HARNESSING ONLINE COMMUNITY AND SOCIAL MEDIA: ORDINARY USERS AND INFLUENTIAL USERS

LI ZHUOLUN

(B.Eng., South China University of Technology)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF INFORMATION SYSTEMS

NATIONAL UNIVERSITY OF SINGAPORE

2015

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Li Zhuolun

Li Zhuolun

12 August 2015

ACKNOWLEDGEMENTS

I would like to thank all the following persons who have guided, supported and encouraged me throughout the dissertation process.

First and foremost, I would like to express my sincere gratitude to my doctoral advisor Prof. Huang Ke-Wei. Without his patient guidance, I would not have made this thesis possible. I benefit a lot from his excellent expertise in mathematics, econometrics and information system, his insightful observations, and most importantly, his sense of integrity in research. Over the past five years, he taught me how to be a well-trained researcher specializing in rigorous research methodologies. In many ways, he inspires me to be a serious researcher like him.

I would also like to thank my thesis committee members, Prof. Goh Khim Yong and Prof. Phan Tuan Quang, for their invaluable suggestions on the research design and positioning of my papers. Special thanks to Prof. Goh for teaching me econometrics in the Department of Information Systems, serving as the committee member in the evaluation of my GRP and writing the referee report for me. I received a lot of help and support from him during my PhD study.

My thanks also go to all professors in the Department of Information Systems, especially those professors who participate in the lunch seminar regularly. Their helpful comments significantly contributed to this thesis.

I am also grateful to my fellow students. It is their company and encouragement that help me to complete this journey. I would like to extend my sincere appreciation to them all.

Last but not least, I would like to thank my family for their unconditional support and love, which encourage me to explore my life. To them I dedicate this thesis.

Contents

SUMMARY

The business value of an online community greatly depends on the number of active users in this online community. However, it is never easy to establish and maintain a large user base. In this thesis, we investigate how to motivate users' voluntary contribution by gamification (Study 1), and how to incorporate the interaction between different groups of users to maintain a stable user base (Study 2). In addition to establish a large user base, it is also important for online community, especially those social media sites, to attract a group of influential users. These influential users are used to be celebrities or experts who post contents consumed by a large number of ordinary users, thus helping to maintain an active user base. In this thesis, we examine whether these influential users can obtain economic return from their social media participation (Study 3).

In Study 1, we investigate the impact of a hierarchical badge system on users' voluntary contribution on four kinds of activities (answering, commenting, revision, and asking) in a Q&A website. Our results confirm that almost all badges motivate users to contribute more in related activities. There is a spillover effect to other activities. Furthermore, our results reveal that gold badges are more powerful than silver badges and silver badges are more influential than bronze badges. Hence, there is a ranked ordering in efficiency of badges corresponding to difficulty levels. Our results present strong empirical evidence that confirms the effectiveness of gamification.

In Study 2, we investigate the network effect in online community. Empirically, we examine how the increase of user participation drives the decrease

of senior users' participation at English Wikipedia. We find that the decrease of senior users' contribution is caused by the negative network effect of the new users who generated low-quality contribution. Our results indicate that it is the quality of user contribution that moderates the direction of direct network effect in online community. Our findings suggest there should be a balance between the quantity and the quality of user contribution.

In Study 3, we quantify the economic value of celebrities' participation and popularity in social media. Specifically, we study whether NBA players' participation and popularity in Twitter help them earn higher salaries. Our results suggest that both NBA players' participation and popularity in Twitter helps them to earn higher salaries. Furthermore, we investigate the impact of social media on salary inequality among NBA players. Our analysis suggests an interesting U-shape effect: above-average and below-average players are benefited more than average players. Our results imply that the salary inequality among NBA players is decreased due to the emergence of social media. This study not only confirms the business value of social media but also reveal the societal impact of social media.

The notable findings from this thesis provide significant contributions to the literature on online community and social media in the field of Information Systems. Our research also offers helpful and practical suggestions to industry practitioners.

List of Tables

List of Figures

1. General Introduction

Online community has become an integral part in our daily life. Today, we are accustomed to engage in a variety of online communities especially social media sites, to meet our different needs. For example, we share funny stories/moments on Facebook, search for promising jobs on LinkedIn, watch intriguing videos on YouTube, and so on. According to web analytics site Statisticbrain.com, as of 2014, there are more than 1.3 billion monthly active users on Facebook and over 645 million active registered users on Twitter. LinkedIn has already acquired over 259 million users from more than 200 countries and territories (Nishar 2013).

With expectation to directly interact with a vast number of online community members and further convert these community members to customers, firms are actively engaging in online community, considering online community especially social media as a valuable marketing channel. It is reported that 69% of small business owners are engaged in some kinds of social media platform (e.g., Twitter, Facebook, and LinkedIn) and about 78% of them plan to allocate more budgets on social media marketing (Protalinski 2011). The expenditure on social media marketing in US is expected to grow 34% yearly and reach 3.1 billion USD in 2014 (Forrester Research 2009).

1.1. Ordinary Users

Despite the immeasurable value embedded in online community, developing and maintaining an active online community is never an easy task. In their business model, online communities provide a virtual platform for users to start conversations across a broad range of topics as well as marketing tools and services for firms to target their potential customers. Firms investing in online community care how many people their advertisings can reach and how effective their advertisings are. Therefore, online community is a multi-sided platform in which its business value origins from the large user base. However, ordinary users in online community generally do not receive monetary rewards in return. It is challenging to incentivize users to participate voluntarily and have active engagement in online community.

In understanding the motivations for ordinary users' voluntary participation and engagement in online community, there exists extensive research conducted by scholars in the field of Information Systems, (Bateman et al. 2011; Faraj and Johnson 2011; Kankanhalli et al. 2005; Ma and Agarwal 2007; Ren et al. 2007; Wasko and Faraj 2000). However, most existing literature is survey-based and little research has been done on investigating the impact of badge system, which is a key element in the gamification framework. In this thesis, Study 1 helps to fill in this research gap and investigate the impact of a hierarchical badge system on ordinary users' voluntary contribution on four kinds of activities (answering, commenting, revision, and asking) in a Q&A website. Our results confirm that almost all badges motivate users to contribute more in related activities. There is a spillover effect to other activities. Furthermore, our results reveal that gold badges are more powerful than silver badges and silver badges are more influential than bronze badges. Hence, there is a ranked ordering in efficiency of badges corresponding to difficulty levels. Overall, our results present strong empirical evidence that confirms the value of the hierarchical badges system and the effectiveness of gamification in stimulating voluntary participation and continued engagement. This study enriches the literature on online community and adds to the growing literature on gamification in Information Systems.

Although we use the objective data from an online Q&A website in this study, our findings is not limited to traditional online communities but also applied to online social network platforms. Social media platform like Foursquare[1] and Weibo[2] also launch their badge systems. The impact of badge system in social media platform is estimated to be larger than in traditional online communities due to the fact that user interaction is more frequent and user identity is more prominent in social media platform.

We also notice that purely emphasizing on the increase of user participation may not always bring a positive outcome. This can be demonstrated by the case of English Wikipedia. In July 2013, English Wikipedia simplified the editing interface to increase the participation of new users but it turned out to drive senior users away (DailyMail 2013). This real-world case serves as an example to challenge the theoretical prediction of the positive network effect which states the incentive for a user to participate increases with the increase of the number of other users (Shapiro and Varian 2013; Zhang and Zhu 2011). In this thesis, Study 2 investigates the network effect in online community. Empirically, we examine how the increase of new users' participation drives the decrease of senior users' participation at English Wikipedia. We find that it is the quality of user contribution that moderates the

[1] http://www.4squarebadges.com/foursquare-badge-list/

[2] http://www.techinasia.com/weibo-badges/

relationship between the size of user base and the level of user participation. Our results suggest that there should be a balance between the quantity and the quality of user contribution. This study extends our understanding of the phenomenon of network externality in online community, therefore contributing to both the literature of online community in Information Systems and the literature of network externality in Economics.

Our conclusions drawn from the case of Wikipedia can be naturally generalized to other online communities, especially those communities such as Q&A websites providing service based on high-quality user contribution.

1.2. Influential Users

Recently, social media make users more connected and facilitate more frequent interaction than traditional online community. The emergence of social media has equipped influential users with the power to express themselves and significantly influence others in social network, which they never before experience in traditional online communities. For example, Lance Armstrong's retweeting the marketing topic of "#ineedanewphone" on Twitter, generated 65 million impressions within 24 hours and leaded to a double-digit sales increase in wireless platform of Radio Shack (Slutsky 2011). The impact of influential users in social media has drawn the attentions of researchers. Research on the impact of influential users is mainly conducted in the field of new product marketing and scholars find that influential users or opinion leaders can significantly influence the adoption of new products such as prescription (Iyengar et al. 2011; Nair et al. 2010). Another stream of research related to influential users is on how to identifying the influential users in social media (Aral and Walker 2012; Li and Du 2011; Nair et al. 2010).

However, there is little research investigating the economic impact of influential users' participation and popularity in social media. In this thesis, Study 3 aims to quantify the economic value of celebrities' participation and popularity in social media. Specifically, we study whether NBA players' participation and popularity in Twitter help them earn higher salaries. Our results suggest that both NBA players' participation and popularity in Twitter helps them to earn higher salaries. Furthermore, we investigate the impact of social media on salary inequality among NBA players. Our analysis suggests an interesting U-shape effect: above-average and below-average players are benefited more than average players. Our results imply that the salary inequality among NBA players is decreased due to the emergence of social media. This study not only confirms the business value of social media but also reveal the societal impact of social media, thus contributing to the literature of social media in Information Systems and the literature of wage inequality of Economics.

Our analyses based on the sample of NBA players can be generalized to other professional sports leagues. The leagues and player association which care players' welfare can encourage players to actively participate and engage in social media. Those bottom players in the league are strongly recommended to engage in social media since their salaries are below the poverty line of the league and they benefit most from social media than other players.

2. Study 1: The Impact of Badge System on Voluntary Contributions

2.1. Introduction

Since the inception of the Internet, the number as well as the size of online communities has proliferated over time. Among the myriads of online communities facilitating private provisioning of public goods, questions and answers (Q&A) sites have a special place because of their contribution to knowledge creation. A Q&A site enables users to post questions and to help each other by answering questions posted by others. Launched in 2006, Yahoo! Answers is an early-mover in this domain. In April 2012, Yahoo! Answers had about 50 million unique visitors per month.[3] Following the success of Yahoo! Answers, a number of Q&A sites have flourished in recent years serving niche areas with unique design elements. One such popular Q&A site is Stack Overflow, created in 2008, enables computer programmers from all around the world to help each other with technical questions. Given the huge success of Stack Overflow, Stack Overflow launched the Stack Exchange in 2010, a network of Q&A sites on topics ranging from computer science to cooking. As of November 2012, the Stack Exchange network has 99 Q&A sites over 64 million monthly unique visitors.[4] On March 24, 2011, Facebook also launched Facebook Questions to catch the wave of Q&A communities.

Most users of online communities participate without getting monetary rewards in return. Therefore, it is a challenging proposition to incentivize users to

[3] http://siteanalytics.compete.com/answers.yahoo.com/
[4] http://stackexchange.com/about

contribute constantly and have active engagement in community sites. Similar to other communities on the Internet, the success of Q&A sites heavily depends on users' voluntary contributions because these sites can survive only if there are meaningful contributions from community members in terms of both quality and quantity of questions and answers. To minimize the adverse effect of free-riding behavior among community members, these sites should also maintain an acceptable ratio of answers to questions. This requires a deeper understanding of the motives of users and subsequently designing proper incentive mechanisms to facilitate engagement in different community activities.

Gamification, which is the use of game mechanisms and elements in non-game contexts (Deterding et al. 2011), can be a solution to alleviate the incentive problem and to motivate large-scale participation at Q&A sites. Game design principles, such as points, badges, levels, status, can be embedded into the incentive structure to drive engagement in various community activities. By effectively employing game mechanisms, voluntary contributions to Q&A sites can be induced.

Despite the adoption of various game design elements in online communities, there is limited scholarly research that empirically quantifies the significance of gamification for community participation and engagement. This paper attempts to fill this gap by examining the value of a hierarchical badges system on user activities in an online Q&A community, namely Stack Overflow. The hierarchical badges system deployed by Stack Overflow embeds several game elements, including badges, levels (i.e., categories of badges), points (i.e., reputation scores), and leader boards, to motivate users take part in community activities. Users contributing on the Q&A site are awarded badges and earn points based on the type as well as the level of their site activities. Earning badges and points sometimes requires acknowledgement of user contributions by other community members (i.e., peers) through up-votes, acceptance, and views. Hence, badges reflect both the quantity and the quality of voluntary contributions. Furthermore, each badge belongs to a category based on the difficulty level in earning it, namely bronze, silver or gold, and there is a hierarchy among badge categories. A gold badge requires more participation than a silver badge, which in turn requires more participation than a bronze badge. We explicitly study the influence of earned badges in voluntary contributions. After controlling for other factors, such as reputation and tenure, we assess the impact of individual badges and badge categories in gamifying contributions and therefore inducing community activities from users.

In this study we consider participation in four major community activities that earned badges seek to incentivize. These activities are (i) asking questions, (ii) answering questions, (iii) making revisions, and (iv) making comments. Our data set consists of a detailed history of 58,479 registered users on the Q&A site from 2009 to 2012, including a complete record of four activities performed and specific badges earned by each user. We perform our data analyses in two levels: micro level and macro level. In the micro-level analysis, we examine the impact of a specific badge on user activities by estimating a difference-in-differences model. We aggregate user activities a week before and after a user receives a particular badge. We use the propensity score matching technique to form a control group of users. Our results reveal that almost all badges motivate users to contribute more in related activities of badges. Furthermore, we find that even activities that are not specified in the requirements of earning a badge are affected by the status of getting the badge. Surprisingly, we find that even a negative badge correlates with more user activities. In the macro-level analysis, we analyze a three-year panel data set and estimate a fixed-effects model to investigate the impact of three categories of badges, namely gold, silver, and bronze, on user activities. We show that the influences of three categories of badges on answering, commenting and revisions activities are qualitatively the same. In addition, there is a ranked-efficiency relationship among the categories of badges. Gold badges provide a powerful stimulus than silver badges and silver badges are more influential than bronze badges, in terms of ability to induce user participation. Furthermore, our analysis shows that the hierarchical badges system induces more voluntarily contribution from users once users obtain a badge from a higher category. Thus, the hierarchical badges system not only facilitates differential influence among badge categories, but also gives rise to positive externalities across them. Hence, we conclude that the hierarchical badges system is highly effective in inducing continuous contributions to the community. Different from other three activities, we demonstrate that badges do not seem to motivate users much to ask questions.

To the best of our knowledge, none of the existing studies investigates the quantitative impact of badges and categories among badges on user activities in an online community using econometrics models. Prior studies either focused on identifying motivations behind user participation (Wasko and Faraj 2000) or studied the impact of network effects in online communities (Gu et al. 2007b; Zhang and Wang 2012; Zhang and Zhu 2011). The main contribution of this paper is two-fold. First, we quantify the influence of a hierarchical badges system on voluntary user contributions. We characterize the significance of rewarding contributions using badges and having a hierarchy among badges in facilitating sustained contributions from users over time. Since online communities endow users with an assortment of badges, our results provide insights into the optimal badges system design. Overall, we add to the literature in online communities by providing evidence that badges enable users to boost the intrinsic motivations for voluntary contributions to Q&A sites, rather than being a source of extrinsic motivation only. In theory, users participate in a Q&A site because of intangible and psychological benefits and badges could bring tangible benefits into the equation. However, if a user only cared about the benefit associated with one badge, he/she would stop contributing after obtaining that badge. On the contrary, our findings suggest that users participate significantly more after earning badges. This implies that awarded badges boost the intrinsic motivations of users, creating a positive reinforcement loop between badges and contributions, thereby facilitating continuous user engagement over time. Second, although gamification has potential applications in every industry, there is limited empirical evidence that quantifies the value of gamification. Our results provide much needed evidence that game design techniques work in promoting incentives to achieve the real world objectives. Specifically, we shed light on the value of the use of gamification techniques in inducing voluntary participation in Q&A community activities. To the best of our knowledge, we present the first large-scale empirical evidence that shows the effectiveness of gamification in increasing engagement in online communities.

2.2. The Relevant Literature

From the early days of the Internet, we have witnessed an explosive growth of information produced in various online communities. This increase is largely attributed to users spending time and effort to share their knowledge voluntarily with other community members without getting any monetary reward in return. Researchers from different disciplines made inquiries to study the motivations behind such large-scale "unselfish" behavior using a wide range of research methodologies. Since we study the role of game elements to motivate unpaid knowledge contributions, our research is relevant to three streams of research in prior studies: voluntary participation in online communities, private provision of public goods, and gamification.

2.2.1. Voluntary Participation in Online Communities

In theory, a person is motivated to act only when the expected benefits associated with posting questions or answers exceed their anticipated costs, such as time and effort. Motivation theories in psychology attempting to explain the drivers of individuals' behavior have distinguished two main types of motivations: intrinsic and extrinsic (Ryan and Deci 2000). Intrinsic motivation exists if an individual is driven to perform an activity due to satisfaction from the activity itself. Intrinsic motivation is based on the pleasure generated by the activity rather than relying on an external reward. In contrast, extrinsic motivation occurs when the activity is performed in order to attain an external or separable outcome. Because there exist few extrinsic motivations for the users of online communities, IS researchers primarily focused on studying the sources of intrinsic motivations to explain voluntary contribution on the Internet. By using behavioral research methods such as surveys and case studies, IS researchers have identified a number of intrinsic motivations that could drive user participation in online communities. This section discusses the motivations identified in prior studies.

In one of the earliest studies Wasko and Faraj (2000) investigated why users share knowledge in three programming communities using a survey and content analysis. They identified three categories of benefits: tangible return (e.g., answers to questions or monetary rewards), intangible return (e.g., entertaining and learning), and community interest (e.g., interacting with users, altruism, reciprocity, and advancing the community). They observed that the community interest is the most important one among three categories of benefits in inducing participation.

With similar methodologies, many researchers have studied related research questions in different contexts. The main contribution of each paper lies in a unique focus on a specific type of intrinsic motivation and/or in studying an unexplored form of online community. For example, Daugherty et al. (2005) surveyed online panel[5] participants and found that learning and gaining information is the strongest motivation. Wasko and Faraj (2005) surveyed participants of a legal professional online community and showed that reputation and social capital are two key motivations in this context. The enjoyment of reputation and social image has also

[5] An online panel is a consortium of registered persons who have agreed to take part in online research on a regular basis.

been documented in a number of other studies (Constant et al. 1994; Constant et al. 1996; Daugherty et al. 2005; Marett and Joshi 2009; Ren and Kraut 2011; Tiwana and Bush 2005). Participants in online communities may also take part in community activities due to self-efficacy, which means users capitalize on their ability to find solutions in order to accomplish challenging goals (Compeau and Higgins 1995). For example, Sun et al. (2011) found that self-efficacy is the most important motivation on TaskCn, a crowd-sourcing site, while Jin et al. (2012) reached the same conclusion on Yahoo! Answers China. A motivation similar to self-efficacy is that individuals may fulfill entertainment needs while participating in online communities, either through communications with other users or interacting with them to solve problems (Kankanhalli et al. 2005; Marett and Joshi 2009; Ren and Kraut 2011; Sutanto et al. 2011).

Because of the popularity of Facebook and other social networking platforms, recent studies about online communities focused more on social aspects. In general, reciprocity, social capital, and community interests were highlighted as three major categories of intrinsic motivations. Reciprocity, a practice of exchanging things with others for mutual benefit, refers to making valuable contributions to the community after reading helpful posts on the same community (Chiu et al. 2006; Faraj and Johnson 2011; Hall and Graham 2004). In the literature, social capital is defined as the resources embedded within networks of human relationships (Nahapiet and Ghoshal 1998). Using the social capital theory in studying online communities, it is hypothesized that users engage in knowledge contribution to build mutual trust and to establish long-term relationship. Raban

(2009) showed the importance of social capital on Google Answers. Other researchers provide similar evidence in other online communities (Chiu et al. 2006; Ren and Kraut 2011; Ren et al. 2007; Ren et al. 2012; Tiwana and Bush 2005). Lastly, members of an online community may feel identified with the group when this group of people shares common interests or characteristics. If members experience a strong sense of community, they may become altruistic and committed to actively contribute knowledge (Bateman et al. 2011; Ma and Agarwal 2007). Oh (2012) showed that altruism is the most important motivation among 10 motivations on Yahoo! Answers for health-care related questions.

As for the influence of extrinsic motivations, particularly monetary rewards, studies have revealed counterintuitive results. Contrary to economic theories, researchers have identified situations in which providing monetary incentives does not necessarily increase contributions to online communities. Conducting an survey based experiment among the members of a German Q&A site, Garnefeld et al. (2012) found that monetary incentives can increase community members' participation only in the short term. This finding is consistent with the real-world business cases. Google Answers, which once offered monetary rewards to incentivize users to contribute actively, could not succeed whereas Yahoo! Answers thrived to be the largest Q&A site without offering any monetary rewards to its contributors.

Our study is different from the studies mentioned above in the following ways. First, in conventional sense, a badge could serve as an extrinsic stimulus/motivation for participation. Users contribute because they want to earn badges. However, if a user only cared about the benefit associated with one badge, he/she would stop contributing after obtaining that badge. In this paper, we adopt the perspective of the emerging theory of gamification and argue that a badge, when it is earned, enhances the intrinsic motivations of users associated with participation activities, instead of being just a source of extrinsic motivation. Users participate significantly more after earning badges because awarded badges boost the intrinsic motivations of users, creating a positive reinforcement loop between badges and contributions, thereby facilitating continuous user engagement over time (Please see section 2.2.3 for more details.) For example, after getting a badge, the user could feel an increased sense of community identity or enjoy more fun in interacting with or helping to other users in the same community. Consistent with our arguments, our findings show that earned badges indeed enhance participants' intrinsic motivations that result in more contributions to community activities in the subsequent period. Second, almost all of the existing studies use surveys to directly solicit participation motivations from the users of online communities. In contrast, this study utilizes field data and econometric methods to quantify how a hierarchical badges system, a unique form of non-monetary incentive mechanism, may affect user activity levels on a Q&A site. In other words, our focus is on badges, not on the aforementioned motivations. Third, our study sheds light on how different categories of badges in a hierarchy vary in terms of influence on participation levels to different site activities. Fourth, our econometric analyses using a large sample from a real Q&A site complement prior studies because survey-based studies may suffer from the sampling bias due to the low response rate and also in survey studies, all measured variables are solicited from the respondents, making it challenging to rule out alternative unobservable covariates for the observed significant relationships (Wang and Noe 2010).

Very little research has been done in an online community setting to investigate when, how much, and which types of reinforcement, such as badges, should be used to increase membership participation in online communities (Tedjamulia et al. 2005). Gazan (2011) summarized the recent research in design science about Q&A sites. Addressing the lack of research in this area from a social psychology perspective, Antin and Churchill (2011) proposed five theoretical lenses to study badges: goal setting, instruction, reputation, status or affirmation, and group identification. To the best of our knowledge, there is no rigorous econometric analysis that explicitly studied the effect of badges in online community participation. There only exist several remotely related studies that examined other design features, such as reputation systems (Chen et al. 2010), feedback systems (Moon and Sproull 2008), and knowledge validation processes (Durcikova and Gray 2009), on the quantity and quality of knowledge contribution. Although badges have become a widely popular mechanism to induce online user engagement, thanks to the success of mobile social networking site Foursquare, there still exists limited scholarly research that investigates the effect of badges on user participation in online communities. Our study is an attempt to fill the gap in this direction.

2.2.2. Private Provision of Public Goods

The earliest research related to voluntary contributions appeared in the literature on the private provision of public goods. Public good is defined as the good that individuals cannot be effectively excluded from its use and use by one individual does not reduce availability to others (Samuelson 1954). Textbook examples of public goods include parks and light houses. Several significant innovations over the Internet created a new set of public goods, produced entirely from free user contributions, such as open source software, Wikipedia, online communities, and YouTube. Since the benefits of public goods are enjoyed by everyone but the production cost is accrued to an individual, the free-rider problem and under-supply of public goods is a typical equilibrium outcome in theoretical models (Andreoni 1988).

One main proposition for empirical testing is that the free-rider problem is aggravated when the group size increases: the average level of individual contribution declines as group size increases. Extant studies focused on finding remedies to alleviate the free-rider problem, such as using government policies or incentive mechanisms to reimburse (or penalize) users to produce more (or less) public goods with positive (or negative) externalities.[6] However, most theoretical models failed to explain the phenomenon of extensive donations in the charitable sector of the economy (Andreoni 1988). A remedy for the inconsistency between theoretical models and empirical evidence is the impure altruism models (e.g.

[6] Interested readers could refer to Chen, Y. 2008. "Incentive-Compatible Mechanisms for Pure Public Goods: A Survey of Experimental Research," *Handbook of Experimental Economics Results* (1), pp. 625-643. and Chaudhuri, A. 2011. "Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature," *Experimental Economics* (14:1), pp. 47-83. for a recent comprehensive literature review.

Andreoni (1989), Andreoni (1990), Cornes and Sandler (1994), and Steinberg (1987)). In these models, individual contributors obtain utility from not only the total provision of public good but also their own private benefits (or warm glow), such as enjoyment of helping others and moral satisfaction. In this case, these private benefits are positively related to group size because individual's enjoyment of helping others is amplified by the number of recipients. As group size increases, the motivation of pure altruism fades away while private benefits increase and lead to a rise of "social effects", sustaining individual contributions in a large group (Zhang and Zhu 2011).

