Yu Jeffrey Hu, D.J. Wu, Hua Yuan (Eds.)

PROCEEDINGS

of The 8th China Summer Workshop on Information Management, (CSWIM 2014)

Chengdu, China, June 21–22, 2014

电子科技大学
University of Electronic Science and Technology of China

Georgia Institute of Technology

ORGANIZING COMMITTEE

Honorary Co-Chairs:

Yongkai Ma
Dean, School of Management and Economics, University of Electronic Science and Technology of China, China

Yong Zeng
Director, Research Center of China Capital Market, School of Management and Economics, University of Electronic Science and Technology of China, China

Workshop Co-Chairs:

Yu Jeffrey Hu
Georgia Institute of Technology, USA

D.J. Wu
Georgia Institute of Technology, USA

Hua Yuan
University of Electronic Science and Technology of China, China

Local Arrangements Co-Chairs:

Yingjin Lu
University of Electronic Science and Technology of China, China

Jingming Pan
University of Electronic Science and Technology of China, China

Yu Qian
University of Electronic Science and Technology of China, China

Web and Internet Support:

Harry Jiannan Wang
University of Delaware, USA

ADVISORY COMMITTEE

Ravi Bapna	University of Minnesota
Guoqing Chen	Tsinghua University
Hsinchun Chen	University of Arizona
Jian Chen	Tsinghua University
Prabuddha De	Purdue University
Paulo Goes	University of Arizona
Alok Gupta	University of Minnesota
Alan Hevner	University of Southern Florida
Robert Kauffman	Singapore Management University
Ramayya Krishnan	Carnegie Mellon University
T.P. Liang	National Sun Yat-Sen University
James Marsden	University of Connecticut
Vijay Mookerjee	University of Texas at Dallas
Sridhar Narasimhan	Georgia Institute of Technology
Sumit Sarkar	University of Texas at Dallas
Yong Tan	University of Washington, Seattle
Bruce Weber	University of Delaware
Andrew Whinston	University of Texas at Austin
Leon Zhao	City University of Hong Kong

PROGRAM COMMITTEE

Jinwei Cao	University of Delaware
Patrick Chau	University of Hong Kong
Ramnath Chellappa	Emory University
Jianqing Chen	University of Texas – Dallas
Rui Chen	Ball State University
Kenny Cheng	University of Florida
Roger Chiang	University of Cincinnati
Honghui Deng	University of Nevada – Las Vegas
Yifan Dou	Beihang University
Wenjing (Wendy) Duan	George Washington University
Ming Fan	University of Washington
Patrick Fan	Virginia Tech
Shaokun Fan	City University of Hong Kong
Xianjun Geng	University of Texas – Dallas
Ram Gopal	University of Connecticut
Bin Gu	Arizona State University
Yi Maggie Guo	University of Michigan – Dearborn
Zhiling Guo	Singapore Management University
Daning Hu	University of Zurich
Xiaorui Hu	Saint Louis University
Kai-Lung Hui	Hong Kong University of Science and Technology
Yonghua Ji	University of Alberta
Zhenhui (Jack) Jiang	National University of Singapore
Yi-Cheng Ku	Providence University
Beibei Li	Carnegie Mellon University
Dahui Li	University of Minnesota, Duluth
Ting Li	Erasmus University
Wenli Li	Dalian University of Technology
Xin Li	City University of Hong Kong
Lihui Lin	Tsinghua University

Zhong Yao	BeiHang University
Qiang Ye	Harbin Institute of Technology
Cheng Yi	Tsinghua University
Fang Yin	University of Oregon
Dezhi (Denny) Yin	University of Missouri
Byungjoon Yoo	Seoul National University
Wei Thoo Yue	City University of Hong Kong
Daniel Zeng	Chinese Academy of Sciences & University of Arizona
Dongsong Zhang	University of Maryland – Baltimore County
Han Zhang	Georgia Institute of Technology
Michael Zhang	Hong Kong University of Science and Technology
Pengzhu Zhang	Shanghai Jiao Tong University
Ping Zhang	Syracuse University
Zhongju Zhang	University of Connecticut
Huimin Zhao	University of Wisconsin – Milwaukee
Rong Zheng	Hong Kong University of Science and Technology
Weijun (Will) Zheng	University of Wisconsin – Parkside
Lina Zhou	University of Maryland – Baltimore County
Yilu Zhou	Fordham University
Bin Zhu	Oregon State University
Dan Zhu	Iowa State University
Kevin Zhu	University of California – San Diego
Qinghua Zhu	Nanjing University

