al., 2014), while crowd diversity refers to the extent of differences among crowd members in terms of background and demographic statistics (Daniel et al., 2013; Harrison & Klein, 2007; Horwitz & Horwitz, 2007; Ren et al., 2015).

Many existing findings related to the performance implication of crowd attributes are contradict to each other. For instance, Boudreau, Lacetera, and Lakhani (2011) found that an increase in crowd size has a negative influence on solvers' efforts and crowd performance due to a reduced chance of winning, but Bockstedt, Druehl, and Mishra (2015) identified a positive association between crowd size and crowd performance. As for the issue of diversity, Jeppesen and Lakhani (2010) argued that solvers who are marginal in terms of technical skills and social background are more likely to develop good solutions, indicating that an increase in crowd diversity can have a positive performance implication (i.e.. the marginality effect). However, Bockstedt, Druehl, and Mishra (2015) claimed that solvers who are similar to the buying companies are more likely to be the winners in crowdsourcing, suggesting that an increase in crowd diversity might have negative influence on performance (i.e., the homophily effect). These opposing findings demonstrate that our understanding on the performance implication of crowd attributes is insufficient and incomplete. Thus, the first objective of our second theory development is to understand the relationships between crowd attributes and crowd performance.

Some scholars argue that the competition mechanism based on tournament theory explains the relationship between crowd attributes and crowd performance (Boudreau et al., 2011), while others suggest that a search process based on innovation search literature can better explain the performance implications of crowd attributes (Afuah & Tucci, 2012; Jeppesen & Lakhani, 2010). These two mechanisms offer different explanations for how crowd attributes relate to crowd performance, which indicates that the true mechanism that links the crowd attributes and crowd performance is not quite clear. Therefore, the second objective of this theory development is to test whether a competition mechanism is more applicable than a search mechanism in explaining the relationship between crowd attributes and crowd performance.

As indicated in the literature review chapter, we identified that different theories (e.g., diffusion theory and tournament theory) or research streams (e.g., search literature) offer opposing views on crowd emergence or crowd performance (e.g., crowd productivity). We thus took the strong inference epistemological approach developed by John R. Platt (1964) to develop alternative hypotheses. The strong inference approach is a model of scientific methodology that encourages a priori specification and the subsequent evaluation of multiple, often competing, hypotheses (Davis, 2006; Jewett, 2005; Platt, 1964). Many scholars believe that this method can avoid the confirmation bias of a single hypothesis and intensify the process of science (Nadler, Thompson, & Boven, 2003; Platt, 1964; Rungtusanatham, Forza, Koka, Salvador, & Nie, 2005).

Understanding the Influence of Event Design on Crowd Emergence

In this section, we develop hypotheses on the relationship between elements of event design and crowd emergence (i.e., crowd growth rate and crowd size). According to our discussion on crowd development in previous chapters, the elements of event design include task complexity, payment size, event length, and the involvement of influentials. Table 2 lists all the constructs used in this theory development. As indicated

Table 2

Constructs in Theory Development for Crowd Emergence

Constructs	Definition	Reference
Crowd growth rate	The relative speed with which suppliers participate in a particular crowdsourcing event	Rogers (2010)
Crowd size	The total number of suppliers that participate in a particular crowdsourcing event	Liu et al. (2014)
Task complexity	The degree to which a crowdsourcing event is perceived is perceived as relatively difficult by suppliers to understand and solver	Rogers (2010); Wood (1986)
Payment size	The amount of money specified for the winners in a tournament-based crowdsourcing event	Lazear & Rosen (1981)
Event length	The amount of time specified for suppliers to solve a particular crowdsourcing event	Connelly et al. (2014)
Influential agents	Suppliers whose behaviors and decision-making are influential to those of others within the same network	Rogers (2010)

by our literature review in chapter 2, both diffusion theory and tournament theory offer different explanations on crowd formation. Specifically, diffusion theory takes the perspective that the crowd will form based on spreading of suppliers' participation behavior within a supplier network (Goldenberg et al., 2009; Keller & Berry, 2003;

Rogers, 2010). This spreading depends on individual participants' evaluation of the crowdsourcing task – i.e. ease of completion, task attractiveness, and participation of others (Boyd & Mason, 1999). On the contrary, tournament theory takes the perspective that the crowd will form based on each participant's evaluation the competition – i.e., chance of winning, the amount of expected returns, and the amount of inputs (Connelly et al., 2014; Lazear & Rosen, 1981; Terwiesch & Xu, 2008). These two theoretical lenses offer different suggestions on how elements of event design in crowdsourcing relate to crowd growth and crowd size. We thus take a strong inference approach (Davis, 2006; Platt, 1964) and develop competing hypotheses in the following sections.

Task Complexity and Crowd Emergence. Task complexity refers to the degree to which a crowdsourced problem is perceived as relatively difficult by agents (i.e., solvers) to understand and solve (Rogers, 2010; Wood, 1986). Crowdsourcing tasks can be classified on the complexity-simplicity continuum. For instance, a mobile screen design challenge focused on creating concepts and visual solutions for customer relationship management application will be more complex than a web design challenge on information search (Topcoder, 2016). This is because the former challenge involves multiple interfaces (e.g., mobile, computers, and technical software) and the interactions among them, and requires solvers to consider and to incorporate more technical features in their designs. Both diffusion theory and tournament theory offer explanations for how task complexity of a crowdsourcing event influences the emergence of a crowd for a particular event, but the predictions on the relationships are opposite to each other.

From the lens of diffusion theory, crowd emergence is the process through a crowdsourced task communicated and diffused over the event cycle time among the solvers in a network. Diffusion theory suggests that complex tasks slow down crowd emergence (i.e., crowd growth rate and crowd size) for a few reasons. First, complex tasks increase solvers' information processing cost and thus reduce the attractiveness of an event to potential solvers (Denis, Hébert, Langley, Lozeau, & Trottier, 2002; Rogers, 2010). A complex task like the mobile screen design challenge normally has several components to consider, which involves more uncertainty for potential agents to solve (Campbell, 1988). Second, complex tasks expand solvers' communication cost, diminishing the chances for solvers to communicate with other potential solvers and to spread the participation behavior within a solver community (Rogers, 1962, 2010). Third, task complexity relates to solvers' information searching costs. Complex tasks increase the cognitive demands placed on the solvers who might be interested in solving this task (Campbell, 1988; Wood, 1986). A crowdsourced task that is perceived to be too complex to understand or solve may result in a state of information overload among potential solvers (Meyer, Johnson, & Ethington, 1997). We thus propose that,

H1a (b): In the context of tournament-based crowdsourcing, crowd growth rate (crowd size) is negatively related to task complexity of a crowdsourcing event.

By contrast to diffusion theory, tournament theory suggests that the relationship between task complexity and crowd emergence is not linear but quadratic. From this theoretical lens, task complexity influences crowd emergence through two mechanisms (e.g., pay disparity and motivation), both of which suggest the existence of a quadratic relationship between task complexity and crowd emergence. First, as empirical evidence from corporate tournaments (Henderson & Fredrickson, 2001) suggests, task complexity is positively related to pay disparity defined as the spread between the winning prize and losing (Lazear & Rosen, 1981). When pay disparity is small, agents are not sufficiently incentivized to compete with other. However, there is a point of diminishing return. A large pay disparity can create tournament inefficiencies as it induces excessive effort on behalf of the agents, for which they must be compensated (Lazear & Rosen, 1981; Wowak et al., 2016).

Second, task complexity can be positively associated with solvers' willingness to participate, but it can also have negative influence once the level of complexity increases to a certain point. An increase in the task complexity can advance the level of challenge and activation, which can stimulate solvers' curiosity and enjoyment (Bendoly, Croson, Goncalves, & Schultz, 2010). Participating in solving complex problems can also strength solvers' social recognition and reputation, send out positive signals about their skills, and escalate their personal advancement (Connelly, Certo, Ireland, & Reutzel, 2011; Leimeister, Huber, Bretschneider, & Krcmar, 2009). However, complex tasks place high cognitive burden on solvers, which might lead to information overload and discomfort to solvers and cause them to lose interest and withdraw their willingness to participate (Zheng et al., 2011), thus slowing down the crowd emergence.

H1c (d): In the context of tournament-based crowdsourcing, crowd growth rate (crowd size) is curvilinear related to task complexity of a crowdsourcing event in an inverted-U shape.

Payment Size and Crowd Emergence. Payment size is an important element of event design when buying companies create crowdsourcing events. It refers to the amount of money specified for a crowdsourced event. Depending on the nature of the task crowdsourced, the payment size varies with a wide range. For instance, statistics show that the payment size for the programming contests hosted by Topcoder between July 2014 and October 2016 fluctuated from zero to $100,000 with a mean of $1,127.83 and a standard deviation of $1,592.26. Both diffusion theory and tournament theory agree that financial payment is an effective mechanism to attract solvers to participate and stimulate the crowd emergence, but they provide opposite implications on the relationship between payment size and crowd emergence.

Diffusion theory suggests that payment size can have a positive effect on the crowd emergence for a particular crowdsourcing event. This is because payment size is positively related to relative advantage of an innovation which is often expressed as economic profitability (Rogers, 2010). Other things being equal, an increase in the payment size can increase the expected return. Studies in the diffusion literature demonstrate that the relative advantage of an innovation is positively related to perceived attractiveness and diffusion rate (Boyd & Mason, 1999). Following this logic, we believe that a crowdsourcing event with a large payment size can stimulate the crowd emergence for this event by increasing the attractiveness of this event to solvers. Moreover, as many new crowdsourcing events occur every day, payment size serves as a signal for the importance of an event. Participating in an event with a large payment size cannot only saves solvers' time in evaluating attractiveness of events, but also increase solvers' social reputations (Connelly et al., 2011; Zheng et al., 2011). For instance, participating in the

Harvard Catalyst's Experiment that is incentivized with $1 million can surely increase a solver's social recognition in the scientific problem-solving community (Guinan et al., 2013; Lakhani et al., 2007).

H2a (b): In the context of tournament-based crowdsourcing, crowd growth rate (crowd size) is positively related to payment size of a crowdsourcing event.

However, tournament theory that is based on agents' utilization maximization thinking suggests that a large payment size can exert a negative influence on the crowd emergence for a crowdsourcing event. This is because a large payment size can potentially attract many agents to participate in a particular crowdsourcing event and increase the competition intensity within a crowd (Bognanno, 2001; Liu et al., 2014). Many research in tournament literature has shown that strong competition for a particular event dilutes the chance of winning and reduces the attractiveness of an event to its potential solvers (Boudreau et al., 2011; Garcia & Tor, 2009), thus slowing down the crowd emergence for an event with large payment size.

H2c (d): In the context of tournament-based crowdsourcing, crowd growth rate (crowd size) is negatively related to payment size of a crowdsourcing event.

Event Length and Crowd Emergence. Another element of event design that firms need to specify before launching a crowdsourcing event is the event length, which refers to the amount of time (in hours, days, weeks, or months) specified for a crowdsourced task. Similar to payment size, event length widely varies, depending on issues such as the nature of the task and time pressure facing the buying firms. For instance, the average length for the programming contests hosted by Topcoder between July 2014 and October

2016 was around two weeks, while the shortest one was a few hours and the longest was almost three months. Depending on the theoretical lens that we draw, event length can have positive or negative implications on the crowd emergence.

From a diffusion perspective, the longer the event length, the better for a crowd to emerge. Looking from the lens of diffusion theory, crowd emergence is a process by which a particular event or task is communicated through a crowdsourcing platform over the event cycle time among the registered online members (Robertson, 1967; Rogers & Shoemaker, 1971). An event with long cycle time can give buying firms more time to broadcast the event to as many solvers as possible and allow solvers to diffuse the information via "word-of-mouth" to other potential solvers (Valente & Rogers, 1995; Van den Bulte & Joshi, 2007). In the meantime, long event length allows agents to search more information and to learn necessary skills, thus reducing the uncertainty associated with problem solving and solution-creation in crowdsourcing. An increase in the event length can then increase the attractiveness of a particular crowdsourcing event to potential solvers, thus facilitating the emergence of a crowd for a particular crowdsourcing event.

H3a (b): In the context of tournament-based crowdsourcing, crowd growth rate (crowd size) is positively related to event length of a crowdsourcing event.

However, tournament theory indicates that an increase in the event length can create a practical situation called problem crowding which means that multiple crowdsourcing events run simultaneously (Piezunka & Dahlander, 2015). For instance, statistics from Topcoder show that there were around eight programming contests, on average, running every day between July 2014 and October 2016. Other things being equal, an increase in the event length can lead to more events running simultaneously, i.e., problem crowding. Problem crowding is challenging and problematic because solvers are limited in terms of their attention and the ability to process information (Hansen & Haas, 2001; Ocasio, 1997). Problem crowding dilutes potential solvers' attention and reduces the perceived attractiveness of a particular crowdsourcing event (Piezunka & Dahlander, 2015). Therefore, long event length is actually detrimental for crowd emergence.

H3c (d): In the context of tournament-based crowdsourcing, crowd growth rate (crowd size) is negatively related to event length of a crowdsourcing event.

Influential Agents and Crowd Emergence. Influential agents in crowdsourcing refer to solvers whose decisions and behaviors influence those of others in the same network. In the social network literature, such people are called opinion leaders, mavens, or hubs, depending on the aspect of influence in question (Van den Bulte & Wuyts, 2007). Scholars in both diffusion literature and tournament literature have been interested in examining the influence of influential agents for a while (Greenhalgh et al., 2004; Rogers, 1962, 2010). In a crowdsourcing context, issues related to influential agents include the number of influential agents involved and the timing that these agents participate in an event.

From a diffusion perspective, influential agents can trigger the imitation mechanism within a crowd (Le Bon, 1897; Roger, 2010). Lower ranking community members aspire to be like influentials and find it useful to resemble powerful leaders.

Participation by influential agents shifts community norm or interaction patterns sufficiently that others might find it hard not to go along (Strang & Soule, 1998). Influential agents involved in a particular crowdsourcing event shows a positive signal of attractiveness of this event to other potential solvers in the same network (Connelly et al., 2011). The more influential agents participate in an event, the more positive influence they can exert on crowd emergence. The earlier influential agents become involved in an event, the sooner they will initiate the imitation of other potential agents. They can also show the positive attractiveness signal earlier, thus increasing the crowd growth.

> *H4a (b): In the context of tournament-based crowdsourcing, crowd growth rate (crowd size) is positively related to the* <u>number</u> *of influential agents participating in a crowdsourcing event.*

> *H5a (b): In the context of tournament-based crowdsourcing, crowd growth rate (crowd size) is negatively related to the* <u>early involvement</u> *of influential agents participating in a crowdsourcing event.*

However, tournament theory suggests that influential agents shrink the winning chance of other potential agents and thus reduce the perceived attractiveness of an event to them, thus slowing the growth of a crowd (Brown, 2011; Garcia & Tor, 2009). The more influential agents are involved in an event, the stronger the negative influence that they impose on other potential agents. The earlier influential agents participate in a crowdsourcing event, the sooner other potential agents realize their diminishing chance of winning due to the participation of influential agents. Under such a situation, the crowd grows slowly. On the contrary, the later influential agents are involved in an event, the less other potential agents perceive the threat from influential agents and have more chance to increase the crowd growth.

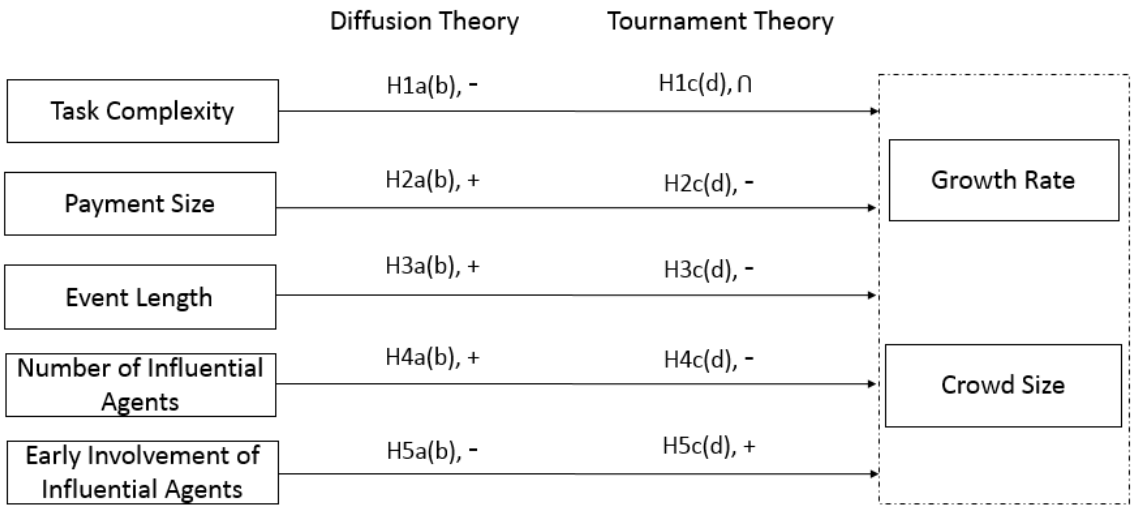
> *H4c (d): In the context of tournament-based crowdsourcing, crowd growth rate is (negatively) related to the <u>number</u> of influential agents participating in a crowdsourcing event.*

> *H5c (d): In the context of tournament-based crowdsourcing, crowd growth rate (crowd size) is positively related to the <u>early involvement</u> of influential agents participating in a crowdsourcing event.*

Summary In this first theory development section, we develop hypotheses on the influence of event design on crowd emergence (i.e., crowd growth rate and crowd size) by drawing from diffusion theory and tournament theory. The following Figure 1 summarize the theoretical model that we propose in this theory development. As indicated by our proposed theoretical model, diffusion theory and tournament theory offers different predictions on the relationships between crowd attributes and crowd emergence. Specifically, diffusion theory mainly offers a contagion view based on individual participant's considerations of the crowdsourcing task – i.e. ease of completion, task attractiveness, and participation of others. Tournament theory provides a competition view that is based by individual participants' evaluation of the chance of winning and the expected returns. Empirical testing on this theoretical model will be addressed in the following chapters.

Figure 9

Theoretical Model for Crowd Emergence



Understanding the Performance Implications of Crowd Attributes

In this section, we develop hypotheses on the relationship between crowd attributes and crowd performance. We focus our performance on crowd productivity and crowd efficiency. The following Table 3 lists all the constructs used in this theory development. As indicated by our literature review in chapter 2, both innovation search and tournament literature offer explanations on how crowd attributes relate to crowd performance. Specifically, innovation search view takes the perspective that crowd performance depends on the extensiveness and effectiveness of recombinant search, which, in turn, relies on changes in crowd attributes (Afuah & Tucci, 2012; Fleming & Sorenson, 2004; Laursen & Salter, 2006). On the contrary, tournament theory argues that crowd performance hinges upon individual participants' competition behavior which is contingent on factors such as chance of winning, expected returns, and competitive social comparison within crowd members (Che & Gale, 2003; Boudreau et al., 2011; Bothner et al., 2007). These two explanations are different to each other. We thus take a strong inference approach (Davis, 2006; Platt, 1964) and develop competing hypotheses on how crowd attributes relate to crowd productivity and crowd efficiency in crowdsourcing in the following sections.

Table 3

Constructs in Theory Development for Performance Implications of Crowd Attributes

Constructs	Definition	Reference
Crowd performance	The quantitative outcomes of a crowd in crowdsourcing	Cohen & Bailey (1997)
Crowd productivity	The amount of outputs produced by a crowd in crowdsourcing	Horwitz & Horwitz (2007)
Crowd efficiency	The relative speed with which a crowd solves a particular crowdsourced task	Horwitz & Horwitz (2007)
Crowd size	The total number participants for a particular crowdsourcing event	Liu et al. (2014)
Crowd diversity	The extent of differences among crowd members in terms of background and demographic statistics	Harrison & Klein (2007)

Crowd Size and Crowd Performance. Scholars conceptualize crowdsourcing as a solution to distant search over a landscape (Afuah & Tucci, 2012). Landscapes offers a useful heuristic for thinking about the space that firms must search when attempting to discover solutions for their crowdsourced tasks (Afuah & Tucci, 2012; Fleming & Sorenson, 2004). Search includes two dimensions: search depth and search scope. Search breadth is defined as the number of external sources or search channels that firms reply on in their innovative activities, while search depth refers to the extent to which firms draw deeply from the different external sources or search channels (Laursen & Salter, 2006). Combing landscape with a search algorithm allows us to make predictions regarding the likely outcomes of a crowd in crowdsourcing.

From an innovation search perspective, an increase in the crowd size has positive implications on crowd productivity and crowd efficiency. First, a large crowd size increases search scope of a problem-solving attempt; that is, the number of external sources or search channels that firms rely upon in their innovative activities or problem solving (Laursen & Salter, 2006). Search with high scope enriches the knowledge pool by adding distinctive new variations. For instance, solvers can use computer language such as JavaScript, HTML, C+, or C++ to design a website. An increase in the crowd size for a programming design contest means that firms can find solvers with different combinations of these skills. New variations are necessary to provide a sufficient amount of solutions for problem-solving (March, 1991).

Second, a large crowd size increases the number of solutions through enhancing the recombination of different searches (Fleming, 2001; Fleming & Sorenson, 2004). There is a limited number of ideas or solutions that can be created by using the same set of knowledge elements (Katila & Ahuja, 2002). An increase in the search scope adds new elements to the solution landscape and increases the chances to find new and useful combinations (Katila & Ahuja, 2002; Laursen & Salter, 2006). Third, an increase in the crowd size can increase the chance for firms to disclose a problem or a task to agents who possess different problem-solving skills and heuristics (Jeppesen & Lakhani, 2010). The idea that differences in perspective and heuristics are the sources of problem solving or solution design has been explored not only in innovation search literature but also at the intersection of economics and behavioral theory of the firm literatures (Dosi, Levinthal, & Marengo, 2003; Fleming & Sorenson, 2004; Nelson & Winter, 2009). A main insight is that multiple sources of perspectives and heuristics contribute to effective and efficient problem-solving (Jeppesen & Lakhani, 2010; Marengo, Dosi, Legrenzi, & Pasquali, 2000)

H1a (b): In the context of tournament-based crowdsourcing, crowd size is positively related to crowd productivity (crowd efficiency).

The tournament perspective suggests that an increase in crowd size can have both positive and negative effects on crowd performance, which means that the relationship between crowd size and crowd performance is not linear but quadratic. An increase in the number of participants, i.e., crowd size, stimulates the competition intensity within a crowd. Research in tournament literature shows that an increase in competition intensity motivates agents to invest more effort in a particular competition (Che & Gale, 2003; C. Harris & Vickers, 1987), which can positively relate to a high crowd performance. In an innovation contest with only one participant, this contestant will have little incentive to exert effort to improve his/her performance because there is no competition. Thus, some level of competition through adding contestants will lead to greater effort to improve overall performance (Harris & Vickers, 1987). Studies on organizational-level competition confirm that threats from frontier entrants induce incumbents in sectors that are initially close to the technology frontier to innovate more, and this triggers productivity growth (Aghion, Blundell, Griffith, Howitt, & Prantl, 2009; Aghion, Harris, Howitt, & Vickers, 2001).

However, excessive crowd size can have negative consequences. Literature identified two negative effects of excessive crowd size: shrinkage of winning and motivation crowd-out. We argue below that these negative effects of size at some point exceed the benefits and, thus, the relationship between crowd size and crowd performance is actually nonlinear. As the number of participants increase, the chance of winning for each contestant shrinks, reducing the motivation effects of competition on agents' efforts in problem solving (Boudreau et al., 2011; Garcia & Tor, 2009). A large increase in crowd size induces excessive effort on the participants in order for them to win a tournament-based crowdsourcing event. As solvers keep increasing their efforts, they may lose interests and enjoyment and hence withdraw their participation (Zheng et al., 2011).

H1c (d): In the context of tournament-based crowdsourcing, crowd size is curvilinear related to crowd productivity (crowd efficiency) in an inverted-U shape.

Crowd Diversity and Crowd Performance. From an innovation search perspective, an increase in crowd diversity means an expansion of search depth in a problem-solving process, which can allow firms to access diverse knowledge sources and approaches (Fleming, 2001; Fleming & Sorenson, 2004; Katila & Ahuja, 2002). An increase in search depth also allows firms to search over a rugged landscape and locate solvers who are "marginal" and possess alternative knowledge and approaches that may be amenable to an effective solution (Jeppesen & Lakhani, 2010). Scholars in the search literature say that agents who are marginal in terms of technical expertise and social establishment are useful for effective and efficient problem solving. This is because these agents are not burdened by prior assumptions and they approach problems with different perspectives and heuristics (Fleming & Sorenson, 2004; Jeppesen & Lakhani, 2010).

They have a unique way of problem-solving called “focused naïveté”, which refers to a useful ignorance of prevailing assumptions and theories that allows them to attack problems generally regarded as impossible or uninteresting by specialists (Gieryn & Hirsh, 1983). To illustrate this, Howe (2008) describes a firm that faced an in tractable chemical engineering problem that had stymied progress for a significant period of time. The ultimate solution came from an external physicist who applied principles of electromagnetism to what was thought to be a chemistry issue (Howe, 2008; Jeppesen & Lakhani, 2010).

H2a (b): In the context of tournament-based crowdsourcing, crowd diversity is positively related to crowd productivity (crowd efficiency).

However, tournament perspective argues that competition is a manifestation of social comparison (Festinger, 1954; Garcia, Tor, & Gonzalez, 2006). An increase in crowd diversity can reduce the extent and the amount of social comparison between crowd members (Bothner et al., 2007; Festinger, 1954), which can be detrimental to crowd performance. The social comparison between participants in a tournament determines agents’ problem-solving efforts and subsequent behaviors (e.g., submission or withdrawal) in a tournament (Bothner et al., 2007). An increase in crowd diversity reduces similarity between crowd members and causes them to lose target for social comparison in a competition (Festinger, 1954; Garcia et al., 2013; Suls, Martin, & Wheeler, 2002). This is because participants in a competition have a tendency to choose a reference person who is close to their own characteristics (Garcia et al., 2013; Goethals, 1986). Empirical evidence from the tournament literature has shown that an increase in homogeneity among competitors increases aggregate effort and overall performance (Konrad, 2009).

The above line of reasoning is consistent with the work on structural equivalence and relative deprivation in social network and sociology literature. Structural equivalence refers to the extent to which the agents (i.e., notes) in a network are similar to each other in terms of network connections, behavioral and demographic statistics (Burt, 1982; Lorrain & White, 1971; Sailer, 1978). Relative deprivation refers to a psychological state that occurs when agents feels that their personal attainments are below their expectations (Forsyth, 2009). Burt (1982) argued that relative deprivation is concentrated between structurally equivalent agents and that the feeling of relative deprivation is most acute precisely when a peer has moved ahead of him or her. Sociologists also believe that agents who feel that their status quo is violated or challenged are more likely to invest more effort in improving their performance (Bothner et al., 2007; Forsyth, 2009). Thus, we believe that an increase in the crowd diversity reduces agents' structural equivalence and relative deprivation, which will demotivate them to invest effort in problem-solving and might lead to poor performance.

H2c (d): In the context of tournament-based crowdsourcing, crowd diversity is negatively related to crowd productivity (crowd efficiency).

Combination of Crowd Size and Crowd Diversity. The above hypotheses focus on the distinct effects of size and diversity on crowd performance. In this section, we propose that these variables have interactive effects. Depending on the literature that we draw, the interaction effect could be positive or negative.

From a search perspective, we argue that crowd size and crowd diversity are mutually beneficial to crowd performance. We suggest two mechanisms that underlie this positive interaction: the rugged solution landscape and uniqueness of combination. Crowd size and crowd diversity are related to each other but not necessarily in a proportional manner (Harrison & Klein, 2007). When they both increase simultaneously, they can greatly magnify the ruggedness and complexity of the solution landscape (Gavetti & Levinthal, 2000; Rivkin, 2001). A rugged solution landscape can allow firms to conduct both local search and distant search and enjoy as many options as possible (Fleming, 2001; Stuart & Podolny, 1996). A combination of depth and scope search can increase the uniqueness of recombination, which can intensity the positive effects on crowd performance due to an increase either in crowd size or crowd diversity (Katila & Ahuja, 2002). By combing agent-specific accumulated understanding of certain knowledge elements (depth) with a large number of agents (scope), the crowd is more likely to generate new, unique combinations that can be commercialized (Winter, 1984).

H3a (b): In the context of tournament-based crowdsourcing, crowd size and crowd diversity positively interact to influence crowd productivity (crowd efficiency).