In Stack Overflow and other Q&A sites, the knowledge base is accessible to all Internet users and thus it is a typical example of public goods. However, to the best of our knowledge, empirical studies in economics have not examined private provisioning in the context of Q&A sites.[7] Furthermore, most of the findings in the private provision of public goods literature may not be applicable to our context due to the following reasons. First, majority of empirical studies examined voluntary contributions using experiments on a small group of users, which is very different from Q&A sites with many users. Second, some theoretical models focused on how to redistribute the costs and profits, which is also not relevant to Q&A sites, where there is no monetary reward and it is difficult to reimburse users for their time and effort. Third, a number of studies investigated how group size correlates with public good provisions, which is also different from

[7] The only exception is Chen, Y., Ho, T.-H., and Kim, Y.-M. 2010. "Knowledge Market Design: A Field Experiment at Google Answers," *Journal of Public Economic Theory* (12:4), pp. 641-664. who conducted a field experiment at Google Answers by manipulating the monetary rewards to users.

the main issue studied in our study. We focus on the effects of a hierarchical badges system. Specifically, we analyze how badges and categories among them may enhance individual contributors' warm-glow effect to induce contributions.

2.2.3. Gamification

Online games are hugely popular among many users because they tap into the drivers of user engagement, such as challenge, achievement, competition, entertainment, and interactivity.[8] Organizations now realize that some elements of game design can be applied to business and social processes to promote desired behavior, leading to the emergence of the concept of gamification. Gamification refers to the use of game elements and techniques in non-game contexts to drive a game-like player behavior (Wu 2011). It can be viewed as a new paradigm for enhancing brand awareness and loyalty, innovation, and user engagement (Werbach 2013).[9] The definition of gamification has three components. The first component of game elements and techniques comprises game design principles, game dynamics, player profiles, and other aspects of games. Users are incentivized for participation and performance using points, badges, leaderboards, levels & status (Burke 2011). The second component is the non-game context which can

---

[8] The motivations behind playing games and game design mechanisms have been widely studied in the literature. The interested readers can refer to Olson, C.K. 2010. "Children's Motivation for Video Game Play in the Context of Normal Development," *Review of General Psychology* (14:2), pp. 180-187., Ryan, R.M., Rigby, C.S., and Przybylski, A. . 2006. "The Motivational Pull of Video Games: A Self-Determination Theory Approach," *Motivation and Emotion* (30:4), pp. 344-360., Yee, N. 2006. "Motivations for Play in Online Games," *Cyber Psychology & Behavior* (9:6), pp. 772-775., and Fogg, B. 2009. "A Behavior Model for Persuasive Design," *Proceedings of the 4th international Conference on Persuasive Technology*: ACM, p. 40..

[9] This trend is predicted to intensify over time. It is estimated that more than 70% of Global 2000 organizations will employ at least one gamified application by 2014 Gartner. 2011. "Gartner Predicts over 70 Percent of Global 2000 Organisations Will Have at Least One Gamified Application by 2014," Barcelona, Spain.. Furthermore, over 50% of organizations that manage innovation processes are expected to gamify these processes by 2015 ibid..

include work, innovation, marketing, education, health and fitness, environment and community participation. The third component refers to game-like player behavior, such as competition, interaction, collaboration, learning, addiction, and engagement (Wu 2011). Gamifying a system or application requires defining its objectives, deriving desired user behaviors, devising extrinsic rewards that should appeal to users, and implementing game mechanics to connect desired behaviors to extrinsic rewards (Zichermann and Cunningham 2011). It is important to note that a gamified application is not a game. Building on the findings about human motivations related to game playing, gamification creates just an environment that induces desired behavior to facilitate continuous engagement in a business context (Werbach 2013). Gamification should not be assessed through the lenses of conventional incentive theory. While the presence of an explicit reward induces a desired behavior in many incentive structures, the induced behavior is typically short-lived or does not occur again after the receipt of the reward. Gamification, on the other hand, aims to achieve the reoccurrence of the desired behavior. Extrinsic rewards presented to users upon the occurrence of the desired behavior, such as points and badges, incentivize continuous engagement with the help of game mechanisms, causing the repetitive behavior-reward combination to occur (Kankanhalli et al. 2012).

Several anecdotal cases suggest the benefits achievable through gamification. For example, a call center company in the US, LiveOps, used game elements to motivate call center agents. Since then, sales have improved by 8-12% and call time has reduced by 15% (Silverman 2011). Yahoo’s gamification of its web-based ethics training module has achieved a 99 % completion rate (Ashraf 2011). DevHub, which enables users to create their own blogs and web sites using site tools, reported that average revenue per user increased four-fold after gamification, with around 80% of users completing their sites (Takahashi 2010). Foursquare, a mobile location-based application, has successfully driven its users to "check-in" at their nearby locations using gamification techniques. Nike+ system, which tracks pace, distance, time, and calories burned while users run or walk using a little sensor in shoes and let users set personal goals and compete with friends, achieved tremendous success by capitalizing on gamification and the underlying game mechanisms.

While case studies and practitioner articles provide useful insights regarding gamification benefits, there is no rigorous academic research to test these results for greater credibility. The hierarchical badges system used by Stack Overflow embeds several game design elements to elicit voluntary participation and continuous engagement, thereby creating a perfect environment to test the effectiveness of gamification. In addition, Stack Overflow is one of the most cited gamified applications in practice (Werbach 2013). First, it offers various badges designed to induce different community activities. Contributors earn badges that reflect their type as well as the level of participation and involvement in the community. This is like earning points and trophies as users accomplish tasks in games. Second, there are categories of badges (bronze, silver, and gold) and a hierarchy among the categories. The ordering among the badge categories is similar to the levels in games, corresponding to difficulty in achievement and advancement.

Third, users engage with each other through social communication channels, such as voting contributions up or down, accepting an answer as the correct answer, and visiting contribution pages, making revisions and leaving comments on others' contributions. This engagement brings interactivity that promotes relatedness to others. Fourth, users are ranked in leaderboards based on their contributions and earned badges, enabling users to compare their own performance with others and stimulating competition, just like games.

Although gamification rewards desired behavior using extrinsic motivators, such as points and badges, continuous engagement in the long run depends heavily on how gamified applications tap into intrinsic motivations of users (Werbach 2013). For instance, gamified marketing applications can initially encourage customers to buy more by allowing them to track their points and accordingly offer discounts. However, long term engagement with the application is likely to hinges on how users internalize these motivations by developing loyalty and identification with the brand. Self-determination theory argues that extrinsic motivations must be integrated into intrinsic motivations for engagement (Ryan and Deci 2000). Similarly, badges offered by Stack Overflow can elicit contributions from users because badges bring tangible benefits in the short run. The effectiveness of the hierarchical badges system in the long run ultimately depends on whether earned badges can enhance the intrinsic motivations of users. Our empirical results reveal that the hierarchical badges system is effective because it helps users internalize the extrinsic motivations and builds a positive reinforcement loop between badges and contributions, thereby facilitating continuous engagement from users over time. To the best of our knowledge, this study presents the first empirical evidence that confirms (i) the effectiveness of gamification in eliciting and sustaining desired behavior and (ii) the value of hierarchical badges system in online community participation and engagement.

2.3. Data Description

2.3.1. Sample Construction

Our data come from Stack Overflow, which is the most popular Q&A site for programmers. Stack Overflow was launched in August 2008. From its establishment till today, Stack Overflow has attracted over 1.1 million users and these users have generated over 3 million questions and 6.1 million answers with an answered rate over 80%. We employed the Stack Exchange APIs to crawl data from August 2008 to January 1st 2012. We excluded the observations in 2008 because user activities were relatively unstable due to the nascence of Stack Overflow. We observe that more than 50% users (29,450 out of the first 50,622) who registered during this period later dropped out. We also excluded users who registered after January 1st 2011 to ensure that all users in our sample have an activity period of more than a year. After this step, there were 354,029 users left in our sample. Our final sample period included three years of data from January 1st 2009 to January 1st 2012.

Our initial assessment of the panel dataset revealed that a large fraction of the registered users were rather inactive during the entire three years: about 40% (128,933/354,029) of users had a reputation score of 1, meaning that they had not engaged in any activity after registration (since 1 reputation score is given directly by the system upon registration).[10] These users with almost no activity after sign up might deteriorate our data analysis. Therefore, we excluded users if a user's cumulative reputation score was never greater than 200 throughout the three year sample period. The threshold of 200 is determined by Stack Overflow and it could be the most objective value for us to define users with basic level of contribution. According to Stack Overflow, users with reputation score less than 200 are not tracked in the reputation league on stackoverflow.com. On average, users who have less than 200 reputation score posted less than one question or one answer per month. In fact, a user can earn 200 reputations in one day by engaging in community activities. Therefore, the 200 rule is a reasonable cutoff to characterize our sample for analysis. Note that this cutoff does not imply that we only consider users with continuous contribution over time. A user might earn necessary reputation by having a burst of activity in one week and being inactive in the rest of the weeks. Alternatively, a user might build the necessary reputation by contributing little by little over weeks during the three year period. In either case, we keep the user in our data set for analysis. With this cutoff rule, the percentage of users dropped from the sample was 83% (295,550/354,029). At the end, our final data set consisted of 58,479 users.

The scraped data include cross-sectional user profiles and user-level panel datasets. A user profile contains a user id, display name, self-reported age, and self-reported location in an unstructured format. A user-level panel dataset consists of a

[10] Reputation points are earned by posting good questions and helpful answers on Stack Overflow. In addition, users can earn reputation points for suggesting edits. For the detailed rules to obtain reputation points on Stack Overflow, please refer to "http://stackoverflow.com/faq#reputation".

detailed history of reputation, types of badges received, and types of activities engaged in with a time stamp, which is at seconds-level.[11]

The unit of analysis in our study is at user-week level. We conduct the analysis at weekly instead of daily level to avoid weekend effects. Weekly analysis is a more reasonable choice compared to longer periods, such as bi-weekly or monthly analysis because the impact of a specific badge may not last longer than one week. In addition, if we aggregated data for more than one week, it would be more difficult, if not impossible, to isolate the overlapping impacts of badges because a user is more likely to receive multiple badges in the same period.

2.3.2. Dependent Variables: User Activities

Stack Overflow records four types of user activities and we use the level of each of these activities as a dependent variable in this study. These four activities are (i) *answering* (the number of questions answered), (ii) *commenting* (the number of comments to posted questions or answers), (iii) *revision* (the number of revisions to posted questions or answers), and (iv) *asking* (the number of questions asked).

Among all activities, "answering" is the most vital activity in our research context. In Stack Overflow, getting users to answer questions is relatively more challenging than getting users to ask questions. Users who submit their questions may not seek, and therefore motivated by, an external reward such as a badge, because they are likely to consider answers to their questions as tangible benefit. However, users who answer questions do not earn much obvious benefit in return

[11] Specifically, Stack Overflow uses the UNIX time stamp. For example, 1335966004 could be converted to "Wed, 02 May 2012 13:40:04 GMT".

for their time and effort. Therefore, it is utmost important to offer them non-monetary psychological stimulus, such as badges.

"Commenting" is another essential activity that enhances user interactions to improve the quality of questions and answers. Users can comment on each other's questions and answers. Comments could be as short as "Thank you!" or could include constructive suggestions. The increased level of interaction among users is likely to influence users' perceived community identity and induce users to engage more on the Q&A site.

"Revision" activity is designed to imitate Wikipedia for building a high-quality knowledge base at Stack Overflow. Stack Overflow encourages users to collaborate in revising their questions and answers. After several rounds of revisions by users, the quality of both questions and answers may improve drastically. In a sense, revisions help Stack Overflow become the most valuable online reference for programmers, which subsequently drives more users to the site.

2.3.3. Badges and Independent Variables

Our independent variables are related to badges. In Stack Overflow, a user is awarded with a badge when his/her participation reaches a threshold in one or several types of activities. The earned badge will appear on the member's user page and user card. Therefore, badges could induce not only direct psychological effects but also indirect effects via the social image.[12]

[12] This is consistent with the quote from Joel Spolsky, the cofounder of Stack Overflow: "The number of badges you've earned is displayed on your user card for everyone to see. Most people claim not to care about badges, but as soon as you think someone is seeing what you're doing, you start to care about it more."

The total number of badges in Stack Overflow is 73.[13] Badges can be categorized along several dimensions. Some badges are awarded based on the cumulative number of activities while others are awarded based on the number of activities within a specific time period. For instance, the badge “Commentator” is awarded when a user leaves 10 comments. In contrast, the badge “Fanatic” is earned if a user visits the site each day for 100 consecutive days. In addition, some badges are awarded based the quality of user activities, not just the number. Peers endorse the quality of a contribution with their votes and/or visits. For instance, the badge “Popular Question” is given when a user asks a question with 1000 views by others, confirming the quality of the asked question.

Badges are classified by Stack Overflow into three categories (Gold, Silver, or Bronze) in terms of the hierarchical difficulty in obtaining them. For instance, “Popular Question” is a bronze badge (Asked a question with 1000 views), “Notable Question” (Asked a question with 2500 views) is a silver badge, and “Famous Question” is a gold badge (Asked a question with 10000 views). In this example, a user must have earned the bronze badge before he could be qualified for the silver or gold badge. The obtained badge is not replaced by the badges in higher levels; instead, the user keeps all the badges. Badges and their hierarchy provide us with an opportunity to examine the relative effectiveness of a hierarchical badges system in inducing continuous user participation. Out of 73 badges, 35 are bronze, 25 are silver, and 13 are gold badges. There are 14 sets of hierarchical badges. Only

[13] Stack Overflow continuously introduces new badges into their badges system. This number is as of May 2012.

5 among them are three-level (Gold, Silver, and Bronze) and the others are two-level (either Bronze and Silver or Silver and Gold).

We also notice that there are a few badges with "negative" connotations, hereafter referred to as negative badges. For example, "Tumbleweed" is earned when a user asks a question with no votes, no answers, no comments, and low views for a week". Another example is "Unsung Hero", which means a user has at least 10 answers with a score of zero, and those answers make up at least 25% of all of the user's accepted answers. Since most of the badges at Stack Overflow dress up users with a positive social image, these negative badges give us a unique opportunity to examine whether they encourage or discourage users' activities.

We carry out our empirical analyses in two steps. In the first step, we consider a micro-level (badge-level) analysis. We analyze the value of 27 badges in detail. We are interested in finding how each of these badges affects user engagement. We calculate the number of activities before and after a user has obtained a specific badge. Then we use the propensity score matching to construct a control group of users who did not obtain badges in the same period. By comparing the average change in activities of treatment and control groups, we quantify the impact of the specific (target) badge. In other words, our approach is similar to the difference-in-differences method. The weakness of this approach is that we cannot fully isolate the effects from multiple badges because users may receive several badges in a given week. Another concern is that the users may truly pursue hard-to-obtain rare badges, and yet they first get easier badges, which are not the ones that motivate the users, on the way to the target badge. Therefore, we complement our first analysis with a macro-level analysis.

In the second step, using standard panel data methods, we investigate how much gold, silver, and bronze badges awarded to users affect the level of their subsequent activities. Further, we conduct analysis on subsamples, which are defined based on the stage of engagement of users on Stack Overflow. Our results support our conjecture that badges have both economically and statistically significant impacts on encouraging user participation and continuous engagement.

The definitions of variables used in this study are given in Table 2-1 while the descriptive statistics reported in Table 2-2 and Table 2-3 summarize the average weekly levels of activities and the numbers of observations used in micro-level and macro-level analysis, respectively.

Table 2-1 Definitions of Variables

	Variable Name	Definition
Dependent Variables	Answering	Number of questions answered
	Commenting	Number of comments posted
	Revision	Number of revisions made
	Asking	Number of questions asked
Independent Variables	Tenure	Number of months since a user has joined Stack Overflow
	Reputation	Natural logarithm of reputation score
	Cum_reputation	Natural logarithm of cumulative reputation score
	Gold	Number of gold badges received
	Silver	Number of silver badges received
	Bronze	Number of bronze badges received
	Cum_gold	Cumulative number of gold badges received
	Cum_silver	Cumulative number of silver badges received
	Cum_bronze	Cumulative number of bronze badges received

Table 2-2 Summary Statistics for the Micro-level Data Analysis

Activity	Mean	Std.dev	Min	Max
Answering	6.401	13.496	0	248
Commenting	11.050	24.588	0	666
Revision	5.028	15.274	0	1326
Asking	0.546	1.371	0	72
Observations	270,761 user-badge pairs from 58,479 users[14]			

[14] For different badges, we have different numbers of user-badge pairs. Therefore, we have different numbers of observations for different badges and we report here the sum of the user-badge pairs over all 27 badges studied.

Table 2-3 Summary Statistics for the Macro-Level Data Analysis

Activity	Mean	Std. dev.	Min	Max
Answering	0.679	3.227	0	204
Commenting	1.199	5.920	0	708
Revision	0.440	3.890	0	3018
Asking	0.232	0.864	0	72
Tenure	11.845	7.997	0	36
Reputation	0.803	1.535	0	9.387
Cum_reputation	4.441	2.339	0	12.320
Gold	0.002	0.049	0	4
Silver	0.033	0.197	0	10
Bronze	0.137	0.515	0	29
Cum_gold	0.086	0.409	0	25
Cum_silver	1.294	3.564	0	265
Cum_bronze	8.001	8.576	0	702
Observations	4,070,427 user-week pairs from 46,571users[15]			

[15] Please note that the sample size for the macro-level analysis is different from that for the micro-level analysis. This is due to the following reasons. First, the number of users considered in both analyses is different because of two variables: "cumulative reputation" and "cumulative number of badges". Since we cannot accurately calculate these variables for users who registered before 2009, we drop these users from the macro-level analysis sample to reduce the potential bias. Second, for the same user, the number of weekly observations is different. In the micro-level analysis, we only include observations one week before and one week after getting any one of 27 badges. In the macro-level analysis, we include all weekly observations throughout the three-year sample period. Since a large fraction of observations in the macro-level panel dataset are zero or a small number, the means of dependent variables at the macro-level are much smaller than those at the micro-level. While the micro-level analysis depicts the short-term impact, the macro-level analysis captures the long-term effect of badges. Therefore, these two analyses complement each other to provide a more complete picture about how badges correlate with contributions.

2.4. Micro-Level Analysis

To investigate the effect of individual badges, we use propensity scoring and regression methods to exclude alternative explanations. We report the results for 27 badges (Please refer to Table A1 in the Appendix for the definitions of these badges). These 27 badges out of 73 badges were selected based on the following two criteria. First, we excluded badges designed to motivate novice users to explore site features. For example, "Editor" is awarded for the first edit. These trivial badges are owned by almost all users. Also, the impacts of these badges are less stable since users who earned those badges are relatively new to Stack Overflow. Second, we included all the badges that are awarded based on four main contribution activities because the levels of these activities are used as dependent variables in the analysis. After these steps, we ended up with 27 distinct badges. There are 9 badges for asking questions, 11 badges for answering questions, 2 badges for commenting, and 5 badges for revisions in our sample, including 3 negative badges.

2.4.1. Propensity Score Matching Analysis

Our analysis in this section is conducted as follows. Take a target badge "Nice Answer" as an example. First, we identify each date that a user gets the badge "Nice Answer". Second, for the user who earned the target badge, we aggregate the number of contributions made by the user to each of the 4 activities 7 days before and 7 days after that date, separately. As a result, for each user-badge pair, we obtain 8 numbers representing the activity levels. The weekly number of contributions to activity *j* before and after the badge awarding date is denoted by

$Contributions_{ij0}$ and $Contributions_{ij1}$, respectively. The subscript *i* refers to the target badge. Third, we can compare the average number of activities 7 days before and after the date a user obtained the target badge. In other words, we test whether

$$H_0\text{: } \mathrm{E}[Contributions_{ij1}] > \mathrm{E}[Contributions_{ij0}].$$

To establish the causal relationship that getting a badge motivates users to contribute more, it is not sufficient to show that each user's activity increases after getting a badge. There are several alternative explanations that may give rise to the same increase in contributions. First, it could be that all users participate more over time. Second, there exist unobservable variables that correlate with getting the target badge and the increased activity level. For instance, having more free-time to spent on Stack Overflow or the altruistic personality could be the true cause of getting a badge and contributing more in activities.

In the literature, there are two popular methods to eliminate these alternative explanations. One can either include control variables in a panel regression setting or use the Propensity Score Matching (PSM) method suggested by Rosenbaum and Rubin (1983). The idea of PSM is to use a set of control variables to select some observations that are most similar to the observations in the treatment group. The matched observations are used to form a control group. If the dependent variable only correlates with those control variables, then this method produces results as good as a randomized experiment while excluding the impact of unobservable heterogeneity. In other words, using PSM we try to show that

$$H_0\text{: } \mathrm{E}[Contributions_{ij1} - Contributions_{ij0} | with\ target\ badge\ i]$$
$$> \mathrm{E}[Contributions_{ij1} - Contributions_{ij0} | without\ target\ badge\ i].$$

Therefore, our hypothesis testing is similar to the well-known difference-in-differences approach for establishing causality.

An attractive feature of PSM is that matching estimates facilitate a causal interpretation (Angrist and Pischke 2008). We choose PSM over regression as the baseline analysis also due to the following reasons. First, observations in our treatment group could be quite different from the majority of observations. Some badges, such as most gold badges, are very difficult to obtain. Only few users can earn those badges and their behavior could be very different from the rest of the population. In a panel regression, we compare frenzy users with the mainstream users who typically contribute much less. By PSM, for each observation in the treatment group, we can construct a control group of users with contribution patterns, similar to the user in the treatment group. Second, badges are awarded to users across three years, creating another layer of complexity of applying panel regression. By PSM, we can isolate the short-term impacts of a badge one week before and after the badge awarding date. In a panel regression, it is much more difficult to transform the data to achieve the same goal.

In this paper, our control group is constructed as follows. For each user-badge pair comprising the treatment group in our sample, we find the ten most similar users who did not receive the target badge 7 days before and 7 days after the treatment group user received the target badge. The similarity between two users is calculated by the Euclidean distance in $\mathfrak{R}^4$ with the 4 dimensions being the numbers of 4 activities 7 days before the badge awarding date. Following Brynjolfsson et al. (2011), we use the 10-nearest neighbor matching algorithm with replacement. In other words, for each user-badge pair in the treatment group, we identify 10 users who did not earn the target badge, but they had the most similar levels of activities during the week before the treatment user earned the target badge. That is, users in our control group exhibited a very similar activities pattern to the user in the treatment group, except receiving the target badge. The average numbers of activities in the control group are used to benchmark against the numbers derived from the treatment group.

Next, for each specific badge i and activity j, we use a simple regression model to investigate the impact of a badge on contributions to a given activity. The dependent variable is the first-differenced, weekly number of activities to eliminate any user-specific, time invariant heterogeneity (Wooldridge 2009). Formally, the dependent variable is

$$\Delta Contributions_{ij} = Contributions_{ij1} - Contributions_{ij0} .$$

The estimation model is given by

$$\Delta Contributions_{ij} = \beta_{0ij} + \beta_{1ij} * Badge_i + u_{ij}, \tag{1}$$

where $Badge_i$ is a dummy variable and equals to one when a user receives the target badge i at t=1, which means the user is in the treatment group. In this estimation, β_{1ij} captures the average treatment effect. As long as the variable $Badge_i$ is uncorrelated with the error term u_{ij}, the OLS estimator of β_{1ij} is an unbiased and consistent estimator of the badge effect on the activity level.

$$\beta_{1ij} = \mathrm{E}[\Delta Contributions_{ij}|Badge_i = 1] - \mathrm{E}[\Delta Contributions_{ij}|Badge_i = 0].$$

In other words, our estimation is as robust as a randomized experiment when $\mathrm{cov}(\mathrm{Badge_i}, u_{ij})=0$.

2.4.2. Micro-level Results and Discussion

Before discussing the results of PSM estimation, we first examine the quality of matching since this is the foundation for us to establish causality. In Table 2-4, we use a set of Question badges (Popular/Notable/Famous Question) to demonstrate the quality of matching.

Table 2-4 Quality of Matching

		Activity Type			
Badge Name		Answering	Commenting	Revision	Asking
Popular Question	Treatment Group	1.780	4.431	1.946	1.053
	Control Group	0.081	0.144	0.225	0.028
	Matched Control Group	1.984	4.473	1.830	0.882
Notable Question	Treatment Group	1.694	4.046	1.934	0.878
	Control Group	0.081	0.144	0.225	0.028
	Matched Control Group	1.910	4.209	1.830	0.823
Famous Question	Treatment Group	1.644	3.803	2.608	0.740
	Control Group	0.079	0.147	0.222	0.027
	Matched Control Group	1.846	4.042	2.077	0.758

As shown in Table 2-4, the imbalance of covariates between treatment group and control group has been dramatically improved after matching. Take the case of “Popular Question” for example. Before matching, there are large differences across all contribution activities between treatment group and control group. The absolute values of the difference of means for answering, commenting, revision, and asking activities are 1.669, 4.287, 1.721, and 1.025, respectively.

After matching, these values have been shrunk to 0.204, 0.042, 0.116, and 0.171, respectively. We also noted that the means of four contribution activities of control group for "Popular Question" and "Notable Question" are identical, with the accuracy to three decimal places. This is because of the fact that these two badges are easy to earn and bestowed almost every day. Therefore, the control group (those who did not receive the target badge at the same day when the subject in the treatment group received the target badge) for these two badges are similar with a large overlap.

In estimation with the matched sample, we suspect potential heteroskedasticity in our data set. Therefore, we employed the Breusch-Pagan test to confirm the existence of heteroskedasticity and corrected our OLS estimator with robust estimate of variance. Estimated coefficients from robust OLS estimator are reported in Table 2-5. All estimated coefficients are positive. Only 5 out of 108 coefficients are not significant and only 10 out of 108 coefficients are not significant at 1% level. Our results clearly suggest that getting a badge is correlated with more participation in all four activities in the week right after the badge is awarded. Some badges seem to be very influential: "Archaeologist" is associated with 44.134 more revisions in the following week. In sharp contrast, on average, users in the sample of our micro-level analysis only contribute 5.028 revisions per week. It is interesting to note that even the activities that are not considered in the definition of a badge are affected. All estimated coefficients are positive for non-related activities. For instance, "Notable Question" is bestowed when a user asked a question with 2,500 views. Obviously, this badge does not have any direct relationship with the other three activities. However, a user engages more in all 4 activities after getting this badge. The findings in Table 2-5 imply that earning a badge changes the user contributions drastically across all types of community activities.