KEYNOTE 1

Hsinchun Chen

Arizona Regents' Professor
Thomas R. Brown Chair in Management and Technology
Director of Artificial Intelligence Laboratory
University of Arizona

Bio: Dr. Hsinchun Chen is University of Arizona Regents' Professor and Thomas R. Brown Chair in Management and Technology in the Management Information Systems (MIS) Department and Professor of Entrepreneurship & Innovation in the McGuire Center for Entrepreneurship at the College of Management of the University of Arizona. He received the B.S. degree from the National Chiao-Tung University in Taiwan, the MBA degree from SUNY Buffalo, and the Ph.D. degree in Information Systems from the New York University. Dr. Chen is director of the Artificial Intelligence Lab and has served as a faculty of the UA MIS department (ranked #3 in MIS) since 1989. He had served as a Scientific Counselor/Advisor of the National Library of Medicine (USA), Academia Sinica (Taiwan), and National Library of China (China).

Keynote #1: Health Big Data and Analytics: Clinical Decision Support and Patient Empowerment

Abstract: In this talk I will discuss the research opportunities in health big data analytics. Based on research partnership in Asia, I will present our ongoing research in disease progression modeling using Electronic Health Record (EHR) and adverse drug event extraction using patient social media. A working prototype system DiabeticLink which aims to support healthcare decision making and patient empowerment for diabetes patients will be introduced. Research partnership in Asia for health analytics research is sought. I will also introduce the NSF Smart and Connected Health (SCH) research program, which I will lead in August 2014.

KEYNOTE 2

Ramayya Krishnan

Dean of Heinz College
William W. and Ruth F. Cooper Professor of Management Science and Information Systems
Carnegie Mellon University

Bio: Ramayya Krishnan holds the John Heinz III Dean's Chair and is the W.W. Cooper and Ruth F. Cooper Professor of Management Science and Information Systems at the H. John Heinz III College at Carnegie Mellon University. The College is home to CMU's School of Information Systems and Management and to its School of Public Policy and Management. He completed his studies in Engineering, Information Systems and Operations Research at the Indian Institute of Technology (Madras) and the University of Texas at Austin. His current areas of research interest are in large scale network analysis and in the development of methods that balance confidentiality and data access. He works on data intensive problems that arise in a variety of business and application domains.

Keynote #2: Accelerating Digitization: Opportunities for Innovation

Abstract: The pace of change and digitization-led transformations enabled by Information technology (IT) in important societal and business domains is accelerating. This has led to significant innovation resulting in the creation of new services and business models (e.g., the latest example being the Ubers and Airbnb's of the sharing economy) and in the highly visible role for IT in major societal initiatives such as the rollout of the Affordable Care Act. Vulnerabilities associated with this digitization have also been recently in the news – from the contentious discussions about surveillance to the widely reported security breaches at companies such as Target. Hand in hand with these developments has been the emergence of data science or big data as a field in its own right taking advantage of two unique capabilities enabled by information technology-based platforms. The first pertains to opportunities to instrument collection and analysis of either large amounts of fine grained data or highly contextualized data about phenomena of interest (e.g., about people and their behaviors). The second is the capacity to conduct near real time, closed loop experiments at the scale required by the analysis to draw causal inferences. As evidenced by the considerable interest shown by Private and Government Research Foundations (Sloan, Moore, NSF, NIH and DOT), there are a number of interesting opportunities for innovation using a research, development and deployment metaphor I have introduced in prior addresses. In this talk, I will highlight using business and societal examples some of the important research and innovation opportunities that lie at the intersection of data science, social science and computing.