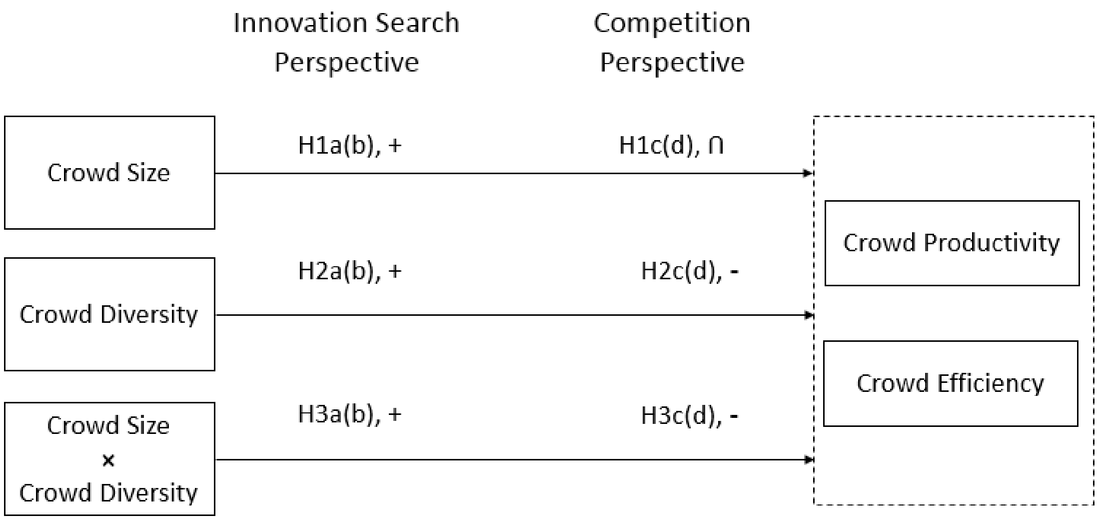
From a tournament perspective, we believe that crowd size and crowd diversity can jointly influence crowd performance in a negative direction. We argue that the mechanisms underlying this negative interaction include the attenuation of winning chance and a loss of social comparison. Increasing the number of participants in a competition not only reduces the chance of winning for each participant but also increases the uncertainty about who might win this competition (Boudreau et al., 2011; Garcia & Tor, 2009). Under such a situation, empirical research has shown that all participants reduce their effort, causing the overall performance to shift down (Boudreau, Lacetera, & Lakhani, 2010; Boudreau et al., 2011).

An increase in both crowd size and crowd diversity can further reduce the similarity of rivals (i.e., solvers) in a crowd for a particular crowdsourcing event (Garcia et al., 2006; Garcia et al., 2013). Research in this line of thinking has already shown that a reduction of similarity among agents causes them to lose propensity to engage in social comparison and thus invest less effort in problem solving (Gibbons & Buunk, 1999). An increase in both crowd size and crowd diversity can also further attenuate agents' relative deprivation and structural equivalence (Burt, 1982; Lorrain & White, 1971; Sailer, 1978), causing participants in a contest to lose motivation to compete and shift down the overall performance.

> *H3c (d): In the context of tournament-based crowdsourcing, crowd size and crowd diversity negatively interact to influence crowd productivity (crowd efficiency).*

Figure 10

Theoretical Model for the Performance Implications of Crowd Attributes



Summary. In this section, we develop hypotheses for the performance implication of crowd attributes in crowdsourcing by drawing from both innovation search literature and tournament theory. Figure 10 summarizes the theoretical model that we propose in the second theory development. The innovation search literature indicates that an increase in crowd attributes (e.g., crowd size and/or crowd diversity) can be positively related to the extensiveness and effectiveness of problem-solving search which, in turn, may lead to positive crowd performance. However, the competition perspective based on tournament theory suggests that an increase in crowd attributes (e.g., crowd size and/or crowd diversity) can reduce either individual participants' chances of winning as well as the expected returns or the competition social comparison among participants, which might lead to negative outcomes. We will empirically test these two perspectives in the following chapter.

Chapter 5: Methodology

Method Design

An ideal empirical setting needs to satisfy a number of nontrivial requirements. One requirement is the temporal consideration for the crowd-development process. Each competition-based crowdsourcing event has a unique cycle time with specific start date and end date. Accordingly, the crowd-development process embodies temporal consideration, which means that we need to take the whole process into consideration when we calculate the growth rate and crowd size for a particular crowdsourcing event. The second requirement is the solvers' precise trace (i.e., participation time and submission time) over the crowd-development cycle time. Solvers' participation and/or submission behaviors can occur at any time during the event cycle time. These first two requirement make cross-sectional survey impossible to capture solvers' precise trace over the event cycle. We thus rule out survey as a method option for this dissertation.

The third requirement is the multi-level information, that is, information at the crowd level (e.g., elements of event setting and crowd performance) and at the individual solver level (e.g., demographic statistics, participation and/or submission behaviors). The final requirement would be the sample size. Through a pre-power analysis via G-power package, we found that the sample size should be at least 150 if we wanted to achieve a medium effect size with 80 percent power. The required sample size would be larger if wanted to detect a small effect size with a high power. The unit analysis for this research is a crowd, and one crowdsourcing event provides only one observation. The large sample size requirement makes lab experiment unrealistic for our empirical testing.

We propose to use secondary data in this dissertation because this method not only meets all the above requirements but also has many obvious advantages that are missing in the other empirical methods such as survey and experiments. For instance, there are many crowdsourcing intermedia (i.e., Topcoder, InnoCentive, and Kaggle) that are specialized in organizing competition-based crowdsourcing contests for companies. These intermedia are very good sources for secondary data (Archak, 2010; Boudreau et al., 2011; Jeppesen & Lakhani, 2010). Table 4 summarizes the main advantages recognized by scholars in the supply chain literature. Because of these advantages, many scholars argue that secondary data can expand academic horizons and deepen the

Table 4

Advantages of Secondary Data Methodology

Relatively large amount of data available
Relatively low amounts of resources (money and time) necessary for data collection
Limited chance to bias in the data collection process due to researchers' preconceptions
Higher internal validity of studies due to measurements and statistics inferences constructed by the third-party and derived from less biased database
Greater opportunity for replication when data is publicly available

Source: adapted from Ellram and Tate (2016), Rabinovich and Cheon (2011)

understanding on the social phenomena and thus are actively calling for research in using secondary data (Ellram & Tate, 2016; Rabinovich & Cheon, 2011). The context and data we describe below allow for a possibility to rely on a real crowdsourcing setting in a natural environment that is characterized by the availability of empirical measures, appropriate identification, and external validity.

Research Setting - Topcoder

Established in 2001, Topcoder creates outsourced software solutions for IT-intensive organizations by specializing in organizing programming crowdsourcing events and soliciting independent solvers (i.e., programmers) from all over the world to compete in programming design and/or software development contests. Topcoder's value proposition to its clients is that it can harness the wisdom of a large number of professional programmers and let the competition determine the best solutions without risking either a wrong hiring or an incorrect solution (Boudreau et al., 2011). Since 2001, Topcoder has served clients such as Best Buy, Eli Lilly, IBM, and GEICO. From 2001 to 2009, Topcoder added an average of 25,000 new computer programmers to its community each year (Lakhani et al., 2010). As of May 2017, over 1 million solvers have registered at Topcoder's website (Topcoder, 2017b). These solvers have the opportunity to win cash prizes, obtain assessments of their skills, and signal their potential in a global sharing economy through participation in thousands of crowdsourcing events (Boudreau et al., 2011; Boudreau & Lakhani, 2013).

Case studies on Topcoder show that Topcoder used to run two types of crowdsourcing competitions on its platform: algorithm and client software development (Archak, 2010; Lakhani et al., 2010). The algorithm competitions served as the primary means for attracting new programmers and retaining existing community members, while the software development competition was targeted at developing software applications for Topcoder's clients (Lakhani et al., 2010). Our latest community review shows that Topcoder nowadays mainly organize two types of crowdsourcing contests: design events and development events (Topcoder, 2017a). Both types of events have very similar structure in which Topcoder or focal buying firms specify elements of event design at the beginning and solvers (i.e., agents) are supposed to convert the specified requirements into usable software (Boudreau et al., 2011; Boudreau, Lakhani, et al., 2016). One notable difference notified by Archak (2010) is that “winning design submissions go as inputs into the development events in which agents are required to submit actual code implementing the provided design” (Archak, 2010, p. 22). We did not find this notable difference in our data, but we found the descriptive statistics of some variables are different between these two types of crowdsourcing event. For instance, the average payment size for the design events is significantly larger than that of develop events, while the average event length of develop events is longer than that of design events.

The general process for these two types of events is essentially similar. TopCoder works with its clients to identify software needs and converts identified needs into design, development, or data science related contests of its community of programmers (Lakhani et al., 2010). According to these needs, TopCoder specifies the detailed requirements for programing design/development, the length of event (i.e., start date, end date, and winner announcement), payment size, payment structure, and evaluation criteria. Before events go alive, Topcoder post them on the “coming events” section for programmers to review. After events go alive online, programmers view the details, evaluate the attractiveness of the contests, and decide whether they compete for an event. Depending on the situation, TopCoder or its clients provide feedback to programmers in their problem solving process. Programmers submit their solutions for a particular programming event before its deadline.

All the software design and development events organized by Topcoder are run in a tournament format. The nature of the events creates a limitation in our data analysis – specific type of crowdsourcing event – but does not deter the generalizability of conclusion. This is because the tournament format of programming events is applicable to all the other competition-based crowdsourcing events. Topcoder uses an open call to attract programmers nested in its online community to compete in different programming events. The companies that sponsor the crowdsourcing events include not only Fortune 500 companies such as Eli Lilly, IBM, and Best Buy but also many small and medium size companies like Mediafly. These companies represent industries such as pharmaceutical, information technology, retailing, electronics, insurance, and others (Boudreau et al., 2011).

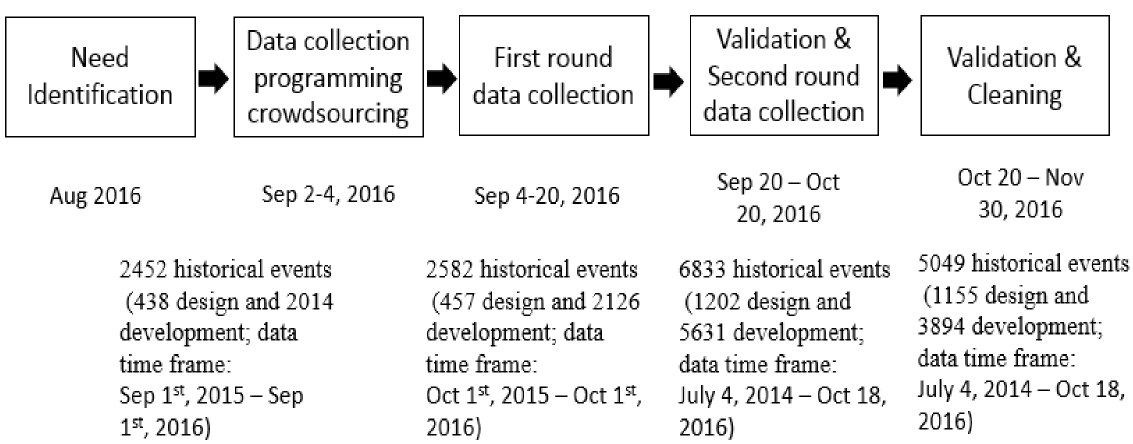
Scholars in the operations and supply chain management field, such as Boudreau and his coauthors, have justified the validity of Topcoder's archival data and used the data between 2001 and 2007 to study the performance implication of tournament design (e.g., number of participants) (Boudreau et al., 2011; Boudreau, Lakhani, et al., 2016). Other scholars in the computer science literature have also used Topcoder's data between 2001 and 2013 to examine programmers' participation and their performance implications. We thus believe that the secondary data from Topcoder is representative.

Data Collection

Topcoder makes its archival data publically available on its website. We used web crawling to automate the assembly of Topcoder's historical programming contests. Web crawling is "the systematic, automated navigation of a series of internet-based references" (Massimino, 2016, p. 35). Using web-crawling techniques to collect research data is relatively new to scholars in the operations and supply chain filed (Massimino, 2016), but this technique is popular in the computer science and information systems literature (Dissanayake et al., 2015; Javadi Khasraghi & Aghaie, 2014). We hired a professional web crawler for the data collection in this dissertation. This data collection process follows the following steps to ensure validity in the data collection process. The following Figure 11 summarizes the data collection and cleaning process.

Figure 11

Data Collection and Cleaning Process



First, we searched Topcoder's website that listed past challenges and identified the specific number of both design and development events. Through this search, we understood the structure of a programming contest (e.g., challenge details, payment size and structure, participants, and results). This understanding was useful when we described our data collection requests to our web crawler in the following step. We started our data collection on Sep 2, 2016. At that time, we found that there were 2,452 historical events (438 design events and 2,014 development events) available on Topcoder website according to the one-year default filtering setting (i.e., Sep 1, 2015 – Sep 1, 2016).

Second, based on our understanding of the structure of an event and our expected needs for data, we developed a data request proposal that described our specific, detailed needs for data collection. In this document, we also described how to access Topcoder's historical data step by step. Since each programming contest was one observation, we specified that all crawled data related to one event should be saved separately in an Excel file for validation purpose. We asked one graduate student in the Accounting Department at W.P.Carey School of Business who had data crawling experience to check the clarity of our proposal.

Third, we crowdsourced our data collection proposal to ten potential Chinese crawlers and identified a qualified crawler. We targeted only Chinese web crawlers due to cost and response considerations. For instance, the price for developing a data crawling program was between RMB800 and RMB1200 (i.e., between $115 and $175), which was cheaper than many crawlers in the US. The lead time for designing program was around one week, and the lead time for program and data collection was around one month. We first identified ten potential suppliers through a Chinese e-commons website (www.taobao.com) based on their online reputation scores and transaction records. We contacted them one by one by sending them our data collection proposal. Three out of ten showed interest in our data collection process. One of three accepted our offer (RMB1100 (i.e., around $160) and two weeks leading time) on Sep 4, 2016.

Fourth, we cross-validated the accuracy our collected data in multiple times. We got our first round of data on Sep 20, 2016. We collected 2,582[5] observations (457 design events and 2,126 development events). We randomly selected 30 observations from our data and compared our collected observations with the original ones on Topcoder's website. This comparison took us almost one month and helped us identify two potential issues. We found that we could customize the time framework filter by tracing backing to July 2014 and that Topcoder's server actually hosted around 6,807 historical events available (1,196 design events and 5,611 development events). The one-year default setting made us miss many observations available on Topcoder's server. Another big issue we found was that we forgot to specify event cycle time (i.e., start time, end time, feedback time, and winner announcement time) in our data collection proposal. Except the cycle time issue, we found that all the other information in our crawled data matched exactly with the observations on Topcoder's website.

We then reported these two issues to our crawler to update our requests on Oct 20, 2016. Since we changed our data collection needs, our crawler charged us another RMB800 (i.e., around $115). We maintained communication during the second round of data crawling process to make sure these two issues were properly taken into consideration. Our crawler finished the second-round data collection on Nov 1, 2016. We got 6,833 observations (1,202 design events and 5,631 development events). In this round, we randomly selected 50 observations from our crawled data and manually compared them with the correspondent observations on Topcoder's website. All the

[5] The discrepancy between 2,582 and 2,452 (our previous identified number of observations) was due to the two-week difference between step one and step four.

information matched well except a minor issue. We found that the time format for some observations was not consistent. Specifically, all the time format shown on Topcoder's website was in "EDT" format (i.e., Eastern Time). In our crawled data, some format was shown as "EDT", but the time format for around 35 percent of our observation was shown at "00.000-05:00" which is equivalent to "EDT". For instance, "2014-11-30T09:00:00.000-05:00" is equivalent to "2014-11-30T09:00EDT". This minor issue did not influence our data analysis (i.e., calculating the correct event length). We thus did not ask our crawler to update our second crawled data.

Data Cleaning

We cleaned the data in several ways. We first removed eight test events created by Topcoder to check its internal system. There were no participants for these test events which were titled with words such as "Test Event" or "Do not register". We had 6,825 useable observations in total (i.e., 6,833 - 8 = 6,825). We then reviewed our observations in detail and removed the observations that were kept "confidential". The detailed descriptions related to these "confidential" events were not publically available on Topcoder's website. Programmers had to register and sign confidentiality agreements before they could view the details of these events. We thus missed the text information to evaluate the complexity of these events.

We totally identified 1,776 confidential observations (i.e., 26 percent of 6,825 useable observations). Because 97.24 percent of the confidential observations came from the development event category, we suspected that the missing data on the confidential events was caused by the avoidance of information leakage in the new product development process. This reason was completely out of the control of our methodology design and data collection. We thus assumed that this missing was completely random and took the 1,776 confidential observations out of our usable data (Cohen, Cohen, West, & Aiken, 1983; Cohen, Cohen, West, & Aiken, 2013). We eventually got 5,049 effective observations (1,155 design events and 3,894 development events) which had 161,735 total participation records from 21,741 community programmers.

Chapter 6: Data Analysis

Data Analysis for the Influence of Event Design on Crowd Emergence

Data Description

We used the design events collected from Topcoder to analyze the relationship between elements of event design and crowd emergence. We chose to use only design events for two practical reasons. First, all design events were organized in competition-based crowdsourcing format. For all these design events, solvers from Topcoder's network competed with each other to provide the best solutions and were rewarded according to the relative rankings of their solutions for a particular event. We thus believe that design events are representative for crowdsourcing events. Second, design events have more complete information compared to the category of crowdsourcing events from Topcoder (i.e., development events). Design events are publicly open to all solvers, which means that we could get complete information about design events. However, we found that many development events were kept confidential, which means that some important information was not available for web-crawling.

All events were saved in a separate csv file with unique event ID number. We used Python, a widely used computer programing language for data processing (Shaw, 2013), to extract and pre-process the raw data that are saved in csv files. The unit level of our data extracting and pre-processing is at event level. First, we extracted all solvers' participation time for a particular design event by using trace extraction code (Appendix A) to build a crowd growth trajectory over the cycle time of this event. Second, we fitted a Bass Diffusion Model to each crowd growth trajectory to identity the two growth parameters (i.e., *p* and *q*). This process was executed over R software by using regression codes attached in Appendix B. It took us almost one month to process the total design events collected from Topcoder. Third, we extracted the description information (i.e., text information) from our raw files to analyze the difficulty of each design event by using text extraction code (Appendix C and D). Fourth, we extracted the basic statistics for each solver for each event (e.g., ID, membership registration data, country origin, and winning records) through extraction codes attached in Appendix E. Finally, we extracted the basic summary information for event level (e.g., event ID, event start date, event ending date, payment size, number of payments, and feedback) through codes attached in Appendix F.

We collected 1,202 design events from Topcoder. Among them, 1,154 were effective observations. During the process of fitting the linear Bass Model (Bass, 1969) to empirical crowd growth trajectory, we found that not all regression models converged. Specifically, 734 out of 1,154 events converged with significant fit, yielding a 63.60 percent effective observations with meaningful and comparable crowd growth rate. We compared the mean differences of each variables between convergence group and non-convergence group through ANOVA analysis. We did not find significant mean difference between these two groups on crowd size and main independent variables (e.g., Fog Readability Index, event length, payment size, and number of influential agents). We assumed that these 734 effective observations are representative for our data analysis and justified this assumption in the robust check sections.

Measurement

Dependent Variables. There are two dependent variables in this analysis: crowd growth rate and crowd size. *Crowd growth rate* is defined as the relative speed with which solvers from a social network form a solution-providing crowd for a particular crowdsourcing event by self-selecting to participate in an event (Rogers, 1962, 2010). We operationalized this crowd growth rate by using the Bass Diffusion Model developed by Frank Bass (1969) to quantitatively measure the growth rate of new product adoption within a social network (Bass, 1969). This model has three basic parameters: innovation coefficient p (i.e., the coefficient for solvers to participate in a particular event due to the influence coming from event itself); imitation coefficient q (i.e., coefficient for solvers to follow other agents' decisions to participate a particular event); and potential market size m (i.e., potential crowd size for a particular crowdsourcing event). Crowd growth rate is operationalized by the sum of innovation coefficient and imitation coefficient, i.e., "$p + q$" (Bass, 1969; Mahajan, Muller, & Bass, 1991). *Crowd size* defined as the number of participants for a crowdsourcing event at the time when the event reaches its deadline (Liu, Yang, Adamic, & Chen, 2011).

Independent Variables. *Task complexity* is defined as the perceived difficulty of a crowdsourced task (Rogers, 2010). We measured this construct by two indicators: number of words and Fog Readability Index. Number of words means the length of task description for a crowdsourcing event (Haas et al., 2015). Fog Readability Index indicates the extent to which a verbal description to a crowdsourcing task is perceived to be difficult, also referred to simply as FOG index (Collins - Thompson & Callan, 2005;

Li, 2008). The FOG index, developed by Robert Gunning, is a well-know and simple formula for measuring text complexity in computational linguistics literature (Gunning, 1969). Assuming that the text is well formed and logical, the FOG index captures text complexity as a function of the number of words per sentence and the number of syllabus per word to crease a measure of readability (Li, 2008). It is calculated as follows:

$$Fog = (word\_per\_sentence + percet\_of\_compex\_words) * 0.4,$$

Complex words are defined as words with three syllables or more. The FOG index indicates the number of years of formal education of a reader of average intelligence would need to read the text and understand that piece of writing with its word-sentence workload (Li, 2008). The relationship between the FOG index and reading ease is as follows: FOG ≥ 18 means that the text is unreadable; 14-18, difficult; 12-14, ideal; 10-12, acceptable; 8-10, childish.

Event length is defined as the amount of time specified for a crowdsourcing event (in hours). *Payment size* is the amount of money specified for a crowdsourcing event (in dollars). *Influential agents* are agents whose decisions and/or behaviors influence those of the others within a same network (Rogers, 2010). Specifically, we defined influential agents as solvers with above average membership length (in months) and at least one winning record. We used membership length as one of main indictors for being influential because research shows that membership length is a significant predictor for being a winner in a tournament-based crowdsourcing event (Bockstedt et al., 2016). We operationalized the influence of influential agents through two indicators: the number of influential agents involved for an event and the early involvement of influential agents defined as the time difference between the event start time and the participation time of the first influential agent.

Control Variables. The purpose of including control variables in empirical study is to increase the power to detect the significance of variables in interests (Becker, 2005; Spector & Brannick, 2011). In this study, we control for two potential influences from a crowdsourcing event: number of payments and checkpoint. *Number of payments* is operationalized as the number of payments specified for a crowdsourcing event (Chen et al., 2011). We observe that many events in our sample have exact the same total payment size but different numbers of payment. Under this situation, the pay gap between different ranks of a tournament (e.g., first prize, second prize, and/or third prize) will be smaller for event with multiple payments. Tournament theory suggests that when the pay gap is small, suppliers are not motived to compete and are less likely to participate in crowdsourcing (Knoeber & Thurman, 1994; Lazear & Rosen, 1981). Number of payments specified for an event could contaminate the effect of payment size in a tournament-based crowdsourcing. We thus control the potential influence of the number of payments in our data analysis.

Checkpoint is defined as a binary variable to capture whether feedback is provided to solvers during crowd formation process. We observed in our data is that not all events were provided feedback on the progress of problem-solving during the crowd formation process. Latest results from filed experiments on innovation tournaments shows that in-process feedback is associated positively with agents' participation (Wooten & Ulrich, 2017). Feedback to participants thus could contaminate the growth of a crowd and final crowd size. We used checkpoint to control for the potential influence of feedback on crowd emergence: 1 means that feedback is provided; zero otherwise.

Measurement Validity

The data used in this research for empirical analysis comes from a reliable crowdsourcing platform company (i.e., Topcoder), whose primary business is organizing programming crowdsourcing events for its customers such IMB, Eli Lily, and Best Buy (Archak, 2010; Lakhani et al., 2010). All events are organized in competition-based crowdsourcing format. Information related to constructs such as crowd size, payment size, event length, number of payment, and checkpoint (i.e., feedback control) comes from objective crowdsourcing events. The content validity of these constructs are thus satisfactory. We operationalized crowd growth rate through regression and calculated task complexity (i.e., number of words and Fog Readability Index) through text analysis package in Perl. We also operationalized influential agents indirectly through calculating the membership length. We justified the validity of our measurement for these three constructs in the following sections.

Inspired by Bass' (1969) work, the Bass Diffusion Model has been adopted by scholars from different disciplines to model the growth rate of innovation diffusion, information diffusion, and technology adoption at the society level (Mahajan et al., 1991; Rand et al., 2015; Van den Bulte, 2000). In this research, we conceptualized the crowd formation process as a process through which a particular crowdsourcing event diffuses across solvers from a social network. This conceptualization allowed us to operationalize crowd growth rate by using the Bass Diffusion Model. In this research, we compared three most common approaches to operationalize the Bass Model (Bass, 1969; Boswijk & Franses, 2005; Van den Bulte, 2002).

As shown in Table 5, each approach has unique advantages and disadvantages. After this comparison, we chose linear regression for this research for two main reasons.

Table 5

Comparison between Different Operationalization of Bass Model

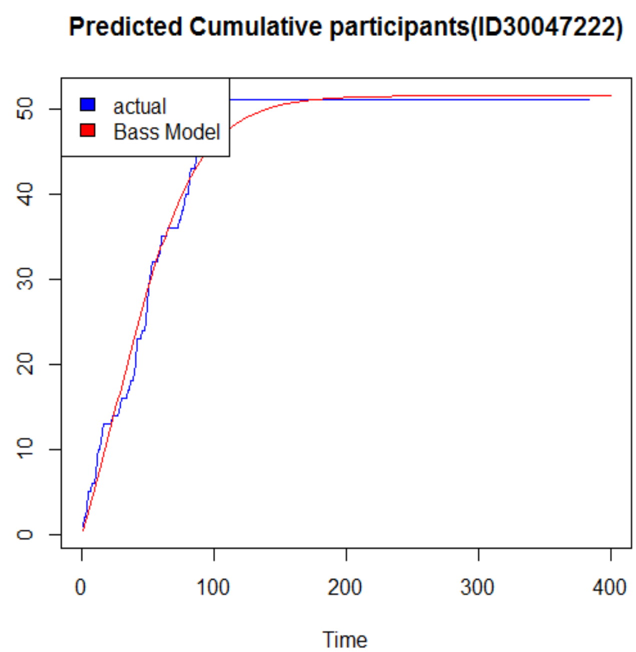
	Model Formation	Advantages	Disadvantages
Linear Regression	$S(T) = pm + (q-p)Y(T) - q/m[Y(T)]^2$ Discrete Analogue: $S_T = a + bY_{T-1} + cY_{T-1}^2$, T=2, 3, … where: S_T, number of agents at T, Y_{T-1}= cumulative participants at T-1. $m = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$, $p = \frac{a}{m}$, $q = b + \frac{a}{m}$	Simple; provide good fit; discrete form	$b^2 - 4ac \geq 0$, otherwise, no solution
Non-Linear Regression	$Y(T) = mF(T) = m(\frac{1-e^{-(p+q)T}}{1+\left(\frac{q}{p}\right)e^{-(p+q)T}})$ $Y(T)$, cumulative participants at time T; $F(T)$, cumulative probability density at time T.	Continuous form; offer better fit	Problems with convergence
Agent-based Simulation	$p(T)$=$\frac{f(T)}{1-F(T)} = p + q/mY(T)$ $p(T)$, probability of participation at time T; $f(T)$, the likelihood of participation at time T	Capture emergence well; new	Assume m = observed “m”.

Source: Bass (1969), Jain and Rao (1990), Rand et al., 2015

First, this approach provides acceptable estimation of growth rate. The amount of regression effort due to large sample size in this research makes this simple approach attractive. Second, the discrete form of linear regression captures the data collection process (i.e., number of participants per each hour). By using this linear regression, we correctly captured the growth rate for many crowdsourcing event with good fit indices. Event 30047222 was a good example. As shown in Figure 12, the fitted Bass curve (in red) captured the growth trend of real crowd trajectory (in blue) and the actual crowd size well (i.e., R-square: 0.27; F-value: 70.72, p<0.001; growth rate (i.e., "$p + q$"): 0.034; predicted crowd size: 52). However, we ran into a convergence issue that caused us to lose 467 observations. We addressed this shrinkage of sample size in the robust check section.

Figure 12

Linear Bass Model Fit for Event 30047222



We adopted Lingua::EN:Fathom package coded in Perl language to analyze text complexity. This package is a well-established tool to analyze the complexity of text information in linguistics, communication, and accounting literature (Collins -

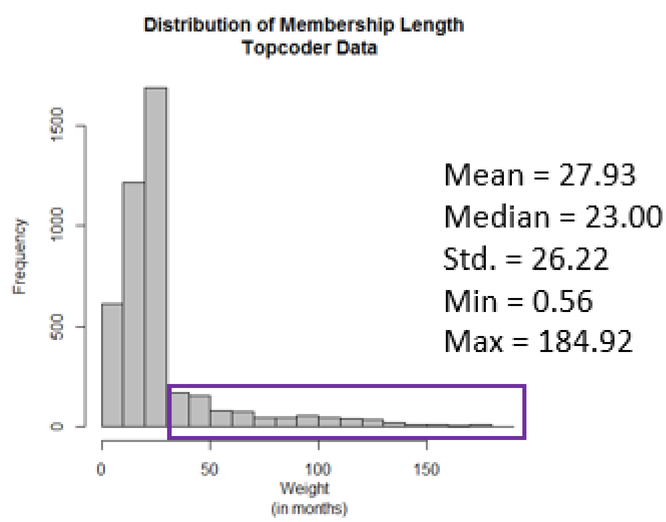
Thompson & Callan, 2005; Li, 2008; Muresan, Cole, Smith, Liu, & Belkin, 2006). Basic description of the Lingua::EN:Fathom package was attached in Appendix D. To our best knowledge, this study will be the first to introduce this Lingua::EN:Fathom package in supply chain management literature.