Our results suggest that even a negative badge could induce more participation. We can explain this surprising result from two different angles depending on whether users view these badges as a carrot or a stick. First, negative badges may prevent users from making contributions that are not appreciated, even considered as scam, by the community at large, thereby discouraging users from posting valueless questions and answers in subject areas that are of interest to many users. Hence, a negative badge may work as a stick to deter participation that dilutes the quality of content in popular areas on Stack Overflow. However, earned badges cannot be revoked on Stack Overflow. Hence, users of negative badges can only make valuable contributions in the subsequent periods to offset their negative social image from negative badges. This argument is also consistent with the findings in prior studies about shame aversion. In a lab experiment, Savikhina and Sheremetab (2010) showed that aversion from shame (avoid being recognized as the least contributor) is a more powerful motivator than being recognized as the largest contributor to public goods. Masclet et al. (2003) documented that nonmonetary sanctions lead to more contributions to public goods. Andreoni and Petrie (2004) and Soetevent (2005) provided empirical evidence that people donate more when being observed by others. Second, negative badges could be interpreted as fun and challenging badges to collect in the eyes of some users. To obtain those badges, users are encouraged to post questions and answers in less popular and/or low visibility subject areas. Hence, a negative badge can be considered as a carrot to reward participation in order to improve content in niche areas. That is, these badges can act as a positive stimulus for some users in the first place. Therefore, it should not be surprising that positive stimulus leads to a positive behavior. To sum up, irrespective of which angle users view the negative badges, they facilitate more contribution.

Table 2-5 PSM Estimation Results

	Activity Type			
Badge Name	Answering	Commenting	Revision	Asking
Popular Question	0.573*** (0.015)	1.530*** (0.033)	0.689*** (0.026)	0.377*** (0.008)
Notable Question	0.553*** (0.024)	1.475*** (0.051)	0.677*** (0.047)	0.350*** (0.012)
Famous Question	0.632*** (0.062)	1.439*** (0.117)	0.421 (0.267)	0.338*** (0.025)
Nice Question	1.294*** (0.045)	3.824*** (0.105)	1.569*** (0.098)	0.523*** (0.016)
Good Question	1.421*** (0.111)	3.644*** (0.245)	1.229*** (0.229)	0.350*** (0.033)
Great Question	1.939*** (0.448)	4.146*** (0.809)	2.017*** (0.420)	0.472*** (0.072)
Favorite Question	1.607*** (0.197)	3.563*** (0.441)	1.602*** (0.282)	0.323*** (0.053)
Stellar Question	2.246*** (0.546)	4.258*** (1.326)	1.804** (0.597)	0.227** (0.113)
Nice Answer	4.879*** (0.047)	7.729*** (0.085)	3.476*** (0.067)	0.177*** (0.004)
Good Answer	4.015*** (0.108)	6.768*** (0.202)	3.223*** (0.197)	0.174*** (0.009)
Great Answer	3.734*** (0.333)	5.694*** (0.545)	1.987*** (0.345)	0.195*** (0.030)
Enlightened	5.959*** (0.102)	9.666*** (0.193)	4.126*** (0.158)	0.146*** (0.007)
Guru	4.148*** (0.238)	6.992*** (0.470)	2.927*** (0.455)	0.168*** (0.016)
Necromancer	1.740*** (0.085)	3.133*** (0.153)	1.604*** (0.185)	0.235*** (0.015)
Populist	3.134*** (0.420)	5.223*** (0.864)	1.623*** (0.505)	0.224*** (0.036)
Reversal	4.995*** (1.777)	14.220*** (5.160)	7.125*** (2.259)	0.180 (0.157)
Revival	1.617*** (0.106)	2.771*** (0.177)	1.677*** (0.226)	0.176*** (0.015)
Commentator	3.853*** (0.060)	7.084*** (0.079)	1.654*** (0.034)	0.666*** (0.013)
Pundit	3.899*** (0.292)	9.505*** (0.699)	3.152*** (0.480)	0.186*** (0.029)
Archaeologist	2.125* (1.168)	7.642** (3.527)	44.134*** (11.745)	0.046 (0.085)
Excavator	1.045*** (0.099)	2.380*** (0.204)	2.732*** (0.292)	0.169*** (0.019)
Strunk & White	5.332*** (0.334)	10.049*** (0.663)	13.065** (1.139)	0.212*** (0.032)
Copy Editor	3.931***	12.244***	2.628	0.072

	(0.861)	(2.023)	(4.424)	(0.056)
Proofreader	2.463*** (0.456)	5.445*** (1.115)	5.754*** (1.733)	0.107*** (0.039)
Tenacious	2.119*** (0.178)	2.731*** (0.290)	1.177*** (0.189)	0.178*** (0.020)
Unsung Hero	3.214*** (0.337)	3.817*** (0.512)	2.147*** (0.587)	0.221*** (0.037)
Tumbleweed	0.815*** (0.053)	1.800*** (0.090)	0.485*** (0.055)	0.473*** (0.023)

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

The results in Table 2-5 show that different earned badges have varying impacts on user activities. This finding has several implications for Stack Overflow for designing an effective badges system. First, comparing the set of badges about asking questions with the set of badges about answering questions, we can conclude that badges designed for answering activity are more impactful. For example, Nice/Good/Great Questions and Answers badges provide a direct comparison for this conclusion. This seems to suggest that Stack Overflow could devise additional badges based on answering activity. Second, we observe that, in general, across four activities, the influence of badges on asking questions is the smallest. This indicates that users post their questions when they are seeking answers and therefore may not post useless questions just to game the system to earn badges. One limitation of our study is that we do not have enough information to analyze the quality of questions asked. Therefore, we cannot conclude whether additional questions induced by badges lead to information overloading problem at Stack Overflow. If the administrators of Stack Overflow observe abusing behavior in posting meaningless questions, Stack Overflow can reduce the number of question badges or introduce more badges related to quality questions. Third, these 8 question badges are designed based on 3 different types of quality metrics (i.e., number of views, number of scores from up-votes, and number of favorites by other users.). Our results suggest that question badges are more effective when awarded based on positive endorsement by community members (i.e., voting up and adding as favorites) than based on the number of views only. All these findings seem to support the adoption of badges in this context; users seem to pursue badges and follow community norms at the same time. Fourth, the influence of badges on posting comments is very large and positive, implying that badges encourage users to be more socially active in terms of commenting on existing questions and answers. With the boost of sociability, users may feel more confident and capable about making contribution to the community through comments and the other three activities. This result is consistent with entertainment needs and sense of community being the key motivations on Q&A sites. Fifth, commenting badges "Commentator" and "Pundit" are different only in one additional criterion, which is about the quality of comments. Comparing the coefficients of these two badges reveals that adding a quality threshold induces more comments, indicating that Stack Overflow could consider introducing more badges with two criteria: one about quantity and one about quality of induced activities. Lastly, not many users have revision badges. Given that editing previous posts improves the quality of content and make Stack Overflow a valuable online information repository similar to Wikipedia, Stack Overflow should encourage more user participation for collaboratively editing active posts and therefore should introduce more revision badges and/or make it easy to earn badges for editing by lowering the threshold of the number of edits needed.

Because of the inclusion of a control group, our research design rules out the possibility that our findings are due to a common covariate, such as time trend. The PSM method also excludes the effects of unobserved covariates that are highly correlated with four weekly activity levels. For example, it is plausible that some users contribute more than others when they have more free time. Because they play more, they may get the target badge and contribute more in the following week. However, "free" users without the target badge should have also contributed more a week before the badge awarding date. Since we use the previous week's activities as the selection variables to form a control group in PSM, our analysis effectively compares the difference in the activity levels of two groups of free users in the following week. Therefore, the difference in activity levels in the following week can be attributed to the target badge, not to "time spent on Stack Overflow". Similarly, our research design also excludes other unobserved covariates, such as altruistic personality.

However, our current approach does not rule out the alternative explanation due to possible correlation among badges. For example, if two badges are usually earned simultaneously, then it could be that one badge, not the other, motivates users to contribute more. A similar alternative explanation is that users might only care about gold badges and because of that zeal, we observe increasing activities for bronze and silver badges. To address these potential issues, we conduct more robustness checks in the next section, and subsequently perform a macro-level analysis.

2.4.3. Robustness Checks for Micro-Level Analysis

To strengthen our causality argument that received badges motivate users to contribute more, we explored a number of alternative models. First, a user might receive multiple badges of the same type on the same day. Therefore, we might have overestimated the effects of badges in Section 2.4.2. We checked our sample to verify that this possibility does not pollute our results. We found that most badges are not awarded daily except few frequently obtained badges like “Nice Question”. In our data set, there are 2,877,348 user-badge pairs (for all 73 badges on Stack Overflow) with only 1.5% (44,172 out of 2,877,348) badges are received by the same user on the same day. Therefore, multiple badges awarded on the same day may only slightly inflate our estimates but do not change the qualitative nature of our results. Second, another concern against causality claim in Section 2.4.2 is that using the levels of four activities may not fully account for all unobservable covariates in selection of the control group. The generic weakness of the PSM approach is that unlike instrumental variables, PSM can only control for the observable covariates. Although levels of four activities right before getting a badge may capture a wide range of unobservable covariates, this approach is not perfect. To alleviate this issue, we included all available control variables in model (1) to verify our findings. Specifically, we added the number of gold, silver, and bronze badges awarded in t=0 and the life-time cumulative gold, silver, and bronze badges awarded upon t=0. The new estimation results reported in Table 2-6 are qualitatively similar to those in Table 2-5. Third, PSM analysis assumes that we can properly construct a control group for each user in the treatment group. Instead of using a control group, we can also compare each user’s activity level before earning a badge to his/her activity level after earning the badge. We used the paired *t*-test to compare the average number of activities 7 days before and after the date a user obtained a target badge. Only 3 out of 108 *t* statistics were negative.[16] However, these negative values were not significant at all. These results confirm

[16] Paired *t*-test results are available from the authors upon request.

the PSM findings that earned badges motivate users to contribute more in community activities.

Table 2-6 PSM Estimation Results with Additional Control Variables

	Activity Type			
Badge Name	Answering	Commenting	Revision	Asking
Popular Question	0.550*** (0.027)	1.506*** (0.049)	0.613*** (0.032)	0.372*** (0.008)
Notable Question	0.447*** (0.053)	1.344*** (0.089)	0.589*** (0.052)	0.360*** (0.013)
Famous Question	0.480*** (0.156)	1.004*** (0.261)	0.521** (0.244)	0.332*** (0.027)
Nice Question	1.107*** (0.061)	3.440*** (0.112)	1.483*** (0.098)	0.508*** (0.017)
Good Question	0.997*** (0.152)	3.061*** (0.280)	0.908*** (0.220)	0.351*** (0.034)
Great Question	1.443*** (0.497)	2.927*** (0.911)	1.611*** (0.384)	0.479*** (0.075)
Favorite Question	1.109*** (0.236)	2.511*** (0.484)	1.204*** (0.233)	0.322*** (0.055)
Stellar Question	2.217*** (0.532)	5.290*** (1.328)	1.584** (0.611)	0.397*** (0.125)
Nice Answer	3.863*** (0.050)	6.191*** (0.095)	2.843*** (0.064)	0.169*** (0.004)
Good Answer	3.222*** (0.110)	5.531*** (0.235)	2.791*** (0.173)	0.172*** (0.010)
Great Answer	2.545*** (0.349)	3.885*** (0.609)	1.503*** (0.364)	0.188*** (0.032)
Enlightened	4.876*** (0.101)	8.082*** (0.198)	3.597*** (0.152)	0.145*** (0.007)
Guru	3.224*** (0.245)	5.535*** (0.508)	2.532*** (0.494)	0.170*** (0.017)
Necromancer	1.553*** (0.095)	2.788*** (0.159)	1.434*** (0.164)	0.228*** (0.015)
Populist	2.365*** (0.445)	3.788*** (0.882)	1.426*** (0.502)	0.222*** (0.039)
Reversal	4.429*** (1.580)	11.424** (4.571)	6.082*** (2.174)	0.148 (0.128)
Revival	1.522*** (0.103)	2.609*** (0.166)	1.246*** (0.157)	0.171*** (0.015)
Commentator	2.513*** (0.050)	4.473*** (0.065)	0.844*** (0.035)	0.591*** (0.012)
Pundit	3.492*** (0.285)	9.023*** (0.695)	2.166*** (0.474)	0.168*** (0.030)
Archaeologist	1.492 (0.970)	4.988* (3.009)	40.542*** (9.412)	0.054 (0.100)
Excavator	0.941*** (0.097)	2.040*** (0.199)	2.074*** (0.273)	0.171*** (0.019)
Strunk & White	4.916*** (0.333)	9.306*** (0.655)	12.023*** (1.070)	0.193*** (0.032)
Copy Editor	3.905***	12.687***	3.567	0.069

	(0.923)	(2.196)	(4.250)	(0.064)
Proofreader	2.488*** (0.412)	5.406*** (0.978)	5.227*** (1.208)	0.111*** (0.040)
Tenacious	1.822*** (0.172)	2.435*** (0.294)	0.967*** (0.185)	0.180*** (0.021)
Unsung Hero	2.848*** (0.354)	3.573*** (0.555)	2.649*** (0.819)	0.213*** (0.038)
Tumbleweed	0.823*** (0.048)	1.687*** (0.086)	0.415*** (0.061)	0.465*** (0.022)

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

2.5. Macro-Level Panel Analysis

In the previous section, we characterized the causal relationship between earning a specific badge and increased incentive to contribute to four activities. Our identification hinged on the relative increase in activity levels in the week following the week after receiving a badge. Although we quantified and showed the significance of earned badges in inducing contributions from users, we could not say much about whether this increase was short-lived (i.e., affecting contributions to the next period only) or whether badges could facilitate continuous user participation. These limitations are caused by the lack of proper methods in assessing contributions in multiple periods and attributing the change in contributions to a specific badge that was awarded. First, a contributing user is likely to earn multiple badges (same kind and/or different kinds) over time. This makes it difficult to tease out the effect of each badge. Second, there might be individual specific factors, such as experience and reputation, that might affect the contributions of users, and these factors themselves change with time. Third, badges that were received in previous periods, i.e., badge history of users, might impact the incentive provided by a badge received in the current period. Fourth, badges are categorized into three groups (bronze, silver and gold) based on the difficulty in earning them, and some badges are interdependent because of this hierarchy. Due to the reasons mentioned above, one cannot accurately estimate the value of a badge in inducing contributions using micro-level analysis only. Therefore we complement our micro-level analysis with a macro-level analysis. Instead of assessing the impact of individual badges in a cross sectional setting, we consider the categories of badges and analyze the average value of a badge based on its category using panel data. Our objective is not only to quantify the influence of badges but also to assess the effectiveness of the hierarchical badges system in inducing continuous user participation.

2.5.1. Baseline Panel Regression

To examine the average effect of a badge based on its badge category (i.e., bronze, silver, or gold) on user participation in each activity, we estimate a fixed effects model as follows

$$\begin{aligned} Contributions_{i,j,t} \\ &= \alpha_{ij} + \beta_{1,j} tenure_{i,t-1} + \beta_{2,j} reputation_{i,t-1} \\ &+ \beta_{3,j} cum\_reputation_{i,t-1} + \beta_{4,j} gold_{i,t-1} + \beta_{5,j} silver_{i,t-1} \\ &+ \beta_{6,j} bronze_{i,t-1} + \beta_{7,j} cum\_gold_{i,t-1} + \beta_{8,j} cum\_silver_{i,t-1} \\ &+ \beta_{9,j} cum\_bronze_{i,t-1} + \varepsilon_{i,j,t} \qquad (2) \end{aligned}$$

where the $Contributions_{i,j,t}$ captures the number of times user *i* contributed to activity *j* in week *t*, and $gold_{i,t-1}$, $silver_{i,t-1}$, and $bronze_{i,t-1}$ are the total numbers of gold, silver, and bronze badges user *i* received in week *t-1*, respectively. Meanwhile, $cum\_gold_{i,t-1}$, $cum\_silver_{i,t-1}$, and $cum\_silver_{i,t-1}$ are the cumulative numbers of gold, silver, and bronze badges received upon week *t-1*, respectively. $tenure_{i,t-1}$, $reputation_{i,t-1}$, and $cum\_reputation_{i,t-1}$ serve as control variables in this panel model. Tenure captures the influence of user experience since the users' activity patterns may change with experience. Reputation score is an aggregate measure of the quality and relevance of a user's involvement and can therefore be considered as a proxy for how much the community trusts the user's contributions. Finally, α_{ij} is the fixed effect that captures the activity-specific individual unobserved heterogeneity.

The results for the answering activity are shown in Table 2-7. The results pertaining to the other three activities are given in Table 2-8, Table 2-9, and Table 2-10. From Model (c), which is the specification given in (2), we can see that the coefficients in front of the number of gold, silver, and bronze badges are all positive and significant across all three activities (answering, commenting, and revision). Since cumulative number of badges (whether gold, silver or bronze) increases at the same rate as the number of badges, we need to add two coefficients when analyzing the marginal impact of a badge in a badge category. Using the estimated coefficients in Table 2-7, Table 2-8, and Table 2-9, we can easily calculate that marginal impacts (i.e., sums of two estimated coefficients) of badges are always positive. For instance, getting a bronze, silver, or gold badge is associated with 0.683, 0.744, or 1.319 more answers, respectively.[17] Similarly, getting a bronze, silver or gold badge is associated with 0.864, 1.462, or 2.011 (0.351, 0.934, or 1.781) more comments (revisions), respectively. This implies that any kind of badge received in the current period increases the level of contribution to any of the three activities in the next period. Hence, badges are effective in motivating user participation irrespective of the users' history of badges. Furthermore, there is a ranked ordering among categories of badges in terms of their influence. We observe a sharp increase in contributions facilitated by categories of badges,

[17] These numbers are calculated from Table 2-6, column (c).as (0.683=0.706-0.023), (0.744=0.845-0.101), and (1.319=1.531-0.212).

consistent with the levels of difficulty in earning them. Across all these three activities, a gold badge received in the current period induces more user participation in the next period than a silver badge. Similarly, a silver badge received in the current period induces more user participation in the next period than a bronze badge. These findings suggest the importance of the hierarchical badges system. This hierarchy is like moving to the next (higher) level in computer games. Having more participation with a higher level of badge also proves that the hierarchical badges system is effective in facilitating continuous user engagement.

As for the influence of the cumulative number of gold, silver, and bronze badges, all the estimated coefficients are significantly negative or non-significant. One exception is the influence of cumulative number of gold badges on revision activity (The estimated coefficient is positive, but significant at the 10% level only). These results indicate that users who have already earned many badges may contribute less. That is, between two users with the same number of badges earned in the current period, the user with less cumulative number of badges contributes more to the Q&A site in the next period. Because each additional badge of any kind still leads to more activities, this finding just suggests a diminishing positive effect of badges over the cumulative number of badges. There are several possible explanations behind this finding. First, users may get bored after having obtained many badges and the stimulus from badges may reduce. Second, users with a higher cumulative number of badges are more likely to have fewer unearned badges left for them to pursue. Third, users may first pursue the badges that are more attractive to them, leading to a diminishing impact over the history of badges earned.

To explore more about this pattern, we included squared terms of cumulative number of badges in Model (d) to see if the diminishing rate is accelerating. Except one, all of the coefficients of these square terms are insignificant. Thus, there is no accelerating diminishing effect of cumulative number of badges. All these results are consistent across badge categories and across contribution activities.

Table 2-7 Results of the Fixed Effects Model (DV = Answering)

Variables	(a)	(b)	(c)	(d)
Tenure	-0.014*** (0.001)	-0.028*** (0.001)	-0.001 (0.003)	0.003 (0.002)
Reputation		0.475*** (0.006)	0.407*** (0.005)	0.409*** (0.005)
Cum_reputation		0.027*** (0.006)	0.058*** (0.012)	0.067*** (0.007)
Gold			1.531*** (0.097)	1.553*** (0.093)
Silver			0.845*** (0.039)	0.858*** (0.039)
Bronze			0.706*** (0.015)	0.706*** (0.015)
Cum_gold			-0.212*** (0.066)	-0.287*** (0.058)
Cum_silver			-0.101*** (0.022)	-0.114*** (0.013)
Cum_bronze			-0.023** (0.011)	-0.033*** (0.005)
Cum_gold^2				0.026* (0.015)
Cum_silver^2				0.000 (0.000)
Cum_bronze^2				0.000 (0.000)
Constant	0.842*** (0.009)	0.506*** (0.020)	0.300*** (0.021)	0.297*** (0.020)
R-squared	0.001	0.054	0.088	0.089

The sample includes 4,070,427 user-week pairs and 46,571unique users.
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2-8 Results of Fixed Effects Model (DV = Commenting)

Variables	(a)	(b)	(c)	(d)
Tenure	0.000 (0.001)	-0.047*** (0.002)	-0.025*** (0.006)	-0.024*** (0.007)
Reputation		0.732*** (0.014)	0.626*** (0.010)	0.627*** (0.010)
Cum_reputation		0.177*** (0.016)	0.246*** (0.024)	0.251*** (0.020)
Gold			2.125*** (0.173)	2.147*** (0.169)
Silver			1.471*** (0.092)	1.470*** (0.089)
Bronze			0.922*** (0.030)	0.922*** (0.029)
Cum_gold			-0.114 (0.117)	-0.214* (0.121)
Cum_silver			-0.009 (0.040)	-0.006 (0.037)
Cum_bronze			-0.058*** (0.016)	-0.063*** (0.010)
Cum_gold^2				0.026 (0.036)
Cum_silver^2				-0.000 (0.000)
Cum_bronze^2				0.000 (0.000)
Constant	1.196*** (0.017)	0.382*** (0.056)	0.206*** (0.050)	0.207*** (0.048)
R-squared	0.000	0.046	0.062	0.062

The sample includes 4,070,427 user-week pairs and 46,571unique users.
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2-9 Results of Fixed Effects Model (DV = Revision)

Variables	(a)	(b)	(c)	(d)
Tenure	0.004*** (0.001)	-0.018*** (0.001)	-0.009*** (0.003)	-0.011*** (0.003)
Reputation		0.305*** (0.007)	0.257*** (0.005)	0.256*** (0.005)
Cum_reputation		0.090*** (0.007)	0.106*** (0.012)	0.105*** (0.009)
Gold			1.563*** (0.160)	1.579*** (0.156)
Silver			0.976*** (0.066)	0.967*** (0.065)
Bronze			0.367*** (0.017)	0.367*** (0.016)
Cum_gold			0.218* (0.114)	0.125 (0.106)
Cum_silver			-0.042 (0.026)	-0.031* (0.018)
Cum_bronze			-0.016* (0.008)	-0.013** (0.006)
Cum_gold^2				0.020 (0.030)
Cum_silver^2				-0.000 (0.000)
Cum_bronze^2				-0.000 (0.000)
Constant	0.394*** (0.010)	0.018 (0.027)	-0.054** (0.023)	-0.053** (0.023)
R-squared	0.000	0.016	0.023	0.024

The sample includes 4,070,427 user-week pairs and 46,571unique users.
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2-10 Results of Fixed Effects Model (DV = Asking)

Variables	(a)	(b)	(c)	(d)
Tenure	-0.005***	-0.006***	-0.002***	0.001*
	(0.000)	(0.000)	(0.001)	(0.000)
Reputation		0.056***	0.050***	0.051***
		(0.001)	(0.001)	(0.001)
Cum_reputation		0.004***	0.027***	0.030***
		(0.001)	(0.003)	(0.002)
Gold			-0.008	-0.007
			(0.012)	(0.012)
Silver			-0.014***	-0.004
			(0.004)	(0.004)
Bronze			0.088***	0.088***
			(0.002)	(0.002)
Cum_gold			-0.032***	-0.017*
			(0.009)	(0.010)
Cum_silver			0.018***	0.006**
			(0.004)	(0.003)
Cum_bronze			-0.023***	-0.027***
			(0.003)	(0.002)
Cum_gold^2				0.001
				(0.003)
Cum_silver^2				0.000***
				(0.000)
Cum_bronze^2				0.000***
				(0.000)
Constant	0.292***	0.244***	0.238***	0.235***
	(0.002)	(0.003)	(0.004)	(0.003)
R-squared	0.002	0.011	0.022	0.024

The sample includes 4,070,427 user-week pairs and 46,571unique users.
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Interestingly, the impact of badges on users' activity of asking questions is different from the other three activities. From Table 2-10 (Model (c)), we can see that getting a bronze, silver, or gold badge is associated with 0.065, 0.004 or -0.04 more questions, respectively. First, these numbers are much smaller than the numbers for the other three activities. Although they are statistically significant, economic significance is lacking. Hence, users do not seem to be motivated much by badges to ask questions. This is intuitive because typically users ask questions when they genuinely seek answers. Users do not need external rewards such as badges to boost their incentive to post their questions. Possible solutions through other users' answers to their questions might provide enough stimuli. This finding is also consistent with the results we derived for individual badges in Section 2.4. Second, the number of gold badges is negatively correlated with the number of questions asked. One possible explanation for this finding is that users might prefer to engage in different activities on the site. While some users mostly seek help by asking questions, some other users enjoy helping community members by answering their questions. Since gold badges are mainly awarded for the answering activity, a small group of users who primarily provide answers also earn the majority of gold badges. Another possible reason is that users with gold badges are more knowledgeable in programming and therefore less likely to ask questions.

Since answering questions is the most vital activity for the continued success of the Stack Overflow, and the contributions from experienced community members are the key to continuous user participation, the result regarding the effect of "tenure" on answering activity is worthy of discussion. With only tenure included in the regression (Model (a)), the coefficient is negative, indicating that users answer less as their experience grows in the community. This effect is qualitatively the same even when we include “reputation” (Model (b)). However, after including badges into the regression, the coefficient of tenure on answering activity becomes insignificant (in both Models (c) and (d)). The change of the significance of the estimated coefficient may suggest that conditional on the number of badges, users do not necessarily contribute less answers as their experience grows, implying that badges help alleviate the negative effect of tenure. This finding provides evidence that the badges system is indeed effective in motivating experienced users to continuously contribute to the community. Furthermore, we can observe that reputation plays a significant role in community involvement. The coefficients of reputation and cumulative reputation are positive and significant in all models and across all types of activities. That is, the greater the reputation of a user, the greater his/her involvement level. Since reputation is a reflection of the quality and relevance of user participation, the finding that users with higher quality submissions contribute more to the Q&A is instrumental in achieving the status of being reliable and trusted information source for programmers. Overall, we can conclude that the use of the hierarchical badges system enables Stack Overflow to facilitate both continuous user engagement and quality user engagement.