PANEL 1

Opportunities for Research Collaborations in China

Guoqing Chen	Tsinghua University
Hsinchun Chen	University of Arizona
Jeffrey Hu (Moderator)	Georgia Institute of Technology
Yongkai Ma	University of Electronic Science and Technology of China
Qiang Ye	Harbin Institute of Technology
Leon Zhao	City University of Hong Kong
Wei Zhang	Tianjin University

PANEL 2

Future of Digital Innovation and Transformation

Ravi Bapna	University of Minnesota
Ramnath Chellappa	Emory University
Prabuddha De (Moderator)	Purdue University
Anindya Ghose	New York University
Ramayya Krishnan	Carnegie Mellon University
Giri Kumar Tayi	University at Albany-SUNY

TABLE OF CONTENT

How Costs and Heterogeneous Consumer Price Sensitivity Interact with Add-On Pricing

Xianjun Geng
The University of Texas at Dallas
geng@utdallas.edu

Jeffrey D. Shulman
University of Washington
jshulman@u.washington.edu

Abstract

Firms often cite cost savings as a reason why they charge separately for add-ons. For example, some airlines claim that they charge check-in luggage fees to discourage passengers from packing too much unnecessary baggage. While such cost savings may directly translate into profit gains in some scenarios, this paper examines the strategic implications of add-on pricing and is the first to suggest that cost savings from add-on pricing may in fact result in profit loss for firms. This is because add-on pricing can trigger a revenue loss that exceeds any cost savings, thus leading to a negative net profit change for competing firms. We further show the possibility that the greater the cost of providing the add-on service (and the greater the cost savings generated from add-on pricing), the worse this profit loss gets.

Keywords: add-on pricing, competitive price discrimination

1. Introduction

Firms often recognize that some consumers are more costly to serve than others and consequently choose to charge extra for certain add-ons that were previously included in the base price. Charging add-on prices can sometimes reduce usage or at least compensate the firm for the costs of providing the add-ons. In discussing the decision to charge checked baggage fees and other add-on prices, Ryanair CEO Michael O'Leary stated, "It was about persuading people to change their travel behavior –to travel with carry-on luggage only…That helps us significantly lower airport and handling costs" (Michaels 2009). Similarly, banks incent customers to use automated processes for transactions by charging add-on fees for making transactions through a customer service representative. In explaining the add-on fee for extra condiments, a manager of a McDonald's franchise said, "We want to control condiment cost."

In addition to reducing costs, add-on prices have created an opportunity to generate differential revenue from the customers who value the add-on products or services. For example, many hotels charge for parking as well as Internet, safe, and pool access. Banks have drawn attention for the long list of add-on fees they charge. This paper allows for both a cost-reducing role and a price discrimination role of add-on pricing and examines the interaction of the two. We develop an analytical model to address the following research questions:

1. When will add-on pricing diminish, increase, or maintain firm profit relative to all-inclusive pricing?

2. How do the cost savings associated with add-on pricing affect the profit of add-on pricing relative to all-inclusive pricing?

3. How does consumer bounded rationality regarding add-on prices affect consumer welfare and the profitability of add-on pricing?

Our first key finding is that competing *firms can earn less profit with add-on pricing* than if prices from both firms are all-inclusive. The key contribution lies in the fact that we not only uniquely identify this possibility, but also give a clear roadmap as to when this result will occur.

The model identifies concrete measures that will lead add-on pricing to reduce, improve, or have no effect on firm profitability. Both cost-reduction factors and heterogeneous price sensitivity play critical roles in this finding.

Our second key finding is that *greater cost savings due to add-on pricing can decrease the profit differential that add-on pricing provides relative to all-inclusive pricing*. Prior to this model, common intuition would suggest that add-on pricing will be more profitable for firms when it has greater potential to decrease costs. However, we find testable market conditions that will counter this intuition.