In order to increase the validity of this package, we conducted a pilot test by comparing the differences on the evaluations of two texts from the Lingua::EN:Fathom package and from undergraduate students. The texts that we used for this pilot test are two cases from a textbook on supply chain management (Johnson & Flynn, 2015). 86 of 99 students who major in global logistics or supply chain management from an undergraduate class participated in this pilot test. Students were asked to evaluate the difficulty of understanding the content and analyzing the two selected cases on a 5-Likert scale. The average student evaluations for these two cases were 2.89 and 3.42, respectively. That is, the second case was 18 percent more difficult than the first one. An ANOVA test confirms that the means of students' evaluations are significantly different. The FOG indices (i.e., text difficulty) generated by the Lingua::EN:Fathom package for these two cases were 15.88 and 19.53. The FOG index of the second case was 23 percent more than that of the first case. The evaluation differences between these two approaches are comparable, demonstrating the validity of the Lingua::EN:Fathom package in evaluating text complexity.

We used membership length instead of winning record to define influential agents in this research. This is because Topcoder automatically updates each solver's participation and winning records. That is to say, a solver with many winning records today might not be an influential agent one or two years ago. In a sense, the data related to the winning records is contaminated with noise. On the contrary, the membership length is objective since each solver has unique membership registration data. Research confirms that membership length is a significant predictor for being a winner in a tournament-based crowdsourcing event (Bockstedt et al., 2016). Our post hoc analysis shows that membership length is positively and significantly related to each solver's number of wins ($r = 0.35, p < 0.01$) and that the effect size of membership to predict the number of wins is 0.20 ($p < 0.01$). We thus believed that it was reasonable to use membership length to identity influential solvers.

Figure 13

Distribution of Solver's Membership Length



We calculated each solver's membership length (in months) from his or her registration data to 12:55 pm on November 1, 2016. This was the last time that we finalized and validated our data collection. The distribution of solver's membership length is right-skewed with a long tail (Figure 13). Specifically, solvers whose membership length is above the average membership length take around 20 percent of the total number of solvers. Through a detailed analysis on the membership length, we found that there were quite a few inactive solvers whose membership length was above average. We then took the number of winning record into considerations. By applying these two criteria, 309 out of 4,315 solvers were classified as influential in this research (i.e., 7.16 percent). Based on this classification, we counted the number of these influential agents for each crowdsourcing event. By calculating the time difference between event start data and the participation time of the first influential agent, we got the variable called early involvement of influential agents. The following Table 6 lists all the variables in this data analysis.

Table 6

Variable Operationalization in First Empirical Study

Variable	Measurement	Reference
Crowd growth rate	"p + q" from the Bass Diffusion Model	Bass (1969)
Crowd size	The total number of registrants for a particular crowdsourcing event	Liu et al. (2014)
Number_words	Total number of words in describing an event	Li (2008)
Fog index	Fog readability index	Li (2008)
Event length	The amount of time specified for suppliers to solve a particular crowdsourcing event	Connelly et al. (2014)
Payment size	The amount of money specified for the winners in a tournament-based crowdsourcing event	Lazear & Rosen (1981)
Number of influentials	The number of agents with above average membership length and at least one winning record	Bockstedt et al. (2016)
Early involvement of influentials	The difference between event starting time and the participation time of the first influential agent	Bockstedt et al. (2016)
Number of payments	The number of payments specified for an event (1,2,…)	Chen et al. (2011)
Checkpoint	Binary: 1 means feedback provided; 0, otherwise	Wooten & Ulrich (2017)

Model Specification

We used the ordinary least square (OLS) regression (Cohen et al., 2013; Kutner, Nachtsheim, & Neter, 2004) to analyze the two dependent variables in this research: crowd growth and crowd size. The use OLS regression analysis allowed us to specify the dependent variable as a linear function of the explanatory and control variables discussed in the construct measurement section in addition to an error term.

$$
\begin{aligned}
y_i = \beta_0 &+ \beta_1 N\_words_i + \beta_2 N\_words_i{}^2 + \beta_3 Fog_i + \beta_4 Fog_i{}^2 + \beta_5 E\_length_i \\
&+ \beta_6 Payment_i + \beta_7 E\_involement_i + \beta_8 N\_agents_i + \beta_9 N\_payments_i \\
&+ \beta_{10} Checkpoint_i + \varepsilon_i
\end{aligned}
$$

The error term ε_i is assumed to be random and normally distributed with a mean of zero and a constant standard deviation (Kutner et al., 2004). In this research, both our dependent variables only takes non-negative values and are right skewed (Figure 14), which generally leads to violation on the normality assumption of OLS regression (Cohen et al., 2013). Following the recommendations made by methodologists in dealing with non-normally distributed data (Kutner et al., 2004; Neter, Kutner, Nachtsheim, & Wasserman, 1996), we took a log transformation and found that this transformation made the distribution of both our dependent variables approximate to be normal (Figure 15), thus supporting our decision to use OLS in this study. Another advantage of using OLS regression is that this method has useful regression diagnostics and sophisticated remedies to deal with any kind of assumption violations (Kutner et al., 2004).

Figure 14

Histogram of Dependent Variables

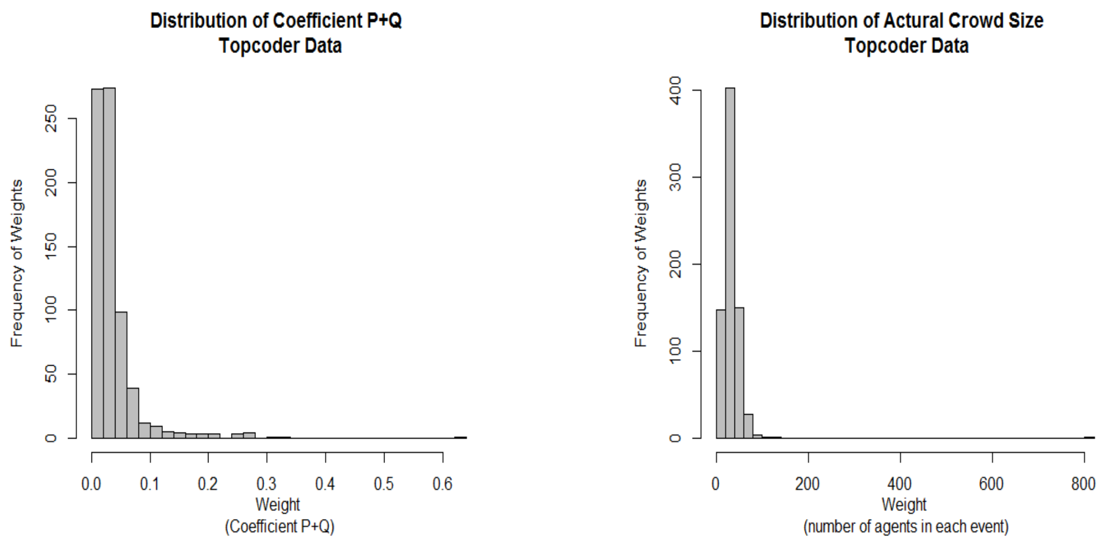
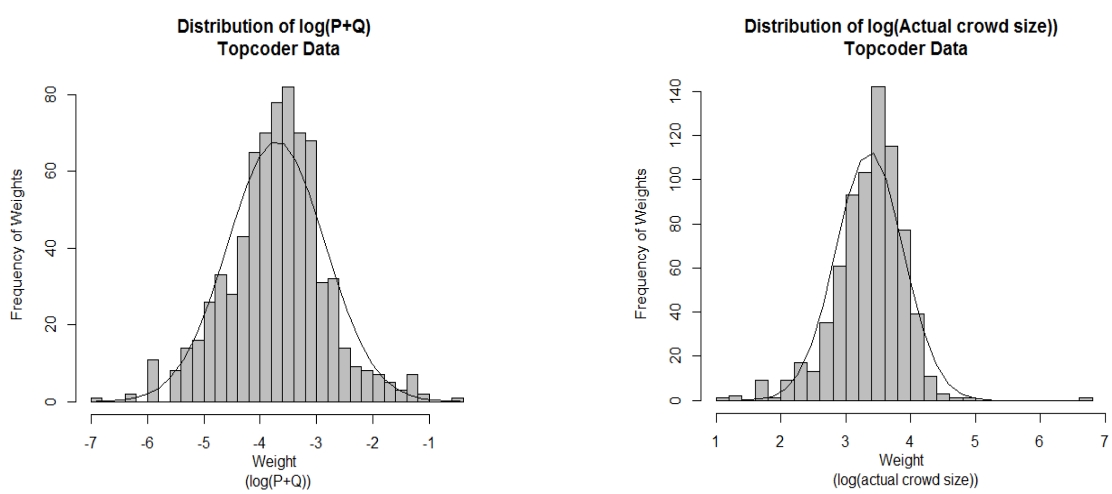


Figure 15

Histogram of Dependent Variables after Log Transformation



Data Analysis and Findings

Process. We used a hierarchical approach as recommended by scholars (e.g., Cohen et al., 2013) to analyze our data. Specifically, we first considered a basic model (i.e., Model 1) in which we separately regressed the dependent variables (i.e., growth rate and crowd size) only upon the control variables (i.e., number of payment and checkpoint). Then, we added the explanatory variables to the basic model and built Model 2, through which we tested our proposed hypotheses in our theoretical model (Figure 9). Finally, we conducted extensive regression diagnostics to check the validity of our OLS regression. This is because both histograms of our dependent variables (Figure 14) suggest the existence of outliers that might cause us to violate the normality assumption for the error term. Besides, the model that we propose to test might not be completely, correctly specified. This can also lead to any other kinds of OLS violations. To increase the validity of our empirical testing, we performed extensive regression diagnostics to check whether we violated assumptions for linear regression (e.g., multicollinearity testing, autocorrelation testing, normality testing, constancy of variance testing, and linearity testing) and took necessary remedies. We further run seemingly unrelated regression (Greene, 2003; Zellner & Huang, 1962) and robust regression (Koller & Stahel, 2011; Yohai, 1987) to justify the validity of our empirical findings.

We ran all analyses in R version 3.3.2. The descriptive statistics (i.e., mean and standard deviation) and simple correlations are reported in Table 7. As indicated by the descriptive statistics, the magnitude of the dependent variables (crowd growth rate in particular) is much smaller than that of a few independent variables (e.g., number of words, event length, and payment size). This difference might lead to trivial coefficient estimations for these independent variables.

Many correlation coefficients among the constructs used in this study are relatively small. A few constructs such as payment size and number of influential agents are significant and are highly correlated ($p = 0.73$). We took several measures to account for the significant relationships among variables that might lead to multicollinearity. First, we used the mean-centered value of all explanatory variables including the quadratic terms to mitigate multicollinearity (Cohen et al., 2013; Neter et al., 1996). Second, we ensured that the variance inflation factor (VIF) scores for each explanatory variable were below the recommended cutoff value of 10.0 typically taken as an indicator of excessive multicollinearity (Neter et al., 1996). Each of the VIFs scores for our dataset met this requirement, suggesting that multicollinearity is not a major issue in our dataset.

Table 7

Descriptive Statistics and Correlations (1)

		Mean	Std	1	2	3	4
1	Crowd size	33.4	32.66	1			
2	Growth rate	0.04	0.05	-0.1**	1		
3	Number_words	1144.30	578.20	0.08*	-0.17***	1	
4	Fog index	13.44	3.06	-0.02	-0.02	0.35**	1
5	Event length	250.47	141.9	0.14***	-0.34***	0.08*	-0.06
6	Payment size	1819.60	916.90	0.22***	-0.14***	0.53**	0.20**
7	E_involvement	2.25	3.04	-0.08*	-0.03	-0.03	-0.08
8	Number_agents	15.19	6.50	0.36***	-0.14***	0.23**	0.10**
9	Number_payment	2.56	1.11	0.25***	-0.03	0.33**	0.21**
10	Checkpoint (0,1)	0.85	0.36	0.16***	-0.40**	0.31**	0.03

Notes: n = 734; ***p<0.001, **p<0.01, *p<0.05

(Table 7 continued)

		5	6	7	8	9	10
1	Crowd size						
2	Growth rate						
3	Number_words						
4	Fog index						
5	Event length	1					
6	Payment size	0.11**	1				
7	E_involvement	-0.03	-0.05	1			
8	Number_agents	0.15**	0.56**	-0.21**	1		
9	Number_payment	-0.02	0.73**	-0.06	0.53**	1	
10	Checkpoint (0,1)	0.34**	0.30**	0.10**	0.31**	0.13***	1

Findings. The results of the OLS regression for crowd growth rate and crowd size are reported in Table 8. The variables are introduced sequentially. For each model, we compared the fit statistics (e.g., R-square and F value) to ensure the validity of our testing. For crowd growth rate, Model 1 only includes the control variables. This base model explains 9.4 percent of variance in growth rate. Binary control variable (i.e., checkpoint) is significantly and negatively related to growth rate, suggesting that the crowd for events provided with feedback (i.e., checkpoint = 1) tends to grow slowly. Model 2 includes all the variables of interests and control variables. The R-square increases from 0.094 in model 1 to 0.283. An F test shows that adding our independent variables significantly increases the gain on the R-square since

$F_{gain} = \frac{r_{all}^2 - r_{set1}^2}{1 - r_{all}^2}\left(\frac{n-k-m-1}{m}\right) = \frac{0.283-0.094}{1-0.283}\left(\frac{734-2-10-1}{10}\right) = 48.15$ is greater than the associated critical F value with the same degree of freedom ($F(10{,}721)_{\alpha=0.05} = 1.84$). A significant increase in the gain of R-square indicates the validity of adding our proposed independent variables. Several coefficients in Model 2 are significant.

Specifically, the coefficient for the second order term of the Fog index is negative ($\beta = -0.814, p < 0.05$), while that of the first order term of the FOG index is positive ($\beta =$

Table 8

OLS Regression for Growth Rate and Crowd Size

	Growth Rate		Crowd size	
	Model 1	Model 2	Model 1	Model 2
Intercept	-3.080***	-8.101*	2.393***	3.493*
	(0.010)	(3.342)	(0.051)	(1.419)
Control variables				
Number_payments	0.001	-0.059	0.191***	0.066***
	(0.028)	(0.038)	(0.014)	(0.016)
Checkpoint (0,1)	-0.739***	-0.014	0.568***	0.167***
	(0.086)	(0.105)	(0.044)	(0.044)
Main variables				
Number_words		0.819		-0.303
		(0.694)		(0.294)
Number_words^2		-0.066		0.027
		(0.053)		(0.022)
Fog		4.372*		-0.671
		(1.972)		(0.837)
Fog^2		-0.814*		0.078
		(0.381)		(0.162)
Payment size		0.009		-0.006
		(0.052)		(0.022)
Event_length		-0.738***		0.162***
		(0.057)		(0.024)
Number_agents		0.009		0.047***
		(0.006)		(0.002)
E_involvement		0.028**		-0.005
		(0.009)		(0.004)
Sample size	734	734	734	734
Fit statistics				
-- R-square	0.094	0.283	0.358	0.646
-- F value	37.77***	28.48***	204.00***	131.80***

Notes: Standard errors in parentheses; *** p<0.001, **p<0.01,*p<0.05.

$4.372, p < 0.05$). This finding suggests that Fog Readability Index (i.e., task complexity) related to crowd growth rate in an inverted U-shape. The competition view based on tournament theory that crowd growth rate relates to task complexity in an inverted U-shape (i.e., H1c) is supported. We did not find evidence to support the alternative hypothesis based on diffusion theory (i.e., H1a). Event length significantly relates to crowd growth rate in a negative way ($\beta = -0.738, p < 0.001$), which supports the competition view on the negative association between event length and crowd growth (H3c) and rejects the alternative hypothesis (H3a). We also found that the early involvement of influential agents (i.e., E_involvement) is positively related to crowd growth ($\beta = 0.028, p < 0.01$). This finding supports the tournament perspective on the positive relationship between early involvement of influential agents and crowd growth rate (H5c) and rejects the alternative hypothesis based on diffusion theory (H5a).

The results based on OLS regression for crowd size are also reported in the above Table 8. The base model (i.e., Model 1) for crowd size is significant ($F = 204, p < 0.001$), which explains 35.8 percent of the variance of the crowd size. In the base model, we found that the number of payments specified for each crowdsourcing event and providing feedback (i.e., checkpoint=1) are both positively associated with crowd size. After adding the main independent variables, the overall R-square increases from 0.358 to 0.646. This increase is strongly significant according to the F test proposed by scholars (Cohen et al., 2013), suggesting that the inclusion of the main independent variables is meaningful in predicting the size of a crowd for a particular crowdsourcing event. Particularly, the coefficient of event length is positive and significant ($\beta = 0.162, p < 0.001$). This finding supports H3b which says that crowd size is positively associated with event length. The diffusion perspective is justified in terms of the implication of event length. We also found the evidence to support H4b since the coefficient for the number of influential agents (i.e., Number_agents) is significantly positive ($\beta = 0.047, p < 0.001$). Diffusion theory is more applicable in explaining the relationship between the number of influential agents and crowd size.

Robust Checks

Check for the Violation of OLS Assumptions. An OLS model makes five basic assumptions about the way in which the observations are generated (1) dependent variables are a linear function of independent variables plus an error term (i.e., linearity assumption); (2) expected value of the error term is zero (i.e., zero-mean disturbance); (3) disturbances have uniform variance and are uncorrelated (i.e., normality assumption and constant variance assumption); (4) observations on independent variables can be considered fixed in repeated samples (i.e., independent observation assumption); (5) no exact linear relationship exists between independent variables and more observations than independent variables (Kennedy, 2003; Kutner et al., 2004). These assumptions are not necessarily independent of each other, which means that one violation might lead to others. Any violation to these assumptions can lead to biased and/or unstable coefficient estimations (Kennedy, 2003; Kutner et al., 2004).

Through Table 7, we assured that there exist no perfect linear relationships between independent variables although a few of them are highly correlated. By mean-centering independent variables and running multicollinearity test, we confirmed that the significant relationships existing between our independent variables did not cause major concerns for using OLS regression. Since the number of observations (i.e., 734) is larger than that of our independent variables (i.e., 10), we concluded that our analyses for both our dependent variables did not violate assumption five. To check whether we violated any other assumptions during our data analysis process, we conducted an extensive post hoc analyses and took remedial measures in case of violations (e.g., re-specifying regression models, running seemingly unrelated regression, and robust regression). Results from our post hoc analysis are reported in the following section.

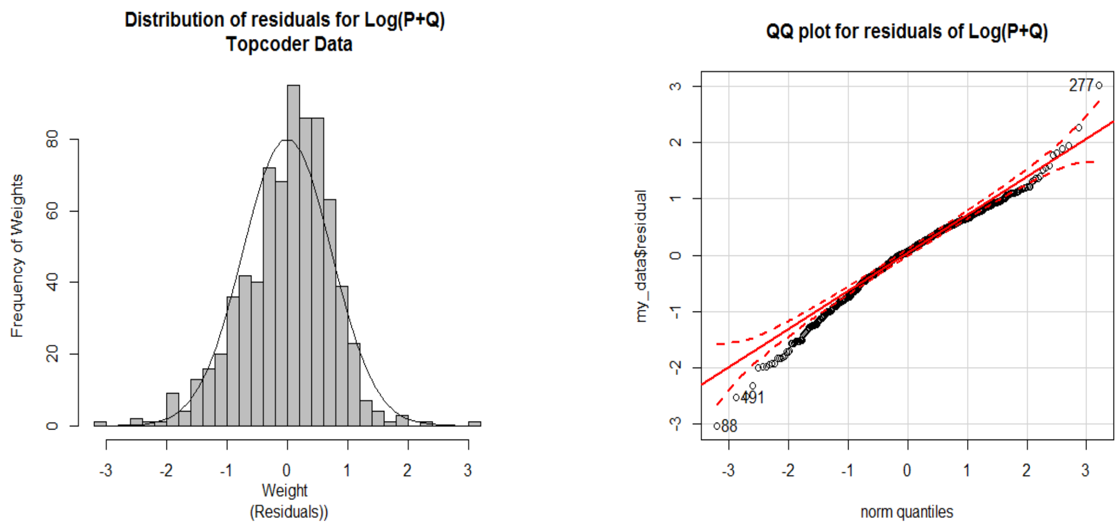
Post Hoc Analysis for Crowd Growth Rate. Following the guidelines on regression diagnostics proposed by scholars (Kutner et al., 2004; Neter et al., 1996), we first generated the residuals of Model 2 for crowd size and then ran a series of recommended tests on them to check the validity of our data analysis (e.g., Shapiro-Wilk test, Breusch-Pagan test, and Durbin-Watson test). The residuals of Model 2 for crowd growth rate has a mean of zero and stand deviation of 0.73, suggesting that we did not violate the zero-mean assumption (i.e., assumption two). We conducted a Durbin-Watson test on the residuals of Model 2 which turned out be non-significant (i.e., D-W statistics = 2.15, p-value = 0.554). The Durbin-Watson test is a test of randomness (i.e., independent observations) (Durbin & Watson, 1951). The null hypothesis of this test is that there exists no autocorrelation in observations. A non-significant Durbin-Watson test makes us fail to reject the null and allows us to conclude that we did not violate assumption four.

We run a Shapiro-Wilk test to check the normality of our residuals. The Shapiro-Wilk test was developed by Samuel Shapiro and Martin Wilk (Shapiro & Wilk, 1965). The null hypothesis of this test is that the population is normally distributed. Our analysis indicates that this test was significant (i.e., $w = 0.98$, $p<0.001$), leading us to reject the null of normal distribution. To check whether our analysis violates the constant variance assumption (i.e., homoscedasticity), we performed a Breusch-Pagan test to the residuals of growth rate. In statistics, the Breusch-Pagan test developed by Trevor Breusch and Adrian Pagan in 1979 is used to test for heteroscedasticity in a linear regression model (Breusch & Pagan, 1979). The null of the Breusch-Pagan test is that the variance of the residuals is constant. The result of this test turned out to be significant (i.e., BP = 21.935, df = 10, p = 0.015). This finding suggests that we should reject the null and conclude the existence of heteroscedasticity. Our Model 2 thus violates the normality assumption and constant of variance assumption.

Although our regression model violates the normality assumption, the distribution of our residuals shown in the following histogram (Figure 16) is approximate to normal. The QQ plot of the residuals suggests the existence of a few outliers, which might be the reason that leads to this violation of normality. We then performed a regression diagnosis to identify these potential outliers by following the guidelines on regression diagnostics proposed by methodologists (Cohen et al., 2013; Kutner et al., 2004). Through our detailed diagnostics, we did identify a few influential outliers. Specifically, we found five most influential outliers whose influence indices obviously exceed the recommended cutoff score ((e.g., $\sqrt[2]{p'/n}$ for DFFIT and $2/\sqrt{n}$ for EFBETA) (Cohen et al., 2013; Kutner et al., 2004). A close examination on the raw data shows that these five outliers take either the maximum or the minimum value of some independent variables (e.g., FOG Readability Index, number of words) or the maximum of the dependent variable. This examination further supports our decision to delete these five outliers in the subsequent regression analysis.

Figure 16

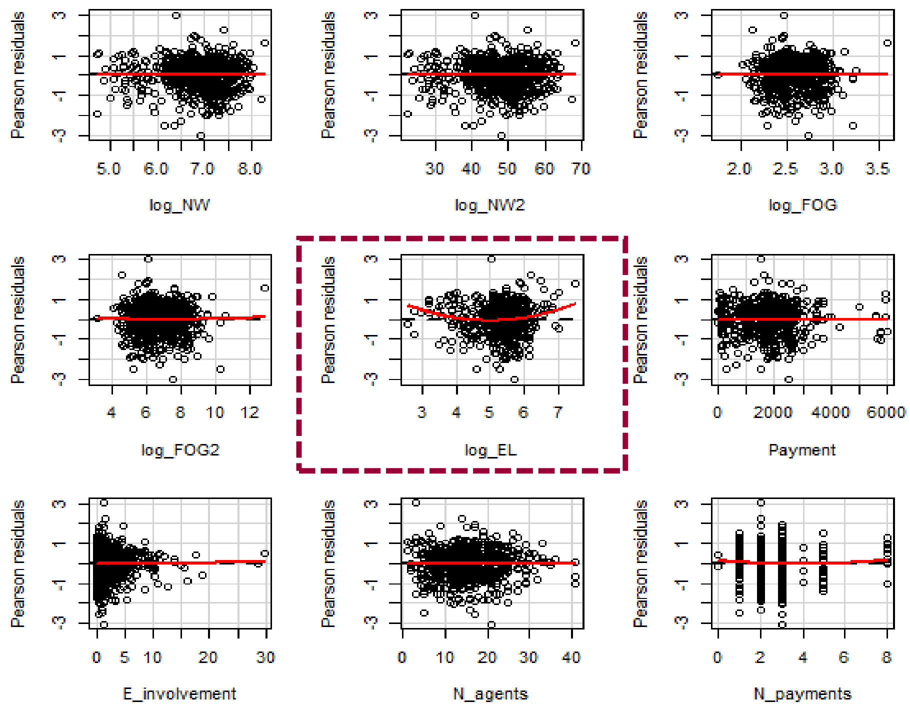
Histogram of Residuals of Crowd Growth Rate in Model 2



We performed the Tukey's test for nonadditivity to check the linearity assumption (i.e., assumption one) (Castle & Hendry, 2010; Fox & Weisberg, 2012; Tukey, 1949). This test is obtained by adding the squared terms of the fitted independent variables to Model 2 and refitting this model. The significant level for the Tukey's test is obtained by comparing the statistics with the standard normal distribution, and the null hypothesis of this test is that the coefficient for the quadratic terms of the fitted variables is zero (Fox & Weisberg, 2012). Through this test, we found that only the coefficient for event length is significant (test statistics = 3.710, $p<0.001$), suggesting that relationship between event length and crowd growth rate is not linear but quadratic. The residual plot against the independent variable also confirms that the relationship between event length and crowd growth is quadratic (Figure 17). It was problematic for us to specify only linear terms for event length in our proposed model.

Figure 17

Residual Plots for Crowd Growth Rate



We thus re-specified our Model 2 by adding a quadratic term for event length. Findings from the re-specified regression analysis are reported in Table 9. We did find a significant quadratic relationship between event length and crowd growth rate ($\beta_{event\ length^2} = 0.151, p < 0.001$), suggesting a U-shape between these two variables. All the other significant coefficients identified previously still hold in our re-specified model (i.e., Model 2' (OLS) in Table 9). The residuals from this re-specified model passed the Breusch-Pagan test (i.e., BP=19.05, df = 11, p-value = 0.06), which suggests the existence of constant variance after we re-specified our regression model. However, we still failed to pass the Shapiro-Wilk test (i.e., normality test) with the re-specified model. Even if we took out the five most influential outliers and reran the re-specified model in Table 3, the Shapiro-Wilk normality test is still significant (i.e., $w=0.98$, $p<0.001$), suggesting the violation of normality assumption.

Table 9

Post Hoc Regression Analysis for Crowd Growth Rate

	Model 2'(OLS)	Model 3(SUR)	Model 4 (Robust)
Intercept	-4.083	-4.168	-4.979
	(3.485)	(3.854)	(3.611)
Control variables			
Number_payments	-0.062	-0.069†	-0.068*
	(0.038)	(0.038)	(0.032)
Checkpoint (0,1)	-0.107	0.144	0.166
	(0.109)	(0.108)	(0.108)
Main variables			
Number_words	0.816	0.868	0.761
	(0.688)	(0.717)	(0.786)
Number_words^2	-0.065	-0.071	-0.063
	(0.052)	(0.054)	(0.059)
Fog	4.172*	4.721*	5.448**
	(1.955)	(2.204)	(1.866)
Fog^2	-0.769*	-0.877*	-1.013**
	(0.378)	(0.428)	(0.361)
Payment size	0.009	0.007	-0.005
	(0.052)	(<0.001)	(<0.001)
Event_length	-2.281***	-2.567***	-2.433***
	(0.420)	(0.422)	(0.513)
(Event length)^2	0.151***	0.176***	0.162**
	(0.041)	(0.041)	(0.051)
Number_agents	0.007	0.009†	0.010*
	(0.005)	(0.005)	(0.005)
E_involvement	0.029**	0.031***	0.027***
	(0.009)	(0.009)	(0.007)
Sample size	734	724	734
Fit statistics			
-- R-square	0.296	0.315	0.336
-- F value	27.60***	n/a	n/a

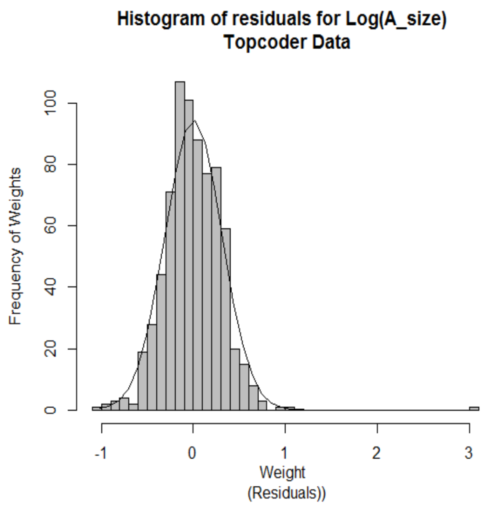
Notes: Standard errors in parentheses; *** p<0.001, **p<0.01,*p<0.05, †<0.1.