2.5.2. Panel Regression with Sub-Samples by Badge History

The analysis in section 2.5.1 assumes that the influence of a badge category remains constant as users go through different phases over time in terms of their contributions. It is possible that the behavior of users may change as they evolve from a novice to a senior community member. Therefore, the motivation provided by a badge, thus contribution induced, can differ depending on the categories of earned badges in a user's badge history. To examine this issue, we estimated the fixed effects model with separated time phases. First, for each user, we partitioned time dimension into four phases based on the badge history of the user: (i) no badges, (ii) only bronze badges, (iii) only silver and bronze badges, and (iv) all categories of badges. Next, we reran the fixed effects model using the sub-samples from the last three phases. In model (e), we included all observations after a user has received his first bronze badge but before the user has received his first silver badge. Similarly, in model (f), we included all observations after a user has received his first silver badge but before the user has received his first gold badge. Finally, model (g) includes all observations after a user received his first gold badge. Table 2-11 presents the results for the fixed effects model with separated phases (when DV = answering). The results pertaining to other three activities are given in Tables Table 2-12, Table 2-13, and Table 2-14.

Table 2-11 FE Estimates with Separated Phases (DV = Answering)

Variables	(e)	(f)	(g)
Tenure	0.017***	0.021***	0.012
	(0.001)	(0.003)	(0.019)
Reputation	0.261***	0.282***	0.512***
	(0.004)	(0.005)	(0.017)
Cum_reputation	-0.001	-0.103***	-0.723***
	(0.004)	(0.021)	(0.192)
Bronze	0.406***	0.654***	0.748***
	(0.008)	(0.027)	(0.067)
Cum_bronze	-0.030***	-0.043***	0.005
	(0.003)	(0.005)	(0.041)
Silver		0.463***	0.769***
		(0.019)	(0.068)
Cum_silver		-0.112***	-0.193***
		(0.014)	(0.064)
Gold			0.928***
			(0.082)
Cum_gold			-0.237*
			(0.131)
Constant	0.265***	1.304***	7.845***
	(0.014)	(0.106)	(1.187)
R-squared	0.053	0.055	0.092
Observations	2,007,205	1,450,082	247,319
Number of users	45,389	35,460	6,050

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

Table 2-12 FE Estimates with Separated Phases (DV = Commenting)

Variables	(e)	(f)	(g)
Tenure	0.002**	0.020***	0.013
	(0.001)	(0.005)	(0.032)
Reputation	0.391***	0.447***	0.834***
	(0.005)	(0.008)	(0.030)
Cum_reputation	0.071***	0.054	-0.735*
	(0.006)	(0.038)	(0.380)
Bronze	0.501***	0.946***	1.353***
	(0.011)	(0.038)	(0.119)
Cum_bronze	0.011**	-0.096***	-0.025
	(0.005)	(0.011)	(0.072)
Silver		0.679***	1.508***
		(0.034)	(0.149)
Cum_silver		-0.083***	-0.208*
		(0.027)	(0.110)
Gold			1.164***
			(0.160)
Cum_gold			-0.282
			(0.225)
Constant	0.162***	1.306***	10.015***
	(0.021)	(0.189)	(2.366)
R-squared	0.047	0.040	0.051
Observations	2,007,205	1,450,082	247,319
Number of users	45,389	35,460	6,050

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

Table 2-13 FE Estimates with Separated Phases (DV = Revision)

Variables	(e)	(f)	(g)
Tenure	-0.002***	0.004	-0.011
	(0.000)	(0.003)	(0.024)
Reputation	0.118***	0.183***	0.461***
	(0.002)	(0.004)	(0.024)
Cum_reputation	0.022***	0.117***	0.317
	(0.002)	(0.023)	(0.355)
Bronze	0.142***	0.460***	0.967***
	(0.005)	(0.025)	(0.106)
Cum_bronze	0.016***	-0.030***	0.012
	(0.002)	(0.006)	(0.033)
Silver		0.352***	1.206***
		(0.026)	(0.163)
Cum_silver		-0.039***	-0.184***
		(0.011)	(0.061)
Gold			0.958***
			(0.128)
Cum_gold			0.001
			(0.279)
Constant	-0.018**	-0.105	-0.141
	(0.009)	(0.121)	(2.444)
R-squared	0.027	0.019	0.018
Observations	2,007,205	1,450,082	247,319
Number of users	45,389	35,460	6,050

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

Table 2-14 FE Estimates with Separated Phases (DV = Asking)

Variables	(e)	(f)	(g)
Tenure	0.004***	0.004***	-0.003**
	(0.000)	(0.001)	(0.002)
Reputation	0.055***	0.034***	0.025***
	(0.001)	(0.001)	(0.002)
Cum_reputation	0.003**	-0.018***	-0.051***
	(0.001)	(0.004)	(0.018)
Bronze	0.085***	0.031***	0.013**
	(0.002)	(0.003)	(0.005)
Cum_bronze	-0.017***	-0.031***	-0.011**
	(0.001)	(0.003)	(0.005)
Silver		0.021***	-0.005
		(0.004)	(0.007)
Cum_silver		0.014***	0.015**
		(0.003)	(0.007)
Gold			0.036***
			(0.011)
Cum_gold			-0.023**
			(0.012)
Constant	0.252***	0.529***	0.806***
	(0.004)	(0.022)	(0.110)
R-squared	0.015	0.013	0.012
Observations	2,007,205	1,450,082	247,319
Number of users	45,389	35,460	6,050

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Our results reveal that the motivation provided by badges gets stronger once users obtain a badge from a higher difficulty category, implying that there are externalities among categories of badges. For instance, the marginal effect of a bronze badge in models (e), (f) and (g) are 0.376, 0.611, and 0.753, respectively. Hence, although earning a bronze badge always incentivizes users to answer more questions, users with at least a silver badge answer more questions than users with bronze badges only. Similarly, users with at least a gold badge answer more questions than the users with bronze and silver badges only. This relationship is also true for silver badges. The marginal effect of a silver badge in models (f) and (g) are 0.351 and 0.576, respectively. Thus, although earning a silver badge always incentivizes users to answer more questions, users with at least a gold badge answer more questions than users with bronze and silver badges only. Furthermore, the increasing marginal impact of a badge category (whether bronze or silver) together with a higher-level earned badge is also true for other two activities: commenting and revision (see Table 2-12 and Table 2-13). One possible explanation is that as users get more difficult badges, they also learn how to contribute to the site more efficiently, and subsequently the effort cost for them to participate gets lower. Another likely reason is that more experienced users become more emotionally attached to the Q&A community; they may have more virtual friends and have established more valuable social identity through higher categories of earned badges. Irrespective of the underlying mechanism for this result, it is clear that there are positive externalities from higher categories of badges to lower categories of badges on Stack Overflow.

Overall, we can conclude that the hierarchical badges system used by Stack Overflow is highly effective in promoting desired behavior. Having different categories of badges not only facilitates a hierarchical ordering in efficiency (as discussed in section 2.5.1), it also boosts the marginal influence of lower badge categories. A gold badge directly induces more activities than a silver badge. Similarly, a silver badge directly induces more activities than a bronze badge. However, there are also externalities. A gold badge indirectly increases the contribution induced by a silver badge. Likewise, a silver badge indirectly increases the contribution induced by a bronze badge. Taken together, these two effects give rise to continuous participation and engagement in the community.

2.5.3. Robustness Checks for Macro-level Analysis

First, although no direct monetary reward exists to reimburse users for their contributions, users can receive indirect benefits accruing from their contributions to the community. In particular, some contributors might use Stack Overflow as a platform to signal their ability and competence to potential employers. Being recognized as a knowledgeable programmer with many badges on the profile can help users get new jobs and/or advance their careers. Hence, users can spend their time and effort to contribute high-quality answers not only because of their motivations fostered by earned badges, but also due to *indirect* monetary rewards resulting from spillover effects to the labor markets. We are aware that one group of participants on Stack Overflow is also users of "Stack Overflow Career," an online labor market for programmers.[18] To be specific, 2,071 users out of 46,571

[18] http://careers.stackoverflow.com

total users in our dataset participate in the Stack Overflow Career. To ensure that our results regarding the impact of earned badges are not driven by the group of career-focused contributors, we excluded users who participated in Stack Overflow Career from our sample and reran our econometric analysis. Our results remain qualitatively the same after excluding the Stack Overflow Career users.

Second, there may be time trends such as seasonality effects or holiday effects which affect users' contribution activities but is currently not captured in the panel regression model. To show that our estimates are not affected by the time trends, we included dummies of calendar time and reran the panel regression models. Our results are qualitatively the same.

2.6. Concluding Remarks

In this paper, we examine the impacts of virtual badges on user contributions at the Stack Overflow Q&A site. Specifically, we assess the extent to which users are incentivized by earned badges in their contributions to four major activities: answering questions, commenting, making revisions, and asking questions. We present strong empirical evidence that confirms the value of the hierarchical badges system in motivating users to participate more and have a continuous engagement. Our data analyses consist of two complementary parts: the micro-level analysis and the macro-level analysis. While the objective of micro-level analysis is to study the individual effect of each specific badge, the objective of macro-level data analysis is to investigate the relative effectiveness of badges with different difficulty levels.

In our micro-level data analysis, we examine the impact of a specific badge on user activities by estimating a difference-in-differences model. We aggregate user activities a week before and after a user receives a specific badge. To eliminate alternative explanations, we utilize propensity score matching to form a control group. Comparing the differences in the numbers of activities before and after the badge awarding date in the control group and the treatment group, we confirm that almost all kinds of badges motivate users to contribute more in all four types of activities. Interestingly, badges stimulate users to participate more in activities that are not specified in the rule set for getting badges. Furthermore, we find that even negative badges could motivate users to engage more in site activities. Overall, our findings imply that receiving any kind of badge might affect one or more latent variables associated with users, positively influencing them to contribute more in all activities. Finally, we show that our results from the micro-level analysis are robust to a number of control checks and alternative explanations.

In our macro-level data analysis, we examine the influences of three categories (gold, silver, and bronze) of badges on user activities by estimating a fixed effects panel model. These badge categories capture the relative difficulty in earning them because a gold badge requires more contribution than a silver badge, and a silver badges requires more contribution than a bronze badge. We show that the impacts of three categories of badges on answering, commenting and revision activities are qualitatively similar. On average, our results show that gold badges are the most impactful while bronze badges are the least impactful. Hence, there is a ranked ordering in efficiency of badges corresponding to difficulty levels. Furthermore, our analysis shows that the hierarchical badges system helps cultivate users' loyalty to the community because the contribution induced by a badge category increases once users obtain a badge from a higher category. Hence, we can conclude that the hierarchical badges system exhibiting both ranked influence and positive externalities across badge categories is the true force behind continuous user engagement and participation. Different from other three activities, we observe that badges do not seem to motivate users much to ask questions. This is intuitive because getting answers to asked questions provide enough incentive, and users do not need an external stimulus, such as badges.

Taking micro-level and macro-level analyses together, our research provides strong empirical evidence that gamification through the use of the hierarchical badges system at Stack Oveflow Q&A site promotes voluntary participation and continuous engagement. Initially, badges provide extrinsic motivations for participation and induce users to take part in community activities. Once users engage in community activities and earn badges for their contributions, users subsequently start to internalize tangible benefits because earned badges tap into and enhance the intrinsic motivations of users. Hence, the hierarchical badges system is highly effective in eliciting and sustaining desired contributions from community users because it facilitates a positive reinforcement loop between badges and contributions.

Our study inevitably has some limitations. Some of these could be fruitful avenues for future research. First, we acknowledge that it is very difficult to isolate the effect of each specific badge even with our sophisticated PSM approach. The effect of one badge may be confounded with the effect from other badges. A different research design is required to better tease out the influence of each badge.

One possibility is to leverage on advanced econometric approaches such as regression discontinuity (Imbens and Lemieux 2008). Second, although our study documents the efficacy of having badges and levels among badges in providing stimuli to users for continuous engagement, this study does not show the detailed link between intrinsic motivational factors and badges. In order to propose an optimal hierarchical badges system, one has to understand the underlying psychological constructs that are triggered by earned badges in motivating users to contribute constantly on the Q&A site. More research using survey and experimental methodologies is needed to map the intrinsic motivational factors, such as a sense of community and relatedness, entertainment needs, competition, to the drivers of activities triggered by badges. Third, because our data are crawled from the registered users on stackoverflow.com, we do not have any information on lurkers, who regularly read but never post anything on stackoverflow.com. Controlling this self-selection issue and analyzing how badges could turn lurkers into contributors could be another future research direction. Fourth, since we observe that users have different behavior patterns on stackoverflow.com, future research can classify users into different groups, such as askers and answerers (Gazan 2011). This research direction could shed more light on the activities patterns of different stereotypes of users. Fifth, our study does not assess the quality of activities. In reality, badges can influence both quantity and quality of activities. Studying the quality dimension requires textual analysis, and therefore beyond the focus of this study. The other interesting research direction is to examine whether users exhibit "variety seeking" behavior when seeking badges, similar to its counterpart in marketing (McAlister and Pessemier 1982). The last but not the least, one can examine how the newly launched Stack Career Exchange may affect the motivation of users to contribute on Stackoverflow.com.

3. Study 2: The Unexpected Outcome of Increased User Participation

3.1. Introduction

In the last decade, we have seen the proliferation of online community, which facilitates the interactive information exchange on a variety of topics in a large group of people. The conventional wisdom suggests that the success of online community greatly depends on the size of user base. To make the process of information exchange smoothly and frequently, an online community has to maintain a large pool of users who are willing to contribute their time and effort. The larger user base of an online community, the more attractive and useful this online community is. Therefore, online communities are eager to acquire more users to gain competitive advantage. They take different actions to attract new users based on their current condition. For example, Facebook, the currently largest social networking website in the world, plans to scale up its free service to offer free mobile network access in developing countries, in aim to provide a more connected virtual world to its users (Facebook 2015). Twitter, another iconic social networking website, introduced the new homepage in April 2015 with the goal to attract more new users by simplifying the website interface (Mohan 2015).

However, is the conventional wisdom always true? Do additional users always bring positive value to the communities? The conventional wisdom is intuitively the hypothesis of positive network effect in Economics (Shapiro and Varian 2013): the more users in an online community, the more utility a user can derive from joining and participating in this online community, thus leading to both higher user acquisition rate and user retention rate. There is an implicit assumption in developing this hypothesis: the homogeneity of users in an online community and the alignment of interests of all users within the same online community (Gu et al. 2007a). On the contrary, we argue that users in an online community could be heterogeneous and there are complex interactions among different groups of users within the same online community. Attracting more new users may have a negative effect on existing active users if users are heterogeneous and their interests are not aligned. In this paper, we test our conjecture with the real-world case of Wikipedia.

In 2013, Wikipedia initiated to replace its old editing tool Wikitext with a new visual editor (Protalinski 2013). The Wikitext is like Latex, a markup language for word processing, which requires an investment of time to learn how it works. It involves programming, compiling, and debugging of the document edited in Wikitext in order to specify the content and format displayed onscreen. Even an experienced user may make simple mistakes in editing the Wikipedia articles by Wikitext. Unlike Wikitext, the new visual editor is similar to Microsoft Word, a WYSIWYG (What You See Is What You Get) editor, in which content and format onscreen during editing appears in a form closely corresponding to its displayed appearance. Figure 3-1 and Figure 3-2 present the differentiated editing interfaces of Wikitext and visual editor to edit the same article of "Information system" at Wikipedia.

Figure 3-1 the Editing Interface of Wikitext

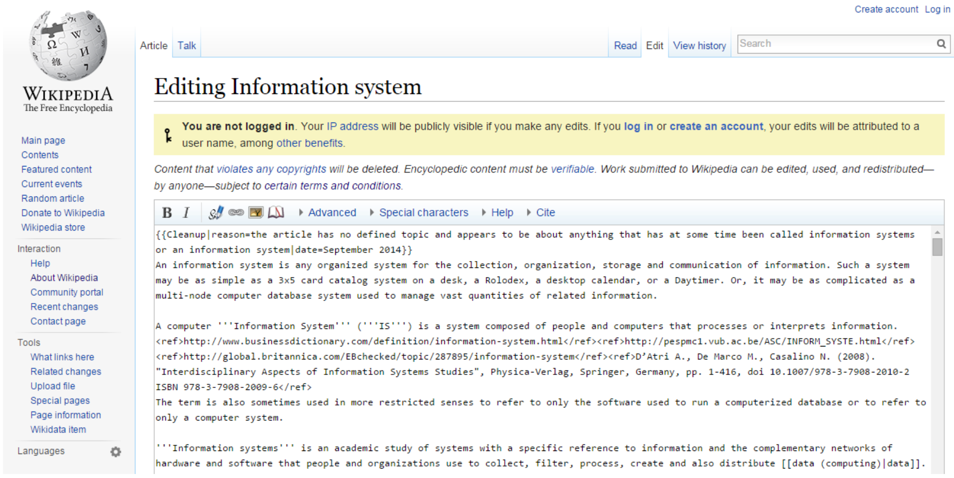
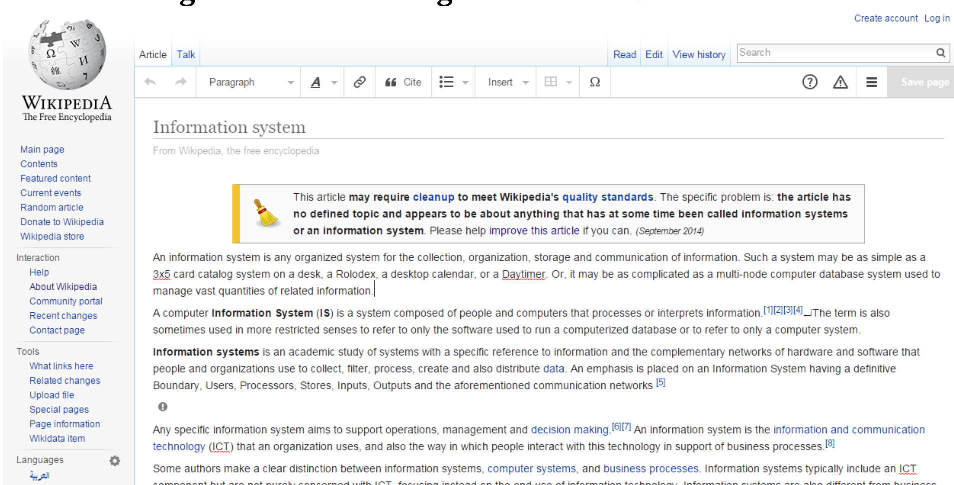


Figure 3-2 the Editing Interface of Visual Editor



The change of editing interface thus greatly simplifies the process of editing Wikipedia articles and encourages new users to try their first-time editing on English Wikipedia. The Wikipedia Foundation anticipated the enhanced user interface would attract more new editors and also help existing senior editors to easily contribute to topics that needed further improvement. In July 2013, English Wikipedia improved their user interface by replacing the originally inconvenient editing tool Wikitext with the new user-friendly visual editor. English Wikipedia observed a significantly increasing number of new users after the change of editor. This policy change happening on Wikipedia brings us an ideal research setting to investigate our research question:

How does an enlarged user base affect senior users' participation in an online community?

In the empirical examination of our research question, we first establish the causality between the launch of visual editor and senior users' participation with the observational data from Wikipedia. In our research context of English Wikipedia, the impact of an enlarged user base is captured by the launch of visual editor because the change of editor suddenly attracted a large number of new users. We collect our data set from the English Wikipedia and examine the effect of the launch of visual editor on senior users' contribution. Our data set is a daily panel data of 5,191 senior users before the launch of visual editor at English Wikipedia. For each user, we obtain his complete editing history data with textual information. We examine the behavior change of senior users' contribution using a fixed-effect panel regression. Our results suggest that, after the launch of visual editor, senior users increased their contribution in a short run (with the time window of 2 weeks) but decreased their contribution in a long run (with the time window of 2.5 months). On average, a senior user submitted 14.9% more in addition in the short run but submitted 22.9% less in addition in the long run. Similarly, a senior user on average submitted 9.9% more in deletion in the short run but submitted 13.9% in deletion in the long run. To establish the causal relationship between the launch of visual editor and the decrease of senior users' contribution, we use the senior users at

German Wikipedia as control group and estimate a Difference-in-Differences (DID) model. We pick users at German Wikipedia as the control group because German Wikipedia is not part of the project of visual editor. The significant results of the DID model indicate that, because of the launch of visual editor, senior users' contribution at English Wikipedia decreased by 2.8%, 3.4%, and 11.2% in the number of submission, the number of characters added, and the number of characters deleted, respectively.

We propose the decrease of senior users' contribution is caused by the negative network effect of the new users who mostly generated low-quality contribution. To capture this negative network effect, we use the anonymous contribution because low-quality contribution of new users is more likely to be submitted anonymously. We estimate the DID model with the treatment being the percentage of anonymous contribution a senior user experienced after the launch of visual editor. Our identification relies on the exogenous shock on the quality level of anonymous users' contribution. After the launch of visual editor, anonymous users, including those who edit the Wikipedia articles in a malicious manner that is intentionally disruptive, can easily submit their contribution. In contrast, anonymous users have to learn the markup language Wikitext to submit their contribution before this policy change. Lowering the bar for anonymous users to contribute consequently lowers the average quality level of anonymous users' contribution. The DID estimates show that, comparing to the contribution of senior users without anonymous collaborators, the contribution of senior users with anonymous collaborators could decrease by up to 22.5%, 62.9%, and 35.7%, in submission, addition, and deletion, respectively. We find that the more a senior user experienced the anonymous contribution after the launch of visual editor, the less he would contribute. We conclude that the negative network effect as a result of lowering the bar of contribution is the major cause driving the senior users away and is harmful to the sustainability of online community. It is the quality dimension of new users' contribution that determines whether the network effect (the relationship between the size of user base and the level of user participation) is positive or negative.

Our research contributes to the literature in two ways. First, in this paper, we find that the interaction between different groups of users (senior uses and new users) plays an important role in sustaining an online community. Our research suggests that incorporating the complex dynamics of different groups of users is important to study the evolvement of online community. Second, we find that an increase of the number of new users can lead to the decrease of senior users' participation. This is different from the theoretical prediction of positive direct network effect which predominantly stated in the literature of network effect (Katz and Shapiro 1985). Our research suggests that the quality dimension of user contribution is the crucial factor in determining the direction of network effect and extend the understanding of network effect in the context of online community (Asvanund et al. 2004; Gu et al. 2007b; Zhang and Zhu 2011).

Our research also provides helpful suggestions to practitioners who manage online community: the administrators of online community should make a tradeoff between the quantity and the quality of user contribution. For every online community, it is critical to acquire new users to build a large user base. However, a larger number of new users do not sufficiently guarantee a better online community with higher quality. Since there are different groups of users in the same online community, making a policy change to please new user may meanwhile make senior or loyal users unsatisfied. Administrators should keep an eye on balancing the needs of different groups of users so as to achieve a better ecosystem in online community.

3.2. Related Literature

Our research is related to two streams of literature: the literature on online community in the field of Information Systems and the literature on network effect in the field of Economics.

3.2.1. Online Community

The sustainability of online community is a fundamental research question in the research area of online community. In online community, most users voluntarily contribute their time and effort without monetary rewards. Therefore, it is important to understand users' motivations of participation so as to induce their continued engagement and establish a sustainable online community.

In the field of Information Systems, researchers primarily focused on identifying intrinsic motivations that drive voluntary user participation in online community. Most extant studies are done by behavioral research methods such as surveys and case studies. The main contribution of each paper lies in a unique focus on a specific type of intrinsic motivation and/or in studying an unexplored form of online community. For instance, Constant et al. (1994) and Constant et al. (1996)

found that reputation and social image play an important role in users' participation in online community. Compeau and Higgins (1995) revealed that self-efficacy, which means users capitalize on their ability to find solutions in order to accomplish challenging goals, could be the source of user engagement in online community. This is further confirmed in the research done by Sun et al. (2011) in the context of a crowd-sourcing website (TaskCn) and Jin et al. (2012) in the context of a Q&A website (Yahoo! Answers China). Other motivations identified in the research on traditional online community includes learning and gaining information (Daugherty et al. 2005), entertainment needs (Kankanhalli et al. 2005; Sutanto et al. 2011), and so on. Recently, the bloom of social network websites has shifted the researcher' attentions to the social aspect of intrinsic motivations. In a broad sense, social capital (Ren et al. 2007; Ren et al. 2012), reciprocity (Chiu et al. 2006; Faraj and Johnson 2011), and community interest (Bateman et al. 2011; Ma and Agarwal 2007) are considered as three major categories of intrinsic motivations for users to develop mutual benefits and maintain a long term relationship within the same online community.

The extant literature investigating users' motivations to participate in online community implicitly assume that the interests of different users in the same community are aligned and therefore an increased number of users will lead to a positive outcome. In this paper, we explore the interaction between two different groups of users (senior users and new users) in online community and view the online community from the perspective of ecosystem. Our research suggests senior users and new users may have conflicted interests and the interaction between different groups of users plays an important role in the sustainability of online community. Researchers who are interested in the evolvement of online community could take into consideration the complex dynamics of different groups of users in their theoretical development.

3.2.2. Network Effect

Network effect (also called network externality or demand-side economies of scale) means that the value of a goods or service depends on the number of other users using it as well (Katz and Shapiro 1985; Katz and Shapiro 1992; Shapiro and Varian 2013). In literature, there are two forms of network effect that has been extensively studied. The first one is direct network effect, suggesting an increase in usage leads to a direct increase in the value of the product or service. The textbook example of direct network effect is the telephone network. The utility of a telephone increases with the number of other users the owner can reach. The more people who own a telephone, the more utility the telephone can bring to each owner. Another one is the indirect network effect, suggesting an increase in the usage of a product or service drives an increase in the value of its complementary product or service, which in turn increases the value of the original. A classic example of complementary product is the software such as the Office Suite to Windows Operating Systems. The more customers use Windows, the more software will be developed for Windows. This in turn leads to higher valuation of Windows to customers.