2. Literature Review

We only review the research stream on add-on pricing due to space limit. A body of research shows add-on pricing results in equivalent profit as an all-inclusive price (Lal and Matutes 1994, Verboven 1999, Gabaix and Laibson 2006). In these models, any gains from the unadvertised price are competed away through the posted prices. Ellison (2005) and Shulman and Geng (2013) identify how add-on pricing can actually lead to improved profitability for firms. Allon, Bassamboo, and Lariviere (2012) find that add-on pricing can increase social welfare via a reduction in costs that benefits the firms and the consumers. In contrast to the aforementioned work, this paper uniquely considers the joint effect of cost-savings from add-on pricing and heterogeneity in consumer price sensitivity. There are several new implications of this modeling framework. This paper is the first to find how and when add-on pricing can reduce profit for symmetric firms. This paper is also the first to show how and when greater cost-savings associated with add-on pricing diminish the profitability of add-on pricing. Thus, this paper proposes a unique model that generates insights new to the add-on pricing literature.

3. The Model

Firms. We employ a standard Hotelling line of unit length to model horizontal competition between two firms. The marginal cost of providing the *base good (add-on)* is c_b (c_a) for each firm. The add-on can only be purchased with the base good.

As is standard in the literature (e.g., Ellison 2005; Gabaix and Laibson 2006) the price for the base good from each firm *j*, p_{jb} is readily observable by all consumers. When firms employ add-on pricing, the add-on price p_{ja} need not be observable by all consumers. We will first consider a model in which there is transparency in add-on prices. In a subsequent extension, we allow the add-on price to be unobserved by a proportion of consumers.

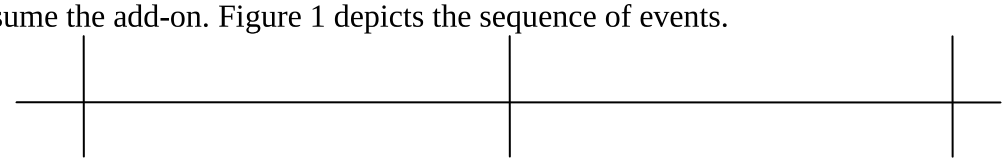
Consumers. Consumer *i* has known reservation utility v_b (which is common to all consumers) of consuming the base good and reservation utility v_{ia} from the add-on. The base good reservation utility is sufficiently high such that all consumers buy from one of the two firms. Consumers are uniformly distributed on the Hotelling line. For consumer *i* at location θ_i, she incurs disutility of $t_i\theta_i$ (or $t_i(1-\theta_i)$) if she buys from firm 1 (or 2), where transportation cost is denoted by t_i. Thus consumer *i's* utility from buying the base good and the add-on from firm *j* is given by $v_b + v_{ia} - t_i\,|\,\theta_i - x_j\,| - p_{jb} - p_{ja}$, and the utility from buying the base good only from firm *j* is given by $v_b - t_i\,|\,\theta_i - x_j\,| - p_{jb}$.

There are two types of segments. *Base consumers* do not value the add-on (i.e., $v_{ia} = 0$) and have transportation cost denoted by $t_i = t_b$. A proportion $(1-\alpha)$ of the consumer population falls into this segment. A proportion α are *non-base consumers* who value the add-on (i.e., $v_{ia} = v_a > 0$) and have transportation cost denoted by $t_i = t_a$. This modeling approach accounts for the possibility that consumers who value the add-on are inherently different and thus have a different sensitivity to price differences. This assumption is analogous to that of Gumus, Li, Oh and Ray (2012) who assume that consumers have heterogeneous price sensitivities over shipping and handling costs in online retailing.

We flexibly allow for both $t_b > t_a$ and $t_b < t_a$. This flexibility will have importance in determining the equilibrium outcome and thus merits further discussion. Note that a consumer's valuation of an add-on product or service depends on its personal relevancy, the usage context, and the availability of alternatives among other factors. Holding these factors constant, one might expect consumers who value the add-ons to be less price-sensitive given a likely lower marginal utility of income. However, it is possible that the personal relevancy of the add-on is systematically correlated with the price sensitivity. For example, business travelers are less price sensitive than families, yet someone traveling for business may not value add-on services such as checking bags, accessing the internet or the pool because they may pack light, have their own internet access, and be short on leisure time.

We also allow for a proportion ρ of the base consumers who mindlessly consume the add-on if it is provided for free and will not consume the add-on if $p_a > 0$. This incorporates the so-called "penny gap" where "in most cases, just a penny - a seemingly inconsequential price - can stop the vast majority of consumers in their tracks" due to the imposed cost of thinking. We refer to this segment as *consumptors* because their use of the add-on when free is wasteful in that it costs the firm money and does not provide utility to the consumer. The remaining ($1-\rho$) base consumers, who we refer to as *base-only consumers*, will not consume the add-on.