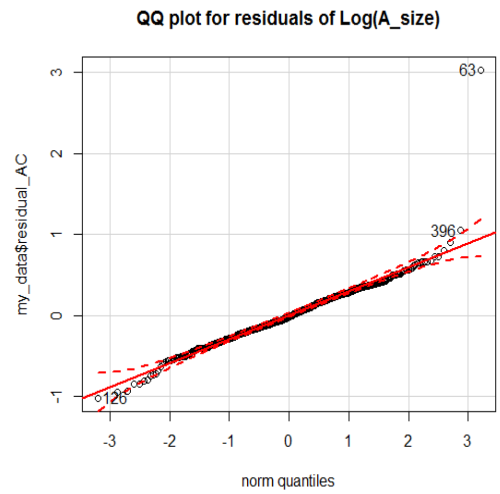
For the normality check, we also compared the actual frequencies of the residuals against expected frequencies under normality. A percent of 69.89 of residuals falls between $\pm\sqrt{\delta_i}$ and 91.28 percent of the residuals falls between $\pm 1.645\sqrt{\delta_i}$ (δ_i is the stand deviation of residuals). These two numbers are above the recommended thresholds recommended by scholars to quantitatively determine the normality of the distribution of residuals, that is, 68 percent of the residuals fall between $\pm\sqrt{\delta_i}$ and 90 percent fall between $\pm 1.645\sqrt{\delta_i}$ (Kutner et al., 2004). Thus, the actual frequencies here are reasonably consistent with those expected under normality. Based on the comparison of frequencies and the histogram of residuals (Figure 16), we concluded that the residuals of our Model 2 fall an approximate normal distribution with a small departure, which does not create any serious problems for our regression (Kutner et al., 2004; Neter et al., 1996).

Post Hoc Analysis for Crowd Size. The residuals of Model 2 for crowd size has a mean of zero and standard deviation of 0.31. We did not violate the zero-mean assumption (i.e., assumption two). We followed the diagnostic procedures that we took previously to check the validity of our analyses for crowd size. By analyzing the residuals of Model 2 for crowd size, we found that they approximately follow a normal distribution (Figure 18) with 91.7 percentage of residuals falling within the expected range of normal distribution (Kutner et al., 2004), although the existence of a few outliers of crowd size shown on the QQ plot in Figure 14 made our regression model fail to pass the Shapiro-Wilk normality test (i.e., w=0.94, p<0.01).

Figure 18

Distribution of Residuals for Crowd Size

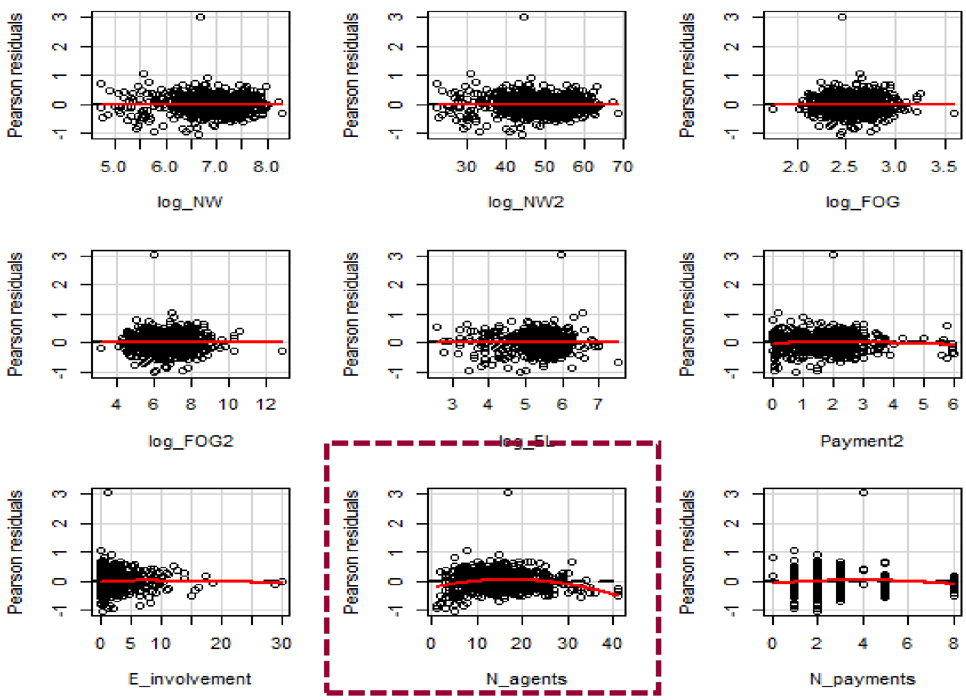




The residuals of crowd size in Model 2 passed the Breusch-Pagan test (i.e., BP=12.408, df = 10, p=0.2587), suggesting that the variance of residuals is constant. Our proposed model for predicting crowd size does not violate the constant of variance assumption. However, the Durbin-Watson test on the residuals of Model 2 for crowd size is significant (D-W = 1.316, p<0.001), suggesting that the residuals of crowd size auto-correlate to each other. Omitting the autocorrelation of residuals might seriously underestimate the true standard deviation of the estimated regression coefficient and lead to ineffective coefficient estimations (Kutner et al., 2004; Neter et al., 1996). In the subsequent analysis, we took the recommended remedy (i.e., the Cochrane-Orcutt procedure) to correct our predictions for crowd size (Beach & MacKinnon, 1978; Betancourt & Kelejian, 1981; Kutner et al., 2004). This procedure was implemented by the “Orcutt” package in R software. Through the Tukey’s test for nonadditivity, we also found that there exists a significant quadratic term for the number of influential agents (i.e., Number_agents) (Figure 19).

Figure 19

Residuals Plots for Crowd Size



We thus reran the regression analysis by specifying a quadratic term of the number of influential agents and taking the Cochrane-Orcutt remedy to account for autocorrelation. Findings for this analysis are reported in Table 10 (i.e., Model 2' (OLS)). We did find a significant and negative quadratic terms for the number of influential agents ($\beta_{Number\_agents^2} = -0.001, p < 0.001$). In this re-specified model (i.e., Model 2' (OLS)), the positive relationship between event length and crowd size still holds. We alas found that the relationship between payment size and crowd size is marginally significant after taking the Cochrane-Orcutt remedy ($\beta_{payment} = 0.033, p = 0.081$). Model 2'

(OLS) passed all the assumption checks except the Shapiro-Wilk normality test (w=0.94, p<0.001). Although Model 2' (OLS) failed to pass the formal normality test, the residuals of this model fall an approximate normal distribution with minor departure due to the existence of a few extreme observations.

Table 10

Post Hoc Regression Analyses for Crowd Size

	Model 2' (OLS)	Model 2" (OLS)	Model 3 (SUR)	Model 4 (Robust)
Intercept	1.434	1.012	3.488*	2.797†
	(1.234)	(1.140)	(1.444)	(1.587)
Controls				
N_payments	0.048***	0.031**	0.053***	0.054***
	(0.014)	(0.012)	(0.015)	(0.014)
Checkpoints	0.023	0.031	0.055	0.084†
	(0.039)	(0.033)	(0.043)	(0.051)
Main effects				
Number_words	-0.031	0.114	-0.214	-0.233
	(0.256)	(0.221)	(0.279)	(0.374)
Number_words^2	0.001	-0.010	0.020	0.021
	(0.019)	(0.017)	(0.021)	(0.027)
Fog	0.137	0.134	-1.008	-0.349
	(0.699)	(0.661)	(0.858)	(0.725)
Fog^2	-0.064	-0.060	0.147	0.018
	(0.135)	(0.128)	(0.167)	(0.143)
Payment size	0.033†	0.045**	0.020	0.017
	(0.019)	(0.016)	(0.019)	(0.021)
Event_length	0.173***	0.153***	0.137***	0.132***
	(0.019)	(0.017)	(0.022)	(0.025)
Number_agents	0.096***	0.096***	0.093***	0.084***
	(0.006)	(0.005)	(0.007)	(0.009)
(Number_agents)^2	-0.001***	-0.001***	-0.001***	-0.001***
	(<0.001)	(<0.001)	(<0.001)	(<0.001)
E_involvement	-0.003	0.002	-0.002	-0.003
	(0.003)	(0.003)	(0.004)	(0.003)
Sample size	734	728	724	734
Fit statistics				
-- R-square	0.707	0.769	0.696	0.673
-- F value	158.40***	215.90***	n/a	n/a

Notes: Standard errors in parentheses; *** p<0.001, **p<0.01,*p<0.05, †<0.1.

By running similar regression diagnostics as we did previously, we identified six influential outliers whose influence indices (e.g., DFFITS and DFBETAS) are beyond the recommended cutoff values. One out of these six outliers (ID=175) was also classified as an influential outlier for crowd growth rate. This observation takes the maximum value of Fog Readability Index and number of words. We reran our regression by taking out the six influential outliers identified by our regression diagnostics. Results are reported in Table 10 under Model 2" (OLS). After taking out the six influential outliers and the Cochrane-Orcutt remedy, the coefficient for payment size became significant ($\beta_{payment\ size} = 0.045, p < 0.01$), suggesting a positive relationship between payment size and crowd growth. We also found that, if we took out influential outliers, the residuals from our re-specified model passed the Shapiro-Wilk normality test (w = 0.997, p = 0.112) but failed to pass the Breusch-Pagan test (BP = 66.36, df = 11, p < 0.001). This finding suggests that the existence of outliers in our data causes our regression model for crowd size to violate either the normality test or the constant variance test. We need to rely on more robust regression techniques to alleviate the influence of outliers, which will be addressed in the following section.

Robust Check on the Regression Coefficients. All our analyses so far for the dependent variables (i.e., crowd growth rate and crowd size) were conducted separately. We assumed that these two models are independent. This assumption might not be true. This is because some unconsidered factors that influence the error term in one equation probably influence the error term in the other equation (Greene, 2003; Henningsen & Hamann, 2007). Particularly, our post hoc analysis confirms that the residuals of our two prediction models are slightly correlated to each other ($r = -0.096, p < 0.001$).

Ignoring this contemporaneous correlation and estimating these two equations separately might lead to biased estimates of the coefficients (Henningsen & Hamann, 2007).

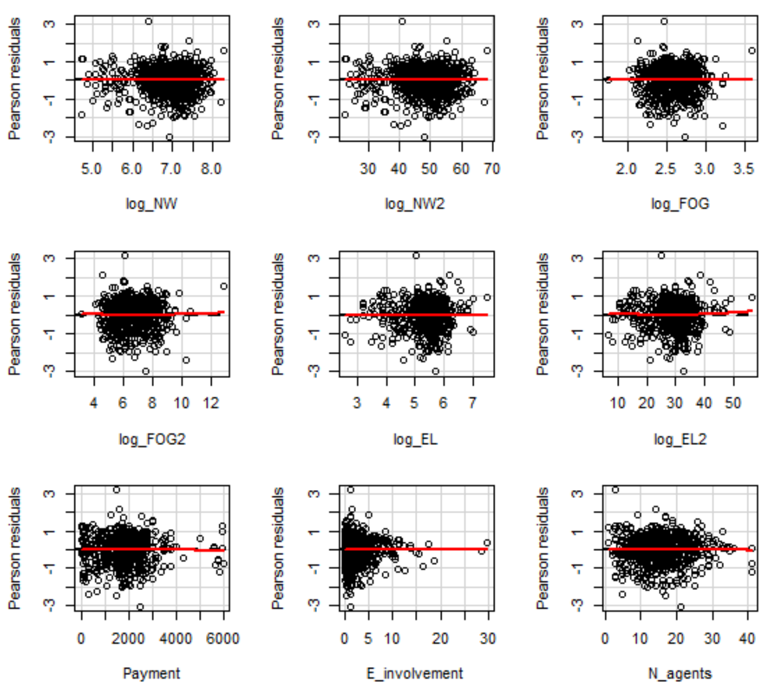
Thus, we have to estimate both our dependent variables (i.e., crowd growth rate and crowd size) simultaneously through seemingly unrelated regression (SUR), which takes the covariance structure of the residuals into account (Moon & Perron, 2006). SUR was developed by econometricians to deal with statistical analyses that are based on models containing structurally related equations (Greene, 2003; Henningsen & Hamann, 2007). We took out the identified ten most influential outliers when we ran the SUR analysis. This is because SUR is not robust to the influence of outliers although it accounts for the potential correlations of error terms (Greene, 2003; Moon & Perron, 2006). Therefore, the sample size for SUR analysis is 724 (i.e., 734-10=724).

Based on the above checks on the regression assumptions, we also took a robust regression (Koller & Stahel, 2011; Kutner et al., 2004; Yohai, 1987) to account for the impact of many other potential outliers other than the ten most influential outliers identified previously. This regression was carried out by the "lmrob" function in the "robust" package in R (Wilcox, 2011). The "lmrob" function computes an MM-type regression estimator as described in Yohai (1987) and Koller and Stahel (2011). By default, this function uses a bi-square redescending score function and returns a highly robust and highly efficient estimator (with 50 percent breakdown point and 95 percent asymptotic efficiency for normal errors) (Wilcox, 2011). Findings from SUR analysis and robust regression for both crowd growth rate and crowd size are reported in Table 9 and Table 10 (i.e., Model 3 (SUR) and Model 4 (Robust)). These findings are consistent with those based on revised OLS regression, demonstrating the robustness of our data analyses.

Endogeneity Check. Endogeneity arises when an independent variable is correlated with the error term, thereby violating the exogeneity condition in OLS specifying that the error terms or the residuals have an expected value of zero given any independent variable (i.e., $E(u|X_1, X_2, \ldots, X_k) = 0$) (Bascle, 2008; Wooldridge, 2015). As shown in the following Figure 20 and Figure 21, the residuals for our two re-specified OLS models for predicting crowd growth rate and crowd size do not correlate with any of our independent variables. Thus, the sufficient condition for endogeneity does not occur.

Figure 20

Final Residual Plots for Crowd Growth Rate



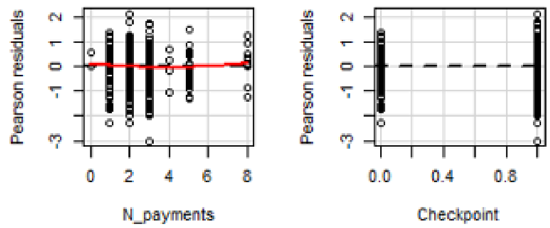
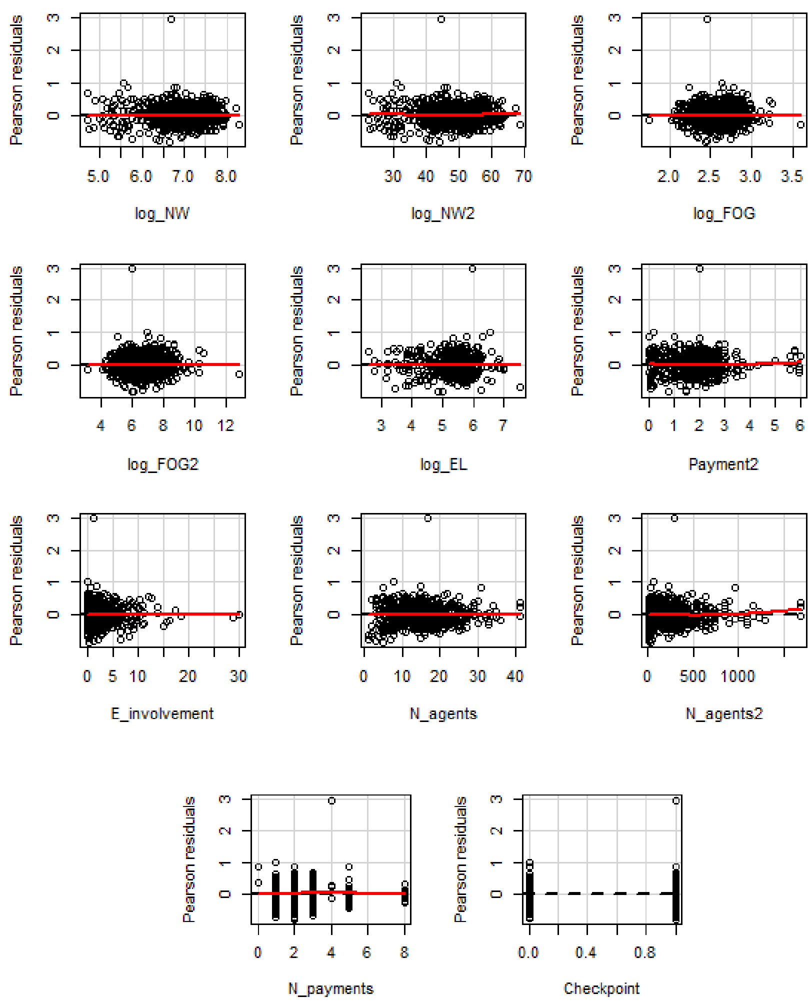


Figure 21

Final Residual Plots for Crowd Size



According to econometricians, three main instances that lead to endogeneity include measurement error, simultaneous causality, and omitted variables (Wooldridge, 2015). In this research, we used objective secondary data from a very reliable crowdsourcing platform to operationalize our variables. Moreover, the dependent and independent variables involved in this research come from different sources. The independent variables such as task complexity (i.e., number of words and Fog readability index), payment size, and event length are fully exogenous to the two dependent variables in this research. This is because these variables are clearly specified by focal buying firms or crowdsourcing platforms at the initiation stage of a crowd development process. We collected the solvers' participation decisions during the crowd formation and realization stages and used this information to operationalize the two dependent variables (i.e., crowd growth rate and crowd size) in this research. The separation of causes (i.e., elements of event design) and effects (i.e., crowd growth rate and crowd size) makes endogeneity not a concern for our data analysis (Wooldridge, 2015).

In this research, we controlled for potential influence of the omitted variable in our analysis by adding number of payments and feedback and adding them into our analysis as control variables. A common source of omitted variable bias is the self-selection of samples that might cause some information related to dependent variables to be unobservable to researchers (Wooldridge, 2002, 2015). In this research, the operationalization of crowd growth rate through the Bass Model made 39.38 percent of our observations "unobservable" due to a convergence issue in regression. We compared the mean differences of each variables between the convergence group and non-convergence group through ANOVA analysis. Results from this analysis are reported in Table 11. We did not find significant mean difference between these two groups on crowd size and four independent variables (e.g., Fog Readability Index, event length, payment size, and number of influential agents). However, we did find some levels of significant differences on number of words and early involvement of influential agents between the convergence group and non-convergence group. We need to be cautious in interpreting the effect of early involvement of influential agents on the crowd growth rate. Based on all these considerations, we conclude that the endogeneity is not a major concern for this research.

Table 11

Mean Differences on Independent Variables between Convergence Group and Non-convergence Group

	Convergence	Non-convergence	Mean difference
Crowd Size	32.7	32.29	n.s
Number_words	1125.58	1033.19	**
Fog	13.45	13.41	n.s.
Event length	245.52	229.33	n.s.
Payment size	1779.76	1759.18	n.s.
Number_agents	14.89	15.45	n.s.
E_involvement	2.25	1.28	***

Notes: ***p<0.001, **p<0.01, n.s.: not significant

Summary

The first purpose of this data analysis was to test the theoretical model on crowd emergence (Figure 9) proposed in chapter 4, that is, to identify the influence of event design on crowd emergence (i.e., crowd growth rate and crowd size). Through our regression analyses and extensive post hoc analyses, we identified a few significant relationships: (1) Fog Readability Index, an indicator of task complexity, relates to crowd growth rate in an inverted U-Shape (H1c); (2) the early involvement of influential agents positively relates to crowd growth rate (H5c); (3) the relationship between event length crowd growth rate is not linear but quadratic, i.e., U-shape; (4) the number of influential agents involved in a crowdsourcing event seems to have a positive influence on crowd growth rate (H4a). As for the relationships between elements of event design and crowd size, we found that the event length is positively relate to crowd size (H3b), while the number of influential agents actually relates to crowd size in a U-shape rather than a linear approach. Detailed discussions on these findings will be addressed later.

The second purpose of this data analysis was to test whether diffusion theory is more applicable than tournament theory in explaining crowd emergence in crowdsourcing. Our findings suggest that neither diffusion theory nor tournament theory can fully explain the relationships between elements event design and crowd emergence. Instead, we need combine these two theoretical lenses to get a full picture on the mechanisms underlying crowd emergence. We propose a new perspective to look at crowd emergence, which will be discussed in the discussion section.

The most unexpected finding is that payment size does not seem to play a significant role in influencing crowd emergence. This finding is contradictory to the claim that financial incentive is a major consideration for suppliers to participate in crowdsourcing events (Brabham, 2010, 2012; Liu et al., 2014). Instead, we found that the number of payments is positively related to both crowd growth rate and crowd size and seems to be more important than the payment size in influencing crowd emergence. The unexpected finding related to payment size suggests that there is more story to tell on how suppliers in crowdsourcing interpret financial information. This will be further discussed in the discussion and conclusion sections.

Data Analysis for the Performance Implications of Crowd Attributes

Data Description

We used 5,049 effective observations from Topcoder through web crawling for this data analysis. Among these observations, 1,155 observations were design events, and 3,894 were development events. There were five observations whose productivity data (i.e., the number of solutions) were more than three standard deviations above the average productivity. We classified these five observations were outliers and then took these five observations out when we analyzed crowd productivity. Thus, the total sample size was 5,044 when we analyzed crowd productivity (i.e., 5,049 – 5 = 5,044). Among the 5,049 effective observations, the efficiency data for 377 observations was missing because these observations had no submission (i.e., crowd productivity was zero). When we analyzed the crowd efficiency, we only used 4,702 observations (i.e., 5,049 – 347 = 4,702).

Variables

Dependent Variables. We considered two main dependent variables in this research: crowd productivity and crowd efficiency. We measured *crowd productivity* by the amount of work done by the solvers for a crowdsourcing event, i.e., the total number of solutions generated by a crowd for a crowdsourcing event (Cohen & Bailey, 1997; Horwitz & Horwitz, 2007; Ren et al., 2015). *Crowd efficiency* is defined as the relative speed with which a crowd solve a particular crowdsourced task (Cohen & Bailey, 1997; Horwitz & Horwitz, 2007). We operationalize crowd efficiency in two ways by calculating the shortest time to solve crowdsourced tasks (i.e., efficiency_1) and the average time to solve a crowdsourced task (i.e., efficiency_2). The following Table 12 summarizes the measurement of the dependent variables, independent variables, and control variables.

Table 12

Variable Operationalization in the Second Empirical Study

Variable	Measurement	Reference
Crowd productivity	The total number of submissions produced by a crowd in crowdsourcing	Horwitz & Horwitz (2007)
Efficiency_1	The shortest time that a crowd takes to complete a crowdsourcing task	Cohen & Bailey (1997)
Efficiency_2	The average time that a crowd takes to complete a crowdsourcing task	Horwitz & Horwitz (2007)
Crowd size	The total number registrants for a particular crowdsourcing event	Liu et al. (2014)
Tenure disparity	The coefficient of variation of crowd members' membership length	Harrison & Klein (2007)
Country variety	The variation in the origin of crowd members, i.e., the Blau's index	Blau (1977)
Task complexity	1) Total number of words in describing an event 2) Fog readability index	Li (2008)
Event length	The amount of time specified for an event	Connelly et al. (2014)
Payment size	The amount of money specified for a crowdsourcing event	Lazear & Rosen (1981)
Number of payments	The number of payments specified for an event (1,2,…)	Lazear & Rosen (1981)
Checkpoint(0,1)	Binary: 1 mean feedback provided; 0, otherwise	Wooten & Ulrich (2017)
Group (0,1)	Binary: 1 means design event; 0, otherwise	

Independent Variables. *Crowd size* refers to the number of participants for a crowdsourcing event (Liu et al., 2014). *Crowd diversity* is defined as the extent of difference among members of a crowd (i.e., solvers) with respect to a common attribute, such as tenure, ethnicity, and knowledge (Harrison & Klein, 2007). Following the common practices on diversity measurement in management literature (Harrison & Klein,

2007) and open source literature (Daniel et al., 2013), we measure crowd diversity from two approaches: tenure disparity and country variety. Disparity diversity reflects the difference in the concentration of valued social assets or resources such as knowledge, status, and reputation among units members (Harrison & Klein, 2007). We measured *tenure disparity* using the coefficient of variation of the membership length of all crowd members. Coefficient of variation is a widely used measure of tenure disparity in diversity research (Harrison & Klein, 2007; Ren et al., 2015). If we denote each member's membership length as L_i and the average membership length of L_{mean}, the coefficient of variation can be calculated using the following formula (Harrison & Klein, 2007):

$$[\sum(L_i - L_{mean})^2/n]^{1/2}/L_{mean}$$

Varity diversity reflects "the number and spread of 'batches' of information content, experience, or unique network ties available across unit members"(Harrison & Klein, 2007, p.1204). Accordingly, *country variety* is defined as the variation in the origin of crowd members. Country origin is a categorical variable. We use the Blau's index to calculate the country variety for a crowd (Blau, 1977; Harrison & Klein, 2007). For a particular crowd, we calculate origin variety by counting the number of crowd members from each country. If we denote the percentage of crowd members in a country as P_i, Blau's index can be calculated as follows (Harrison & Klein, 2007):

$$1 - \sum P_i^2$$

The highest level of origin variety occurs when a crowd has members with origin evenly distributed in all country categories. The lowest level of origin variety occurs when all crowd members come from the same country. In this situation, the origin variety is zero. A moderate level of origin variety arises when a crowd has its members with country origin in some of the categories – some uniqueness and some overlapping (Daniel et al., 2013; Ren et al., 2015).

Control Variables. In order to increase the validity of our empirical analysis, we control for the potential influence on the variations of the dependent variables that might be caused by the following factors. Studies in the diffusion literature and communication literature show that the complexity of an idea or a product influences how consumers interpret and perceive the attractiveness of the idea or the product, which might influence consumers' subsequent decision-making (e.g., adoption) (e.g., (Boyd & Mason, 1999; Rogers, 2010). Similarly, we believe that the complexity of a crowdsourcing event can influence agents' perceptions on the attractiveness of an event, which may determine agents' investment and the final outcome. The first control variable in this research is task complexity defined as the perceived difficulty of a crowdsourcing task. Following the practices adopted in our previous empirical study, we measure task complexity in crowdsourcing by using the number of words used in each description (Haas et al., 2015) and task readability index, i.e., Fog Readability Index (Collins‐Thompson & Callan, 2005; Li, 2008).

We control the potential influence of the length of a crowdsourcing event. This is because studies in psychology literature demonstrate that time is an important factor for individuals' creativity and team productivity (Baer & Oldham, 2006; Pepinsky, Pepinsky, & Pavlik, 1960). In this research, *event length* is operationalized as the amount of time (in days) that is specified by a focal buying firm or a crowdsourcing platform for a particular crowdsourcing event. Field experiments conducted by Liu and her colleagues (2014) show that a higher reward induces more submissions in crowdsourcing events that are related to translation and programming (Liu et al., 2014). This study indicates that the size of reward might be related to crowd productivity in crowdsourcing. We then control *payment size* defined as the amount of money specified for a crowdsourcing event. We also control for the *number of payments* since many crowdsourcing events in our sample are set up as rank-order tournaments (Chen et al., 2011; Lazear & Rosen, 1981). In such situation, multiple payments are specified when buying companies create the crowdsourcing events at the beginning. For instance, the maximum number of payments in our sample is eight. One empirical study shows that feedback offered to a crowd in crowdsourcing influences crowd members' participation behaviors and potential outcomes (Wooten & Ulrich, 2017). We then use a binary variable, i.e., *Checkpoint*, to capture whether a focal buying firm offers feedback for agents who participate in a crowdsourcing event.