We focus on the direct network in this paper. In the context of online community, it works in the same way as the telephone network. The utility of an online community increases with the number of people a user can communicate. Zhang and Zhu (2011) examined the relationship between group size and users' incentive to contribute with a natural experiment at Wikipedia. In their seminal paper, they find that the positive network effect (or "social effects" in their paper) dominates the free-riding incentives and helps to sustain online community. Their results are intuitive and consistent with the theoretical prediction of positive network effect. On the other hand, there exist a handful of papers documenting an interesting phenomenon: the negative network effect in online community. Butler (2001) found that as the user size increases in an online community, although the user gain (measured by the available resource in an online community) increases, this online community experiences a significant "churn" rate. Consistent with the observation of Butler (2001), Gu et al. (2007a) found that an increase of sharers leads to exits of existing sharers in an Peer-to-Peer (P2P) music sharing community. In the same context of P2P music-sharing network, Asvanund et al. (2004) stated that there are both positive and negative network effects. Their results suggest users contribute additional value to the network at a decreasing rate with the size of network (the decreasing positive network effect) and users incur additional costs on the network at an increasing rate with the size of network (the increasing negative network effect) due to the congestion on shared resources. Understanding the direction and the moderator of network effect is important because it helps online community managers to maintain an optimal size of user base. However, in the papers documenting the negative network effect in online community, scholars related population changes at the aggregate level with population status in the community (Gu et al. 2007a). The lacking of individual level data limits scholars' capability to establish causality and explore the underlying mechanism of negative network effect.

In this paper, we emphasize that, in addition to the quantity of user contribution, it is also important to take into consideration the quality of user contribution so as to leverage on the network effect to sustain an online community. We examine how the positive network effect becomes negative by a natural experiment at English Wikipedia to establish the causality and investigate the underlying mechanism. We point out the quality of new users' contribution determines whether the network effect is positive or negative in the context of online community. Our research also suggests that contribution cost or learning cost, which can be implemented as the difficulty of user interface, could be applied as an effective way to achieve the balance between the quantity and quality of user contribution.

3.3. Background and Data

Wikipedia is a multilingual, web-based and free-content encyclopedias project supported by the Wikipedia Foundation and based on the model of openly editable content.[19] Wikipedia is created in 2001 and has become one of the largest reference websites on Internet. There are more than 73,000 active contributors working on more than 35,000,000 articles in 290 languages. As of June 2015, the number of monthly unique visitors of Wikipedia has reached 439 million.

[19] Please refer to the official webpage (https://en.wikipedia.org/wiki/Wikipedia:About) for further information.

Among all the Wikipedia websites in different languages, English Wikipedia is the largest website with 4,939,075 articles and on average 750 new articles per day (Wikimedia Foundation 2015). However, after reaching its peak in 2007, the number of new users (or “Wikipedians”) at English Wikipedia dropped and the retention rate remained low ever since (Wikimedia Foundation 2011). To acquire more new contributors, Wikipedia foundation introduced the visual editor to replace the original markup language Wikitext as the default editing tool and make Wikitext as an “opt-in” feature to edit Wikipedia articles.[20] Before the launch of visual editor, new users have to learn the markup language Wikitext to edit Wikipedia articles. The Wikipedia foundation consider the markup language Wikitext as the entry obstacle for new users to contribute and will drive some potentially good contributors away. Therefore the Wikipedia foundation developed the visual editor, a WYSIWYG (What You See Is What You Get) editor. The difference between Wikitext and visual editor at Wikipedia is analogous to the difference between Latex and Word. The visual editor is made as the default editor for all users on July 15th 2013 at English Wikipedia for all users. The introduction of visual editor at English Wikipedia is considered as a successful move to acquire new contributors. It is described as "the most significant change in Wikipedia's short history" (The Economist 2011) and “the best update in the years” (Softpedia 2013). However, on September 24th 2013, English Wikipedia reverted to using Wikitext as the default editing tool and kept visual editor as the “opt-in” feature

[20] Please refer to https://en.wikipedia.org/wiki/Wikipedia:VisualEditor for more information on visual editor.

instead after observing the significant drop of senior users' participation and receiving the widespread complaints from users. [21]

To study our research questions, we collect our data set from the English Wikipedia (http://en.wikipedia.org). This data set contains the users' complete editing history data with textual information. To investigate how the launch of visual editor affects contribution level of senior users, we pick the set of senior users who are active at English Wikipedia before the launch of visual editor as our sample. The visual editor at English Wikipedia is made as the default editing tool on July 15th 2013 and this is reverted on September 24th 2013. The time period for visual editor to be the default editing tool is around 2.5 months. Therefore, we set the time window of 2.5 months and focus on the user contribution in the time period 2.5 months before and after the launch date of visual editor (from May 1st 2013 to September 23rd 2013). To construct our sample, we identify 5,191 senior users who are active before the policy change made in July 2013. The sample of active users (also called "very active wikipedians" at English Wikipedia) includes those who contribute 100 times or more in May and June 2013. The threshold of 100 is determined by Wikipedia Foundation.[22]

At Wikipedia, users are identified by the user ID if they are registered. For unregistered users, they are identified by the IP address when they are connected to Wikipedia. Because the same user can connect to Wikipedia with different IP addresses at different time and the same IP address can be mapped to different users,

---

[21] Interesting readers can refer to https://www.mediawiki.org/wiki/VisualEditor for more information on the roll-out timeline of visual editor at Wikipedia.

[22] Please refer to https://stats.wikimedia.org/EN/TablesWikipediansEditsGt100.htm for more details.

we only focus on those registered users in this paper. We also filter out administrators and "bots" (the automated or semi-automated robot at Wikipedia) since they may exhibit different contribution pattern from normal users. Therefore, the 5,191 subjects included in our sample are registered users without administrator privileges.

For each contribution submitted by each user, we have the submission ID, the page ID of the Wikipedia article the contribution is submitted to, the user ID, the timestamp of the submission, the number of characters added or deleted in this submission and the hash-coded textual information of this submission. Our data includes all observations of detailed submission information from May 1st 2013 to September 23rd 2013.

We use this completed editing history data to generate three dependent variables for our econometric analysis: the number of times a user submitted his contribution, the number of characters added and the numbers of characters deleted. We also follow literature to include tenure and the squared term of tenure as control variable.

We use Table 3-1 to present the definition of variables and Table 3-2 to summarize the descriptive statistics of variables.

Table 3-1 Definition of Variables

	Variable Name	Definition
Dependent Variables	$Submission_{it}$	Natural logarithm of the number of times a user i submits his contribution to Wikipedia articles on day t.
	$Addition_{it}$	Natural logarithm of the number of characters a user i added to Wikipedia articles on day t.
	$Deletion_{it}$	Natural logarithm of the number of characters a user i deleted on Wikipedia articles on day t.
Independent Variables	$After_t$	A dummy variable to indicate whether the time period is after the launch of visual editor.
	$Tenure_{it}$	The number of weeks since a user i has joined Wikipedia on day t.
	$PercentageAnonymous_i$	The daily average of the percentage of anonymous contribution (measured by the total number of characters added and deleted) to the Wikipedia articles edited by the user i after the launch of visual editor.

Table 3-2 Descriptive Statistics of Active Users at English Wikipedia

Variable Name	Mean	S.D.	Min	Max
$Submission_{it}$	1.248	1.332	0.000	5.497
$Addition_{it}$	3.689	3.791	0.000	14.711
$Deletion_{it}$	2.165	3.001	0.000	14.779
$After_t$	0.479	0.500	0	1
$Tenure_{it}$	229.230	147.608	0	645
$PercentageAnonymous_i$	0.050	0.066	0.000	0.994
Observations	708,591 observations from 5,191 users			

3.4. Empirical Analysis

To empirically examine our research question, we first establish the causality between the launch of visual editor and senior users' participation with the data from Wikipedia. In the context of English Wikipedia, the impact of an enlarged user base is captured by the launch of visual editor because the change of editor suddenly attracted a large number of new users. We estimate the fixed-effect panel regression model with different time windows to examine both the short-term and long-term change of senior users' contribution after the change of editor. We then proceed to construct a control group for the active users at English Wikipedia with the similar users at German Wikipedia and estimate the Difference-in-Differences model. Finally, to further verify it is the negative network effect of the new users who generated low-quality contribution that causes the decrease of senior users' contribution, we rely on the exogenous shock on the quality level of anonymous contribution as a result of the launch of visual editor. We estimate the Difference-in-Differences model with the treatment being the average intensity a senior user is influenced by the unqualified contribution of anonymous users.

3.4.1. The Behavior Change of User Contribution

We now examine the behavior change of user contribution with the fixed-effect panel regression model (Wooldridge 2010):

$$Contribution_{it} = \beta_0 + \beta_1 After_t + Controls_{it} + \alpha_i + \varepsilon_{it}, (1)$$

where i indexes the users and t indexes the days.

The dependent variable $Contribution_{it}$ is the daily contributions of active users to Wikipedia articles. We use three different measures to capture a user's contribution: the number of submissions, the number of characters added, and the number of characters deleted. The numbers of characters added and deleted are considered differently because the editing efforts involved in these two forms of contribution are different (Zhang and Zhu 2011). To account for the skewness of the distribution of users' contribution, we take the logarithmic form of users' contribution as our dependent variables. $After_t$ is a dummy variable which equals 1 if the time period is after the launch of Visual editor, and 0 otherwise. As for control variables, we include tenure and the squared term of tenure, as suggested by the literature (Zhang and Zhu 2011). Tenure is measured by the number of weeks since a user has joined the online community of Wikipedia. The squared term of tenure is also included to capture the potential non-linear effect. Finally, the fixed effect α_i captures time-invariant unobserved player-specific effects, and ε_{it} is the residual error term.

We examined the behavior change of user contribution by applying different time windows. The visual editor is launched at English Wikipedia on July 15th 2013. To investigate the short-term behavior change, we estimated the panel regression with a time window of 2 weeks. In other words, we include those observations 14 days before and after the launch date of Visual editor at English Wikipedia (from July 1st 2013 to July 29th 2013). Since the launch of Visual editor is reverted on September 23rd 2013, we have observations of around 2.5 months after the launch of Visual editor. To investigate the long-term behavior change, we estimated the panel regression with a time window of 2.5 months. In investigation of the long-term behavior change, we include the observations from May 1st 2013 to September 23rd 2013.

Table 3-3 summarized the regression results. The estimates with a time window of 2 weeks are summarized in column (1) to (3) while the estimates with a time window of 2.5 months are summarized in column (4) to (6). The estimated coefficients of tenure are all negative across all models, indicating a user's contribution decreases as he continually participates in the online community of Wikipedia. As for the estimated coefficient of the squared term of tenure, we find that it is not significant in the results estimated with the time window of 2 weeks but is significantly positive in the results estimated with the time window of 2.5 months. This suggests that users' contribution decreases at a decreasing rate with tenure in a long run.

In Model (1) and (4), we use the number of submission as dependent variable. We find that users significantly submitted more in a short run but significantly submitted less in a long run. Since the dependent variable is in the form of natural logarithm, the estimated coefficient should be interpreted as semi-elasticity. For the short-term behavior change, users submitted 5.2% more in the 2 weeks after the launch of Visual editor, in comparison to his own submission in the 2 weeks before the launch. Similarly, for the long-term behaviour change, users submitted 7.7% less in the 2.5 months after the launch of visual editor, in comparison to his own submission in the 2.5 months before the launch.

Intuitively speaking, senior users' contribution should increase after Wikipedia replaced the inconvenient markup language Wikitext with the convenient visual editor as the default editing tool: the cost for users to edit drops significantly and thus it induce senior users to contribute more at Wikipedia. This positive effect consists of two components. First, the simplified and more convenient editing tool helps senior users to edit Wikipedia articles more easily than ever. Second, the reduced cost to edit Wikipedia articles will attract more new users to contribute. Due to the positive network effect, the more new users participate in Wikipedia, the more senior users will contribute. The latter component of positive effect is the motivation for Wikipedia to introduce visual editor (The Economist 2011). However, the drop of contribution cost may also have the negative side effect: it will induce low quality contribution to Wikipedia articles by unqualified new users. In the past, a new user must learn the complex and tedious markup language Wikitext before he obtains the capability to edit the Wikipedia articles. Hence, only those new users who are more willing to improve the content quality of Wikipedia articles will incur the learning cost to master Wikitext and contribute their knowledge. After Wikipedia adopted visual editor, suddenly everyone including those who want to make casual changes for fun can edit Wikipedia articles. In this case, senior users have to spend more effort to maintain the content quality by revising the Wikipedia articles continually. Senior users may become tired after many revisions or reversion and become less motivated to contribute their knowledge. The net effect of visual editor on user contribution depends on the dominant effect between the positive effect and the negative effect, leading to different behavior patterns of user contribution in a short run and in a long run. If the positive effect dominates the negative effect, senior users' contribution will increase in both a short run and a long run. On the other hand, if the negative effect dominates the positive effect, senior users' contribution will increase in a short run but decrease in a long run. Our results with the dependent variable as the number of submission reveal that the negative effect is the dominant effect.

We observe the same behaviour pattern in the results with dependent variables as the number of characters added and deleted. The addition and deletion, different measures to capture the amount of content a user contributed and his incurring effort, significantly increased in the short run but significantly decreased in the long run. On average, a senior user submitted 14.9% more in addition in the short run but submitted 22.9% less in addition in the long run. Similarly, a senior user on average submitted 9.9% more in deletion in the short run but submitted 13.9% in deletion in the long run. These consistent results help us to rule out an alternative explanation for the decrease of the number of senior users' submission in the long run: after the introduction of the visual editor, it is possible that a senior user could utilize the user-friendly interface and edit more content in one submission other than submitting the same amount of content multiple times. However, we find that senior users not only contribute less in terms of the number of times they submit their contribution (measured by submission), but also contribute less in terms of the content they contribute (measured by addition and deletion). Our results also suggest that addition, the form of contribution relatively involving more time and effort, is more affected by the introduction of visual editor than deletion.

Table 3-3 Results of Fixed-Effect Panel Regression Model

Specification	Short-term Behavior Change			Long-term Behavior Change		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Submission	Addition	Deletion	Submission	Addition	Deletion
After	0.052***	0.149***	0.099***	-0.077***	-0.224***	-0.139***
	(0.009)	(0.028)	(0.023)	(0.005)	(0.015)	(0.012)
Tenure	-0.047***	-0.131***	-0.092***	-0.022***	-0.060***	-0.033***
	(0.005)	(0.015)	(0.012)	(0.001)	(0.002)	(0.001)
$Tenure^2$	-0.000	-0.000	-0.000	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	12.295***	34.772***	23.178***	5.348***	14.981***	8.936***
	(0.886)	(2.632)	(2.162)	(0.098)	(0.284)	(0.227)
Observations	141,357	141,357	141,357	708,591	708,591	708,591
R-squared	0.002	0.002	0.001	0.015	0.014	0.008
Number of id	5,191	5,191	5,191	5,191	5,191	5,191

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

We also conduct sensitivity analysis to test the robustness of our results regarding different time window sizes. We tried three different time window sizes: 1 week, 1 month, and 1.5 months. We find that senior users contribute significantly more in a short run (using 1-week and 1-month samples) but contribute significantly less in a long run (using 1.5-months sample). These results are consistent with the findings we mentioned above.

The results of this sensitivity analysis are summarized in Table 3-4. The estimates with a time window of 1 week are summarized in column (1) to (3) while the estimates with a time window of 1.5 months are summarized in column (4) to (6). The estimates with a time window of 1 month are quantitatively the same to that with a time window of 1 week. Since it is more convincing to examine the short-term behavior using the 1-week sample, we choose to report the estimates with 1-week sample.

Table 3-4 Results of Sensitivity Analysis Regarding Time Window Size

Specification	Time Window of 1 Week			Time Window of 1.5 Months		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Submission	Addition	Deletion	Submission	Addition	Deletion
After	0.058***	0.161***	0.109***	-0.022***	-0.051***	-0.039***
	(0.011)	(0.032)	(0.027)	(0.006)	(0.018)	(0.014)
Tenure	-0.055***	-0.145***	-0.117***	-0.036***	-0.103***	-0.058***
	(0.011)	(0.033)	(0.027)	(0.001)	(0.003)	(0.002)
$Tenure^2$	-0.000	-0.000**	-0.000	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	15.632***	44.335***	31.252***	7.766***	22.539***	13.359***
	(1.880)	(5.637)	(4.660)	(0.190)	(0.555)	(0.447)
Observations	73,105	73,105	73,105	443,998	443,998	443,998
R-squared	0.001	0.001	0.001	0.011	0.010	0.006
Number of id	5,191	5,191	5,191	5,191	5,191	5,191

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

3.4.2. Identification of the Impact of Visual Editor

Our results of the fixed-effect panel regression model show that senior users' contribution decreased significantly in a long run after the launch of visual editor. Nevertheless, this fixed-effect panel regression model compares the user behavior before and after the launch of visual editor. One potential concern is that the significant drop of senior users' contribution may be attributed to the effect of time trend. That is, a senior user tends to contribute less as time goes by. Including tenure and its squared term in regression to some extent solves this problem. To establish a causal relationship that the launch of visual editor causes the drop of senior users' contribution, we construct a control group from the pool of users at German Wikipedia and estimate a Difference-in-Differences model.

We pick users at German Wikipedia as control group for two reasons. First and foremost, German Wikipedia did not introduce the visual editor.[23] Second, German Wikipedia is comparable with English Wikipedia, in terms of size. According to the multilingual statistics of Wikipedia, German Wikipedia is the second largest Wikipedia next to English Wikipedia (the largest Wikipedia), ordering by the number of articles.[24] In this paper, we include the active senior users (those who contribute 100 times or more in May and June 2013) at German Wikipedia as our control group. In our extended sample, there are 5191 users in the treatment group and 1611 users in the control group with observations from May 1st 2013 to September 23rd 2013.

[23] https://www.mediawiki.org/wiki/VisualEditor (accessed August 2015).

[24] https://en.wikipedia.org/wiki/Wikipedia:Multilingual_statistics (accessed August 2015).

After constructing the control group for users at English Wikipedia, we estimate the following Difference-in-Differences (DID) econometric model (Wooldridge 2010) with the observations from May 1st 2013 to September 23rd 2013 to quantify the long-term impact of visual editor on user contribution:

$$Contribution_{it} = \beta_0 + \beta_1 After_t + \beta_2 Treated_i \times After_t + \beta_3 Treated_i + Controls_{it} + \alpha_i + \varepsilon_{it}, (2)$$

where $Treated_i$ is the treatment dummy variable which equals 1 for English Wikipedia users and 0 for German Wikipedia users. Meanwhile, $Treated_i \times After_t$ is the interaction term of $Treated_i$ and $After_t$ and captures the treatment effect in the DID model. The DID model is the classical econometric model to estimate the treatment effect and establish causal relationship and therefore is widely applied in IS researches (Chan and Ghose 2013; Ghose et al. 2014).

We summarized the results of DID estimation in Table 3-5. The estimated results of DID model with dependent variables as submission, addition, and deletion, are presented in columns (1), (2), and (3), respectively. The results of DID model are consistent with those of the baseline fixed-effect panel regression model.

As reported in Table 3-5, the estimated coefficients of the interaction term capturing the treatment effect are significantly negative across all models with different measures of user contribution, suggesting senior users' contribution drops significantly after the launch of visual editor at English Wikipedia. More specifically, senior users' contribution decreased by 2.8%, 3.4%, and 11.2% in the number of submission, the number of characters added, and the number of characters deleted, respectively.

Table 3-5 Results of Difference-in-Differences Model

VARIABLES	(1) Submission	(2) Addition	(3) Deletion
After	-0.064***	-0.221***	-0.062***
	(0.006)	(0.017)	(0.014)
Treated × After	-0.028***	-0.034**	-0.112***
	(0.005)	(0.015)	(0.012)
Tenure	-0.020***	-0.057***	-0.030***
	(0.000)	(0.001)	(0.001)
$Tenure^2$	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Constant	5.151***	14.718***	8.309***
	(0.088)	(0.258)	(0.204)
Observations	936,300	936,300	936,300
R-squared	0.014	0.013	0.007
Number of id	6,802	6,802	6,802

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

3.4.3. Identification of the Negative Network Effect

We have shown the causal relationship between the launch of visual editor and the decrease of senior users' contribution at English Wikipedia by the Difference-in-Differences model. However, the decrease of senior users' contribution can be attributed to other alternative explanations relevant to visual editor other than the negative network effect of the new users who generated low-quality contribution. For instance, even though Wikipedia anticipate the new visual editor is user-friendly and designed to induce more input from contributors, contributors may actually find the visual editor is horrible because of the immature nature of visual editor. In other words, there is a gap between the anticipation of Wikipedia and the actual user experience of visual editor. Furthermore, assuming that the visual editor is indeed more convenient for senior users to edit Wikipedia articles, senior users need to incur a certain cost in the switching of editing tools from the original markup language Wikitext to the new visual editor. Some senior users who mastered the Wikitext may be not willing to incur the switching cost and therefore contribute less.

The English Wikipedia offers us an opportunity of natural experiment to identify the negative network effect. We use the anonymous contribution to capture this negative network effect because low-quality contribution of new users is more likely to be submitted anonymously. It is plausible to assume that the contribution by anonymous users on average is of lower quality than that by senior users because (1) anonymous or new users may not have the same level of expertise as senior users to provide added value of the content of Wikipedia articles; (2) anonymous or new users may not have the same level of experience as senior user to maintain the organization and readability of Wikipedia articles; and (3) anonymous users may include those who edit the Wikipedia articles in a malicious manner that is intentionally disruptive.

In Table 3-6, we present the results in the report of Wikipedia article rating (Wikimedia Foundation 2010) to justify our assumption. In the project of article rating, Wikipedia asked both anonymous users and registered users to rate the same set of Wikipedia articles in the period from September 22nd 2010 to October 4th 2010. The quality of a Wikipedia article is rated in four dimensions: well-sourced, neutral, complete, and readable. As shown in Table 3-6, anonymous users are easier to give high rating than registered users, across four different quality measures. This indicates anonymous users on average has a lower standard than registered users. We also find that the standard deviation of the rating of anonymous users are larger than that of registered users, across four different quality measures. This suggests there is a wider dispersion of quality among anonymous users than among registered users. Based on these descriptive statistics, we think it is reasonable to assume that the contribution by anonymous users on average is of lower quality than that by senior users since all senior users are registered users. Hence, senior users need to spend additional efforts to maintain the quality level of Wikipedia articles by revising anonymous contribution, in comparison to their own contribution.

Table 3-6 Report of Wikipedia Article Rating

	User Type	Well-sourced	Neutral	Complete	Readable
Mean	Anonymous	3.8	3.7	3.6	3.9
	Registered	2.5	3.3	2.4	3.1
S.D	Anonymous	1.45	1.48	1.48	1.4
	Registered	1.32	1.15	1.23	1.13
Num. of Rating	Anonymous	2086	1967	2013	2041
	Registered	540	517	537	531

Our identification of the negative network effect relies on the exogenous decrease of the quality level of anonymous users' contribution. Before the launch of visual editor, the cost of learning markup language Wikitext serve as a bar to prevent low-quality anonymous users as well as associated low-quality anonymous contribution. However, the launch of visual editor suddenly lower the bar and let these low-quality anonymous users freely edit the Wikipedia article, leading to a drop of the quality level of anonymous contribution. If the decrease of senior users' contribution is really driven by the negative network externality of unqualified contribution, we expect that, after the launch of visual editor, senior users who are exposed more to the anonymous contribution will have higher possibility to contribute less. Hence, we estimate the following Difference-in-Differences model:

$$\begin{aligned} Contribution_{it} &= \beta_0 + \beta_1 After_t + \beta_2 PercentageAnonymous_i \times After_t \\ &+ \beta_3 PercentageAnonymous_i + Controls_{it} + \alpha_i + \varepsilon_{it}, (3) \end{aligned}$$

where $PercentageAnonymous_i$ is the daily average of the percentage of anonymous contribution (measured by the total number of characters added and deleted) to the Wikipedia articles edited by the user i after the launch of visual editor. Thus, $PercentageAnonymous_i$ captures the daily average intensity a user is influenced by the unqualified contribution. $PercentageAnonymous_i \times After_t$ is the interaction term of $PercentageAnonymous_i$ and $After_t$ and captures the effect of unqualified contribution after the launch of visual editor.

Table 3-7 summarized the results of DID estimates of the effect of visual editor on user contribution with different percentage of anonymous contribution. Columns (1), (2), and (3) present the results estimated by Ordinary Least Square (OLS) estimator while columns (4), (5), and (6) show the results estimated Fixed-Effect (FE) estimator.

The estimated coefficient of $PercentageAnonymous_i$ is significantly positive across Models (1), (2), and (3), indicating users who edit Wikipedia articles with more anonymous contribution contribute more in terms of submission, addition, and deletion.

The percentage of anonymous contribution could also capture a user's preference or expertise in contribution to Wikipedia articles. For example, if a contributor consistently works on Wikipedia articles related to trivial topics like current affairs that everyone can easily understand and contribute, he need to cooperate with a large number of contributors including anonymous contributors. On the contrast, if he prefers to work on Wikipedia articles related to relatively advanced topics such as econometrics that only someone who has the domain knowledge can contribute while the crowd is less interested in modification, he is likely to collaborate with a small number of contributors. Therefore, the significantly positive coefficient of $PercentageAnonymous_i$ suggests that users who edit popular Wikipedia articles with more unknown collaborators would contribute more. This is the evidence of positive network effect which is documented in Zhang and Zhu (2011).

Interestingly, the estimated coefficient of the interaction term is significantly negative across Models (1), (2), and (3), suggesting that users edit Wikipedia articles with more anonymous contribution contribute significantly less AFTER the launch of visual editor. Our results showed the evidence of negative network effect.