Timing of the Model. The sequence of the game follows conventions of previous literature (e.g., Gabaix and Laibson 2006; Muir, Seim, Vitorino 2013; Shulman and Geng 2013). Firms first choose their posted and add-on prices simultaneously. Consumers then choose from which firm to buy and subsequently cannot switch firms. Finally, consumers decide whether or not to consume the add-on. Figure 1 depicts the sequence of events.



Stage 1	*Stage 2*	*Stage 3*
Firms 1 and 2 simultaneously set p_{1b}, p_{1a} , p_{2b} and p_{2a}.	Consumers choose which base good to buy	Consumers choose whether or not to buy add-on

Figure 1: Sequence of Events

4. Results

In this section, we identify the profit implications of firms having the choice to use add-on pricing. In this light, we compare the equilibrium profit when firms may choose an add-on price to the equilibrium profit to when firms are restricted to an all-inclusive posted price with $p_{ja} = 0$. We also examine how the difference in profit between these two cases changes with the marginal cost of the add-on, c_a, and the number of consumptors, ρ.

Using backward induction, we first determine demand for each firm from the $(1-\rho)(1-\alpha)$ base-only consumers, the $\rho(1-\alpha)$ consumptors (these two segments together are base consumers) and the α non-base consumers. To this end, we solve for the locations of the consumer in each segment who is indifferent between purchasing from firm 1 and purchasing from firm 2 by comparing utilities of each action. Consequently, demands for firm j's base good, γ_{jb}, and add-on, γ_{ja}, can be written as functions of prices:

$$\gamma_{jb} = \frac{1}{2} + \frac{(1-\alpha)(p_{-jb} - p_{jb})}{2t_b} + \frac{\alpha(p_{-jb} - p_{jb} + p_{-ja} - p_{ja})}{2t_a}, \qquad (1)$$

$$\gamma_{ja} = \frac{\alpha}{2} + \frac{\alpha(p_{-jb} - p_{jb} + p_{-ja} - p_{ja})}{2t_a} + \Phi\rho(1-\alpha)(\frac{1}{2} + \frac{(p_{-jb} - p_{jb})}{2t_b}) \qquad (2)$$

where $\Phi = 1$ if $p_{ja} = 0$ and $\Phi = 0$ otherwise. Appendix A provides detailed derivations of all results. Revenue (and cost) for firm *j* is given by $r_j = p_{jb}\gamma_{jb} + p_{ja}\gamma_{ja}$ (and $C_j = c_b\gamma_{jb} + c_a\gamma_{ja}$), therefore profit for firm *j* is given by

$$\pi_j = (p_{jb} - c_b)\gamma_{jb} + (p_{ja} - c_a)\gamma_{ja} \qquad (3)$$

Under all-inclusive pricing, each firm chooses a single posted price (i.e., $p_{1a} = p_{2a} = 0$) to maximize its own profit. The unique equilibrium is symmetric, under which firm *j*'s optimal all-inclusive price is $\hat{p}_{jb} = c_b + \frac{t_b t_a + c_a(\alpha t_b + (1-\alpha)\rho t_a)}{\alpha t_b + (1-\alpha)t_a}$. Revenue, cost and profit are respectively $\hat{r}_j = \frac{t_a t_b + c_b((1-\alpha)t_a + \alpha t_b) + c_a((1-\alpha)\rho t_a + \alpha t_b)}{2(1-\alpha)t_a + 2\alpha t_b}$, $\hat{C}_j = \frac{1}{2}((\alpha + \rho - \alpha\rho)c_a + c_b)$ and $\hat{\pi}_j = \frac{(1-\alpha)(1-\rho)\alpha c_a(t_b - t_a) + t_a t_b}{2(1-\alpha)t_a + 2\alpha t_b}$.

Under add-on pricing, each firm chooses its own posted price and add-on price to maximize its own profit. The lemma below identifies the equilibrium under add-on pricing.