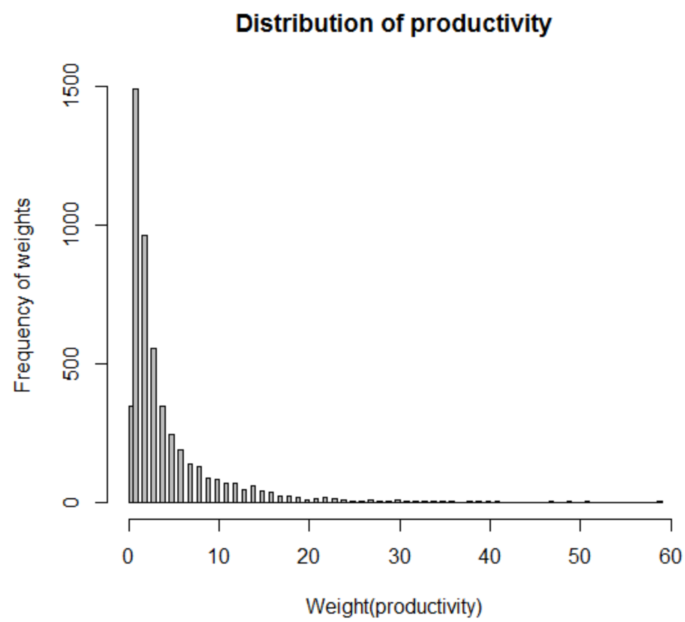
There are two types of crowdsourcing events in our sample that are organized by Topcoder: design programming events and development programing events. Both two types of events have very similar structure in which Topcoder or focal buying firms specify elements of event design at the beginning and solvers (i.e., agents) are supposed to convert the specified requirements into usable software (Boudreau et al., 2011; Boudreau, Lakhani, et al., 2016). One notable difference notified by Archak (2010) is that "winning design submissions go as inputs into the development events in which agents are required to submit actual code implementing the provided design" (Archak, 2010)(p.22). We did not find this notable difference in our data, but we found the descriptive statistics of some variables are different between these two types of crowdsourcing event. For instance, the average payment size for the design events is significantly larger than that of development events, while the average event length of development events is longer than that of design events. We thus create another binary variable named *group* to control for the potential influence of these differences on our dependent variables (i.e., group = 1 means design event; otherwise, development events.

Model Specification

In this research, we have two dependent variables: crowd productivity and crowd deficiency. By definition, crowd productivity is a count variable which takes only non-negative integer values (i.e., 0, 1, 2 …). The distribution of crowd productivity is right-skewed with a mean of 4.32 and a variance of 122.32 (Figure 22), which indicates the existence of overdispersion (Cameron & Trivedi, 1986, 2013). A linear regression model is inappropriate for analyzing right-skewed, overly dispersed data since this distribution violates the basic assumptions of homoscedastic, normally distributed residuals in linear regression (Kutner et al., 2004). Following the recommendations made by many statisticians and econometricians on analyzing count data, we adopt the negative binomial regression to analyze crowd productivity in this research, i.e., crowd productivity (Cameron & Trivedi, 1986, 2013; Greene, 2003; Zeileis, Kleiber, & Jackman, 2008). The negative binomial model not only accounts for overdispersion but also helps avoid high levels of significance due to coefficients whose standard errors might be underestimated (Bellamy, Ghosh, & Hora, 2014; Cameron & Trivedi, 1986, 2013).

Figure 22

Histogram of Crowd Productivity

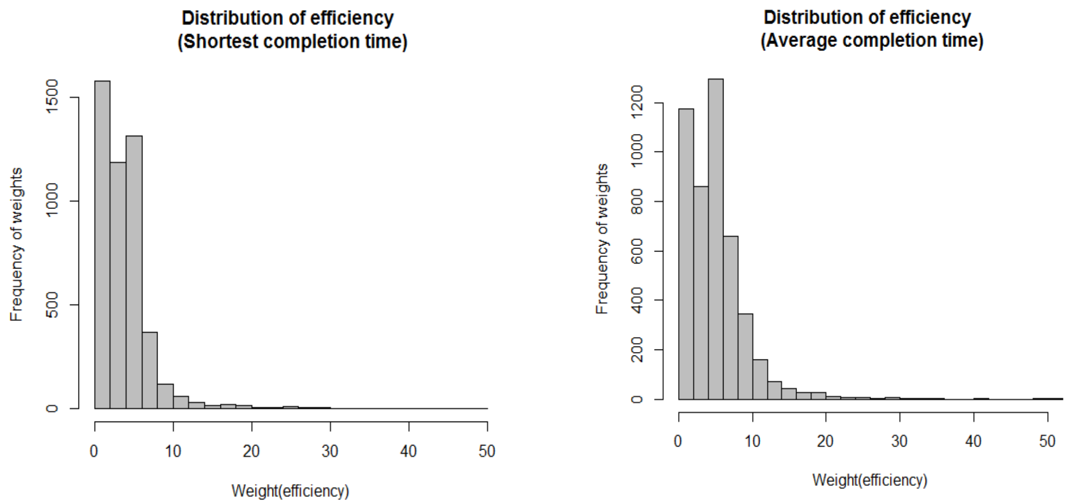


We operationalize the other dependent variable (i.e., crowd efficiency) by using (1) the shortest solution time (i.e., Efficiency_1) and (2) the average solution time within a crowd (i.e., Efficiency_2). Thus, a short task completion time means a high efficiency. We use a Generalized Linear Model (GLM) with a gamma distribution to analyze the two time-related variables (i.e., Efficiency_1 and Efficiency_2). This approach is suitable for analyzing these two variables for two reasons. First, both our time-related variables are continuous and take only positive values. Second, the distributions of these two variables are right-skewed (Figure 23). By probing the relations between dependent variables and the covariates, we found that the variance of these two variables increases with the mean.

These two reasons are consistent with the basic assumptions of a GLM with a gamma distribution (Ballinger, 2004; Crawley, 2012; Dobson & Barnett, 2008).

Figure 23

Histogram of Crowd Efficiency



Data Analysis and Findings

Process. We ran all analyses in R version 3.3.2. The descriptive and simple correlations are reported in Table 13. The standard deviation of crowd productivity is more than twice its mean, which further supports the judgement that this variable is over-dispersed (Cameron & Trivedi, 1986, 2013). The magnitude of two control variables (i.e., Number_words and Pay_size) is obviously greater than that of our dependent variables. We thus scale down these two variables by 10 and 100, respectively, to avoid the occurrence of very small coefficient estimations in the following data analysis.

Quite a few constructs such as crowd size and productivity are significantly highly correlated with each other. We took several measures to account for the significant relationships among variables that might lead to multicollinearity. First, we used the mean-centered value of all explanatory variables including interactions terms to mitigate multicollinearity (Cohen et al., 2013; Neter et al., 1996). Second, we ensured that the variance inflation factor (VIF) scores for each explanatory variable were below the recommended cutoff value of 10.0 typically taken an indicator of excessive multicollinearity (Neter et al., 1996). Each of the VIFs scores for our dataset met this requirement, indicating that multicollinearity is not an issue in the give dataset.

Our theoretical models involve curvilinear term (i.e., crowd size^2), linear interaction (e.g., crowd size × tenure disparity), and quadratic interaction terms (e.g., crowd size^2 × tenure disparity). As recommended by some methodologists (Cohen et al., 2013), we took a hierarchical approach to test the significance of main explanatory variables including quadratic term of crowd size, linear interaction terms, and quadratic interactions terms separately. Variables were introduced sequentially to ensure model stability and to make sure that any significant relationship is robust to the inclusion of other variables. Specifically, we first considered a basic model in which we regressed the dependent variables (i.e., crowd productivity and crowd efficiency) only upon the control variables (Model 1). We then added the main explanatory variables (e.g., crowd size, crowd size^2, tenure disparity, and country variety) to develop Model 2, through which we tested the first two hypotheses in our theoretical model. After step two, we tested the linear interaction terms in Model 3 and the quadratic interaction terms in Model 4. For each model, we conducted the Chi-square likelihood test based on the null hypothesis that all the estimated coefficients that were not present in the previous model are zero. The Chi-square statistics and significance levels are presented in the following Table 14. For each step, we also compared the fit statistics to control for the validity of our data analysis (e.g., AIC and Deviance).

Table 13

Descriptive Statistics and Correlations (2)

		Mean	Std. Dev.	1	2	3	4	5
1	Productivity	4.32	11.06	1.00				
2	Efficiency_1	3.65	3.23	-0.06**	1.00			
3	Efficiency_2	4.95	4.03	0.21**	0.81**	1.00		
4	Crowd size	24.83	25.25	0.71**	0.18**	0.42**	1.00	
5	Tenure disparity	0.95	0.26	0.19**	0.05**	0.21**	0.30**	1.00
6	Country variety	0.75	0.14	-0.01	0.23**	0.18**	0.22**	-0.06**
7	Number_words	660.26	486.36	0.16**	0.15**	0.32**	0.20**	0.09**
8	Fog_index	13.97	5.46	0.02	0.10**	0.07**	0.06**	-0.05**
9	Event length	12.85	11.01	-0.02	-0.03*	0.01	-0.07**	0.16**
10	Pay_size	1252.30	1060.04	0.23**	0.30**	0.38**	0.36**	-0.07**
11	Number_pay	1.84	0.91	0.19**	0.13**	0.25**	0.26**	-0.02
12	Checkpoint (0,1)	0.19	0.39	0.26**	0.08**	0.41**	0.20**	0.22**
13	Group (0, 1)	0.23	0.42	0.27**	0.01	0.33**	0.17**	0.29**

	6	7	8	9	10	11	12	13
Productivity								
Efficiency_1								
Efficiency_2								
Crowd size								
Tenure disparity								
Country variety	1.00							
Number_words	0.01	1.00						
Fog_index	0.09**	0.02	1.00					
Event length	-0.18**	-0.22**	-0.10**	1.00				
Pay_size	0.20**	0.49**	0.11**	-0.45**	1.00			
Number_pay	0.10*	0.47*	0.10**	-0.46**	0.69**	1.00		
Checkpoint (0,1)	-0.11**	0.55**	-0.04*	-0.08**	0.33**	0.44**	1.00	
Group (0, 1)	-0.15**	0.52**	-0.05**	-0.14**	0.28**	0.45**	0.86**	1.00

Notes: n = 5049; $^{**}p<0.01$, $^{*}p<0.05$

Table 14

Native Binomial Regression Model - Crowd Productivity

Variables		Model 1	Model 2	Model 3	Mode 4
Controls					
Number of words		0.002	-0.002	-0.002	-0.003
		(0.003)	(0.002)	(0.002)	(0.002)
Fog readability index		-0.006*	-0.005*	-0.005*	-0.005**
		(0.002)	(0.002)	(0.002)	(0.002)
Event length		-0.013***	-0.017***	-0.017***	-0.016***
		(0.001)	(0.001)	(0.001)	(0.001)
Payment size		-0.007***	-0.020***	-0.020***	-0.021***
		(0.002)	(0.002)	(0.002)	(0.002)
Number of payment		0.21***	0.178***	0.175***	0.180***
		(0.018)	(0.013)	(0.013)	(0.013)
Checkpoint (0,1)		0.631***	0.455***	0.443***	0.428***
		(0.052)	(0.042)	(0.042)	(0.042)
Group (0, 1)		0.638***	0.635***	0.619***	0.600***
		(0.050)	(0.043)	(0.044)	(0.044)
Main effects					
Crowd size			0.280***	0.282***	0.306***
			(0.007)	(0.008)	(0.009)
Crowd size^2	H1c		-0.007***	-0.007***	-0.019***
			(0.001)	(0.001)	(0.002)
Tenure disparity	H2a		0.131***	0.099*	0.037
			(0.043)	(0.045)	(0.046)
Country variety	H2c		-0.413**	-0.503***	-0.934***
			(0.085)	(0.094)	(0.113)
Linear interactions					
Crowd size × Tenure disparity	H3a			0.039*	0.055*
				(0.019)	(0.023)
Crowd size × Country variety				-0.177***	-0.291***
				(0.004)	(0.052)
Curvilinear interactions					
Crowd size^2 × Tenure disparity	H3a				0.004*
					(0.002)
Crowd size^2 × Country variety					0.079***
					(0.011)
AIC		22368	20577	20433	20390
Deviance		4989.5	4754.7	4745.4	4771.1
Chi-square likelihood ratio test		2976.82***[a]	1926.17***	150.22***	47.07***
Over-dispersion (Theta)		3.026	7.364	7.424	7.813
N		5044	5044	5044	5044

Notes: Standard errors in parentheses; *** p<0.001, **p<0.01,*p<0.05. a: The AIC and Deviance for the null model in which no variables are included are 25971 and 5263.5.

Findings. The results of the negative binomial regression for crowd productivity are presented in Table 14. Model 1 includes only control variables. Some of the control variables are significant. Specifically, number of payments, checkpoint (i.e., feedback = yes), and group (i.e., event type = design) are positively related to crowd productivity. Conversely, Fog Readability Index (i.e., task complexity), event length, and payment size are negatively associated with crowd productivity. All the significant relationships between control variables and crowd productivity remain consistent across four testing models, suggesting the robustness of these relationships.

Model 2 includes only the main explanatory variables. Compared with Model 1, Model 2 demonstrates a good fit with substantial deductions on both AIC and Deviance. The Chi-square likelihood ratio test demonstrates that these deductions are significant. In this model, crowd size significantly relates to crowd productivity in a negative, curvilinear way ($\beta_{crowd\ size} = 0.280\ , p < 0.001;\ \beta_{crowd\ size^2} = -0.007, p < 0.001$), supporting the competition view that crowd size is related to crowd productivity in an inverted U-shape (H1c). Tenure disparity displays a significant, positive relationship with crowd productivity ($\beta_{tenure\ disparity} = 0.131,\ p < 0.01$). This finding supports the search view that crowd diversity is positively related to crowd diversity (H2a). On the contrary, country variety shows a significant, negative relationship to crowd productivity ($\beta_{country\ variety} = -0.413,\ p < 0.001$), thus supporting the competition view that crowd diversity is negatively related to crowd productivity (H2c).

Both Model 3 and Model 4 in Table 14 test the significance of linear and curvilinear interaction effects between crowd size and crowd diversity (i.e., tenure disparity and country variety). Model 4 fits the data better than Model 3 in terms of the AIC index but slightly increases the residual deviance. We got first punishment by fitting a complicated model to our data. However, the Chi-square likelihood ratio test in Model 4 is significant, suggesting the existence of significant curvilinear interactions terms. Both the linear and curvilinear interaction terms between crowd size and tenure disparity are significantly positive ($\beta_{crowd\ size \times tenure\ disparity} = 0.039,\ p < 0.05;\ \beta_{crowd\ size2 \times tenure\ disparity} = 0.004, p < 0.05$). This finding supports a search view on the positive interaction between crowd size and crowd diversity (H3a). The positive linear and curvilinear interactions between crowd size and tenure disparity suggest that the negative quadratic relationship between crowd size and productivity is less concave under situations of high tenure disparity. This finding also suggests that the axis of symmetry for the quadratic function between crowd size and crowd productivity is moved rightward under situations of high tenure disparity. In a sense, a high level of tenure disparity cancels out some negative influence of crowd size on crowd productivity.

As shown in Table 14, we found mixed interaction effects between crowd size and country variety. Specifically, the linear interaction is significant and negative ($\beta_{crowd\ size \times country\ variety} = -0.177, p < 0.001$). The significant linear interaction supports a competition view on the negative interaction between crowd size and crowd diversity (H3c). However, the curvilinear interaction is positive ($\beta_{crowd\ size2 \times country\ variety} = 0.079, p < 0.001$), which indicates that the negative quadratic relationship between crowd size and productivity is less concave under situation of high country variety. The significant curvilinear interaction supports an innovation search view on the positive interaction between crowd size and crowd diversity (i.e., H3a).

The results of the GLM for crowd efficiency are presented in Table 15 (i.e., shortest task completion time) and Table 16 (average task completion time). We took the same hierarchical approach as we did previously for crowd size. Unlike the negative binomial regression, the generalized linear regression offers F statistics on the overall fit of our regression models. As shown in Table 15 and Table 16, both the basic models (i.e., model 1_1 and Model 1_2) only includes control variables. The fit indices show that these two models fits the data better than their null models. Statistics from both models show that Fog Readability Index (i.e., task complexity), event length, payment size, and checkpoint (yes=1) significantly and positively relate to task completion time, suggesting that these control variables are negatively related to crowd efficiency. The number of payments and design events (i.e., group =1) are negatively associated with task completion time, suggesting that they have a positive implication on crowd efficiency.

The inclusion of the variables of interests (i.e., crowd size, tenure disparity, and country variety) in Model 2_1 and 2_2 significantly improves the fit indices. As indicated by the significant, negative coefficients for the quadratic term in Table 15 and Table 16 (i.e.,$\beta_{crowd\ size^2} = -0.001, p < 0.001; \beta'_{crowd\ size^2} = -0.001, p < 0.001$), crowd size relates to task completion time in an inverted U-shape, which suggests that crowd size is associated with crowd efficiency in a U-shape. This finding means that as crowd size increases, crowd efficiency first declines and then increases. This finding does not support a positive association between crowd size and crowd efficiency (H1b) nor an inverted U-shape between crowd size and crowd efficiency (H1d).

The coefficients for tenure disparity and country variety from both Model 2_1 in Table 15 and Model 2_2 in Table 16 are all significantly positive. These findings suggest that tenure disparity and country variety are positively related to task completion time, indicating the existence of a negative relationship between crowd diversity and crowd efficiency. This finding supports a competition view that there exists a negative association between crowd diversity and crowd efficiency (H2d).

The fit indices of Model 3_1 in Table 15 and those of Model 3_2 in Table 16 indicate that the inclusion of the linear interaction terms marginally improves the model fit (i.e., $F = 3.16, p < 0.05$; $F' = 5.52, p < 0.01$). However, Model 4_1 and Model 4_2 are not significant. None of the coefficients for the curvilinear interaction terms are significant. We conclude that there exists only linear interactions between crowd size and crowd efficiency. Specifically, the coefficients for the linear interaction between crowd size and country variety across Model 3_1 and Model 3_2 are consistent and significantly negative ($\beta_{crowd\ size \times country\ variety} = -0.144, p < 0.01$; $\beta'_{crowd\ size \times country\ variety} = -0.174, p < 0.01$). This finding suggests that crowd size and country variety negatively interact to influence task completion time, thus supporting a search view on positive interaction for crowd efficiency (H3b). As for the influence of interaction between crowd size and tenure disparity on crowd efficiency, we found significant evidence from only Model 3_2 ($\beta'_{crowd\ size \times tenure\ disparity} = -0.014, p < 0.001$), which supports H3b in a similar approach.

Table 15

Generalized Linear Regression Model – Efficiency_1 (Shortest Task Completion Time)

Variables	Model 1_1	Model 2_1	Model 3_1	Model 4_1
Controls				
Number of words	0.005 (0.004)	0.003 (0.004)	0.003 (0.004)	0.003 (0.004)
Fog readability index	0.011*** (0.003)	0.009*** (0.003)	0.009*** (0.003)	0.009*** (0.002)
Event length	0.012*** (0.001)	0.009*** (0.001)	0.008*** (0.001)	0.008*** (0.001)
Payment size	0.046*** (0.002)	0.040*** (0.002)	0.040*** (0.002)	0.040*** (0.002)
Number of payment	-0.123*** (0.023)	-0.117*** (0.022)	-0.122*** (0.023)	-0.123*** (0.023)
Checkpoint (0,1)	0.256*** (0.069)	0.343*** (0.068)	0.325*** (0.069)	0.328*** (0.069)
Group (0, 1)	-0.332*** (0.064)	-0.385*** (0.066)	-0.390*** (0.067)	-0.389*** (0.067)
Main effects				
Crowd size		0.044** (0.009)	0.057*** (0.011)	0.058*** (0.013)
Crowd size^2		-0.001*** (<0.001)	-0.001*** (<0.001)	-0.001*** (<0.001)
Tenure disparity	H2d	0.445*** (0.056)	0.388*** (0.059)	0.390*** (0.066)
Country variety	H2d	1.399*** (0.106)	1.118*** (0.149)	1.210*** (0.168)
Linear interaction				
Crowd size × Tenure disparity			-0.002 (0.005)	-0.002 (0.029)
Crowd size × Country variety	H3b		-0.144** (0.068)	-0.134† (0.074)
Quadratic interactions				
Crowd size^2 × Tenure disparity				-0.0001 (0.0005)
Crowd size^2 × Country variety				-0.003 (0.008)
AIC	20988	20628	20625	20629
Deviance	4404.9	4112.9	4107.5	4107.5
F value	93.07***	96.60***	3.26*	0.04
Dispersion parameter	0.869	0.843	0.845	0.845
N	4702	4702	4702	4702

Notes: Standard errors in parentheses; *** p<0.001, **p<0.01,*p<0.05, †<0.1.

Table 16

Generalized Linear Regression Model – Efficiency_2 (Average Task Completion Time)

Variables		Model 1_2	Model 2_2	Model 3_2	Model 4_2
Controls					
Number of words		0.003	-0.001	-0.001	-0.001
		(0.003)	(0.003)	(0.003)	(0.003)
Fog readability index		0.009***	0.006**	0.006**	0.006**
		(0.002)	(0.002)	(0.002)	(0.002)
Event length		0.012***	0.005***	0.005***	0.005***
		(0.001)	(0.001)	(0.001)	(0.001)
Payment size		0.038***	0.030***	0.030***	0.030***
		(0.002)	(0.002)	(0.002)	(0.002)
Number of payment		-0.120***	-0.109***	-0.112***	-0.109***
		(0.019)	(0.019)	(0.019)	(0.020)
Checkpoint (0,1)		0.598***	0.671***	0.658***	0.654***
		(0.060)	(0.059)	(0.059)	(0.059)
Group (0, 1)		-0.063	-0.116*	-0.138*	-0.139*
		(0.055)	(0.056)	(0.057)	(0.057)
Main effects					
Crowd size			0.100***	0.115***	0.107***
			(0.008)	(0.010)	(0.011)
Crowd size^2			-0.001***	-0.001***	-0.001***
			(<0.001)	(<0.001)	(<0.001)
Tenure disparity	H2d		0.460***	0.427***	0.462***
			(0.048)	(0.051)	(0.056)
Country variety	H2d		1.396***	1.114***	1.135***
			(0.091)	(0.128)	(0.144)
Linear interaction					
Crowd size × Tenure disparity	H3b			-0.014***	0.035
				(0.004)	(0.025)
Crowd size × Country variety	H3b			-0.174**	-0.205**
				(0.059)	(0.064)
Quadratic interactions					
Crowd size^2 × Tenure disparity					-0.001
					(0.001)
Crowd size^2 × Country variety					0.0001
					(0.006)
AIC		23212	22439	22431	22431
Deviance		3425.7	2948.8	2942	2939.7
F value		179.66***	192.61***	5.52**	1.79
Dispersion parameter		0.641	0.619	0.624	0.624
N		4702	4702	4702	4702

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

Robust Checks

Robust Checks for Negative Binomial Regression. We used a negative binomial regression to analyze crowd productivity. Developed by statisticians to analyze non-negative count data with over-dispersion, a negative binomial regression model includes three basic components: overly-dispersed error structure, a link function, and linear predictor(s) (Ballinger, 2004; Hardin, Hilbe, & Hilbe, 2007). In this research, crowd productivity is a count variable. As shown in Table 14, the over-dispersion parameters range from 3.026 to 7.424, justifying the validity of choosing negative binomial regression (Hofer, Cantor, & Dai, 2012) and supporting the assumption of this method on the distribution of error (Crawley, 2012; Hardin et al., 2007).

As suggested by scholars, a miss-specified link could lead to biased coefficient estimations (Crawley, 2007). In our data analysis, we used the default log link for our proposed negative binomial regression. To test the validity of this selection, we compared our findings with those from the other two available link functions (i.e., identity link and square root link) (Crawley, 2007; Hardin et al., 2007; Nelder & Baker, 1972). Our comparison demonstrated that the log link is more robust and effective than the other two link functions for our data for two reasons. First, we ran into convergence issues frequently when using the identity link or square root link to analyze our data, but the convergence was not an issue for a log link. Second, for those models did converge by using either identity link or square root link, the log link provided much better model fit in terms of residual deviance.

The linearity in generalized linear models means that the conditional mean of a response variable (i.e., dependent variable) is equal to a linear combination of the predictors (Fox & Weisberg, 2012). To test the linearity of predictors in our negative binomial models, we performed the Tukey's test for nonadditivity (Castle & Hendry, 2010; Tukey, 1949) by adding the squared terms of the fitted main variables (e.g., tenure disparity and country variety) to our proposed models. The significance level for the Turkey's test is obtained by comparing the statistics with the standard-normal distribution (Fox & Weisberg, 2012). All our test results are not significant, suggesting the non-existence of quadratic terms for these two variables.

To further validate our findings, we performed a regression diagnosis according to the procedures and guidelines proposed by methodologists (Cohen et al., 2013; Kutner et al., 2004; Neter et al., 1996). Our regression diagnoses support our decision to delete the five most influential outliers that are distributed more than three standard deviations beyond the mean of crowd productivity. All these post hoc analyses that we performed demonstrate the robust of our data analysis.

Robust Checks for Generalized Linear Regression. We used a generalized linear model with a Gamma distribution to analyze crowd efficiency which is a positive and right skewed variable (Figure 2). By design, the generalized linear model with a Gamma distribution was developed by scholars to analyze continuous variance with skewed distribution (Ballinger, 2004; Dobson & Barnett, 2008). Similar to a negative binomial regression model, a generalized linear model also has three components: error structure, link function, and linear predictor(s) (Crawley, 2012). The variance of a Gamma distribution is defined by $v(\mu_i = \mu_i^2/v)$ (Dobson & Barnett, 2008; Hardin et al., 2007), i.e., the variance is proportional to the squared mean. The dispersion parameters shown in Table 15 and Table 16 range from 0.619 to 0.869, indicating that the variance is proportional to the mean. These parameters are very close to the results that we calculated using the descriptive statistics in Table 13 (i.e., 0.66 and 0.78), which justifies our selection on the generalized linear model with a Gamma distribution.

The default link function in R for Gamma distribution is an inverse function, and there are other two alternatives: identity function and log function (Fox & Weisberg, 2012). In our data analysis, we compared the performance of all three link functions. All three links offered convergent results for the base model only with control variable. The fit indices based on Model 4_1 are reported in the following Table 17. For the basic model, the identity link offers the best fit, but as we add more variables of interests to the basic model, we ran into convergence issues for both the identity link function and the inverse link function. The log link consistently converges for all models, suggesting the robust of the Gamma distribution with a log link.

Table 17

Fit Indices Comparison for Model 4_1

	Identity link	Log link	Inverse link
AIC	20824	20988	21356
Deviance	4272.4	4404.9	4714.7
F value	106.07***	93.07***	41.18***
Dispersion parameter	0.94	0.869	0.888

To check the linearity issue, we performed the Tukey's test for nonadditivity (Castle & Hendry, 2010; Tukey, 1949) by adding the squares of the fitted main variables (e.g., tenure disparity and country variety) to our proposed models. Test results were not significant. We failed to reject the null that the coefficients for the second order terms of tenure disparity and country variety were zero. This result suggests the non-existence of quadratic terms for these two variables that are in question. Our proposed models meet the linearity requirement for a generalized linear model. To further validate our findings, we performed a regression diagnostics according to the procedures and guidelines proposed by methodologists (Cohen et al., 2013; Kutner et al., 2004; Neter et al., 1996). Our analysis identified a few influential outliers that are beyond the recommended cutoff values. By taking out these influential outliers, our findings still hold and become more significant, further justifying the robust of our data analysis.

Endogeneity Test. In our second data analysis, crowd-level attributes are endogenous variables because they are contingent on the setting of a crowdsourcing event (e.g., payment size, number of payments, and event length). As indicated by tournament theory (e.g., Boudreau et al., 2011, Garcia & Tor, 2009), there exists a possibility that suppliers observing a relatively large number of submissions may choose to submit their solutions or to withdraw their participation. This possibility suggests the existence of simultaneity between crowd attributes and crowd performance, which might lead to threats to the internal validity of our data analysis (Antonakis, Bendahan, Jacquart, & Lalive, 2010). Our post hoc analyses discovered that only crowd size correlated with the error terms when we used crowd attributes (i.e., crowd size, tenure disparity, and country variety) to predict crowd productivity. This finding further supports the existence of endogeneity problem associated with crowd size in our data analysis.

To address the endogeneity of crowd size, we ran a two-stage least squares (2SLS) regression. Before the 2SLS was executed, we had to identify instrument variables that met validity requirements according to the criteria proposed by scholars (e.g., Antonakis et al., 2010; Wooldridge, 2002): (1) an instrument variable should be significantly and strongly related to the endogenous variable(s); (2) an instrument variable should be uncorrelated with the error terms of the main regression model; (3) an instrument variable must be related to the dependent variable(s) but less strongly than is the endogenous variable(s). Among all the exogenous variables controlled in our data analysis, we found that only Fog Readability Index met all these three requirements when predicting crowd productivity. However, we did not find any exogenous variable that could meet these criteria when analyzing crowd efficiency. We thus performed an endogeneity test for crowd productivity.

In the first stage, crowd size was regressed on all exogenous variables in order to obtain predicted values and error terms (ε_i) for this endogenous variable. In the second stage, the predicted values from the first stage were included as independent variables to replace the actual values of crowd size. The error terms (ε_i) from the first stage were also included as a predictor for crowd productivity in the second stage. The significance of the coefficient for error terms indicates the existence of endogeneity of crowd size (Wooldridge, 2015). Other exogenous variables (e.g., payment size, number of payments, and event length) and two dummy control variables (i.e., checkpoint and group) were also included. Because crowd size and crowd productivity are count variables, we adopted the negative binomial regression instead of the ordinary least square (OLS) regression in the both stages of 2SLS regression analysis (Antonakis et al., 2010; Wooldridge, 2015).