We conclude that it is the quality dimension of new users' contribution that determines whether the network effect is positive or negative in the context of Wikipedia. Before the launch of visual editor, senior users enjoy the positive network effect. If a senior user perceives a large number of unknown collaborators are constructively contributing to the Wikipedia articles he is editing, he feels more motivated to contribute. After this launch, the average quality of unknown collaborators' contribution drops and senior users suffer from the negative network effect. If a user experiences a large number of unqualified contributions from unknown collaborators, he has less motivation to contribute.

Our results are also related to the notable finding of "conditional cooperator" in the field of behavioral economics. In the literature of private provision of public good, the theoretical prediction is the free-riding behavior. However, the analytical models assuming the rationality of agents fail to explain the phenomenon of extensive donations in the charitable sector of the economy (Andreoni 1988). To adjust the inconsistency between the theory and reality, as one of the promising remedies, analytical models incorporating the heterogeneity of social preference are proposed. These models give rise to the "conditional cooperators", whose contribution to the public goods is positively correlated either with his ex-ante belief about the contribution made by his peers (Chaudhuri 2011). Field experiments to test the theory of conditional cooperator are conventionally conducted in the context of public donation (Frey and Meier 2004; Heldt 2005; Shang and Croson 2009). In this paper, we reveal the phenomenon of conditional cooperators in the new context of online community. Our results are consistent with the theory of conditional cooperators, suggesting a senior user's contribution is conditional on his belief about the contribution quality made by other users. Moreover, we show that a subject's belief about the contribution quality made by others is more important than his belief about the contribution quantity made by others, at least in the case of English Wikipedia.

In Table 3-7, column (4), (5), and (6), we present the results estimated by fixed-effect estimators. The estimated coefficient of the interaction term capturing the treatment effect is qualitatively unchanged. Our estimated results showed that, comparing to the contribution of those editing Wikipedia articles without anonymous contribution, the contribution of users editing Wikipedia articles with anonymous contribution could decrease by up to 22.5%, 62.9%, and 35.7%, in submission, addition, and deletion, respectively.[25]

[25] These numbers are interpreted as upper bound because the percentage of anonymous contribution to the Wikipedia articles edited by the target user cannot be 1.

Table 3-7 DID Estimates with Percentages of Anonymous Contribution

Model Specification	OLS			FE		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Submission	Addition	Deletion	Submission	Addition	Deletion
After	-0.228***	-0.635***	-0.404***	-0.067***	-0.191***	-0.123***
	(0.003)	(0.010)	(0.008)	(0.006)	(0.016)	(0.013)
PercentageAnonymous	-0.217***	-0.602***	-0.346***	-0.225***	-0.629***	-0.357***
× After	(0.040)	(0.115)	(0.092)	(0.040)	(0.115)	(0.092)
PercentageAnonymous	0.938***	1.686***	5.166***			
	(0.169)	(0.459)	(0.358)			
Tenure	-0.000	0.001	0.002***	-0.020***	-0.056***	-0.031***
	(0.000)	(0.001)	(0.000)	(0.001)	(0.002)	(0.001)
$Tenure^2$	0.000***	0.000***	-0.000	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	1.253***	3.625***	1.762***	5.192***	14.560***	8.671***
	(0.026)	(0.071)	(0.056)	(0.102)	(0.298)	(0.239)
Observations	708,591	708,591	708,591	708,591	708,591	708,591
R-squared	0.012	0.010	0.006	0.014	0.012	0.008
Number of id	5,191	5,191	5,191	5,191	5,191	5,191

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

3.5. Concluding Remarks

The business value of an online community is rooted in the number of active users in this online community. An intuitive idea for online communities to grow their business value is to acquire more users. However, the policy changes made to acquire new users as well as the group of new users may have a negative impact on existing users. This can be illustrated by the case of the launch of visual editor at English Wikipedia: increased participation of new users turns out to drive seniors away. In this paper, we use econometric analysis to investigate the underlying mechanism.

We first employed the fixed-effect panel regression model to investigate the behavior change of senior users after the launch of visual editor. We estimated the fixed-effect model with different time windows to examine the behavior change of senior users in the short run and in the long run. We find that, after the launch of visual editor, senior users' contribution significantly increased in the short run but significantly decreased in the long run, in terms of the number of submissions, the number of characters added, and the number of characters deleted.

To establish the causal relationship between the launch of visual editor and the drop of senior users' contribution, we construct a control group for senior users at English Wikipedia with the similar senior users at German Wikipedia and estimate a difference-in-differences model. The estimates of difference-in-differences model are consistent with those estimated by the baseline fixed-effect panel regression model. Our results suggest that the policy change related to visual editor indeed caused the decrease of senior users' contribution.

To verify whether the negative network effect of low-quality contribution is the underlying mechanism, we estimated the DID model with the treatment being the percentage of anonymous contribution a senior user experienced after the launch of visual editor. We find that senior users who edit Wikipedia articles with more anonymous contribution contribute more before the launch of visual editor. However, after the policy change that lowered the bar for anonymous contribution, senior users become less motivated if they experience more anonymous contribution. We conclude that the quality of user contribution plays an important role in determining the direction of network effect in the context of online community.

Our study also has limitations and provides research opportunities for future research. First, in this paper, our findings suggest the quality of user contribution determines the direction of network effect and Wikipedia can achieve a balance between the quantity and the quality of user contribution by the learning cost for new users to participate. Our findings in the context of English Wikipedia can be naturally generalized to other online communities such as Q&A website. However, the optimal learning cost for different online communities may depend on the classification of online communities. Therefore, it is interesting to explore how to design the entry barrier based on the differentiation and competition of online communities. Furthermore, online community such as Wikipedia is a special type of public goods (Samuelson 1954), it would be interesting to explore whether our findings can be generalized to other types of public goods.

4. Study 3: The Monetary Value of Twitter Followers

4.1. Introduction

The inception of social media has greatly changed the way we communicate with each other and consequently our daily life has been reshaped. We observe that an increasingly large number of new social network sites, such as Twitter and Facebook, are emerging to meet different kinds of humans' needs. According to web analytics site Statisticbrain.com, as of 2014, there are over 645 million active registered users on Twitter and more than 1.3 billion monthly active users on Facebook. The boom of social media platforms brings firms new marketing opportunities. As an on-going transformation, firms start learning to reach a vast number of audiences and directly interact with their potential consumer on social media. It is reported that 69% of small business owners are engaged in some kinds of social media platform (e.g., Twitter, Facebook, and LinkedIn) and about 78% of them plan to allocate more budgets on social media marketing (Protalinski 2011). The expenditure on social media marketing in USA is expected to grow 34% yearly and reach 3.1 billion USD in 2014 (Forrester Research 2009).

An inspiring example of social media marketing can be illustrated by the successful story of Radio Shack, an electronic retailer in USA (Slutsky 2011). By posting a trending topic using the hash tag "#ineedanewphone" on Twitter, Radio Shack quickly gained 65 million impression counts for this marketing tweet within 24 hours, ending up with a double-digit sales increase in the three days that followed the promoted trend. "The ROI on this social-media initiative was stratospheric for us," The CMO of Radio Shack said.

Similar to firms, individuals also benefit from the emerging self-promotion avenue. Take professional athletes for another example. Athletes' popularity is one of the key factors in determining their court time and salary. However, before the introduction of Twitter, only superstars and some middle-class players had access to mainstream media as the marketing channel for promoting their image. The majority of ordinary players have few chances to reach out to fans to improve their popularity. Thanks to social media, there is a new avenue for ordinary athletes to manage fan loyalty: nowadays they can interact with fans on social media and cultivate a larger fan base by posting entertaining contents on Twitter or Facebook. For example, National Basketball Association (NBA) is a business valued over 19 billion USD. 75% of NBA players have a Twitter account with 214,539 Twitter followers on average.

Despite abundant examples of social media marketing, there is little scholarly research supporting the business value of social media. To fill this gap, we strive to prove that celebrities' participation and popularity on Twitter can be translated into real-world monetary outcomes, thus providing evidence of economic return to their social media marketing efforts. In this study, we examined the economic value of celebrities' participation and popularity in social media in the context of NBA. We choose NBA as our research context because the sample of NBA players is well-defined, in comparison to other celebrities such as actors and musicians. In addition, NBA players' performance and income can be measured more objectively than other professions and their demographics are also publicly available. This provides us an adequate set of control variables in data analysis. Our research questions are thus:

1. Do NBA players earn more salary after they participate in Twitter?
2. Do NBA players earn more salary if they have more Twitter followers?

Popularity is an important asset for celebrities (e.g. professional athletes, musicians, and movie stars) whose income (business value) is highly dependent on their fan base. That's why we observe celebrities such as NBA players are eager to create a Twitter account and conduct self-promotion by communicating with Twitter followers to cultivate a large fan base.

Furthermore, wage/salary inequality is a critical social issue and is continuously increasing since the 1980s. In the 2014-15 season of NBA, the highest-paid player (Kobe Bryant) earns 23.5 million USD while the lowest-paid player (Orlando Sanchez) earns only 15,000 USD, with the average salary in NBA being 4,203,105 USD. Social media can potentially help not so well-known athletes to reach out to their fans and gain increased exposure. However, it is not easy to predict who can benefit more from the increased exposure on social media: the more famous ones or the relatively less famous ones? Therefore, we are also interested in how social media changes the income inequality.

3. Does players' participation in Twitter affects salary inequality in NBA?

Our data set is a yearly panel data of 539 NBA players who are in the rosters of NBA from 2005 to 2014 with their associated salary, performance statistics, demographics, the registration date of Twitter account, and historical number of Twitter followers. We first investigate the impact of a player's participation in Twitter on his salary using a difference-in-differences (DID) model. Our results suggest that participation in Twitter helps a player to increase his salary by 861,628 US dollars on average. As a robustness check, we performed full covariate matching by Coarsen Exact Matching (CEM) to account for potential sample selection issues. The results are consistent with the estimates in the DID model. In quantifying the economic value of popularity, which is measured by the number of Twitter followers, we estimate a Heckman two-stage model with matched sample to alleviate the self-selection into Twitter bias. We find that doubling the number of Twitter followers of an NBA player is correlated with an increase of 682,122 US dollars in his salary. In sum, this study presents rigor empirical evidence of the economic value of players' participation and popularity on Twitter.

To investigate the change of salary inequality caused by Twitter, this study employs quantile regression to compare the relative gain from Twitter for players at different quantiles of salary. Our analysis reveals an interesting polarization pattern: above-average players and below-average players benefited more than average players. We also notice that the return of social media efforts is largest for players at the lowest 10% quantile (i.e. bottom players). Moreover, considering that players at the lower quantiles (those at the 0.10 quantile) benefit much more from Twitter than those at higher quantiles and the number of players at lower quantiles is much larger than those at higher quantiles, we conclude income inequality in the NBA has been reduced due to the emerging of social media.

Our study differs from prior research on social media in following ways. (1) The contribution of this paper lies in demonstrating the economic value of microblogging platform. Most existing evidence relevant to the business value of social media is founded in Facebook and other social media sites while microblogging platform remains untouched and the business model of microblogging platform is still questionable. Our study provides compelling empirical evidence to support the business model of microblogging platform, thus contributing to the literature on the business value of social media in Information Systems. (2) This paper focuses on the added value that social media brings to individuals while extant research mainly focus on the benefits of social media to firms. In this paper, we rigorously quantify the monetary value of celebrities' participation and popularity in microblogging social media. Our study proves the effectiveness of individual users' self-promotion in social media. (3) To the best of our knowledge, our study presents the first tested societal value of social media, while existing research only emphasize the business value of social media. We revealed the fact that salary inequality among NBA players is decreased because of player's adoption of social media. Our study identifies a completely new source of the change of wage inequality and contributes to the literature on wage inequality in the field of labor economics. Our study also provides meaningful and practical suggestions to professional sports leagues and player associations which care players' welfare: they can encourage players to actively participate and engage in social media, especially for those bottom players.

4.2. Related Literature

Microblogging platform has its appealing features to attract more celebrities than other social media sites. For example, Twitter is reported to be a better social media platform than Facebook for celebrities who pursue visibility and popularity because of its openness (Motwani 2013). Twitter is an open platform in which everyone can view each other's tweets, giving the content generated by celebrities the boost of visibility they desire. In contrast, Facebook is relatively closed and designed for communication among a group of people with closer relationship and most conversations are private, thus limiting celebrities' self-promotion. Consequently, celebrities such as NBA players prefer to engage in Twitter instead of Facebook and we decide to investigate the impact of celebrities' social media participation and popularity in the context of microblogging platform instead of other social media sites such as Facebook.

To the best of our knowledge, there is no econometric analysis with modern identification strategies that examines the economic value of celebrities' participation and popularity in microblogging platforms although microblogging platforms such as Twitter have become an integral part in today's business landscape. We also attempt to study whether social media affects income inequality. This can extend the current understanding on the "polarization phenomenon of wage distribution" and the societal impact of information technology. Our research is thus relevant to three streams of prior studies: microblogging platform, the business value of social media, as well as wage inequality.

4.2.1. Microblogging Platform

Microblogging is a broadcast medium that similar to blogging. A microblog differs from a traditional blog in that its content is typically smaller (Kaplan and Haenlein 2011). Twitter is the most famous microblogging platform. On Twitter, users can share tweets (text messages up to 140 characters long) with their followers. Followers are those who subscribe to a users' timeline of tweets. This is like using Rich Site Summary (RSS) feed service to receive timely updates from favorite blogs. The following behavior on Twitter creates a directed social network in which there is a link from user A to user B if user A chooses to follow user B. With over 645 million active registered users, Twitter is already part of everyone's daily life in the mobile era.

The popularity of microblogging platform has attracted increasing attentions from academic research. Extant literature, mostly in Computer Science, focuses on studying the structure and nature of the Twitter's social network. For example, one popular topic is information diffusion in this network (Bakshy et al. 2011; Goel et al. 2012; Kwak et al. 2010; Nair et al. 2010; Romero et al. 2011; Weng et al. 2010; Wu et al. 2011). However, academic research on Twitter from social science is still scant (Toubia and Stephen 2013). There exist only few studies related to Twitter or microblogging platform published in top journals in Information Systems. The only exception we found is the research conducted by Ghose et al. (2012). In their study, Ghose et al. (2012) investigated the difference of Internet browsing behavior between mobile phones and personal computer users using data from a Twitter-like microblogging platform. They found that ranking effects are higher and preferences for geographically proximate brands are also higher on the mobile Internet. Although their study is conducted in the context on microblogging platform, their focuses are different from our research questions. In this paper, we aim to empirically investigate the economic value of NBA players' participation and popularity in Twitter.

4.2.2. The Business Value of Social Media

Our study is closely related to the IS literature on the business value of social media. Aral et al. (2013) proposed a framework for social media research. Research in this area can be characterized by the level of analysis: user and society, platform and intermediaries, and firms and industries. In their framework, social media's value and strategy describe how users, platforms, and firms create the value from using social media and how they can create strategies that best satisfy their needs. Aral et al. (2013) also pointed out that social media may have both a direct effect (improving the outcomes of decisions) and a strategic effect (changing the decisions). Our paper is relevant to research conducted at the firm level.

At the firm level, researchers investigated the outcome of social media after firms or industries adopted social media. These studies mainly focus on the direct effect of social media. For instance, Goh et al. (2013) demonstrated that firm's engagement in social media helps to promote sales and concluded that user generated content is more impactful than marketer generated content. Rishika et al. (2013) found that a firm' social media effort is effective to induce customer participation, which leads to an increase of customers' visit frequency and profitability. Luo et al. (2013) shown that firm equity value can be predicted by social media and social media metrics have substantially stronger predictive power than those conventional online behavioral metrics (Google searches and Web traffic). On the other hand, Wu (2013) revealed the strategic effect of social media: the introduction of social network tool not only improved employees' productivity (the outcome of decisions) but also changed employees' network position (changing the decisions) and found that social communication is more correlated with reduced risk of layoff than with information diversity. None of these studies are conducted in the context of microblogging platforms. In fact, we found that the effectiveness of firm's social media marketing in microblogging platforms has not yet been touched in extant literature, particularly in top IS journals.

Our study is conducted at the level of users. Very little research has been done to investigate the effectiveness of social media and how individuals can participate and engage in social media to meet their needs. To the best of our knowledge, there is no empirical research quantifying the value of popularity in microblogging platform like Twitter.

4.2.3. Wage Inequality

Our research is also related to the issue of wage inequality, which has been widely studied in economics. In a broad sense, the change of wage inequality boils down to the change in wage-setting institutions and skill biased technical changes.

Institutional factors such as unionization have been shown to affect wage inequality. Empirical research mainly focuses on the 1980s during which the increase of wage inequality in the United States is accelerated and economists are eager to find out the reasons driving such phenomenon. Freeman (1991) shown that de-unionization explains about 20 percent of the increase in the overall wage inequality for men during the 1980s. DiNardo et al. (1996) further found that a large proportion (around 40 percent) of the increase in the higher tail inequality for men in the 1980s can be attributed to the decline in unionization. However, due to the secular decline in unions, unionization does not has a significant impact on wage inequality later in the 1990s (Card et al. 2004).

The other line of research of wage inequality focuses on skill biased technical change (SBTC). The overall wage inequality has ceased growing since the late 1980s while the upper tail inequality rise as rapidly in the 1990s as in the 1980s (Autor et al. 2008). SBTC is identified by economists as the key source of upper tail inequality (Acemoglu 2002; Katz and Auctor 1999). Economists state that the wide application of computer helps college educated labor to establish increasing productivity advantages over those non-college educated labor, thus lead to a demand shift and employment change. One shortcoming for the traditional model in previous literature is that it assume there is a one-to-one match between skills (or education background) and tasks (or working activity) but the dynamics between skills, tasks and technologies is complex in reality. In recent years, the role of task or routine offers a new perspective to evaluate the impact of SBTC. Acemoglu (2011) proposed a task-based model which incorporates the endogenous assignment of skills to tasks and the substitution for certain tasks resulted from technology change. Similarly, Autor et al. (2006) and Autor and Dorn (2013) proposed the model of computerization as well as the hypothesis of routinization to characterize the pattern of "polarization" in the U.S labor market, in which employment polarizes into high-wage and low-wage jobs at the cost of middle-wage jobs. They argued that the middle-wage jobs are associated with high level of routine tasks and significantly replaced by automating routines from technology advances. On the other hand, technology advances like the wide application of information technology, complement abstract cognitive tasks, which are related to high-wage jobs, and have little impact on non-routine manual tasks, which are related to low-wage jobs.

Our study aims to investigate the impact of social media on wage inequality. To the best of our knowledge, the effect of social media has not been touched in existing literature on wage inequality. In our context, social media may result in a pattern of polarization in the wage distribution in NBA. In NBA, in addition to his performance, a player’s popularity is also a key factor in determining his salary. For instance, Ertug and Castellucci (2013) found that, while there is no significant effect of the average status of players on the teams’ season performance (measured by the qualification for the Playoffs as well as the advancement to Conference Semifinals, Conference Finals, and NBA Finals), there is a positive and significant effect of the average status of players on the teams’ revenue (measured by the teams’ ticket income). The popularity of a superstar (a player with high status), is an intangible asset to his team and help his team to increase revenue through ticket sales and merchandise. Therefore, an NBA team is willing to give a superstar player a higher salary than ordinary players not only for his superior performance but also for his high popularity. In their paper, Ertug and Castellucci (2013) also found that when an NBA team's revenue is low relative to its aspiration, this team will display a preference for recruiting high-popularity players than high-performance players.

Upon its emergence, social media may amplify the existing superstar effect. Superstar players enjoy their salary premium over other players because of widened reach to fans and increased popularity among fans. Superstar players with professional marketing team can conduct effective marketing activities on social media, thus benefiting more from social media than average players without marketing team. On the other hand, social media serves as a new marketing avenue for bottom players. In the past, superstars and some middle-class players can obtain promotion by league and sponsorship from brands while bottom players have few chances to improve their popularity. Today, all bottom players can have the equal opportunity as superstars and middle-class players to interact with their fans and establish a larger fan base by social media. Therefore, bottom players may also benefit more from social media than average players. If these two effects work concurrently, the overall outcome would be the polarization in the salary distribution in NBA.

4.3. Data Description

In this study, we conducted our research in the context of NBA. We choose NBA as our research context for several reasons. First, compared to other celebrities on Twitter such as politicians, actors and musicians, the sample (the set) of NBA athletes is well-defined and the sample size is large enough yet manageable. Otherwise, we may face the sampling issue about who should be included or excluded into our sample. Second, an NBA player's performance can be measured more objectively than other professions. All major professional sports leagues regularly publish a large number of performance metrics for fans to read over the Internet. Each NBA player's demographics are also publicly available. This provides us an adequate set of control variables in regressions. Third, we choose NBA among the four largest professional sports leagues in USA (NFL, MLB, NBA, and NHL) since the adoption rate of Twitter is the highest in NBA than the other three sports leagues. We found that many NBA players have opened a Twitter account and actively interact with their fans while only a few of NFL/MLB/NHL players have registered a Twitter account. As a consequence, NBA players have cultivated a larger fan base on Twitter even though NFL and MLB have larger offline fan bases than NBA. This can be illustrated by the number of followers of Most Valuable Player (MVP) in these four professional sports leagues: the MVP of NBA (LeBron James) has 12.6 million followers whereas the MVP of NHL (Alex Ovechkin) and MLB (Miguel Cabrera) have 0.766 and 0.437 million followers, respectively. In contrast, the MVP of NFL (Peyton Manning) does not have a Twitter account. Because we analyze the impact of Twitter by DID and Heckman selection model in a panel regression, it is better to analyze a sample in which most players, especially famous players, have adopted Twitter so that we can compare their salaries before and after the adoption of Twitter. Hence, we decide to choose NBA instead of other three sports leagues as our research context.

Our sample consists of 539 NBA players who are in the rosters of NBA from 2005 to 2014. We scraped their salary, performance statistics, and demographics from NBA's official website and ESPN.com to construct a yearly panel data set. To make sure the scraped data are accurate, we also scraped relevant data from Basketball-reference.com and check the consistency of data from different sources for triangulation. For players who have a Twitter account, we collected their social media information including the date they registered account on Twitter and their historical number of Twitter followers.

Unlike the information related to players' salary, performance, and demographics, players' social media information is difficult to collect. The date a player opened a Twitter account can be collected from his profile page on Twitter. However, some players (e.g., Kobe Bryant) do not want to disclose their registration date on his profile page. For such players, we cannot directly collect their registration date but need to rely on Twitter API (Application Programming Interface) and develop a program to fetch data from Twitter. The statistics of historical number of followers are even more difficult to collect. Because the historical follower data is critical for social media analysis and thus of high business value, Twitter do not allow researchers or developers to access the historical follower data by their designed API. We have to obtain the historical follower data from third-party data providers (e.g. Twittercounter.com and Gnip.com) that have granted data access privileges from Twitter. These data providers usually employ a freemium pricing model. For example, Twittercounter.com allows us to track the historical number followers of every Twitter account but it is only free for historical follower data recorded in the last three months. For historical follower data beyond the period of three months, we need to pay 29 USD for each account queried. To collect the historical follower data for NBA players in our sample, we developed a program to fetch data from Twittercounter.com and ran it regularly on a three-month basis. This technical difficulty to collect social media information partially explains why there exist few empirical studies conducted in the context of Twitter although microblogging platform is an important and interesting topic in the field of Information Systems.

4.3.1. Dependent variable

Our dependent variable is the annual salary of NBA players. In professional sports leagues, salary is the most important source of income for most players. Only a few superstars in the league can receive endorsement from brands. Even for those superstars, the ratio of endorsement to players' total income drops dramatically from the most famous one to less popular ones. We can observe this trend in the list of 10 NBA's highest-paid players (Badenhausen 2013). Kobe Bryant tops the list with his impressive income of 59.8 million USD dollars. Within his total income, 32 million dollars come from endorsement. On the other hand, in the bottom of this list, Pau Gasol earned 21.5 million USD dollars in income which largely driven by his playing salary of 19 million USD dollars. To account for the skewness of the distribution of NBA players' salaries, we take the logarithmic form of salary as our dependent variable.

4.3.2. Independent variables

Twitter Account

To indicate whether an NBA player has opened a Twitter account, we construct a dummy variable based on the registration date of the focal player's Twitter account. For players who own a Twitter account, this dummy variable equals 1 after (including) the year during which a player joined Twitter and 0 for previous years. For those who do not have a Twitter account, this dummy variable is 0 across all years. This variable is the main treatment variable in our DID analysis.

Twitter Followers

In literature, popularity is associated with the notion of acceptance (Bukowski and Hoza 1989; Coie et al. 1982; Newcomb and Bagwell 1995). On Twitter, we generally follow the ones we accept or appreciate so that we can be informed about everything related to the people we are interested in. In this sense, users who follow a celebrity on Twitter can be considered as fans of the focal celebrity. The number of followers of a celebrity is the direct measure of his fan base on Twitter and can be used to proxy his popularity on Twitter.

4.3.3. Control variables

We included a set of control variables suggested by literature (Ertug and Castellucci 2013; Massey and Thaler 2013): efficiency rating, team-selection honor, demographics , minutes on court, position dummy, to control for confounding factors that may affect an NBA player's salary.

Efficiency Rating

It is intuitive that a player's better performance can help his team to win more games and to improve the probability to win championship. Since winning percentage and championships are highly correlated with the ticket sales and the market value of teams, a player's salary is largely determined by his on-court performance (Massey and Thaler 2013). There are many performances metrics used in NBA. To alleviate the multi-collinearity issue among performance metrics, we need to pick one variable and the most useful predictor of performance is Player Efficiency Rating (PER) proposed by Hollinger (2005).[26] PER is a per-minute rating for a player's performance and is standardized by the league average in each year. Furthermore, to fully control for a player's performance, including PER of one single season is not sufficient since a player's performance may fluctuate from season to season. Instead, we follow the literature to use a weighted average of PER to control for a player's performance.[27]

Team-Selection Honor

In NBA, it is an honor for a player to be selected into the All-NBA teams (first/second/third teams). Only the few elite players in the league can be bestowed this honor. In each NBA season, the All-NBA teams are selected based on a voting conducted by a panel of sportswriters and broadcasters in United States and Canada. For each position, the player who receives the most votes is selected into the first team, the player who receives the second most votes is selected into the second team, and so on. Being selected into the All-NBA teams is an indicator for a player to be a star. Therefore, we include a player's team-selection honor to control for potential superstar effect. To allow for lagged effect for such honors, we followed the literature and construct a discrete variable to indicate whether the focal player has been selected into first/second/third team in the previous three seasons (Ertug and Castellucci 2013).