Table 18

Endogeneity Test for Crowd Productivity

Variables	(1) Crowd size (NBR) [a]	(2) Crowd productivity (2SLS) [b]
Controls		
Checkpoint (0,1)	0.232***	0.537***
	(0.046)	(0.042)
Group (0, 1)	-0.038	0.534***
	(0.043)	(0.043)
Number of words	0.006*	-0.004†
	(0.002)	(0.002)
Fog readability index [c]	0.004**	
	(0.002)	
Event length	0.001	-0.018***
	(0.001)	(0.001)
Payment size	0.025***	-0.035***
	(0.001)	(0.003)
Number of payments	0.022	0.168***
	(0.015)	(0.014)
Main effects		
Crowd size		0.039***
		(0.003)
Crowd size^2		-1.357e-05***
		(<0.001)
Tenure disparity		0.027
		(0.044)
Country variety		-0.729***
		(0.089)
$\varepsilon_{Crowd\ size}$		0.458***
		(0.010)
Constant	3.106***	-0.018
	(0.011)	(0.067)
AIC	39960	20344
Deviance	5362.0	4698.3
N	5044	5044

Notes: Standard errors in parentheses; *** p<0.001, **p<0.01,*p<0.05, †<0.1. Models (1) depicts the results of the first-stage regression considering the endogenous crowd size. Model (2) is the second stage regression incorporating the predicted values and residuals from the first stage, other exogenous variables, and dummy control variables. a) and b): negative binomial regression. c): instrument variable.

Table 18 shows the results of the first stage and second stage of regression analyses. In the first stage regression, Fog Readability Index (i.e., the instrument variable) is a significant predictor for crowd size ($\beta = 0.004, p < 0.01$). However, a Chi-square test shows that Fog Readability Index is a weak instrument because a 7.504 Chi-square value ($p = 0.006$) is below the recommended threshold of 9 for strong instrument (Staiger & Stock, 1997; Wooldridge, 2015). A weak instrument might create a limitation in our endogeneity test (Murray, 2006), which suggests that we should interpret our findings on endogeneity test with caution. Besides Fog Readability Index, we found that number of words, payment size, and checkpoint were also significant predictors for endogenous crowd size.

In the second stage, we found that the coefficient for the residual terms of the first stage was statistically different from zero ($\beta_{\varepsilon} = 0.458, p < 0.001$). We rejected the null that crowd size was exogenous and concluded that it was indeed endogenous. This finding suggests that we should use the predicted crowd size from the first stage of 2SLS analysis instead of the actual crowd size in our data analysis. Accordingly, we updated our data analysis and found that the significant findings identified before still held, which demonstrated that the endogeneity of crowd size did not jeopardize the validity of findings after we used the instrumental regression approach. The existence of the endogeneity between crowd size and crowd productivity suggests the simultaneity between these two variables. The managerial implication of this simultaneity will be further addressed in the discussion section.

Summary

The main objective of the second data analysis was to reveal the associations between crowd attributes (e.g., crowd size and crowd diversity) and crowd performance (e.g., crowd productivity and crowd efficiency) and to discover the performance implications of crowd attributes in crowdsourcing. By analyzing secondary data from a crowdsourcing platform, we identified significant relationships proposed in our theoretical model in chapter 4 (Figure 10).

Specifically, we found that (1) crowd size relates to crowd productivity in an inverted U-shape that supports a competition view on the performance implication of crowd size; (2) tenure disparity is positively associated with crowd productivity, which means that a crowd with diversified members in terms of membership length can generate more solutions for buying firms, thus supporting a search view on the performance implication of crowd diversity; (3) the relationship between country variety and crowd productivity is significantly negative, which supports a competition view on the performance implication of crowd diversity; (4) there exists a significant and positive interaction between tenure disparity and crowd size, which justifies a search view on the complementary effect of these two crowd attributes. As for the interactions between country variety and crowd size, we found that a negative linear interaction term and a positive quadratic interaction term. Detailed explanations and implications of these findings will be further discussed in chapter 7.

As for the implications of crowd attributes on crowd efficiency, we found that (1) both tenure disparity and country variety are negatively related to crowd efficiency, supporting a competition view on the performance implication of diversity in crowdsourcing; and (2) both tenure disparity and country variety positively interact with crowd size to make a crowd function more efficiently (i.e., reducing task completion time). This positive linear interaction between crowd size and crowd diversity supports an innovation search view on the complementary effects for crowd attributes on crowd efficiency. We will discuss the implications of these findings in detail in the following chapter.

Another purpose of this data analysis was to test whether innovation search view is more applicable than competition view in explaining the associations between crowd attributes and crowd performance. Based on the above findings, we found that neither an innovation search view nor a competition view can fully explain the performance implications of crowd attributes. We need to combine these two views together to get a full understanding on the mechanisms that underlie the associations between crowd attributes and crowd performance. We thus propose a competition-search view in this dissertation and address the application of this new perspective in the discussion section.

This research also discovers an unexpected finding that calls for further exploration. For instance, we found that crowd size relates to crowd efficiency in a U-shape, which is different from the proposed positive relationship based on innovation search view (e.g., Afuah & Tucci, 2012; Laursen & Salter, 2006) and the inverted U-shape relationship based on competition view (e.g., Boudreau et al., 2011; Che & Gale, 2003). This U-shape relationship between crowd size and crowd efficiency suggests the existence of other mechanism(s), other than innovation search and competition that influence the operations of a crowd in crowdsourcing. We will talk about this issue more in the following two chapters.

Chapter 7: Discussion

This chapter discusses the theoretical contributions of this dissertation and managerial implications of our findings. With a purpose to explain the phenomenon of performance variation puzzle in crowdsourcing, we systematically examine the issue of crowd development and its performance implications. Specifically, we first propose a double-funnel crowd development framework to elaborate the process of a crowd development. We then uncover the mechanisms that underline the relationship between crowd development and crowd performance through two empirical studies based on secondary data analysis. Accordingly, we discuss the theoretical contributions and managerial implications separately in the following sections.

Theoretical Contributions

Contributions of the Crowd Development Framework. The proposed double-funnel crowd development framework is the first theoretical model that describes an emerging process – crowd development – in current crowdsourcing literature and supply chain management literature. As crowdsourcing becomes more popular for executives and supply chain managers to solve their innovation-related problems, crowd management becomes an important salient issue for managers and scholars in the supply chain field (Felin et al., 2015; Wooten & Ulrich, 2017). Increasing the understanding of a crowd in crowdsourcing and crowd development process then becomes the first and foremost for scholars in the supply chain field. We believe that the double-funnel crowd development framework proposed in this dissertation makes a few significant theoretical contributions.

First, this framework establishes a developmental perspective to look at the concept of a crowd and crowd attributes. The double-funnel crowd development framework demonstrates that a crowd in crowdsourcing is contingent on a particular crowdsourcing task. This demonstration challenges the dominant view in existing crowdsourcing literature that a crowd is a population outside of the focal firms' organizational boundaries and exists before the start of a crowdsourcing event (Afuah & Tucci, 2012; Malhotra & Majchrzak, 2014). Due to the emergence nature of a crowd in crowdsourcing, crowd-level attributes (e.g., pervasiveness of problem-solving know-how in a crowd) are not what scholars assumed to be recognized inputs for a crowdsourcing decision making (e.g., Afuah & Tucci, 2012). Instead, these attributes remain unknown to managers before the start of an event and are actually outcomes of a crowd development process. These new understanding on a crowd based on the proposed crowd development framework suggests an emergence perspective to understand and examine the human crowd in crowdsourcing.

Second, the proposed double-funnel crowd development framework enriches current supply chain literature by identifying crowd development as a new process from a supply chain management perspective. As many companies apply crowdsourcing to solve their innovation-related problems, the human crowd has emerged as an alternative supplier in supply network. Crowd management thus becomes an emergent issue (Kaganer et al., 2013). However, little research in current crowdsourcing literature has systematically described crowd management, and no framework on crowd engagement has been provided in current crowdsourcing literature (Wooten & Ulrich, 2017). Our proposed double-funnel crowd development framework fills this void by providing a holistic view on the whole crowd development process and partitions this process into four stages. The proposed double-funnel crowd development framework demonstrates the emergence and transient nature of a crowd development process. This is very different from the controlled, systematic, and deliberate nature of a supplier development in traditional sourcing literature (Hahn et al., 1990; Krause & Ellram, 1997). These differences suggest that the criterion-based, performance-oriented supplier development does not work for crowd development in crowdsourcing. Scholars and professionals need to take an emergence perspective to look at crowd development. Our framework thus introduces a new research topic – crowd development – to the supply chain management literature.

Finally, the proposed crowd development framework offers a structural view on how event design in crowdsourcing might influence the development process of a crowd that, in turn, determines the performance outcomes of a crowd. This framework thus creates many new research opportunities for scholars to examine crowdsourcing from supply chain perspective. The direct theoretical contributions of framework are to help us partition crowd development into two testable phases (e.g., crowd formation and crowd evaluation). Based on this partition, we develop two empirical studies in this dissertation that examine the influence of event design on the emergence of a crowd and the performance implication of crowd attributes. As such, this framework offers support to reveal the underlying mechanisms that influence crowd performance and increase academic understanding on suppliers' participation behavior in crowdsourcing and crowd performance variation.

Beyond direct theoretical implications in this dissertation, this framework also maps out a new landscape for academic research on crowdsourcing in supply chain literature. For instance, this framework suggests that crowd attributes mediate the relationships between crowdsourcing event design and crowd performance. Testing and revealing mediation mechanisms can be useful for scholars and professionals to better understand the emergence of a crowd in crowdsourcing. Scholars can further examine individual participant's interactions within the crowd development process from a system dynamics perspective (Größler et al., 2008). It is also worthy of applying more interactive methodology (e.g., agent-based simulation or behavioral experiments) to further understand the implications of event setting (e.g., payment adjustment) on solvers' interaction behaviors (Delre, Jager, Bijmolt, & Janssen, 2010; Delre, Jager, & Janssen, 2007).

Contributions of Contatition Perspective on Crowd Emergence. Our dissertation uncovers the contatition (i.e., contagious competition) mechanism underlying a crowd development in this first empirical study that examines the relationships between elements of event design and crowd emergence. The contagious competition mechanism means that imitation and competition are the two coexisting forces that influence solvers' participation behaviors that, in turn, form a crowd for a particular crowdsourcing event. These two forces are not exclusive to each other as indicated by tournament theory and diffusion theory. Instead, they jointly influence the crowd-development process and create variations on crowd growth rate and realized crowd size for different crowdsourcing events. The discovery of the contagious competition mechanism contributes to the understanding on crowd development in several nontrivial ways.

First, the contatition perspective on crowd emergence offers a complete view on the considerations for suppliers' participation behavior in crowdsourcing. The proposed new perspective suggests that suppliers' interactive participation in crowdsourcing is driven not only by economic considerations from tournament theory (e.g., winning chance and expected returns) but also by social and technical reasons from diffusion theory (e.g., imitation of others and ease of completion). The blended nature of participation behavior determines that both tournament theory and diffusion theory cannot fully explain the spreading of participation behavior within a crowd and crowd emergence in crowdsourcing. This is the main reason we need to combine elements from both theoretical lenses to provide a full story on crowd emergence.

Second, the proposed contatition perspective implies that subgroups within a crowd serve different roles in crowd emergence. Our empirical findings suggest that some senior members with winning record(s) can facilitate the spreading of participation behavior within a crowd that, in turn, can be beneficial for crowd emergence. These crowd members whose behaviors and decision-making are influential to others are referred to as influential agents. Our findings also indicate that if many influential agents participate in a crowdsourcing event, they might exert a negative influence on crowd emergence by overinflating the competition within a crowd. The existence of influential agents in a crowd development process indicates that we need to take a less homogenous view toward participants in a crowd in crowdsourcing. As the strategic supplier management in sourcing literature indicates (e.g., Dyer, Cho, & Chu, 1998; Yan, Choi, Kim, & Yang, 2015), the contatition perspective suggests that we need to take a strategic view to segment crowd members so as to better understand the human crowd in innovation processes.

Third, the discovery of the contatition mechanism answers a research call on crowd management. As the human crowd plays an increasingly important role in innovation processes, scholars call for research that can help organizations better engage, utilize, and organize both internal and external crowds when innovating (Felin et al., 2015). Our literature review indicates that crowd development is an under-explored topic in current crowdsourcing literature. Through examining the underlying mechanisms for crowd development, this study fills a void in current crowdsourcing and supply chain literature. This contatition mechanism offers direct implications for future academic research and creates new conversations on crowd development for scholars in these two research streams. For instance, scholars need to relax the rational behavioral assumption and incorporate social motivations when using tournament theory to understand crowdsourcing.

Last but not least, many empirical findings from this research are helpful for scholars to better understand the influence of crowdsourcing event design. Specifically, we identified a quadratic relationship between task complexity and crowd growth rate. This finding suggests the existence of optimal complexity for a crowdsourcing task, which means that increasing the complexity of a crowdsourcing task by extending the problem scope and making the description slightly difficult to understand may not be a bad thing. We found that early involvement of influential agents can stimulate the growth rate of a crowd, which means that influential agents can increase the attractiveness of a crowdsourcing event and send out positive signals to other potential solvers. We also found that the longer an event lasts, the more solvers participate. This finding implies that time is a constraint for problem solving and is one important factor when solvers evaluate the attractiveness of an event and decide whether to participate. Extending the length of an event might increase the attractiveness of a crowdsourcing event.

Our post hoc analysis identified some thought-provoking conclusions. For instance, we found that the number of payments for each event seems to be more influential than payment size in influencing crowd emergence and more useful in interpreting solvers' participation behavior in crowdsourcing. This finding suggests that a sense of winning and the likelihood of winning might be more important to a solver than how much he or she actually wins. The fact that many suppliers come from low income countries such as Malaysia and India might also dilute the influence of payment size on crowd emergence. We discovered a positive quadratic relationship between event length and crowd growth rate (i.e. U-shape), which suggests the existence of "non-optimality" of event length. This finding means that giving more time for suppliers to solve a particular crowdsourcing event may not necessarily increase the attractiveness of this event. Moreover, we revealed a negative quadratic relationship between the number of influential agents and crowd size (i.e., inverted U-shape), which indicates the existence of optimal number of influential agents for a particular crowdsourcing event. Influential agents can increase the attractiveness of an event and attracts other solvers to participate. However, they can trigger and even intensify the competition mechanism, which reduces the chance of winning for all participants and slow down the crowd growth.

Contributions of Competition-Search View on Crowd Performance. Our second empirical study examined the performance implications of crowd attributes. In this study, we drew from a competition perspective based on tournament theory and a search perspective based on innovation search literature. Through this study, we discovered a competition-search mechanism beneath the variations of crowd performance. The discovery of the competition-search view contributes to the understanding of crowd performance variation in crowdsourcing in several significant ways.

First, the competition-search mechanism indicates that the logic linkage between crowd attributes and crowd performance includes not only a competition process driven by solvers' utility maximization but also a search process over a solution landscape. These two forces are not necessarily exclusive in explaining crowd performance. Instead, they are complementary to each other. This view is different from the predominant thinking in either crowdsourcing literature that mainly conceptualizes crowdsourcing as a solution to distant search (Afuah & Tucci, 2012) or tournament literature that mainly claims crowdsourcing as tournament (Boudreau et al., 2011; Fullerton & McAfee, 1999; Lazear & Rosen, 1981). It explains the puzzle in current crowdsourcing literature that although many scholars argue against increasing the number of participants in a tournament (e.g., Che & Gale, 2003; Garcia & Tor, 2009), the application of a crowd in solving innovation-related problems becomes even more popular nowadays (Roth et al., 2015). This is because these scholars overlook the performance implications of the search mechanism. As crowd size for a particular crowdsourcing event increases, firms can obtain the benefits of distant search over a rugged solution landscape although they might experience some losses due to increased competition within a crowd. The competition-search mechanism thus better explains the application of crowdsourcing than either the tournament theory or innovation search literature does.

Second, the proposed competition-search mechanism explains performance variation puzzle in crowdsourcing and thus answers the grand research question proposed in this dissertation. This mechanism suggests that the crowd level attributes such as size and diversity have direct implications for the crowd performance variations. Although many of the event design elements such as payment size and payment structure are similar, the crowd performance can vary if the emergent crowd level attributes are different. For instance, we mentioned two similar data search programming contests in the introduction sections (Topcoder, 2014a & 2014b). The main reason that one programming contest was more productive than the other is that the crowd for the crowd for the first contest was more diversified in terms of membership tenure. As suggested by the proposed competition-search mechanism, a diversified crowd in terms of knowledge, skills, and experience allows firms to search over a rugged solution landscape and locate participants who are at a marginal position with available problem-solving skills. The crowd for the first crowdsourcing event thus generate more solutions although other conditions remains very similar those of the second event.

Third, the competition-search mechanism demonstrates that we need to take a less homogenous view toward the diversity issue in a crowd. In this research, we mainly examined two sources of differences that exist in a crowd: experience difference measured by crowd members' tenure disparity and participants' origin difference measured by crowd members' country variety. Our empirical findings show that these two sources of differences have very opposite implications for crowd productivity. Specifically, tenure disparity exhibits a positive association with productivity, while country variety shows a negative association. The competition-search mechanism also indicates the different logic behind these associations between crowd diversity and crowd performance. In particularly, tenure disparity follows a distant search perspective to influence crowd performance, while country variety relies on competitive social comparisons among crowd members to impact crowd performance. Due to these different implications and different underlying mechanisms, scholars need to treat the crowd diversity issue differently.

Fourth, this research reveals the complexity of performance implication of crowd attributes and contributes to a better understanding of crowd performance in crowdsourcing. For instance, the inverted U-shape relationship between crowd size and crowd performance (i.e., productivity and efficiency) suggests the existence of optimality of crowd size. A moderate level of competition within a crowd in crowdsourcing can motivate all participant to increase their effort and thus improve the overall performance. However, very strong competition due to a large increase in crowd size can be detrimental. This inverted U-shape relationship between crowd size and crowd performance also suggests that the wisdom of a crowd is not unlimited in tournament-based crowdsourcing situation as what scholars claim (Howe, 2006; Surowiecki, 2005), but actually depends on the number of participants. Moreover, the positive interaction effect between crowd size and tenure disparity on productivity indicates that the effects of crowd size and tenure disparity are complementary. This finding suggests that up-ward comparison (i.e., more tenure -disparity) matters to solvers when they compete with many agents. Furthermore, there exists a negative interaction effect between crowd size and country variety on crowd performance (i.e., productivity and efficiency). This finding implies that the effects of crowd size and country variety are incompatible. When these two factors meet each other, they tend to increase the spatial distance and "psychological distance" among crowd members and might lead to "over-search", which can be detrimental for performance (Laursen & Salter, 2006).

Last but not least, this study creates new research topic related to the influence of crowd size on crowd efficiency. Our data analysis shows that crowd size relates to crowd efficiency in a U-shape, which is different from either the proposed positive relationship based on innovation search view (e.g., Laursen & Salter, 2006) or the inverted U-shape relationship based on competition view (e.g., Boudreau et al., 2011; Che & Gale, 2003). This unexpected finding suggests that the mechanism underlying the relationship between crowd size and crowd efficiency remains unclear. It is possible that an increase in the crowd size might create a shared sense of responsibility toward solving a particular crowdsourcing task within a crowd which might lead to reduced crowd efficiency (Forsyth, 2009; Zimbardo, 2007). The negative effect of shared responsibility could be intensified by the anonymity of crowd members in crowdsourcing and the loosely coupled relationship between crowd members and focal buying firm. It could also be possible that the crowd size has to be large enough for the competition mechanism to be effective to improve the crowd efficiency. Anyway, the implication of crowd size on crowd efficiency requires further exploration.

Managerial Implications

Insights of the Double-Funnel Crowd Development Framework. The purpose of this dissertation was to explain the performance variation puzzle in crowdsourcing and to answer managers' concern and doubts on the application of a crowd in innovation processes. Based on four qualitative cases and structural thinking, we proposed a double-funnel crowd development framework in this dissertation. We believe that our proposed framework offers many strategic and operational implications for managers.

First, this framework indicates that managers cannot use criterion-based supplier development to manage crowd development and need to develop new skills to engage a crowd in crowdsourcing. This is because crowd development is an emergent process in the supply chain field. Unlike a supplier development, a crowd development involves an open call through which suppliers make their own decision to form a loosely coupled relationship with focal buying companies. The limited information visibility within a crowd and suppliers' self-selection make a crowd development full of uncertainty. All these significant differences suggests that managers need to switch from a control mentality in outsourcing to an emergence mentality toward crowd development. As indicated by our double-funnel crowd development, managers need to focus their attention on creating crowdsourcing event design so as to provide a beneficial structure for a crowd to emerge by itself, that is, for suppliers to self-select themselves to join the crowd formation process.

Second, our framework offers an indirect approach for managers to exert their influence on crowd performance. Although our framework suggests that managers need to switch from a controlled mentality in outsourcing context to an emergence in crowdsourcing, this does not mean that managers have no influence on crowd performance. Our proposed framework provides a structural view on how the event setting impacts the crowd formation process and realized crowd attributes that, in turn, impact the crowd performance. In a sense, managers can indirectly exercise their influence by managing the crowdsourcing event design and by supervising the crowd formation process. An understanding of the associations between event design and crowd development and the mechanisms underlying crowd performance thus becomes very critical for managers to organize and engage a crowd in crowdsourcing.

Third, our proposed framework demonstrates the challenges associated with crowd engagement and crowd performance management. A crowd in crowdsourcing is a collective of suppliers that are nested in a virtue network. Our framework suggests that the boundary of a crowd in crowdsourcing is vaguely and loosely determined. Suppliers can self-select to join a crowd and withdraw their participation at any time without any contract liability. Managing a collective of suppliers with fuzzy boundary and no specific organization structure is thus very challenging. The proposed framework also show the uncertainty associated with the crowd development. The crowd for a particular crowdsourcing event realizes at the end of a crowd development process. Who gets involved in this process and how a crowd emerges remain unclear to managers at the beginning of a crowdsourcing event. This is the main reason that we develop this dissertation to uncover the mechanisms underlying crowd development and crowd performance.

Insights of the Contatition Perspective. The contagious competition mechanism proposed in the first empirical study in this dissertation suggests that that imitation and competition are the two coexisting forces that influence solvers' participation behaviors that, in turn, form a crowd for a particular crowdsourcing event. These two forces are not exclusive to each other but jointly influence crowd development process and create variations on crowd growth rate and realized crowd size for different crowdsourcing events. In reality managers concern about crowd growth rate and crowd size for a particular crowdsourcing event when managing the crowd development process. This is because both variables are quantitative indicators for understanding a process that is out of managers' direct control due to agents' self-selection and endogenous participation entry. As indicated by our double-funnel crowd development framework, these two variables could have strong performance implications. The proposed contatition perspective can allow professionals to manage crowd emergence better in several significant ways.

The contatition perspective on crowd emergence offers many options for managers to increase the growth rate of a crowd. Specifically, managers can increase the problems scope of a task and make the task a little bit more challenging to motivate suppliers to participate in solving a crowdsourced problem. Managers can also attract influential agents (i.e., agents with winning records and above average membership length) to participate in a particular event as early as possible to leveraging their positive influence of attracting other suppliers to join the crowd formation. At the same time, manager can increase the number of influential agents by sending customized invitations to these influentials. As indicated by our empirical findings, there exists a U-shape relationship between event length and crowd growth rate. This finding means that events with relative short or long duration are more attractive to suppliers. Managers can create a crowdsourcing event with relative short event length in which suppliers are motivated to participate due to time pressure. They can also think of give suppliers enough time and extend the event length.

If the goal is to increase the crowd size (i.e., the number of participants for a crowdsourcing event), managers can increase the number of payments for each event instead of the payment size since our study suggests that the number of payments seems to be more attractive to suppliers than the total payment size. They can also extend the event length to allow suppliers have more time work on designing their solutions, thus increasing the attractiveness of an event to potential suppliers. Furthermore, managers can attract influential agents to participate in an event with caution. Because the relationship between number of influential agents and crowd size is inverted U-shape, managers should avoid attracting to many influential suppliers for an event and over-inflating the competition within a crowd.

However, our proposed contatition perspective also indicates the challenges associated with crowd management. Our findings suggest that there exists opposing implications of some elements of event settings on growth rate and crowd size. For instance, event length positively relates to crowd size but relates to crowd growth rate in a positive U-shape. This finding suggests that a large increase in the event length might increase the crowd size, but it might have a negative influence on the crowd growth rate. Another element of event design that might have opposing implications for crowd growth rate and crowd size is the number of influential agents. Our data analysis discovered that a large increase in the number of influential agents involved in a crowdsourcing event might stimulate more people to join the crowd formation but it might slow down the growth by over-inflating the competition within a crowd. This finding suggests that managers need to take a less homogenous view towards the crowd members and keep a close eye on the influential suppliers in the crowd formation process.

Insights of the Competition-Search View. Productivity and efficiency are two of the main managerial focuses in crowdsourcing. High productivity means that managers can get multiple solutions for one particular task. Comparing the traditional internal development (e.g., hiring engineers) or contract with suppliers (i.e., outsourcing), crowdsourcing allows managers to harness the wisdom of a crowd and leverage competition mechanism to determine the best outcomes. High efficiency means that managers can reduce the cycle time for their innovation processes and achieve competitive advantages on market competition. An understanding of the competition-search perspective can allow managers to administer crowdsourcing event more efficiently and effectively and achieve better outcomes.

Managers can achieve the objective to acquire multiple solutions or designs for a particular crowdsourced problem by directly adjusting the elements of event setting. As indicated by our data analysis in the second empirical study, the number of payment and checkpoint (i.e., feedback = yes) are positively related to crowd productivity, while Fog Readability Index (i.e., task complexity), event length, and payment size are negatively associated with crowd productivity. Managers thus can take utilize these significant findings to create a beneficial context in which suppliers are motivated to compete with each other to provide best solutions. For instance, managers can increase the number of payments to increase each suppliers' sense of winning. They can provide active feedback for participants and build an information loop between buying firms and the crowd which can increase participants' engagement and sense of belonging. Managers can also increase the attractiveness of an event by narrowing the scope of problems and using simple words to describe events.

Managers can also achieve the goal of increasing submissions by leveraging the influence of crowd attributes. Unlike adjusting elements of event settings, managing crowd attributes to increase crowd productivity is more challenging for two reasons. First, crowd attributes in crowdsourcing has dynamic nature since the crowd for an event is evolving in the crowd development process. The crowd realizes in the last minute and dissolves after the deadline of an event. By that time, solutions have been submitted. Second, all the main relationships between crowd attributes and crowd productivity are not linear but quadratic. For instance, crowd size relates to crowd productivity in an inverted U-shape, suggesting that managers can increase the productivity by increasing the number of participants or attracting participants with a wide range of skills. There also exists a positive interaction. This finding suggests that if managers can take these two actions simultaneously, they have more chances to achieve a productive crowd. In terms of country variety, this variable relates to crowd productivity in a U-shape and negatively interacts with crowd size, indicating that managers need to address where the solvers come from. If they can narrow the geographical distance or spatial distance, it's likely that they can get a productive crowd.

Our endogeneity test on crowd size revealed the existence of simultaneity between crowd size and crowd productivity. This simultaneity suggests that crowd size (i.e., the number of participants) have an influence on the performance of a crowd. It also suggests that crowd productivity (i.e., the number of submitted solutions) influences suppliers' participation and submission behaviors, which demonstrates the dynamic nature of crowd performance in crowdsourcing. The simultaneity between crowd size and crowd productivity indicates that managers can attract more participants for a particular crowdsourcing event by making the number submissions public available to all potential suppliers. Since crowd size relates to crowd productivity in an inverted U-shape, managers should be cautious about the "N-effect" caused by over competition.

If the objective for a crowd is to solve a crowdsourced project more efficiently (i.e., using less time to complete a task), manages can consider reducing task complexity and event length, choosing a relative small payment size while increasing the number of payments, and providing no feedback during the crowd formation process. In terms of managing crowd attributes to increase crowd efficiency, managers can think of either reducing the diversity or increasing the crowd size. This is because both tenure disparity and country variety are positively related to shortest task completion time and average task completion time. That is, attracting participants with similar skills and geographical background can be associated with shorter task competition time (i.e., high crowd efficiency). Crowd size relates to task completion time in an inverted U-Shape and negatively interacts with crowd diversity (i.e., tenure disparity and country variety), suggesting that increasing the crowd size might lead to short completing time, i.e., high crowd efficiency.

Chapter 8: Conclusion

Overview

The grand research question that we addressed in this dissertation is how a crowd development impacts crowd performance in crowdsourcing. We attempted to understand the mechanisms that cause the performance of a crowd to vary across different situations and to generate knowledge through which managers can use to increase crowd performance. To achieve this objective, we first took a structural thinking view and proposed a double-funnel model on crowd development based on four anecdotal crowd development examples. We argued that the elements of event setting (e.g., payment size, payment structure, and event length) created a virtue structure within which agents (e.g., solvers) interact through mechanisms such as competition and contagious imitation to influence the outcomes of a crowdsourcing events.

Through this proposed double-funnel process framework, we not only acquired the basic knowledge about crowd development but also partitioned crowd development into two main stages: crowd emergence and crowd evaluation. This partition allowed us to develop two empirical tests to answer our grand research question. In this first test, we studied whether a competition mechanism or a contagious imitation mechanism can better explain the relationships between elements of event setting and crowd emergence. In the second test, we examined which a distant search mechanism or a competition mechanism can better describe the relationships between crowd attributes (e.g., crowd size and crowd diversity) and crowd performance. Through these two empirical analyses based on secondary data from a crowdsourcing platform, we found that none of the proposed mechanisms could explain our proposed relationships. Instead, we had to combine pieces from two seemingly opposing mechanisms to fully understand the performance implications of crowd development. We thus concluded that this dissertation revealed two actual mechanisms, i.e., contagious competition (i.e., contatition) mechanism and competition-search mechanism through which crowd development influences crowd performance in crowdsourcing.

Limitations

Just like any other studies, this dissertation has limitations. The first limitation comes from the singular secondary data source. In this dissertation, all our data came from one crowdsourcing platform (i.e., Topcoder). Our observations are limited to programing crowdsourcing events. This single data source might create limitations on the interpretation of our findings. However, we believe that the singular data source does not deter the generalizability of our findings. This is mainly because all the programming events are set up in a tournament-based format. Our findings are thus generalizable to all tournament-based crowdsourcing events. In this dissertation, we takes a meaningful first attempt by using the web crawling techniques to assemble secondary data from Topcoder. There are other crowdsourcing platforms available (e.g., InnoCentive, Kaggle, and Eyeka). There are also other forms of crowdsourcing formats (e.g., cooperation-based crowdsourcing and coopetition-based crowdsourcing). Future research on crowd development can apply multiple data sources to examine crowd development under different situations.

The second limitation of this dissertation is that we do not have qualitative performance outcomes when we operationalize crowd performance. We only included productivity and efficiency due to lack of availability of qualitative data. This is not necessary an inefficiency or drawback of our research design because research shows that managers have obvious selection biases when they evaluate the quality of solutions in crowdsourcing (Bockstedt et al., 2015; Boudreau, Guinan, et al., 2016; Piezunka & Dahlander, 2015). However, if we could incorporate quality into our analysis and compare the findings across productivity, efficiency, and quality, that would increase the granularity of our data analyses.

Third, our data analysis might suffer from potential threats from endogeneity. In our first empirical analysis, we used the Bass Model to operationalize the crowd growth rate. We run into a convergence issue when we applied regression techniques to identify the parameters of the Bass Model. This convergence issue caused a shrinkage on the sample size. Although an ANOVA analysis confirmed that this shrinkage did not cause obvious threats to our data analysis, we could not completely eliminate the sample selection biases associated with the measurement of crowd growth rate.

In our second empirical study, we could not find any effective instrument variable in our data to address the potential threats of endogeneity for crowd efficiency. Besides, there was a small percentage of observations missing efficiency data due to unobservable submission (i.e., no submission before the deadline). We conducted extensive post hoc analysis to justify our findings and to make sure these potential threats did not bias our findings. We were confident about our findings; however, we could not completely eliminate this potential influence of endogeneity in our data analysis.

Future Research Ideas

In order to reduce the limitation of data source on the interpretation of our empirical findings, we believe that it’s worthwhile to further validate our findings using archival data from other sources such as InnoCentive and Local Motors in the future. Besides the secondary data, we believe that we can consider other methods to further understand interactive crowd behavior and to avoid the potential sample selection biases and endogeneity issue. For instance, we can consider using behavioral experiment or agent-based simulation to understand how suppliers respond to the different events setting and interact with each other in the crowd development process. We can also use these interactive methodologies to understand the performance implication of event design, especially the influence on the qualitative crow-level performance. As suggested by our discussions on limitations, we need design or collect more exogenous variables that might be used as instrument variables to address the potential endogeneity of crowd-level attributes when we study the performance implications of crowd attributes in the future.

Our empirical findings in this dissertation suggest a few further research ideas. For instance, our data analysis discovered an unexpected U-shape relationship between crowd size and crowd efficiency. This finding suggests that the mechanism underlying this U-shape relationship remains unclear. We suggest to use qualitative case study or behavior experiment to further explore the underlying mechanisms. Our proposed double-funnel crowd development suggests a mediation model in which crowd level attributes and crowd level state variables mediate the relationships between elements of event design and crowd performance. Revealing the significant mediation mechanism can generate many meaningful implications for managers.

Through our informal qualitative conversations with executives and managers in our data collection process, we found that they were struggling with the growth of crowd member community. We believe that scholars can also extend this crowd development research stream by considering other potential research topics such as membership retention and "make vs. buy" crowd development. These topics are similar to the supply base management and product or service "make vs buy" in traditional sourcing literature. Because of the context difference, it is interesting to see whether the knowledge gained from sourcing literature holds in crowdsourcing situation.

Publication Plan

We plan to publish three peer-reviewed papers out of this dissertation: one conceptual paper and two empirical papers. We will work on publishing the two empirical papers first and then the conceptual piece. The potential timeline for each paper is as follows: (1) submit the first empirical to JOM (*Journal of Operations Management*) by October 2017; (2) submit the second empirical to POM (*Production and Operations Management*) by December 2017; (3) submit the conceptual paper to JSCM (*Journal of Supply Chain Management*) by May 2018.

REFERENCES

Afuah, A., & Tucci, C. L. (2012). Crowdsourcing as a solution to distant search. *Academy of management review, 37*(3), 355-375.

Aghion, P., Blundell, R., Griffith, R., Howitt, P., & Prantl, S. (2009). The effects of entry on incumbent innovation and productivity. *The Review of Economics and Statistics, 91*(1), 20-32.

Aghion, P., Harris, C., Howitt, P., & Vickers, J. (2001). Competition, imitation and growth with step-by-step innovation. *The Review of Economic Studies, 68*(3), 467-492.

Ales, L., Cho, S.-H., & Körpeoğlu, E. (2017). Optimal award scheme in innovation tournaments. *Operations Research*, 65(3), 693-702.

Anabtawi, I. (2005). Explaining pay without performance: The tournament alternative. *Emory LJ,* 54(8-03), 1557-1602.

Angst, C. M., Devaraj, S., & D'Arcy, J. (2012). Dual role of IT-assisted communication in patient care: a validated structure-process-outcome framework. *Journal of Management Information Systems, 29*(2), 257-292.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly, 21*(6), 1086-1120.

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

Ashenbaum, B., Salzarulo, P. A., & Newman, W. (2012). Organizational structure, entrepreneurial orientation and trait preference in transportation brokerage firms. *Journal of Supply Chain Management, 48*(1), 3-23.

Baddeley, M. (2010). Herding, social influence and economic decision-making: socio-psychological and neuroscientific analyses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1538), 281-290.

Baer, M., & Oldham, G. R. (2006). The curvilinear relation between experienced creative time pressure and creativity: moderating effects of openness to experience and support for creativity. *Journal of Applied Psychology, 91*(4), 963-970.

Bain, J. S. (1956). *Barriers to New Competition: Their Character and Consequences in Manufacturing Industries*: Cambridge: Harvard University Press.

Ballinger, G. A. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational research methods, 7*(2), 127-150.

Bascle, G. (2008). Controlling for endogeneity with instrumental variables in strategic management research. *Strategic organization, 6*(3), 285-327.

Bass, F. M. (1969). A new product growth for model consumer durables. *Management science, 15*(5), 215-227.

Bayus, B. L. (2013). Crowdsourcing new product ideas over time: An analysis of the Dell IdeaStorm community. *Management science, 59*(1), 226-244.

Beach, C. M., & MacKinnon, J. G. (1978). A maximum likelihood procedure for regression with autocorrelated errors. *Econometrica: journal of the Econometric Society*, 46(1),51-58.

Becker, B. E., & Huselid, M. A. (1992). The incentive effects of tournament compensation systems. *Administrative science quarterly*, 37(2), 336-350.

Becker, T. E. (2005). Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods, 8*(3), 274-289.

Bell, R. M., & Koren, Y. (2007). Lessons from the Netflix prize challenge. *Acm Sigkdd Explorations Newsletter, 9*(2), 75-79.

Bellamy, M. A., Ghosh, S., & Hora, M. (2014). The influence of supply network structure on firm innovation. *Journal of Operations Management, 32*(6), 357-373.

Bendoly, E., Croson, R., Goncalves, P., & Schultz, K. (2010). Bodies of knowledge for research in behavioral operations. *Production and operations management, 19*(4), 434-452.

Bennett, J., & Lanning, S. (2007). *The netflix prize.* Paper presented at the Proceedings of KDD cup and workshop.

Betancourt, R., & Kelejian, H. (1981). Lagged endogenous variables and the cochrane-orcutt procedure. *Econometrica: journal of the Econometric Society*, 49(4), 1073-1078.

Billington, C., & Davidson, R. (2013). Leveraging open innovation using intermediary networks. *Production and operations management*, *22*(6), 1464-1477.

Bjelland, O. M., & Wood, R. C. (2008). An inside view of IBM's' Innovation Jam'. *MIT Sloan management review, 50*(1), 32-40.

Blau, P. M. (1977). *Inequality and heterogeneity: A primitive theory of social structure* (Vol. 7): Free Press New York.

Bloom, M., & Michel, J. G. (2002). The relationships among organizational context, pay dispersion, and among managerial turnover. *Academy of management journal, 45*(1), 33-42.

Bockstedt, J., Druehl, C., & Mishra, A. (2015). Problem-solving effort and success in innovation contests: The role of national wealth and national culture. *Journal of Operations Management,* 36(3), 187-200.

Bockstedt, J., Druehl, C., & Mishra, A. (2016). Heterogeneous submission behavior and its implications for success in innovation contests with public submissions. *Production and operations management*, 25(7), 1157-1176.

Bognanno, M. L. (2001). Corporate tournaments. *Journal of Labor Economics, 19*(2), 290-315.

Boswijk, H. P., & Franses, P. H. (2005). On the econometrics of the Bass diffusion model. *Journal of Business & Economic Statistics, 23*(3), 255-268.

Bothner, M. S., Kang, J.-h., & Stuart, T. E. (2007). Competitive crowding and risk taking in a tournament: Evidence from NASCAR racing. *Administrative science quarterly, 52*(2), 208-247.

Boudreau, K., Lacetera, N., & Lakhani, K. R. (2010). The Effects of Increasing Competition and Uncertainty on Incentives and Extreme-Value Outcomes in Innovation Contests. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1531368.

Boudreau, K. J., Guinan, E. C., Lakhani, K. R., & Riedl, C. (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management science, 62*(10), 2765-2783.

Boudreau, K. J., Lacetera, N., & Lakhani, K. R. (2011). Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management science, 57*(5), 843-863.

Boudreau, K. J., & Lakhani, K. R. (2013). Using the crowd as an innovation partner. *Harvard Business Review, 91*(4), 60-69.

Boudreau, K. J., Lakhani, K. R., & Menietti, M. (2016). Performance responses to competition across skill levels in rank-order tournaments: field evidence and implications for tournament design. *The RAND Journal of Economics, 47*(1), 140-165.

Boyd, T. C., & Mason, C. H. (1999). The link between attractiveness of “extrabrand” attributes and the adoption of innovations. *Journal of the Academy of Marketing Science*, *27*(3), 306-319.

Brabham, D. C. (2008). Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence, 14*(1), 75-90.

Brabham, D. C. (2010). *Crowdsourcing as a model for problem solving: leveraging the collective intelligence of online communities for public good*: The University of Utah.

Brabham, D. C. (2010). Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application. *Information, Communication & Society, 13*(8), 1122-1145.

Brabham, D. C. (2012). Motivations for participation in a crowdsourcing application to improve public engagement in transit planning. *Journal of Applied Communication Research*, *40*(3), 307-328.

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: journal of the Econometric Society*, 47(5), 1287-1294.

Brown, J. (2011). Quitters never win: The (adverse) incentive effects of competing with superstars. *Journal of Political Economy, 119*(5), 982-1013.

Burt, R. S. (1982). *Toward a structural theory of action: network models of social Structure, Perception, and Action*: New York: Academic Press.

Cameron, A. C., & Trivedi, P. K. (1986). Econometric models based on count data. Comparisons and applications of some estimators and tests. *Journal of applied econometrics,* 1(1), 29-53.

Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (Vol. 53): Cambridge university press.

Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of management review, 13*(1), 40-52.

Campbell, S. M., Roland, M. O., & Buetow, S. A. (2000). Defining quality of care. *Social science & medicine, 51*(11), 1611-1625.

Castle, J. L., & Hendry, D. F. (2010). A low-dimension portmanteau test for non-linearity. *Journal of Econometrics, 158*(2), 231-245.

Caves, R. E. (1987). *American industry: Structure, conduct, performance*: Prentice Hall.

Chandrasekaran, D., & Tellis, G. J. (2007). A critical review of marketing research on diffusion of new products *Review of marketing research* (pp. 39-80): Emerald Group Publishing Limited.

Chapman, A. J. (1973). Funniness of jokes, canned laughter and recall performance. *Sociometry*, 36(4), 569-578.

Che, Y.-K., & Gale, I. (2003). Optimal design of research contests. *The American economic review, 93*(3), 646-671.

Chen, H., Ham, S. H., & Lim, N. (2011). Designing multiperson tournaments with asymmetric contestants: An experimental study. *Management science, 57*(5), 864-883.

Chesbrough, H. (2012). GE's ecomagination Challenge. *California management review, 54*(3), 140-154.

Chesbrough, H. W. (2006). *Open innovation: The new imperative for creating and profiting from technology*: Harvard Business Press.

Choi, T. Y., Dooley, K. J., & Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: control versus emergence. *Journal of Operations Management, 19*(3), 351-366.

Christakis, N. A., & Fowler, J. H. (2013). Social contagion theory: examining dynamic social networks and human behavior. *Statistics in medicine, 32*(4), 556-577.

Clough, P., Sanderson, M., Tang, J., Gollins, T., & Warner, A. (2013). Examining the limits of crowdsourcing for relevance assessment. *IEEE Internet Computing, 17*(4), 32-38.

Cohen, J., Cohen, P., West, S., & Aiken, L. (1983). Missing data *Applied Multiple Regression; Correlation Analysis for the Behavioral Sciences* (pp. 275-300): Erlbaum, Hillsdale, NJ.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*: Routledge.

Cohen, S. G., & Bailey, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management, 23*(3), 239-290.

Collins-Thompson, K., & Callan, J. (2005). Predicting reading difficulty with statistical language models. *Journal of the American Society for Information Science and Technology, 56*(13), 1448-1462.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management, 37*(1), 39-67.

Connelly, B. L., Tihanyi, L., Crook, T. R., & Gangloff, K. A. (2014). Tournament theory: Thirty years of contests and competitions. *Journal of Management, 40*(1), 16-47.

Crawley, M. J. (2007). Generalized linear models. *The R book*, 511-526.

Crawley, M. J. (2012). *The R book*: John Wiley & Sons.

Daniel, S., Agarwal, R., & Stewart, K. J. (2013). The effects of diversity in global, distributed collectives: A study of open source project success. *Information Systems Research, 24*(2), 312-333.

Davis, R. H. (2006). Strong inference: Rationale or inspiration? *Perspectives in biology and medicine, 49*(2), 238-250.

DeCanio, S. J., Dibble, C., & Amir-Atefi, K. (2000). The importance of organizational structure for the adoption of innovations. *Management science, 46*(10), 1285-1299.

Delre, S. A., Jager, W., Bijmolt, T. H., & Janssen, M. A. (2010). Will it spread or not? The effects of social influences and network topology on innovation diffusion. *Journal of Product Innovation Management, 27*(2), 267-282.

Delre, S. A., Jager, W., & Janssen, M. A. (2007). Diffusion dynamics in small-world networks with heterogeneous consumers. *Computational and Mathematical Organization Theory, 13*(2), 185-202.

Denis, J.-L., Hébert, Y., Langley, A., Lozeau, D., & Trottier, L.-H. (2002). Explaining diffusion patterns for complex health care innovations. *Health care management review, 27*(3), 60-73.

Devaraj, S., Fan, M., & Kohli, R. (2006). Examination of online channel preference: using the structure-conduct-outcome framework. *Decision Support Systems, 42*(2), 1089-1103.

DeViney, N., Sturtevant, K., Zadeh, F., Peluso, L., & Tambor, P. (2012). Becoming a globally integrated enterprise: Lessons on enabling organizational and cultural change. *IBM Journal of Research and Development, 56*(6), 2: 1-2: 8.

Dirks, K. T. (1999). The effects of interpersonal trust on work group performance. *Journal of Applied Psychology, 84*(3), 445-455.

Dissanayake, I., Zhang, J., & Gu, B. (2015). Task division for team success in crowdsourcing contests: resource allocation and alignment effects. *Journal of Management Information Systems, 32*(2), 8-39.

Dobson, A. J., & Barnett, A. (2008). *An introduction to generalized linear models*: CRC press.

Donabedian, A. (1966). Evaluating the quality of medical care. *The Milbank memorial fund quarterly, 44*(3), 166-206.

Donabedian, A. (1988). The quality of care: how can it be assessed? *Jama*, *260*(12), 1743-1748.

Dooley, K., & Corman, S. (2002). Agent-based, genetic, and emergent computational models of complex systems. *Encyclopedia of Life Support Systems (EOLSS). UNESCO/EOLSS Publishers, Oxford, UK*. http://www.eolss.net/sample-chapters/c15/e1-29-02-00.pdf.

Dooley, K. J. (1997). A complex adaptive systems model of organization change. *Nonlinear dynamics, psychology, and life sciences, 1*(1), 69-97.

Dosi, G., Levinthal, D. A., & Marengo, L. (2003). Bridging contested terrain: linking incentive-based and learning perspectives on organizational evolution. *Industrial and Corporate Change, 12*(2), 413-436.

Dowlatshahi, S. (1998). Implementing early supplier involvement: a conceptual framework. *International Journal of Operations & Production Management, 18*(2), 143-167.

Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika, 38*(1-2), 159-178.

Dyer, J. H., Cho, D. S., & Chu, W. (1998). Strategic supplier segmentation: The next" best practice" in supply chain management. *California management review, 40*(2), 57-77.

Dyer, J. H., & Singh, H. (1998). The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of management review, 23*(4), 660-679.

Economist, T. (2008). Following the crowd. *Technology Quarterly, 3*.

Ellram, L. M., & Tate, W. L. (2016). The use of secondary data in purchasing and supply management (P/SM) research. *Journal of Purchasing and Supply Management, 22*(4), 250-254.

Euchner, J. A. (2010). The limits of crowds. *Research Technology Management, 53*(5), 7.

Felin, T., Lakhani, K. R., & Tushman, M. (2015). Special issue of Strategic Organization:"Organizing Crowds and Innovation". *Strategic organization, 13*(1), 86-87.

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

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management science, 47*(1), 117-132.

Fleming, L., & Sorenson, O. (2004). Science as a map in technological search. *Strategic management journal, 25*(8-9), 909-928.

Forsyth, D. R. (2009). *Group dynamics*: Cengage Learning.

Fox, J., & Weisberg, S. (2012). *An R Companion to Applied Regression*: Thousand Oaks, Calif: SAGE Publications

Franke, N., Keinz, P., & Klausberger, K. (2013). "Does this sound like a fair deal?": Antecedents and consequences of fairness expectations in the individual's decision to participate in firm innovation. *Organization science, 24*(5), 1495-1516.

Freedman, J. L., & Perlick, D. (1979). Crowding, contagion, and laughter. *Journal of Experimental Social Psychology, 15*(3), 295-303.

Frick, B. (2003). Contest theory and sport. *Oxford Review of Economic Policy, 19*(4), 512-529.

Fritz, R. (1989). *The path of least resistance: learning to become the creative force in your life*: Fawcett Columbine.

Fritz, R. (1996). *Corporate tides: The inescapable laws of organizational structure*: Berrett-Koehler Publishers.

Fritz, R. (1999). *The path of least resistance for managers: Designing organizations to succeed*: Berrett-Koehler Publishers.

Fullerton, R. L., & McAfee, R. P. (1999). Auctionin entry into tournaments. *Journal of Political Economy, 107*(3), 573-605.

Galt, J., & Dale, B. (1991). Supplier development: a British case study. *Journal of Supply Chain Management, 27*(1), 16-22.

Garcia, S. M., & Tor, A. (2009). The N-effect: More competitors, less competition. *Psychological Science, 20*(7), 871-877.

Garcia, S. M., Tor, A., & Gonzalez, R. (2006). Ranks and rivals: A theory of competition. *Personality and Social Psychology Bulletin, 32*(7), 970-982.

Garcia, S. M., Tor, A., & Schiff, T. M. (2013). The psychology of competition a social comparison perspective. *Perspectives on Psychological Science, 8*(6), 634-650.

Gavetti, G., & Levinthal, D. (2000). Looking forward and looking backward: Cognitive and experiential search. *Administrative science quarterly, 45*(1), 113-137.

Gerth, R. J., Burnap, A., & Papalambros, P. (2012). Crowdsourcing: A primer and its implications for systems engineering: DTIC Document.

Geyskens, I., Steenkamp, J.-B. E., & Kumar, N. (1999). A meta-analysis of satisfaction in marketing channel relationships. *Journal of marketing Research*, 223-238.

Gibbons, F. X., & Buunk, B. P. (1999). Individual differences in social comparison: development of a scale of social comparison orientation. *Journal of personality and social psychology, 76*(1), 129-142.

Gieryn, T. F., & Hirsh, R. F. (1983). Marginality and innovation in science. *Social studies of science, 13*(1), 87-106.

Girotra, K., & Netessine, S. (2013). OM forum—business model innovation for sustainability. *Manufacturing & Service Operations Management, 15*(4), 537-544.

Gladwell, M. (2006). *The tipping point: How little things can make a big difference*: Little, Brown.

Goethals, G. R. (1986). Social comparison theory: Psychology from the lost and found. *Personality and Social Psychology Bulletin, 12*(3), 261-278.

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

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic management journal, 17*(S2), 109-122.

Greene, W. H. (2003). *Econometric analysis*: Pearson Education India.

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Quarterly, 82*(4), 581-629.

Größler, A., Thun, J. H., & Milling, P. M. (2008). System dynamics as a structural theory in operations management. *Production and operations management, 17*(3), 373-384.

Guinan, E., Boudreau, K. J., & Lakhani, K. R. (2013). Experiments in open innovation at Harvard Medical School. *MIT Sloan Management Review, 54*(3), 45-52.

Gunning, R. (1969). The fog index after twenty years. *Journal of Business Communication*, *6*(2), 3-13.

Gupta, N., Conroy, S. A., & Delery, J. E. (2012). The many faces of pay variation. *Human Resource Management Review, 22*(2), 100-115.

Gustave, L. B. (1982). The Crowd: A Study of the Popular Mind: Fraser Publishing Company.

Haas, M. R., Criscuolo, P., & George, G. (2015). Which problems to solve? Online knowledge sharing and attention allocation in organizations. *Academy of management journal, 58*(3), 680-711.

Hahn, C., Watts, C., & Kim, K. (1989). *Supplier development program at Hyundai Motor.* Paper presented at the 1989 NAPM Conference Proceedings, Tallahassee, Florida.

Hahn, C. K., Watts, C. A., & Kim, K. Y. (1990). The supplier development program: a conceptual model. *Journal of Supply Chain Management*, *26*(2), 2-7.

Hansen, G. S., & Wernerfelt, B. (1989). Determinants of firm performance: The relative importance of economic and organizational factors. *Strategic management journal*, *10*(5), 399-411.

Hansen, M. T., & Haas, M. R. (2001). Competing for attention in knowledge markets: Electronic document dissemination in a management consulting company. *Administrative science quarterly, 46*(1), 1-28.

Hardin, J. W., Hilbe, J. M., & Hilbe, J. (2007). *Generalized linear models and extensions*: Stata press.

Harper, C. (2015). *Organizations: Structures, processes and outcomes*: Routledge.

Harris, C., & Vickers, J. (1987). Racing with uncertainty. *The Review of Economic Studies, 54*(1), 1-21.

Harris, M. (2015). How a Lone Hacker Shredded the Myth of Crowdsourcing. *Backchannel*. https://backchannel.com/how-a-lone-hacker-shredded-the-myth-of-crowdsourcing-d9d0534f1731.

Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of management review, 32*(4), 1199-1228.

Henderson, A. D., & Fredrickson, J. W. (2001). Top management team coordination needs and the CEO pay gap: A competitive test of economic and behavioral views. *Academy of management journal, 44*(1), 96-117.

Henningsen, A., & Hamann, J. D. (2007). systemfit: A package for estimating systems of simultaneous equations in R. *Journal of statistical software, 23*(4), 1-40.

Hofer, C., Cantor, D. E., & Dai, J. (2012). The competitive determinants of a firm's environmental management activities: Evidence from US manufacturing industries. *Journal of Operations Management, 30*(1), 69-84.

Hoffer, E. (1951). *The True Believer: Thoughts on the Nature of Movements*: New York NY: HarperCollins.

Holland, J. H. (2000). *Emergence: From chaos to order*: OUP Oxford.

Horwitz, S. K., & Horwitz, I. B. (2007). The effects of team diversity on team outcomes: A meta-analytic review of team demography. *Journal of Management, 33*(6), 987-1015.

Howe, J. (2006). The rise of crowdsourcing. *Wired magazine, 14*(6), 1-4.

Howe, J. (2008). *Crowdsourcing: How the power of the crowd is driving the future of business*: Random House.

Humphreys, P. K., Li, W., & Chan, L. (2004). The impact of supplier development on buyer–supplier performance. *Omega, 32*(2), 131-143.

Jain, D. C., & Rao, R. C. (1990). Effect of price on the demand for durables: Modeling, estimation, and findings. *Journal of Business & Economic Statistics, 8*(2), 163-170.

Javadi Khasraghi, H., & Aghaie, A. (2014). Crowdsourcing contests: understanding the effect of competitors' participation history on their performance. *Behaviour & Information Technology, 33*(12), 1383-1395.

Jeppesen, L. B., & Lakhani, K. R. (2010). Marginality and problem-solving effectiveness in broadcast search. *Organization science, 21*(5), 1016-1033.

Jewett, D. L. (2005). What's wrong with single hypotheses?: why it is time for strong-inference-PLUS. *Scientist (Philadelphia, Pa.), 19*(21), 10.

Johns, T., Laubscher, R., & Malone, T. (2011). The age of hyper specialization. *Harvard Business Review, 89*(7/8), 56-65.

Johnson, P. F., & Flynn, A. (2015). *Purchasing and supply management*: McGraw-Hill College.

Jouret, G. (2009). Inside Cisco's search for the next big idea. *Harvard Business Review, 87*(9), 43-45.

Kaganer, E., Carmel, E., Hirschheim, R., & Olsen, T. (2013). Managing the human cloud. *MIT Sloan Management Review,* 54(2), 23-32.

Kale, J. R., Reis, E., & Venkateswaran, A. (2009). Rank-order tournaments and incentive alignment: The effect on firm performance. *The Journal of Finance, 64*(3), 1479-1512.

Katila, R., & Ahuja, G. (2002). Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of management journal, 45*(6), 1183-1194.

Kay, L. (2012). *Technological innovation and prize incentives: The Google Lunar X Prize and other aerospace competitions*: Edward Elgar Publishing.

Keller, E., & Berry, J. (2003). *The influentials: One American in ten tells the other nine how to vote, where to eat, and what to buy*: Simon and Schuster.

Kennedy, P. (2003). *A guide to econometrics*: MIT press.

Kiesling, E., Günther, M., Stummer, C., & Wakolbinger, L. M. (2012). Agent-based simulation of innovation diffusion: a review. *Central European Journal of Operations Research, 20*(2), 183-230.

King, A., & Lakhani, K. R. (2013). Using open innovation to identify the best ideas. *MIT Sloan management review, 55*(1), 41-48.

Knoeber, C. R. (1989). A real game of chicken: contracts, tournaments, and the production of broilers. *Journal of Law, Economics, & Organization, 5*(2), 271-292.

Knoeber, C. R., & Thurman, W. N. (1994). Testing the theory of tournaments: An empirical analysis of broiler production. *Journal of Labor Economics, 12*(2), 155-179.

Koller, M., & Stahel, W. A. (2011). Sharpening wald-type inference in robust regression for small samples. *Computational Statistics & Data Analysis, 55*(8), 2504-2515.

Konrad, K. A. (2009). *Strategy and dynamics in contests*: Oxford University Press.

Körpeoğlu, E., & Cho, S.-H. (2017). Incentives in contests with heterogeneous solvers. *Management science,* 1-7.

Kräkel, M. (2000). Relative deprivation in rank-order tournaments. *Labour Economics, 7*(4), 385-407.

Krause, D. R. (1997). Supplier development: current practices and outcomes. *Journal of Supply Chain Management, 33*(1), 12-19.

Krause, D. R., & Ellram, L. M. (1997a). Critical elements of supplier development The buying-firm perspective. *European Journal of Purchasing & Supply Management, 3*(1), 21-31.