---

[26] John Hollinger is an NBA analyst since 1996 and is recognized as a leader in basketball's rising statistical analysis movement in the past decade. He was an analyst and writer for ESPN and is currently in the position of Vice President for NBA team Memphis Grizzlies.

[27] Interested readers can refer to Ertug, G., and Castellucci, F. 2013. "Getting What You Need: How Reputation and Status Affect Team Performance, Hiring, and Salaries in the Nba," *Academy of Management Journal* (56:2), pp. 407-431. for details on how to compute a weighted average of PER with a decay function.

Other Control Variables

A player's demographics (tenure) as well as game-related variables (minutes on courts and position dummies) are also included. In literature (Ertug and Castellucci 2013; Massey and Thaler 2013), researchers used to include tenure as control for experience. Tenure captures the impact of a player's experience since an NBA player can improve his professional skill with his increased experience. The squared-term of tenure is included to control for potential diminishing effect of experience and also the performance of players may degenerate as they become too old.

A player's minutes on courts is included to control for the extent to which he is utilized by his team and is a direct measure of the amount of the player's contribution to his team. Similar to efficiency rating, we use a weighted average of minutes on courts as a control variable. We also apply the logarithmic transformation to handle the skewness of distribution.

Last but not least, we include a set of position dummies. In a basketball game, each team on court is a five-man lineup, therefore leading to five different positions: Center (C), Power Forward (PF), Small Forward (SF), Shooting Guard (SG), and Point Guard (PG). Players in different positions bear different responsibilities and on average receive different amount of salary. Hence, there exist position-level heterogeniety. Also, some NBA players can play two positions. Therefore, we manually consolidate the position coding from different data sources (NBA, ESPN, and Basketball-reference.com), ending up with a set of 16 position dummies to control for players' position-level fixed-effects.

Table 4-1 presents the definition of variables and Table 4-2 summarizes the descriptive statistics of variables.

Table 4-1 Definition of Variables

	Variable Name	Definition
Dependent Variable	Log_salary	Natural logarithm of a player's salary
Independent Variables	Efficiency_rating	Average performance. A player's performance is measured by Player Efficiency Rating (PER).
	Team_selection	A flag to indicate whether the focal player has been selected as first/second/third team in the previous three seasons.
	Tenure	Number of years since a player has entered NBA
	Join	A dummy variable to indicate whether the focal player has registered a Twitter account.
	Log_minutes_played	The average number of minutes a player played on court
	Log_followers	Natural logarithm of the number of Twitter followers
	Age	A player's age

Table 4-2 Descriptive Statistics

Variable Name	Obs.	Mean	S.D.	Min	Max
Log_salary	2283	15.259	0.919	10.650	17.137
Efficiency_rating	2283	8.541	2.459	-6.569	18.929
Team_selection	2283	0.117	0.475	0	3
Tenure	2283	8.302	3.351	4	21
Join	2283	0.238	0.426	0	1
Log_minutes_played	2283	1.948	0.228	0.920	2.309
Log_followers	365	11.657	1.529	7.578	16.353
Age	2283	28.916	3.706	21	42

4.4. Empirical Research Design

4.4.1. The Impact of Twitter Account

We first present results of Difference-in-Differences model (Wooldridge 2010) to show how a player's participation in Twitter helps him to obtain a higher salary. We model the logarithm of player i's salary in year t as

$$\begin{aligned} logsalary_{it} = {} & \beta_1 efficiency\_rating_{it-1} + \beta_2 team\_selection_{it-1} \\ & + \beta_3 tenure_{it-1} + \beta_4 {tenure_{it-1}}^2 + \beta_5 join_{it-1} \\ & + \beta_6 log\_minutes\_played_{it-1} + \beta_7 position_{it-1} + \theta_t + \alpha_i \\ & + \varepsilon_{it}, \end{aligned} \tag{1}$$

where independent variables are in the previous time period (t-1) because the salary is determined before each NBA season in year t and depends on the performance in the previous year. The variable of interest $join_{it-1}$ is a treatment dummy variable which equals 1 after (including) the year during which a player joined Twitter and 0 for previous years. Meanwhile, $efficiency\_rating_{it-1}$ measures player i's performance and $team\_selection_{it-1}$ is a proxy for the potential super-star effect. We also include $tenure_{it-1}$ and ${tenure_{it-1}}^2$ to control for the influence of experience. To further control game-related factors, the minutes a player has played in the previous season and a set of position dummies are included. Finally, θ_t captures the effect of time trend, the fixed effect α_i captures time-invariant unobserved player-specific effects, and ε_{it} is the residual error term.

The DID model is the canonical econometric model to estimate the treatment effect and establish causal relationship (Chan and Ghose 2013; Ghose et al. 2014). It helps us to control for unobservable individual heterogeneity (e.g.

personality) and rule out alternative explanations stemming from unobservable individual heterogeneity.

To establish a stronger causal relationship, we further use matching as our robustness check to handle the potential sample-selection issue. One most important self-selection issue is that superstar players may be more willing to engage in social media to interact with their fans. To account for this issue, we apply matching to construct a control group. In this paper, we construct our control group as follows. For each subject in the treatment group, we find the player who does not join Twitter in the same year but with most similar attributes as the treated subject (i.e. one-to-one matching). We expect that a player's decision of joining Twitter is related to $efficiency\_rating$ and $team\_selection$ because a high-profiled player (which can be proxied by $efficiency\_rating$ and $team\_selection$) has more incentives to join Twitter. We also expect a player's decision to join Twitter is related to his age since an older player is less likely to be a tech-savvy user of any IT products, and thus having fewer interests in joining Twitter.

We use the state-of-the-art matching algorithm in the literature. We need to find a doppelganger in the control group for each treated subject with similar attributes under the constraint that matching must be conducted in the same season/year. The seminal paper about matching (Rosenbaum and Rubin 1983) applies semi-parametric matching techniques such as Propensity Score Matching (PSM). In this paper, we follow Azoulay et al. (2013) and Malter (2014) to apply Coarsen Exact Matching (Blackwell et al. 2009; Iacus et al. 2011), a faster yet accurate non-parametric matching (full covariate matching) approach to construct our control group. Coarsen Exact Matching (CEM) has several advantages over traditional PSM such as automatic restriction to common empirical support and guaranteed balance of covariates ex ante. The algorithm in CEM coarsen the joint distribution of covariates into a finite number of bins and then perform matching if and only if both treated subjects and control subjects can be found in the same bin, thus automatically satisfying the common empirical support.[28] Researchers can rely on the default binning algorithm or manually specify the number of bins used in matching, therefore ensuring the degree of balance of covariates ex ante.[29] CEM is also proved to outperform other matching approaches including PSM, Mahalanobis distance matching, and genetic matching in terms of bias, standard deviation, root mean square error, and computational speed (Iacus et al. 2012). With such advantages and conveniences to perform matching, CEM is getting increasingly recognized and adopted in recent publications in top journals (e.g. Azoulay et al. (2013) and Malter (2014)).

Next, with the matched player data sample, we rerun the DID model given in Equation (1) with the matched sample. As documented in literature, the combination of DID analysis and matching sample can significantly reduce the bias stemming from both observable and unobservable confounding factors and can enhance the consistency of estimates (Rishika et al. 2013; Stewart and Swaffield 2008).

---

[28] Common support requires overlap in the covariate distribution between treatment group and control group.

[29] The degree of balance of covariates is generally measured by the mean difference of covariates between treatment group and control group.

4.4.2. The Impact of Twitter Followers

We further proceed to quantify the economic value of a player's popularity on Twitter. This is equivalent to estimating the monetary value of the number of Twitter followers. The number of Twitter followers of a player can be observed only if he has joined Twitter. This can lead to biased estimate if a player's decision to join Twitter (i.e. treatment assignment) is not uncorrelated with the explanatory variables. To tackle the potential selection bias, we estimate a Heckman selection model (Heckman 1979) which is the most commonly employed econometric tool to correct selection bias when the dependent variable is salary and is widely applied in the field of Information Systems (Bapna et al. 2013; Gu and Ye 2014; Hui et al. 2013; Lin et al. 2013). The bias is corrected through a two-step procedure. In step 1, we estimate a Probit model and compute the selection bias correction term from the estimate. The dependent variable of selection equation is a latent variable $join_i^*$, which can be considered as the propensity to be assigned to treatment group. In step 2, the correction term is included in a regression model to obtain unbiased estimate of the effect of Twitter followers of a player on his salary. The specification of Heckman selection two-stage model is given as follows:

Selection equation

$$join_i^* = \gamma_1 efficiency\_rating_i + \gamma_2 team\_selection_i + \gamma_3 age_i + \alpha + \varepsilon_{1i} \quad (2)$$

$$join_i = \begin{cases} 1, & if\ join_i^* > 0 \\ 0, & otherwise \end{cases} \quad (3)$$

$$Prob(join_i = 1|z_i) = \Phi(z_i{}'\gamma), \quad (4)$$

where z_i is a vector of the set of covariates included in Equation (2).

Regression equation

$$
\begin{aligned}
log\_salary_i = \beta_1 efficiency\_rating_i &+ \beta_2 team\_selection_i + \beta_3 tenure_i \\
&+ \beta_4 tenure_i{}^2 + \beta_5 log\_follower_i + \beta_6 log\_minutes\_played_i \\
&+ \beta_7 position_i + \theta_t + \alpha_i \\
&+ \varepsilon_{2i}\,, \qquad (5)
\end{aligned}
$$

where $\varepsilon_1 \sim N(0,1)$, $\varepsilon_2 \sim N(0, \sigma)$, and $corr(\varepsilon_1, \varepsilon_2) = \rho$.

In the Heckman's first-stage model, as described above, a player's decision to join Twitter is expected to be related to $efficiency\_rating$, $team\_selection$, and age. In the second-stage model, a player's salary is regressed on the variable of interest ($log\_follower$) and other control variables.

In literature, both matching and Heckman selection model can be applied to alleviate the potential biases caused by sampling issues. Matching can tackle the sample selection on observable covariates to enhance the proof of causality. It has been proven to be an effective alternative to including those observable covariates as control variables in the regression. Parallelly, Heckman selection model can correct the bias when the unobservable subjects/samples exhibit systematically different patterns of the dependent or independent variables. In this paper, we employ Heckman selection model with matched sample to take advantages of both approaches (Goh et al. 2013), thus facilitating to establish stronger causal relationship. Furthermore, we also estimate the Heckman selection model with full sample to prove the robustness of our estimates.

4.4.3. The Impact of Twitter Account on Salary Inequality

The aforementioned empirical models aim to measure the average effects of Twitter account and followers on players' salaries. But we are also interested in the changes in the distribution of salary among NBA players. We would like to explore which kind of player (superstar player, average player, or bottom player) can benefit more from the participation in Twitter. In econometrics literature, quantile regression (Koenker 2005) is a powerful tool to model distribution even though the underline story is complex and multidimensional (Angrist and Pischke 2008). In this study, we employ quantile regression with matched sample to investigate whether players' participation in Twitter affects the salary inequality in NBA. The model specification is given as follows:

$$
\begin{aligned}
Q\left[log_{salary_{it}}\middle|x_{it-1}, q\right] \\
&= \beta_{q1} efficiency\_rating_{it-1} + \beta_{q2} team\_selection_{it-1} \\
&+ \beta_{q3} tenure_{it-1} + \beta_{q4} tenure_{it-1}{}^{2} + \beta_{q5} join_{it-1} \\
&+ \beta_{q6} log\_minutes\_played_{it-1} + \beta_{q7} position_{it-1} + \theta_t + \alpha_i \\
&+ \varepsilon_{it} \qquad (6)
\end{aligned}
$$

such that $Prob\left[logsalary_{it} \leq x_{it-1}{}'\beta_q \middle| x_{it-1}\right] = q,$ (7)

where is x_{it-1} is a vector of the set of covariates included in Equation (6) and $0 < q < 1$.

In quantile regression, no assumption is needed on the distribution of dependent variable conditional on covariates or about its conditional variance, thus making the nonparametric specification flexible (Greene 2011).

4.5. Model Estimation and Results

4.5.1. The Impact of Twitter Account

We summarized the results of DID estimates in Table 4-3. The estimates of a fixed-effect panel regression model with all control variables are reported in column (1).

The estimates of the DID model in Equation (1) are shown in column (2). Meanwhile, column (3) present results of DID model in combination with CEM using matched sample.

As reported in column (1), it is intuitive that a player's efficiency rating and minutes played on court in previous season have significantly positive impacts on his salary. In addition, a player's tenure has a significantly positive but diminished effect. The estimated coefficient of $team\_selection$ is positive but not significant. If we only include a player's efficiency rating and his team-selection honor, the estimated coefficients of $team\_selection$ is significantly positive. However, we found that the impact of $team\_selection$ become insignificant once we control for $efficiency\_rating$ and $tenure$. Specifically speaking, the estimated coefficient of $team\_selection$ become insignificant after we add the squared-term of $tenure$. Existing literature seems to ignore the diminishing effect of $tenure$.

As shown in column (2), the DID estimate of the impact of a player's decision to join Twitter is 0.144. Because our dependent variable (log_salary) is in the form of natural logarithm, the estimated coefficient should be interpreted as semi-elasticity: participation in Twitter on average helps a player to increase his salary by 14.4%. The mean of salary in our sample is 5,983,530 US dollars, thus suggesting participation in Twitter increases a player's salary by 861,628 US dollars on average.

The DID estimates with matched sample are presented in Table 4-3, column (3). In our data set, 102 subjects in the treatment group are matched one-to-one by CEM. The results of CEM are summarized in Table 4-4. The measures of matching performance include the mean difference of covariates between treatment and control group as well as the mean differences in different quantiles of covariates between treatment and control group. Take the variable of $efficiency\_rating$ for example. The mean difference of $efficiency\_rating$ between treatment and control group before matching is 0.6703 while that after matching is -0.00938. Similarly, the mean difference in the 0.95 quantile of $efficiency\_rating$ between treatment and control group before matching is 1.6422 while that after matching is -0.176. We also obtain good matching for $team\_selection$ and age. As indicated in Table 4-4, the mean differences of $team\_selection$ and age are both 0 after matching. In all, after matching by CEM, the imbalance of covariates between treatment and control group has been substantially reduced.

In DID model with matched sample, the estimated coefficients of $join$ is still positively significant. The magnitude of 0.128 is also close to the estimated coefficient (0.144) in the DID model in column (2). This robustness check confirms that a player's participation in Twitter indeed causes the increase of his salary. By comparing the R-squared for columns (2) and (3), we also notice that the R-squared increased dramatically when applying DID with the matched sample, suggesting matching did increase the explanatory power of DID model. Following the calculation above, we conclude that participation in Twitter on average increases a player's salary by 12.8%, which is equivalent to approximately 765,891 US dollars.

Table 4-3 Results of DID Estimates

VARIABLES	(1) FE Control	(2) DID Panel	(3) DID, CEM Panel
Join		0.144***	0.128**
		(0.052)	(0.063)
Efficiency_rating	0.064***	0.064***	0.046**
	(0.014)	(0.014)	(0.023)
Team_selection	0.014	0.010	0.128
	(0.045)	(0.045)	(0.090)
Tenure	0.254***	0.247***	0.334***
	(0.023)	(0.023)	(0.036)
$Tenure^2$	-0.014***	-0.014***	-0.019***
	(0.001)	(0.001)	(0.002)
Log_minutes_played	1.601***	1.587***	1.560***
	(0.135)	(0.135)	(0.214)
Position dummies	-included-	-included-	-included-
Time dummies	-included-	-included-	-included-
Constant	10.582***	10.643***	10.353***
	(0.255)	(0.255)	(0.403)
Observations	2,283	2,283	751
R-squared	0.299	0.302	0.419
Number of player_id	539	539	204

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4-4 Results of CEM

Variable		*Efficiency_rating*	*Team_selection*	*Age*
Mean difference	Before	.6703	.01482	-.59427
	After	-.00938	0	0
0.05 quantile difference	Before	6.8711	0	1
	After	0.0217	0	0
0.25 quantile difference	Before	0.6989	0	-1
	After	0.0494	0	0
0.5 quantile difference	Before	0.9444	0	-1
	After	-0.009	0	0
0.75 quantile difference	Before	0.4711	0	-1
	After	-0.102	0	0
0.95 quantile difference	Before	1.6422	0	-2
	After	-0.176	0	0

4.5.2. The Impact of Twitter Followers

The results of Heckman selection model are summarized in Table 4-5. The estimates of Heckman selection model with matched sample are presented in columns (1) and (2) where the results of second-stage model are shown in column (1) and results of first-stage model are shown in column (2). Similarly, the estimates of Heckman selection model with full sample are presented in columns (3) and (4).

As shown in Table 4-5, column (1), the estimated coefficient of $log\_follower$ is 0.114 and significantly positive. Since both dependent variable and independent variable are in logarithmic form, this estimated coefficient should be interpreted as elasticity: if a player's Twitter follower increases by 100%, his salary increases by 11.4%. That is, if a player can double his Twitter followers by active engagement in social media, he can achieve 11.4% increase in salary. Considering that the mean of salary in our sample is 5,983,530 US dollars, this 11.4% increase in salary can be translated to an average increase of 682,122 US dollars.

Furthermore, the parameter ρ is the correlation between the unobserved determinants of a player's propensity to join Twitter (i.e. the error term in the first stage) and unobserved determinants of a player's salary (i.e. the error term in the second stage). As presented in Table 4-5, column (1), $\rho = 0$ in our case, suggesting that there is no sample-selection issues after using a matching sample and our estimates are unbiased (Wooldridge 2010). The estimated coefficient of $log\_follower$ with full sample, which is shown in column (3), is also significantly positive, a result that is consistent with that in column (2) and further confirms the robustness of our results.

Since players who engage more by posting more entertaining contents can attract more followers, popularity measured by the number of Twitter followers can be considered as the outcome of the social media efforts spent by the players. Our analysis quantifies the value of Twitter account and the number of Twitter followers and our results can be used as a yardstick for the players to decide how much efforts and money they should spend on Twitter to maximize their benefit. Our findings also provide evidence of the effectiveness of Twitter as communication channel for managing loyal fans or loyal customers in the NBA context. As a result, other celebrities or firms should have more confidence in investing in marketing efforts on microblogging media.

Table 4-5 Results of Heckman Selection Model

VARIABLES	(1) Log_salary	(2) Join	(3) Log_salary	(4) Join
Efficiency_rating	0.138***	-0.046	0.037*	0.087***
	(0.027)	(0.034)	(0.022)	(0.016)
Team_selection	0.035	-0.238	0.028	-0.065
	(0.246)	(0.240)	(0.076)	(0.074)
Tenure	0.212***		0.209***	
	(0.051)		(0.035)	
$Tenure^2$	-0.011***		-0.010***	
	(0.003)		(0.002)	
Log_minutes_played	1.564***		1.972***	
	(0.270)		(0.195)	
Log_follower	0.114***		0.070***	
	(0.040)		(0.026)	
Position dummies	-included-	-included-	-included-	-included-
Time dummies	-included-	-included-	-included-	-included-
Fixed-effect	-included-	-included-	-included-	-included-
age		0.091***		-0.020**
		(0.017)		(0.009)
Constant	8.912***	-2.800***	10.729***	-1.143***
	(0.575)	(0.590)	(0.465)	(0.297)
Observations	653	653	2,097	2,097
Selection ρ	0.000			-0.883
Wald χ^2	319.94		392.59	
Log-likilihood	-441.412		-1180.992	

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

4.5.3. The Impact of Twitter Account on Salary Inequality

The results of quantile regression with matched sample are summarized in Table 4-6. As shown in Table 4-6, the 0.50 quantile coefficient (with magnitude 0.028) is smallest and is not significant among all quantile estimates. Next, we find that the 0.75 quantile coefficient (with magnitude 0.066) and 0.90 quantile coefficient (with magnitude 0.089) are much larger than the 0.50 quantile coefficient (with magnitude 0.028). The 0.75 quantile coefficient is two times larger than the 0.50 quantile coefficient, implying above-average players can benefit more from participation in Twitter than average players. The impact of participation in Twitter is even larger for those players in the 0.90 quantile (i.e., superstar players). The 0.90 quantile coefficient is three times larger than the 0.50 quantile coefficient. This is not surprising since superstar players generally have agents to build a good public relationship with media and their fans while those average and above-average players do not. Also, after joining Twitter, superstars can broadcast to millions of followers whereas ordinary players only have tens thousands of followers. The huge fan base provides various "branding" opportunities of the focal athlete.

On the other hand, although the 0.25 quantile coefficient is not significant, its magnitude is slightly larger than the 0.50 quantile coefficient. This is a signal that below-average players might gain more from Twitter than average players. It can be further illustrated by the 0.10 quantile coefficient. The most interesting finding in our results is that the 0.10 quantile coefficient (with magnitude 0.195) is the largest, implying that bottom players gain most from participation in Twitter and catch up with players in higher quantiles in terms of salary. This effect for players with the lowest salary is striking in that it is two times larger than that for superstar players. The reason could be that Twitter helps a number of low-paid players get increased exposure to media and improved popularity among fans, thus leading to the increases in the salaries of these players.

If we consider players at the 0.50 quantile as the baseline case, we can infer that social media helps the players at both lower quantiles and higher quantiles to earn more. In other words, social media benefit both bottom players and superstar player at the cost of average players. We borrow the terminology from Autor et al. (2006) and characterize this pattern as polarization in the salary distribution in NBA. Moreover, given that players at the lower quantiles (those at the 0.10 quantile) benefit much more from Twitter than those at higher quantiles and the number of players at lower quantiles is much larger than those at higher quantiles, income inequality in the NBA has been reduced due to the emerging of social media.

Table 4-6 Results of Quantile Regression

VARIABLES	(1) 0.10 quantile Log_salary	(2) 0.25 quantile Log_salary	(3) 0.50 quantile Log_salary	(4) 0.75 quantile Log_salary	(5) 0.90 quantile Log_salary
Join	0.195*	0.037	0.028	0.066*	0.089**
	(0.108)	(0.054)	(0.051)	(0.035)	(0.043)
Efficiency_rating	0.154***	0.141***	0.106***	0.097***	0.083***
	(0.032)	(0.017)	(0.016)	(0.010)	(0.012)
Team_selection	-0.007	-0.018	-0.024	0.002	-0.062
	(0.079)	(0.107)	(0.081)	(0.045)	(0.042)
Tenure	0.335***	0.279***	0.256***	0.262***	0.231***
	(0.080)	(0.040)	(0.036)	(0.023)	(0.028)
Tenure2	-0.020***	-0.015***	-0.014***	-0.014***	-0.011***
	(0.004)	(0.002)	(0.002)	(0.001)	(0.001)
Log_minutes_played	2.542***	2.482***	2.082***	1.820***	1.584***
	(0.241)	(0.144)	(0.145)	(0.099)	(0.117)
Position dummies	-included-	-included-	-included-	-included-	-included-
Time dummies	-included-	-included-	-included-	-included-	-included-
Fixed-effect	-included-	-included-	-included-	-included-	-included-
Constant	7.227***	7.752***	9.489***	10.243***	11.074***
	(0.470)	(0.272)	(0.278)	(0.200)	(0.233)
Observations	751	751	751	751	751

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

4.6. Concluding Remarks

In this paper, we applied econometric analysis to empirically investigate the economic value of celebrities' participation and popularity in social media within the context of NBA. Specifically, we studied whether NBA players' participation and popularity in Twitter help them earn higher salaries and players with higher-income or lower-income can benefit more from Twitter. We present empirical evidence of the business value of Twitter for NBA players at the individual level and Twitter could reduce the salary gap between lowest-paid players and average players, a surprising finding about the positive impact of the use of Twitter.

We examined the value of celebrities' participation in social media by estimating a difference-in-differences (DID) model with a panel data set. The DID estimates suggested NBA players' participation in social media has a positive impact on their salaries. To control for the potential sample selection of players' participation in Twitter, we used Coarsen Exact Matching (CEM) to construct a control group and rerun the DID model with matched sample. The DID model in combination with matched sample helps us to control potential bias stemming from both observable and unobservable factors. The DID estimate with matched sample confirmed that players' participation in Twitter helps them to gain higher salaries. The robustness of our results indicates participation in social media indeed brings huge economic value to celebrities such as NBA players.

In investigating the value of celebrities' popularity in social media, we measure a player's popularity in Twitter by his number of Twitter followers. We applied the Heckman two-stage model to control for potential sample selection issues since Twitter followers are only observable for players who decided to participate in Twitter. We estimated the Heckman two-stage with both matched sample and full sample. Both estimates suggest that higher level of popularity in Twitter helps NBA players to gain higher salaries.

To find out which type of players can benefit more from Twitter, we employ quantile regression and investigate the impact of players' participation in social media on salary inequality among players. We found that: (i) superstar players (and above-average players) benefit more than average players and (ii) bottom players (and below-average players) can also benefit more than average players. Our results exhibit a pattern of polarization in the salary distribution in NBA. Moreover, the first effect increases salary inequality while the second effect decreases salary inequality. Thus, the net effect depends on both the relative impact of these two distinct effects and NBA players' salary distribution. In reality, since players at the lower quantiles benefit much more from Twitter than those at higher quantiles and the number of players at lower quantiles is much larger than those at higher quantiles, income inequality in the NBA has been reduced due to the proliferation of social media.

Our study also has limitation and provides research opportunity for future study. First, the sample in our study is a group of NBA players. Therefore, our findings could be generalized to other professional athletes but the stronger generalizability to musicians, movies stars, or politicians may need further empirical analyses. Furthermore, the monetary value of Twitter followers of small, medium, and large companies in different industries may require customized analysis case by case and is clearly beyond the scope of this paper. Second, a number of existing studies of Twitters analyze the textual contents of tweets for building predictive models. A combination of those text mining studies with our econometric analysis could provide more insightful business implications regarding enhancing the business value of using microblogging sites.