Krause, D. R., & Ellram, L. M. (1997b). Success factors in supplier development. *International Journal of Physical Distribution & Logistics Management, 27*(1), 39-52.

Krause, D. R., Handfield, R. B., & Scannell, T. V. (1998). An empirical investigation of supplier development: reactive and strategic processes. *Journal of Operations Management, 17*(1), 39-58.

Kutner, M. H., Nachtsheim, C., & Neter, J. (2004). *Applied linear regression models*: McGraw-Hill/Irwin.

Lakhani, K. R., Garvin, D. A., & Lonstein, E. (2010). Topcoder (a): Developing software through crowdsourcing. http://www.hbs.edu/faculty/Pages/item.aspx?num=38356.

Lakhani, K. R., Jeppesen, L. B., Lohse, P. A., & Panetta, J. A. (2007). *The value of openess in scientific problem solving*: Division of Research, Harvard Business School.

Lascelles, D. M., & Dale, B. G. (1989). The buyer-supplier relationship in total quality management. *Journal of Supply Chain Management, 25*(2), 10-19.

Laursen, K., & Salter, A. (2006). Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms. *Strategic management journal, 27*(2), 131-150.

Lawrence, K. D., & Lawton, W. H. (1981). Applications of diffusion models: some empirical results (pp. 529-541): Lexington Books, Lexington, MA.

Lazaer, E. P. (1999). Personnel economics: Past lessons and future directions. *Journal of Labor Economics, 17*(2), 199-236.

Lazear, E. P., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of political Economy, 89*(5), 841-864.

Le Bon, G. (1897). *The crowd: A study of the popular mind*: Fischer.

Le Bon, G. (1960). *The crowd: A study of the popular mind*. New York: Viking.

Leimeister, J. M., Huber, M., Bretschneider, U., & Krcmar, H. (2009). Leveraging crowdsourcing: activation-supporting components for IT-based ideas competition. *Journal of Management Information Systems, 26*(1), 197-224.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics, 45*(2), 221-247.

Lin, Y.-F., Yeh, Y. M. C., & Shih, Y.-T. (2013). Tournament theory's perspective of executive pay gaps. *Journal of Business Research, 66*(5), 585-592.

Liu, T. X., Yang, J., Adamic, L. A., & Chen, Y. (2014). Crowdsourcing with all-pay auctions: A field experiment on Taskcn. *Proceedings of the American Society for Information Science and Technology, 48*(1), 1-4.

Local Motors. (2016). Challenge: Airbus Cargo Drone Challenge. https://launchforth.io/localmotors/airbus-cargo-drone-challenge/brief/.

Lorrain, F., & White, H. C. (1971). Structural equivalence of individuals in social networks. *The Journal of mathematical sociology, 1*(1), 49-80.

Mahajan, V., & Muller, E. (1979). Innovation diffusion and new product growth models in marketing. *The Journal of Marketing*, 43(4), 55-68.

Mahajan, V., Muller, E., & Bass, F. M. (1991). New product diffusion models in marketing: A review and directions for research *Diffusion of technologies and social behavior* (pp. 125-177): Springer.

Malhotra, A., & Majchrzak, A. (2014). Managing crowds in innovation challenges. *California Management Review, 56*(4), 103-123.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization science, 2*(1), 71-87.

Marengo, L., Dosi, G., Legrenzi, P., & Pasquali, C. (2000). The structure of problem-solving knowledge and the structure of organizations. *Industrial and Corporate Change, 9*(4), 757-788.

Massimino, B. (2016). Accessing online data: Web-crawling and information-scraping techniques to automate the assembly of research data. *Journal of Business Logistics, 37*(1), 34-42.

Mathieu, J., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. *Journal of Management, 34*(3), 410-476.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology, 27*(1), 415-444.

McWilliams, A., & Smart, D. L. (1993). Efficiency v. structure-conduct-performance: Implications for strategy research and practice. *Journal of Management, 19*(1), 63-78.

Meyer, M., Johnson, J. D., & Ethington, C. (1997). Contrasting attributes of preventive health innovations. *Journal of Communication, 47*(2), 112-131.

Milliken, F. J., & Martins, L. L. (1996). Searching for common threads: Understanding the multiple effects of diversity in organizational groups. *Academy of management review, 21*(2), 402-433.

Modi, S. B., & Mabert, V. A. (2007). Supplier development: Improving supplier performance through knowledge transfer. *Journal of Operations Management, 25*(1), 42-64.

Molm, L. D. (1990). Structure, action, and outcomes: The dynamics of power in social exchange. *American Sociological Review*, 55(3), 427-447.

Monczka, R. M., Trent, R. J., & Callahan, T. J. (1993). Supply base strategies to maximize supplier performance. *International Journal of Physical Distribution & Logistics Management*, *23*(4), 42-54.

Moon, H. R., & Perron, B. (2006). Seemingly unrelated regressions. *The New Palgrave Dictionary of Economics*, 1-9.

Morgan, J., Orzen, H., & Sefton, M. (2012). Endogenous entry in contests. *Economic Theory, 51*(2), 435.

Morgan, J., & Wang, R. (2010). Tournaments for ideas. *California Management Review, 52*(2), 77-97.

Moritz, M., Redlich, T., & Wulfsberg, J. (2016). Collaborative competition or competitive collaboration? Exploring the user behavior paradox in community-based innovation contests. *Wulfsberg• Redlich• Moritz*, 233.

Mukherjee, K., & Hogarth, R. M. (2010). The N-Effect possible effects of differential probabilities of success. *Psychological Science*. 21(5), 745-747.

Muresan, G., Cole, M., Smith, C. L., Liu, L., & Belkin, N. J. (2006). *Does familiarity breed content? taking account of familiarity with a topic in personalizing information retrieval.* Paper presented at the System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference.

Murray, M. P. (2006). Avoiding invalid instruments and coping with weak instruments. *The journal of economic perspectives, 20*(4), 111-132.

Nadler, J., Thompson, L., & Boven, L. V. (2003). Learning negotiation skills: Four models of knowledge creation and transfer. *Management science*, *49*(4), 529-540.

Neal, C. O. (1993). Concurrent engineering with early supplier involvement: a cross-functional challenge. *Journal of Supply Chain Management, 29*(2), 2-9.

Nelder, J. A., & Baker, R. J. (1972). Generalized linear models. *Journal of the Royal Statistical Socitey,* 135(3), 370-384.

Nelson, R. R., & Winter, S. G. (2009). *An evolutionary theory of economic change*: Harvard University Press.

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models* (Vol. 4): Irwin Chicago.

Netflix. (2009). Grand Prize awarded to team BellKor's Pragmatic Chaos. http://www.netflixprize.com/community/topic_1537.html.

Newton, J. W., & Mann, L. (1980). Crowd size as a factor in the persuasion process: A study of religious crusade meetings. *Journal of personality and social psychology, 39*(5), 874-883.

Nieken, P., & Sliwka, D. (2010). Risk-taking tournaments–Theory and experimental evidence. *Journal of Economic Psychology, 31*(3), 254-268.

Nirwan, V. S. (2014). Interpersonal trust and team performance: a QuantItatIve study. *Journal of Organisation and Human Behaviour, 3*(4), 10-14.

Ocasio, W. (1997). Towards an attention‐based view of the firm. *Strategic management journal*, 18, 187-206.

Pepinsky, P. N., Pepinsky, H. B., & Pavlik, W. B. (1960). The effects of task complexity and time pressure upon team productivity. *Journal of Applied Psychology, 44*(1), 34-38.

Peres, R., Muller, E., & Mahajan, V. (2010). Innovation diffusion and new product growth models: A critical review and research directions. *International Journal of Research in Marketing, 27*(2), 91-106.

Piezunka, H., & Dahlander, L. (2015). Distant search, narrow attention: How crowding alters organizations' filtering of suggestions in crowdsourcing. *Academy of management journal, 58*(3), 856-880.

Piller, F. T., & Walcher, D. (2006). Toolkits for idea competitions: a novel method to integrate users in new product development. *R&d Management, 36*(3), 307-318.

Platt, J. R. (1964). Strong inference. *science, 146*(3642), 347-353.

Prassler, E. (2016). Robotics Academia and Industry: We Need to Talk. *IEEE Robotics & Automation Magazine, 23*(3), 11-14.

Prpić, J., Shukla, P. P., Kietzmann, J. H., & McCarthy, I. P. (2015). How to work a crowd: Developing crowd capital through crowdsourcing. *Business Horizons, 58*(1), 77-85.

Raafat, R. M., Chater, N., & Frith, C. (2009). Herding in humans. *Trends in cognitive sciences, 13*(10), 420-428.

Rabinovich, E., & Cheon, S. (2011). Expanding horizons and deepening understanding via the use of secondary data sources. *Journal of Business Logistics, 32*(4), 303-316.

Ralston, P. M., Blackhurst, J., Cantor, D. E., & Crum, M. R. (2015). A structure–conduct–performance perspective of how strategic supply chain integration affects firm performance. *Journal of Supply Chain Management, 51*(2), 47-64.

Rand, W., Herrmann, J., Schein, B., & Vodopivec, N. (2015). An agent-based model of urgent diffusion in social media. *Journal of Artificial Societies and Social Simulation, 18*(2), 1-24.

Randall, R., Ramaswamy, V., & Ozcan, K. (2013). Strategy and co-creation thinking. *Strategy & Leadership, 41*(6), 5-10.

Raynor, M. E., & Panetta, J. A. (2008). A better way to R&D. *Harvard Business Review,* https://hbr.org/2008/02/a-better-way-to-rd.html.

Reagans, R., & McEvily, B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. *Administrative science quarterly, 48*(2), 240-267.

Reicher, S. D. (2001). *The psychology of crowd dynamics* (Vol. 44): Blackwell handbook of social psychology: Group processes.

Ren, Y., Chen, J., & Riedl, J. (2015). The impact and evolution of group diversity in online open collaboration. *Management science, 62*(6), 1668-1686.

Rivkin, J. W. (2001). Reproducing knowledge: Replication without imitation at moderate complexity. *Organization science, 12*(3), 274-293.

Roberts, J. A., Hann, I.-H., & Slaughter, S. A. (2006). Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the Apache projects. *Management science, 52*(7), 984-999.

Robertson, T. S. (1967). The process of innovation and the diffusion of innovation. *The Journal of Marketing*, 14-19.

Rogers, E. M. (1962). *Diffusion of innovations*: Simon and Schuster.

Rogers, E. M. (2010). *Diffusion of innovations*: Simon and Schuster.

Rogers, E. M., & Shoemaker, F. F. (1971). *Communication of Innovations; A Cross-Cultural Approach*: New York, London, Collier MaCmillan Publishers: Free Press.

Rosenfeld, E. (2012). Mountain Dew's 'Dub the Dew' Online Poll Goes Horribly Wrong. *Time*. http://newsfeed.time.com/2012/08/14/mountain-dews-dub-the-dew-online-poll-goes-horribly-wrong/.

Roth, Y., Pétavy, F., & Céré, J. (2015). The state of crowdsourcing in 2015. *eYeka Analyst Report*. https://en.eyeka.com/resources/reports.

Rungtusanatham, M., Forza, C., Koka, B., Salvador, F., & Nie, W. (2005). TQM across multiple countries: convergence hypothesis versus national specificity arguments. *Journal of Operations Management*, *23*(1), 43-63.

Ryan, B., & Gross, N. C. (1943). The diffusion of hybrid seed corn in two Iowa communities. *Rural sociology, 8*(1), 15-24.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, *25*(1), 54-67.

Sailer, L. D. (1978). Structural equivalence: Meaning and definition, computation and application. *Social Networks, 1*(1), 73-90.

Samaddar, S., Nargundkar, S., & Daley, M. (2006). Inter-organizational information sharing: The role of supply network configuration and partner goal congruence. *European journal of operational research*, *174*(2), 744-765.

Schwartz, R. L., & Phoenix, T. (2001). *Learning perl*: O'Reilly & Associates, Inc.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3-4), 591-611.

Shaw, Z. A. (2013). *Learn Python the hard way: A very simple introduction to the terrifyingly beautiful world of computers and code*: Addison-Wesley.

Simula, H. (2013). *The rise and fall of crowdsourcing?* Paper presented at the System Sciences (HICSS) 46th Hawaii International Conference.

Sloane, P. (2012). How Dell and Starbucks crowdsource high volumes of ideas. http://www.destination-innovation.com/how-dell-and-starbucks-crowdsource-high-volumes-of-ideas/.

Spector, P. E., & Brannick, M. T. (2011). Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods, 14*(2), 287-305.

Srinivasan, V., & Mason, C. H. (1986). Technical note—nonlinear least squares estimation of new product diffusion models. *Marketing science, 5*(2), 169-178.

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

Strang, D., & Soule, S. A. (1998). Diffusion in organizations and social movements: From hybrid corn to poison pills. *Annual review of sociology, 24*(1), 265-290.

Stuart, T. E., & Podolny, J. M. (1996). Local search and the evolution of technological capabilities. *Strategic management journal, 17*(S1), 21-38.

Suls, J., Martin, R., & Wheeler, L. (2002). Social comparison: Why, with whom, and with what effect? *Current directions in psychological science, 11*(5), 159-163.

Sultan, F., Farley, J. U., & Lehmann, D. R. (1990). A meta-analysis of applications of diffusion models. *Journal of marketing Research*, 27(1), 70-77.

Surowiecki, J. (2005). *The wisdom of crowds*: Anchor.

Tang, J. C., Cebrian, M., Giacobe, N. A., Kim, H.-W., Kim, T., & Wickert, D. B. (2011). Reflecting on the DARPA red balloon challenge. *Communications of the ACM, 54*(4), 78-85.

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

Topcoder. (2014a). Asteroid Data Hunter Design Challenge. https://www.topcoder.com/challenge-details/30045374/?type=design

Topcoder. (2014b). NASA Enterprise Search Portal Design Concepts Challenge. https://www.topcoder.com/challenge-details/30045231/?type=design

Topcoder. (2017a). Topcoder Challenges Review. https://www.topcoder.com/challenges/

Topcoder. (2017b). Topcoder Community Overview. https://www.topcoder.com/community/members/

Tukey, J. W. (1949). One degree of freedom for non-additivity. *Biometrics, 5*(3), 232-242.

Valente, T. W. (1993). Diffusion of innovations and policy decision-making. *Journal of Communication, 43*(1), 30-45.

Valente, T. W. (1995). *Network models of the diffusion of innovations*: Cresskill, N.J.: Hamption Press.

Valente, T. W., & Rogers, E. M. (1995). The origins and development of the diffusion of innovations paradigm as an example of scientific growth. *Science communication, 16*(3), 242-273.

Van den Bulte, C. (2000). New product diffusion acceleration: Measurement and analysis. *Marketing Science, 19*(4), 366-380.

Van den Bulte, C. (2002). Want to know how diffusion speed varies across countries and products? Try using a Bass model. *PDMA visions*, *26*(4), 12-15.

Van den Bulte, C., & Joshi, Y. V. (2007). New product diffusion with influentials and imitators. *Marketing Science, 26*(3), 400-421.

Van den Bulte, C., & Stremersch, S. (2004). Social contagion and income heterogeneity in new product diffusion: A meta-analytic test. *Marketing Science*, *23*(4), 530-544.

Van den Bulte, C., & Wuyts, S. H. K. (2007). *Social networks in marketing:* Marketing Science Institue.

Van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: a quantitative research synthesis of three socio-psychological perspectives. *Psychological bulletin, 134*(4), 504-535.

Venkatesan, R., Krishnan, T. V., & Kumar, V. (2004). Evolutionary estimation of macro-level diffusion models using genetic algorithms: An alternative to nonlinear least squares. *Marketing Science, 23*(3), 451-464.

Warwick, G. (2016). Airbus, Local Motors move to enable cargo drone services. *Aviation Daily,* http://aviationweek.com/awincommercial/airbus-local-motors-move-enable-cargo-drone-services.

Watts, C. A., & Hahn, C. K. (1993). Supplier development programs: an empirical analysis. *Journal of Supply Chain Management*, *29*(1), 10-17.

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic management journal, 5*(2), 171-180.

Wheeler, L. (1966). Toward a theory of behavioral contagion. *Psychological Review, 73*(2), 179-192.

Wilcox, R. R. (2011). *Introduction to robust estimation and hypothesis testing*: Academic Press.

Winter, S. G. (1984). Schumpeterian competition in alternative technological regimes. *Journal of Economic Behavior & Organization, 5*(3-4), 287-320.

Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational behavior and human decision processes, 37*(1), 60-82.

Wooldridge, J. (2002). *Econometric analysis of cross section and panel data*: MIT Press.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*: Nelson Education.

Wooten, J. O., & Ulrich, K. T. (2017). Idea generation and the role of feedback: Evidence from field experiments with innovation tournaments. *Production and operations management, 26*(1), 80-99.

Wowak, K. D., Craighead, C. W., Ketchen, D. J., & Hult, G. T. M. (2016). Toward a "Theoretical Toolbox" for the supplier-enabled fuzzy front end of the new product development process. *Journal of Supply Chain Management, 52*(1), 66-81.

Yan, T., Choi, T. Y., Kim, Y., & Yang, Y. (2015). A theory of the nexus supplier: A critical supplier from a network perspective. *Journal of Supply Chain Management, 51*(1), 52-66.

Yin, X., & Zajac, E. J. (2004). The strategy/governance structure fit relationship: Theory and evidence in franchising arrangements. *Strategic management journal, 25*(4), 365-383.

Yohai, V. J. (1987). High breakdown-point and high efficiency robust estimates for regression. *The Annals of Statistics*, 15(2), 642-656.

Zabel, A., & Roe, B. (2009). Optimal design of pro-conservation incentives. *Ecological Economics, 69*(1), 126-134.

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of statistical software, 27*(8), 1-25.

Zellner, A., & Huang, D. S. (1962). Further properties of efficient estimators for seemingly unrelated regression equations. *International Economic Review, 3*(3), 300-313.

Zheng, H., Li, D., & Hou, W. (2011). Task design, motivation, and participation in crowdsourcing contests. *International Journal of Electronic Commerce, 15*(4), 57-88.

Zhou, Y., Wilkinson, D., Schreiber, R., & Pan, R. (2008). *Large-scale parallel collaborative filtering for the netflix prize.* Paper presented at the International Conference on Algorithmic Applications in Management.

Zimbardo, P. G. (2007). *Lucifer Effect*: Wiley Online Library.

Zsidisin, G. A., & Smith, M. E. (2005). Managing supply risk with early supplier involvement: a case study and research propositions. *Journal of Supply Chain Management*, *41*(4), 44-57.

APPENDIX A

SOLVERS' TRACE EXTRACTION CODE

The following is the solvers' trace extraction syntax coded in Python. This syntax is used to transfer solvers' unique participation records (i.e., registration time) for a crowdsourcing event to a crowd emergence trajectory over the event cycle time.

```
i = list(range(1,1831))
f = pd.DataFrame(index=i)
print 'starting now'
t=time.time()
for fn in glob.glob('*.xlsx'):
   if fn != '~$30047472.xlsx':
      print fn
          d = pd.read_excel(root_dir+fn)
          d2 = d.ix[:,0:3]
      reg = pd.DataFrame(d2.ix[:,2].dropna())
      reg = reg[1:]
      #start = d2['Unnamed: 1'][0]
      dates = list(reg['Unnamed: 2'])
      dates.append(d2['Unnamed: 1'][0])
      # Create a df, rename column, and sort by date. Convert to datetime format
      a = pd.DataFrame(dates)
      a.columns = ['timestamp']
      a = a.sort_values('timestamp')
      a['datetime'] = pd.to_datetime(a['timestamp'])
      # Calculate the time between each row and the row before it. Convert to
      # hours and get rid of the start date, keeping only registration dates.
      a['time_bw'] = a['datetime'] - a['datetime'].shift(1)
      a['time_bw2'] = (a['time_bw']/np.timedelta64(1,'D'))*24
      a = a[1:]
      # Calculate the cumulative sum of the time difference. Round up.
```

```
        a['cumul_time'] = a['time_bw2'].cumsum()
        a['cumul_time2'] = np.ceil(a['cumul_time'])
        # Create a count by rounded up cumulative sum
        a2 = pd.DataFrame(a.groupby(['cumul_time2']).size())
        # Naming the column the event number
        a2.columns = [fn.split('.')[0]]
        # Add to the dataframe
        f = f.join(a2)
f2 = f.fillna(0)
f2.to_csv(root_dir+'design_events0301.csv',sep=',')
print 'finished importing files'
```

APPENDIX B

BASS DIFFUSION MODEL IN R

The following is the linear regression of Bass Diffusion syntax coded in R. This syntax is used to generate the three coefficients of the Bass Diffusion Model (i.e., p, q, m). The input for this regression analysis is the output of the APPENDIX A.

```
#plot cumulative number of participants
Y=cumsum(data)
plot(Y,type="l", lty=2,col="red", ylab="", xlab="")
points(Y,pch=20,col="blue")
title("Cumulative participants over time")
#title("Cumulative participants over time(ID30051064)")

#fit bass regression and compute m, p,q
Y_lag=c(0,Y[1:(length(Y)-1)])  # we want Y_t-1 not Y_t. Y_0=0
Ysq=Y_lag**2
out=lm(data~Y_lag+Ysq)
summary(out)
a=out$coef[1]
b=out$coef[2]
c=out$coef[3]
mminus=(-b-sqrt(b**2-4*a*c))/(2*c)
m=mminus
mplus=(-b+sqrt(b**2-4*a*c))/(2*c)
p=a/m
q=b+p

#create a bass diffusion by using m,p,and q.
bassModel=function(p,q,m,T=300)
{
  S=double(T)
```

```
  Y=double(T+1)
  Y[1]=0
  for(t in 1:T)
  {
    S[t]=p*m+(q-p)*Y[t]-(q/m)*Y[t]**2
    Y[t+1]=Y[t]+S[t]
  }
  return(list(data=S,cumdata=cumsum(S)))
}
#compute
Spred=bassModel(p,q,m,T=300)$data
ts.plot(data,Spred,col=c("blue","red"))
legend("topleft",legend=c("actual","Bass Model"),fill=c("blue","red"))

#now do this for cumulateive participants
Spred=bassModel(p,q,m)$data
CumSpred=ts(cumsum(Spred))
CumData=ts(cumsum(data))
ts.plot(CumData,CumSpred,col=c("blue","red"))
legend("topleft",legend=c("actual","Bass Model"),fill=c("blue","red"))
title("Predicted Cumulative participants")
```

APPENDIX C

TEXT EXTRACTION PYTHON CODE

The following syntax is designed to extract the textual information on the description of each programming contest from the saved excel files. It is coded in Python. The extracted textual information is saved in separate txt files for subsequent text analysis in APPENDIX D in the following page.

```
for xlsx in xlsxlist:
    workbook = xlrd.open_workbook(xlsx)
    sheet = workbook.sheet_by_index(0)
    pos = 5
    if sheet.cell_value(pos,0) == 'Detail':
        # print sheet.cell_value(pos+1,0)
        with open(xlsx.split('.')[0]+'.txt','w') as f:
            f.write(sheet.cell_value(pos+1,0))
        content += '\n\n'
        content += xlsx
        content += '\n'
        content += sheet.cell_value(pos+1,0)
    else:
        pos += 1
        # print sheet.cell_value(pos+1,0)
        with open(xlsx.split('.')[0]+'.txt','w') as f:
            f.write(sheet.cell_value(pos+1,0))
        content += '\n\n'
        content += xlsx
        content += '\n'
        content += sheet.cell_value(pos+1,0)
with open('final.txt','w') as f:
    f.write(content)
```

APPENDIX D

LINGUA::EN::FATHOM TEXT ANALYSIS CODE IN PERL

The following is the Lingua::EN::Fathom Text Analysis syntax[6], which is designed to analyze text complexity by calculating the length of a document and generating the FOG readability index. This syntax is coded and run in Perl which is a general-purpose programming language originally developed for text analysis and now used for a wide range of tasks including web development (Schwartz & Phoenix, 2001).

```
use Lingua::EN::Fathom;
my $text = Lingua::EN::Fathom->new();
$text->analyse_file("300496369.txt");
$accumulate = 1;
$text->analyse_block($text_string,$accumulate);

$num_chars             = $text->num_chars;
$num_words             = $text->num_words;
$percent_complex_words = $text->percent_complex_words;
$num_sentences         = $text->num_sentences;
$num_text_lines        = $text->num_text_lines;
$num_blank_lines       = $text->num_blank_lines;
$num_paragraphs        = $text->num_paragraphs;
$syllables_per_word    = $text->syllables_per_word;
$words_per_sentence    = $text->words_per_sentence;
%words = $text->unique_words;
foreach $word ( sort keys %words )
   {
     print("$words{$word} :$word\n");
   }
$fog     = $text->fog;
```

[6] http://search.cpan.org/dist/Lingua-EN-Fathom/lib/Lingua/EN/Fathom.pm

```
$flesch  = $text->flesch;
$kincaid = $text->kincaid;
print($text->report);
```

APPENDIX E

SOLVERS' BACKGROUND STATISTICS EXTRACTION CODES

The following syntax is designed to extract solvers' background statistics (e.g., country origin, membership registration time, winning records, and participation records). This syntax is coded in Python. This program loops through all the excel files that save solvers' information and prints out a grand excel file that includes all solvers' background statistics in one excel file for further data analysis.

```
for xlsx in xlsxlist:
        try:
          missiontime = {}
          regtime = {}
          workbook = xlrd.open_workbook(xlsx)
          sheet0 = workbook.sheet_by_index(0)
          sheet1 = workbook.sheet_by_index(1)
          if sheet0.cell_value(1,0).strip() == "Start Date":
                startdate = sheet0.cell_value(1,1)
          for i in range(5):
            if sheet0.cell_value(6+i,1) == "Username":
                        i = i+1
                        break
          for line in range(6+i,sheet0.nrows):
                missiontime[sheet0.cell_value(line,1)] = sheet0.cell_value(line,2)
          for pos in range(1,sheet1.nrows):
                        if not sheet1.cell_value(pos, 0):
                          pos += 1
                          continue
                        else:
                          eid = xlsx.split('.')[0]
                          pid = sheet1.cell_value(pos, 0)
```

```
                country = sheet1.cell_value(pos, 1)
                since = sheet1.cell_value(pos, 2)
                win = sheet1.cell_value(pos, 3)
                skills = sheet1.cell_value(pos, 4)
            contentstr = [comma(i) for i in
[eid,pid,startdate,missiontime.get(pid),country,since,win, skills]]
                cf.write(','.join(contentstr)+'\n')
    except:
        print xlsx
```

APPENDIX F

EVENT SUMMARY EXTRACTION CODES

The following syntax is designed to extract information at event level (e.g., event ID, total number of participants, total number of submission, and payment size). This syntax is coded in Python. This program loops through all the excel files that save event level information and prints out a grand excel file that includes all event-related statistics in one excel file for further data analysis.

```
import xlrd
import os
xlsxlist = [i for i in os.listdir('./') if i.split('.')[1] == 'xlsx']
cf = open('summary.csv','w')
cf.write('"Event ID","Start Date","Checkpoint","End Date","Total number of Registor","Total number of submission","Payment size","number of payments","Checkpoints(yes,no)"\n')
for xlsx in xlsxlist:
    workbook = xlrd.open_workbook(xlsx)
    sheet0 = workbook.sheet_by_index(0)
    eid = xlsx.split('.')[0]
    if sheet0.cell_value(1,0).strip() == "Start Date":
        startdate = sheet0.cell_value(1,1)
    else:
        startdate = ""
    if sheet0.cell_value(2,0).strip() == "End Date":
        enddate = sheet0.cell_value(2,1)
        checkpoint = 0
    else:
        if sheet0.cell_value(2,0).strip() == "Checkpoint":
            checkpoint = sheet0.cell_value(2,1)
        enddate = sheet0.cell_value(3,1)
    for detailpos in range(8):
```

```
    if sheet0.cell_value(detailpos,0).strip() == 'Detail':
        break
if sheet0.cell_value(detailpos+1,1).strip() == "Username":
    TotalReg = 0
    while 1:
        try:
            if sheet0.cell_value(detailpos+1+TotalReg,2):
                TotalReg += 1
            else:
                break
        except:
            break
if sheet0.cell_value(detailpos+1,5).strip() == "submissionId":
    TotalSub = 0
    while 1:
        try:
            if sheet0.cell_value(detailpos+1+TotalSub,5):
                TotalSub += 1
            else:
                break
        except:
            break
if sheet0.cell_value(detailpos+1,10).strip() == 'prize':
    TotalPay = 0
    npay = 1
    while 1:
        try:
            if sheet0.cell_value(detailpos+1+npay,10):
                TotalPay += int(sheet0.cell_value(detailpos+1+npay,10))
```

```
                npay += 1
            else:
                break
        except:
            break
    npay -= 1
else:
    print(sheet0.cell_value(detailpos+1,10).strip())
if checkpoint:
    checkstr = 'yes'
else:
    checkstr = 'no'
csvline = [comma(i) for i in [eid,startdate,checkpoint,enddate,TotalReg,TotalSub,TotalPay,npay,checkstr]]
csvline = ','.join(csvline)+'\n'
cf.write(csvline)
```