BIBLIOGRAPHY

Acemoglu, D. 2002. "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature* (40:1), pp. 7-72.

Acemoglu, D. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics,* O.C. Ashenfelter and D. Card (eds.). North-holland: Elsevier, pp. 1043-1171.

Andreoni, J. 1988. "Privately Provided Public Goods in a Large Economy: The Limits of Altruism," *Journal of Public Economics* (35:1), pp. 57-73.

Andreoni, J. 1989. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *The Journal of Political Economy* (97:6), pp. 1447-1458.

Andreoni, J. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *The Economic Journal* (100:401), pp. 464-477.

Andreoni, J., and Petrie, R. 2004. "Public Goods Experiments without Confidentiality: A Glimpse into Fund-Raising," *Journal of Public Economics* (88:7), pp. 1605-1623.

Angrist, J.D., and Pischke, J.S. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Antin, J., and Churchill, E.F. 2011. "Badges in Social Media: A Social Psychological Perspective," in: *CHI*. Vancouver, BC: ACM.

Aral, S., Dellarocas, C., and Godes, D. 2013. "Introduction to the Special Issue-Social Media and Business Transformation: A Framework for Research," *Information Systems Research* (24:1), pp. 3-13.

Aral, S., and Walker, D. 2012. "Identifying Influential and Susceptible Members of Social Networks," *Science* (337:6092), pp. 337-341.

Ashraf, N. 2011. "Turning Work into Play: Gamifying Your Comms," in: *Strategic Communication Management*. pp. 2-5.

Asvanund, A., Clay, K., Krishnan, R., and Smith, M.D. 2004. "An Empirical Analysis of Network Externalities in Peer-to-Peer Music-Sharing Networks," *Information Systems Research* (15:2), pp. 155-174.

Autor, D., and Dorn, D. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the Us Labor Market," *American Economic Review* (103:5), pp. 1553-1597.

Autor, D., Katz, L., and Kearney, M. 2006. "The Polarization of the Us Labor Market," *American Economic Review* (96:2), pp. 189-194.

Autor, D.H., Katz, L.F., and Kearney, M.S. 2008. "Trends in Us Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics* (90:2), pp. 300-323.

Azoulay, P., Stuart, T., and Wang, Y. 2013. "Matthew: Effect or Fable?," *Management Science* (60:1), pp. 92-109.

Badenhausen, K. 2013. "Kobe and Lebron Top List of the Nba's Highest-Paid Players," in: *Forbes*.

Bakshy, E., Hofman, J.M., Mason, W.A., and Watts, D.J. 2011. "Everyone's an Influencer: Quantifying Influence on Twitter," *Proceedings of the 4th ACM International Conference on Web Search and Data Mining*: ACM, pp. 65-74.

Bapna, R., Langer, N., Mehra, A., Gopal, R., and Gupta, A. 2013. "Human Capital Investments and Employee Performance: An Analysis of It Services Industry," *Management Science* (59:3), pp. 641-658.

Bateman, P.J., Gray, P.H., and Butler, B.S. 2011. "The Impact of Community Commitment on Participation in Online Communities," *Information Systems Research* (22:4), pp. 841-854.

Blackwell, M., Iacus, S.M., King, G., and Porro, G. 2009. "Cem: Coarsened Exact Matching in Stata," *Stata Journal* (9:4), pp. 524-546.

Brynjolfsson, E., Hu, Y.J., and Simester, D. 2011. "Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales," *Management Science* (57:8), pp. 1373-1386.

Bukowski, W.M., and Hoza, B. 1989. "Popularity and Friendship: Issues in Theory, Measurement, and Outcome," in *Peer Relationships in Child Development*, T. Berndt and G.W. Ladd (eds.). New York: Wiley, pp. 15-45.

Burke, B. 2011. "Maverick Research: Motivation, Momentum and Meaning: How Gamification Can Inspire Engagement."

Butler, B.S. 2001. "Membership Size, Communication Activity, and Sustainability: A Resource-Based Model of Online Social Structures," *Information Systems Research* (12:4), pp. 346-362.

Card, D., Lemieux, T., and Riddell, W.C. 2004. "Unions and Wage Inequality," *Journal of Labor Research* (25:4), pp. 519-559.

Chan, J., and Ghose, A. 2013. "Internet's Dirty Secret: Assessing the Impact of Online Intermediaries on Hiv Transmission," *MIS Quarterly* (Forthcoming).

Chaudhuri, A. 2011. "Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature," *Experimental Economics* (14:1), pp. 47-83.

Chen, Y. 2008. "Incentive-Compatible Mechanisms for Pure Public Goods: A Survey of Experimental Research," *Handbook of Experimental Economics Results* (1), pp. 625-643.

Chen, Y., Ho, T.-H., and Kim, Y.-M. 2010. "Knowledge Market Design: A Field Experiment at Google Answers," *Journal of Public Economic Theory* (12:4), pp. 641-664.

Chiu, C.M., Hsu, M.H., and Wang, E.T.G. 2006. "Understanding Knowledge Sharing in Virtual Communities: An Integration of Social Capital and Social Cognitive Theories," *Decision Support Systems* (42:3), pp. 1872-1888.

Coie, J.D., Dodge, K.A., and Coppotelli, H. 1982. "Dimensions and Types of Social Status: A Cross-Age Perspective," *Developmental Psychology* (18:4), p. 557.

Compeau, D.R., and Higgins, C.A. 1995. "Computer Self-Efficacy: Development of a Measure and Initial Test " *MIS Quarterly* (19:2), Jun, pp. 189-211.

Constant, D., Kiesler, S., and Sproull, L. 1994. "What's Mine Is Ours, or Is It? A Study of Attitudes About Information Sharing," *Information Systems Research* (5:4), Dec, pp. 400-421.

Constant, D., Sproull, L., and Kiesler, S. 1996. "The Kindness of Strangers: The Usefulness of Electronic Weak Ties for Technical Advice," *Organization Science* (7:2), pp. 119-135.

Cornes, R., and Sandler, T. 1994. "The Comparative Static Properties of the Impure Public Good Model," *Journal of Public Economics* (54:3), pp. 403-421.

DailyMail. 2013. "Is This the Decline of Wikipedia? A Third of Editors Have Quit Complaining Site Bosses Have 'Lowered the Bar' on Quality." Retrieved August 9 2015, from http://www.dailymail.co.uk/sciencetech/article-2477432/Is-decline-Wikipedia-A-staff-quit-site-thanks-dumbed-software-auto-delete-tools.html

Daugherty, T., Lee, W.N., Gangadharbatla, H., Kim, K., and Outhavong, S. 2005. "Organizational Virtual Communities: Exploring Motivations Behind Online Panel Participation," *Journal of Computer-Mediated Communication* (10:4), pp. 00-00.

Deterding, S., Khaled, R., Nacke, L.E., and Dixon, D. 2011. "Gamification: Toward a Definition," *In CHI 2011 Gamification Workshop Proceedings*, Vancouver, BC, Canada.

DiNardo, J., Fortin, N.M., and Lemieux, T. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica* (64:5), pp. 1001-1044.

Durcikova, A., and Gray, P. 2009. "How Knowledge Validation Processes Affect Knowledge Contribution," *Jouranl of Management Information Systems* (25:4), pp. 81-107.

Ertug, G., and Castellucci, F. 2013. "Getting What You Need: How Reputation and Status Affect Team Performance, Hiring, and Salaries in the Nba," *Academy of Management Journal* (56:2), pp. 407-431.

Facebook. 2015. "One Year In: Internet.Org Free Basic Services." Retrieved August 9 2015, from https://www.internet.org/press/one-year-in-internet-dot-org-free-basic-services

Faraj, S., and Johnson, S.L. 2011. "Network Exchange Patterns in Online Communities," *Organization Science* (22:6), pp. 1464-1480.

Fogg, B. 2009. "A Behavior Model for Persuasive Design," *Proceedings of the 4th international Conference on Persuasive Technology*: ACM, p. 40.

Forrester Research. 2009. "Us Interactive Marketing Forecast, 2009 to 2014," Forrester Research, Cambridge, MA.

Freeman, R.B. 1991. "How Much Has De-Unionisation Contributed to the Rise in Male Earnings Inequality?," in *Uneven Tides: Rising Inequality in America*, S. Danziger and P. Gottschalk (eds.). New York: Russell Sage Foundatio, pp. 133-163.

Frey, B.S., and Meier, S. 2004. "Social Comparisons and Pro-Social Behavior: Testing" Conditional Cooperation" in a Field Experiment," *American Economic Review* (94:5), pp. 1717-1722.

Garnefeld, I., Iseke, A., and Krebs, A. 2012. "Explicit Incentives in Online Communities: Boon or Bane?," *International Journal of Electronic Commerce* (17:1), pp. 11-38.

Gartner. 2011. "Gartner Predicts over 70 Percent of Global 2000 Organisations Will Have at Least One Gamified Application by 2014," Barcelona, Spain.

Gazan, R. 2011. "Social Q&A," *Journal of the American Society for Information Science and Technology* (62:12), pp. 2301-2312.

Ghose, A., Goldfarb, A., and Han, S.P. 2012. "How Is the Mobile Internet Different? Search Costs and Local Activities," *Information Systems Research* (24:3), pp. 613-631.

Ghose, A., Ipeirotis, P.G., and Li, B. 2014. "Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue," *Management Science* (60:7), pp. 1632-1654.

Goel, S., Watts, D.J., and Goldstein, D.G. 2012. "The Structure of Online Diffusion Networks," *Proceedings of the 13th ACM Conference on Electronic Commerce*: ACM, pp. 623-638.

Goh, K.-Y., Heng, C.-S., and Lin, Z. 2013. "Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User-and Marketer-Generated Content," *Information Systems Research* (24:1), pp. 88-107.

Greene, W.H. 2011. *Econometric Analysis*. United States: Pearson Education.

Gu, B., Huang, Y., Duan, W., and Whinston, A. 2007a. "Online Peer-to-Peer Communities: An Empirical Investigation of a Music Sharing Network as a Dynamic Two-Sided Network," in: *INFORMS Conference on Information Systems and Technology*.

Gu, B., Konana, P., Rajagopalan, B., and Chen, H.W.M. 2007b. "Competition among Virtual Communities and User Valuation: The Case of Investing-Related Communities," *Information Systems Research* (18:1), pp. 68-85.

Gu, B., and Ye, Q. 2014. "First Step in Social Media: Measuring the Influence of Online Management Responses on Customer Satisfaction," *Production and Operations Management* (23:4), pp. 570-582.

Hall, H., and Graham, D. 2004. "Creation and Recreation: Motivating Collaboration to Generate Knowledge Capital in Online Communities," *International Journal of Information Management* (24:3), pp. 235-246.

Heckman, J.J. 1979. "Sample Selection Bias as a Specification Error," *Econometrica* (47:1), pp. 153-161.

Heldt, T. 2005. "Conditional Cooperation in the Field: Cross-Country Skiers," *Sweden." Mimeo, Uppsala University*: Citeseer.

Hollinger, J. 2005. *Pro Basketball Forecast, 2005-06*. Potomac Books.

Hui, K.-L., Kim, S.H., and Wang, Q.-H. 2013. "Marginal Deterrence in the Enforcement of Law: Evidence from Distributed Denial of Service Attack." Working paper.

Iacus, S.M., King, G., and Porro, G. 2011. "Multivariate Matching Methods That Are Monotonic Imbalance Bounding," *Journal of the American Statistical Association* (106:493), pp. 345-361.

Iacus, S.M., King, G., and Porro, G. 2012. "Causal Inference without Balance Checking: Coarsened Exact Matching," *Political Analysis* (20:1), pp. 1-24.

Imbens, G., and Lemieux, T. 2008. "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics* (142:2), pp. 615-635.

Iyengar, R., Van den Bulte, C., and Valente, T.W. 2011. "Opinion Leadership and Social Contagion in New Product Diffusion," *Marketing Science* (30:2), pp. 195-212.

Jin, X.-L., Zhou, Z., Lee, M.K., and Cheung, C.M. 2012. "Why Users Keep Answering Questions in Online Question Answering Communities: A Theoretical and Empirical Investigation," *International Journal of Information Management* (33:1), pp. 93-104.

Kankanhalli, A., Tan, B.C.Y., and Wei, K.K. 2005. "Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation," *MIS Quarterly* (29:1), pp. 113-143.

Kaplan, A.M., and Haenlein, M. 2011. "The Early Bird Catches the News: Nine Things You Should Know About Micro-Blogging," *Business Horizons* (54:2), pp. 105-113.

Katz, L.F., and Auctor, D., H. 1999. "Changes in the Wage Structure and Earnings Inequality," in *Handbook of Labor Economics,* O.C. Ashenfelter and D. Card (eds.). North-holland: Elsevier, pp. 1463-1555.

Katz, M.L., and Shapiro, C. 1985. "Network Externalities, Competition, and Compatibility," *American Economic Review* (75:3), pp. 424-440.

Katz, M.L., and Shapiro, C. 1992. "Product Introduction with Network Externalities," *The Journal of Industrial Economics* (40:1), pp. 55-83.

Koenker, R. 2005. *Quantile Regression*. Cambridge University Press.

Kwak, H., Lee, C., Park, H., and Moon, S. 2010. "What Is Twitter, a Social Network or a News Media?," *Proceedings of the 19th International Conference on World Wide Web*: ACM, pp. 591-600.

Li, F., and Du, T.C. 2011. "Who Is Talking? An Ontology-Based Opinion Leader Identification Framework for Word-of-Mouth Marketing in Online Social Blogs," *Decision Support Systems* (51:1), pp. 190-197.

Lin, M., Prabhala, N.R., and Viswanathan, S. 2013. "Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending," *Management Science* (59:1), pp. 17-35.

Luo, X., Zhang, J., and Duan, W. 2013. "Social Media and Firm Equity Value," *Information Systems Research* (24:1), pp. 146-163.

Ma, M., and Agarwal, R. 2007. "Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities," *Information Systems Research* (18:1), pp. 42-67.

Malter, D. 2014. "On the Causality and Cause of Returns to Organizational Status Evidence from the Grands Crus Classés of the Médoc," *Administrative Science Quarterly* (Forthcoming).

Marett, K., and Joshi, K. 2009. "The Decision to Share Information and Rumors: Examining the Role of Motivation in an Online Discussion Forum," *Communications of the Association for Information Systems* (24:1), p. 4.

Masclet, D., Noussair, C., Tucker, S., and Villeval, M.-C. 2003. "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism," *American Economic Review* (93:1), pp. 366-380.

Massey, C., and Thaler, R.H. 2013. "The Loser's Curse: Decision Making and Market Efficiency in the National Football League Draft," *Management Science* (59:7), pp. 1479-1495.

McAlister, L., and Pessemier, E. 1982. "Variety Seeking Behavior: An Interdisciplinary Review," *Journal of Consumer Research* (9:3), pp. 311-322.

Mohan, P. 2015. "Twitter Seeks to Attract More Users with New Homepage." Retrieved August 9 2015, from http://www.fastcompany.com/3049293/fast-feed/twitter-seeks-to-attract-more-users-with-new-homepage

Moon, J.Y., and Sproull, L.S. 2008. "The Role of Feedback in Managing the Internet-Based Volunteer Work Force," *Information Systems Research* (19:4), pp. 494-515.

Motwani, H. 2013. "10 Reasons Why You Should Choose Twitter over Facebook." Retrieved September 9, 2014, from https://smallbusiness.yahoo.com/advisor/10-reasons-why-choose-twitter-over-facebook-000038313.html

Nahapiet, J., and Ghoshal, S. 1998. "Social Capital, Intellectual Capital, and Organizational Advantage," *Academy of Management Review* (23:2), pp. 242-266.

Nair, H.S., Manchanda, P., and Bhatia, T. 2010. "Asymmetric Social Interactions in Physician Prescription Behavior: The Role of Opinion Leaders," *Journal of Marketing Research* (47:5), pp. 883-895.

Newcomb, A.F., and Bagwell, C.L. 1995. "Children's Friendship Relations: A Meta-Analytic Review," *Psychological Bulletin* (117:2), p. 306.

Nishar, D. 2013. "200 Million Members!" *LinkedIn Blog* Retrieved June 20, 2013

Oh, S. 2012. "The Characteristics and Motivations of Health Answerers for Sharing Information, Knowledge, and Experiences in Online Environments," *Journal of the American Society for Information Science and Technology* (63:3), pp. 543-557.

Olson, C.K. 2010. "Children's Motivation for Video Game Play in the Context of Normal Development," *Review of General Psychology* (14:2), pp. 180-187.

Protalinski, E. 2011. "Two Thirds of Small Business Owners Use Facebook for Marketing." Retrieved October 14, 2014, from http://www.zdnet.com/blog/facebook/two-thirds-of-small-business-owners-use-facebook-for-marketing/1529

Protalinski, E. 2013. "Wikimedia Rolls out Wysiwyg Visual Editor for Logged-in Users Accessing Wikipedia Articles in English." Retrieved August 9

2015, from http://thenextweb.com/insider/2013/07/02/wikimedia-rolls-out-its-wysiwyg-visual-editor-for-logged-in-users-accessing-wikipedia-articles-in-english/

Raban, D.R. 2009. "Self-Presentation and the Value of Information in Q&a Websites," *Journal of the American Society for Information Science and Technology* (60:12), pp. 2465-2473.

Ren, Y., and Kraut, R.E. 2011. "A Simulation for Designing Online Community: Member Motivation, Contribution, and Discussion Moderation," Working Paper.

Ren, Y., Kraut, R.E., and Kiesler, S. 2007. "Applying Common Identity and Bond Theory to Design of Online Communities," *Organization Studies* (28:3), pp. 377-408.

Ren, Y.Q., Harper, F.M., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., and Kraut, R.E. 2012. "Building Member Attachment in Online Communities: Applying Theories of Group Identity and Interpersonal Bonds," *MIS Quarterly* (36:3), Sep, pp. 841-864.

Rishika, R., Kumar, A., Janakiraman, R., and Bezawada, R. 2013. "The Effect of Customers' Social Media Participation on Customer Visit Frequency and Profitability: An Empirical Investigation," *Information Systems Research* (24:1), pp. 108-127.

Romero, D.M., Galuba, W., Asur, S., and Huberman, B.A. 2011. "Influence and Passivity in Social Media," in *Machine Learning and Knowledge Discovery in Databases*. Springer, pp. 18-33.

Rosenbaum, P.R., and Rubin, D.B. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* (70:1), pp. 41-55.

Ryan, R.M., and Deci, E.L. 2000. "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being," *American Psychologist* (55:1), pp. 68-78.

Ryan, R.M., Rigby, C.S., and Przybylski, A. . 2006. "The Motivational Pull of Video Games: A Self-Determination Theory Approach," *Motivation and Emotion* (30:4), pp. 344-360.

Samuelson, P.A. 1954. "The Pure Theory of Public Expenditure," *The Review of Economics and Statistics* (36:4), pp. 387-389.

Savikhina, A., and Sheremetab, R.M. 2010. "Visibility of Contributions and Cost of Information: An Experiment on Public Goods," Working Paper.

Shang, J., and Croson, R. 2009. "A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods," *The Economic Journal* (119:540), pp. 1422-1439.

Shapiro, C., and Varian, H.R. 2013. *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business Press.

Silverman, R. 2011. "Latest Game Theory: Mixing Work and Play," in: *Wall Street Journal*.

Slutsky, I. 2011. "Twitter to Offer Marketers More Tools to Target, Track Followers."   Retrieved September 9, 2014, from http://adage.com/article/special-report-digital-conference/twitter-offer-marketers-geo-relevant-ads-tools/226846/

Soetevent, A.R. 2005. "Anonymity in Giving in a Natural Context—a Field Experiment in 30 Churches," *Journal of Public Economics* (89:11), pp. 2301-2323.

Softpedia. 2013. "Wikipedia's New Visualeditor Is the Best Update in Years and You Can Make It Better." Retrieved August 9 2015, from http://news.softpedia.com/news/Wikipedia-s-New-VisualEditor-Is-the-Best-Update-in-Years-and-You-Can-Make-It-Better-365072.shtml

Steinberg, R. 1987. "Voluntary Donations and Public Expenditures in a Federalist System," *American Economic Review* (77:1), pp. 24-36.

Stewart, M.B., and Swaffield, J.K. 2008. "The Other Margin: Do Minimum Wages Cause Working Hours Adjustments for Low-Wage Workers?," *Economica* (75:297), pp. 148-167.

Sun, Y., Fang, Y., and Lim, K.H. 2011. "Understanding Sustained Participation in Transactional Virtual Communities," *Decision Support Systems* (53:1), pp. 12-22.

Sutanto, J., Kankanhalli, A., and Tan, B.C.Y. 2011. "Eliciting a Sense of Virtual Community among Knowledge Contributors," *ACM Transactions on Management Information Systems (TMIS)* (2:3), p. 14.

Takahashi, D. 2010. "Website Builder Devhub Gets Users Hooked by "Gamifying" Its Service," in: *VentureBeat*.

Tedjamulia, S.J.J., Dean, D.L., Olsen, D.R., and Albrecht, C.C. 2005. "Motivating Content Contributions to Online Communities: Toward a More Comprehensive Theory," *System Sciences, 2005. HICSS'05. Proceedings of the 38th Annual Hawaii International Conference on*: IEEE, pp. 193b-193b.

The Economist. 2011. "Changes at Wikipedia: Seeing Things."   Retrieved August 9 2015, from http://www.economist.com/blogs/babbage/2011/12/changes-wikipedia

Tiwana, A., and Bush, A.A. 2005. "Continuance in Expertise-Sharing Networks: A Social Perspective," *IEEE Transactions on Engineering Management* (52:1), pp. 85-101.

Toubia, O., and Stephen, A.T. 2013. "Intrinsic Vs. Image-Related Utility in Social Media: Why Do People Contribute Content to Twitter?," *Marketing Science* (32:3), pp. 368-392.

Wang, S., and Noe, R.A. 2010. "Knowledge Sharing: A Review and Directions for Future Research," *Human Resource Management Review* (20:2), pp. 115-131.

Wasko, M., and Faraj, S. 2000. "“It Is What One Does”: Why People Participate and Help Others in Electronic Communities of Practice," *The Journal of Strategic Information Systems* (9:2–3), pp. 155-173.

Wasko, M.M., and Faraj, S. 2005. "Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice," *MIS Quarterly* (29:1), pp. 35-57.

Weng, J., Lim, E.-P., Jiang, J., and He, Q. 2010. "Twitterrank: Finding Topic-Sensitive Influential Twitterers," *Proceedings of the 3rd ACM International Conference on Web Search and Data Mining*: ACM, pp. 261-270.

Werbach, K. 2013. "Gamification. Introductory Video to Online Course.." Coursera.

Wikimedia Foundation. 2010. "An Update of the Article Feedback as of October 4, 2010." Retrieved December 2015, from https://www.mediawiki.org/wiki/Article_feedback/Research#October_4.2C_2010

Wikimedia Foundation. 2011. "Editor Trends Study/Results." Retrieved August 9 2015, from https://strategy.wikimedia.org/wiki/Editor_Trends_Study/Results

Wikimedia Foundation. 2015. "Wikipedia:Statistics." Retrieved August 9 2015, from https://en.wikipedia.org/wiki/Wikipedia:Statistics

Wooldridge, J.M. 2009. *Introductory Econometrics: A Modern Approach*. South-Western Pub.

Wooldridge, J.M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT press.

Wu, L. 2013. "Social Network Effects on Productivity and Job Security: Evidence from the Adoption of a Social Networking Tool," *Information Systems Research* (24:1), pp. 30-51.

Wu, M. 2011. "Gamification: From Behavior Model to Business Strategy." *Lithium*

Wu, S., Hofman, J.M., Mason, W.A., and Watts, D.J. 2011. "Who Says What to Whom on Twitter," *Proceedings of the 20th international conference on World wide web*: ACM, pp. 705-714.

Yee, N. 2006. "Motivations for Play in Online Games," *Cyber Psychology & Behavior* (9:6), pp. 772-775.

Zhang, X., and Wang, C. 2012. "Network Positions and Contributions to Online Public Goods: The Case of Chinese Wikipedia," *Journal of Management Information Systems* (29:2), pp. 11-40.

Zhang, X.M., and Zhu, F. 2011. "Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia," *American Economic Review* (101:4), pp. 1601-1615.

Zichermann, G., and Cunningham, C. 2011. *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. O'Reilly Media.

APPENDIX

Table A1. Definitions of the Badges used in the Micro-Level Analysis

Badge Name	**Definition / Rule**
Popular Question	Asked a question with 1,000 views
Notable Question	Asked a question with 2,500 views
Famous Question	Asked a question with 10,000 views
Nice Question	Question score of 10 or more
Good Question	Question score of 25 or more
Great Question	Question score of 100 or more
Favorite Question	Question favorited by 25 users
Stellar Question	Question favorited by 100 users
Nice Answer	Answer score of 10 or more
Good Answer	Answer score of 25 or more
Great Answer	Answer score of 100 or more
Enlightened	First to answer and accepted with at least 10 upvotes
Guru	Accepted answer and score of 40 or more
Necromancer	Answered a question more than 60 days later with score of 5 or more
Populist	Highest scoring answer that outscored an accepted answer with score of more than 10 by more than 2x
Reversal	Provided answer of +20 score to a question of -5 score
Revival	Answered more than 30 days later as first answer scoring 2 or more
Commentator	Left 10 comments
Pundit	Left 10 comments with score of 5 or more
Archaeologist	Edited 100 posts that were inactive for 6 months
Excavator	Edited first post that was inactive for 6 months
Strunk & White	Edited 80 posts
Copy Editor	Edited 500 posts
Proofreader	Approved or rejected 100 suggested edits
Tenacious	Zero score accepted answers: more than 5 and 20% of total
Unsung Hero	Zero score accepted answers: more than 10 and 25% of total
Tumbleweed	Asked a question with no votes, no answers, no comments, and low views for a week