Lemma 1 *Under add-on pricing, if* $v_a - c_a > t_a - t_b$ *then each firm* j *charges posted price* $p^*_{jb} = c_b + t_b$ *and add-on price* $p^*_{ja} = c_a + (t_a - t_b)$. *Firm revenue, cost and profit are respectively* $r^*_j = \frac{1}{2}t_b - \frac{\alpha}{2}(t_b - t_a) + \frac{1}{2}(\alpha c_a + c_b)$, $C^*_j = \frac{1}{2}(\alpha c_a + c_b)$ *and* $\pi^*_j = \frac{1}{2}t_b - \frac{\alpha}{2}(t_b - t_a)$.

If $v_a - c_a \le t_a - t_b$ *then each firm j charges posted price* $p^*_{jb} = c_b + \frac{t_b(t_a - \alpha(v_a - c_a))}{(1-\alpha)t_a + \alpha t_b}$ *and add-on price* $p^*_{ja} = v_a$. *Firm revenue, cost and profit are respectively* $r^*_j = \frac{1}{2}(c_b + \frac{(\alpha c_a + t_a)t_b + (1-\alpha)\alpha(t_a - t_b)v_a}{(1-\alpha)t_a + \alpha t_b})$, $C^*_j = \frac{1}{2}(\alpha c_a + c_b)$ *and* $\pi^*_j = \frac{(1-\alpha)\alpha(v_a - c_a)(t_a - t_b) + t_a t_b}{2(1-\alpha)t_a + 2\alpha t_b}$.

Before we present the results on how add-on pricing affects firm profits (as compared to all-inclusive pricing), we first present results on cost-saving and revenue change.

Lemma 2 *Compared to all-inclusive pricing:*

i). Each firm saves a cost of $(1-\alpha)\rho c_a / 2$ *by add-on pricing.*

ii). If $t_b < t_a$ *and* $\rho < \alpha(t_a - t_b) \bullet \min\{v_a, (c_a + t_a - t_b)\} / (c_a t_a)$, *add-on pricing increases each firm's revenue; otherwise, add-on pricing decreases each firm's revenue.*

When firms lose revenue under add-on pricing (as compared to all-inclusive pricing), can this loss overwhelm their cost savings? In the following proposition, we compare firm profit when practicing add-on pricing to when the posted price is all-inclusive.

Proposition 1 *When* $t_b > t_a$ *and if the proportion of consumptors is not too great (i.e.,* $\rho < 1 - (t_b - t_a)/c_a$*), the negative revenue-loss effect dominates the positive cost-saving effect and implies profit is lower with add-on pricing. If either* $\rho > 1 - (t_b - t_a)/c_a$ *or* $t_b < t_a$*, add-on pricing results in greater profit than all-inclusive pricing.*

Proposition 1 is the first in the literature to show that add-on pricing by competing firms -- even if it has a clear cost-saving effect -- can result in diminished profitability for both firms. While much of the literature find a profit-irrelevance effect of add-on pricing (e.g., Lal and Matutes 1994, Gabaix and Laibson 2006), and Shulman and Geng (2013) find that a vertically differentiated firm can gain from add-on pricing at its competitor's expense, this result shows that both competing firms can be disadvantaged by the advent of add-on pricing. Though both firms are free to choose $p_{ja}=0$ to effectively reproduce all-inclusive pricing, in equilibrium both firms choose $p_{ja}>0$ even though the net result of add-on pricing is a reduction in profit.

One immediate managerial implication of Proposition 1 is that, when firms in a competitive market consider charging separately for add-ons, they need to consider both the cost-saving effect and the (possible) strategic revenue-loss effect. Specifically, when consumers who value add-ons have greater price-sensitivity, firms should be cautious that add-on pricing could result in a significant revenue reduction. Thus measures to reduce costs by charging for add-ons can have an unintentional and surprising consequence of actually reducing profit.

If consumers who value the add-on are less price sensitive than consumers who only value the base offering (i.e., $t_a>t_b$), firms are able to charge high add-on fees because non-base consumers are the least price sensitive type -- a dynamic opposing the one under the first scenario. As a result, add-on pricing improves firm profits as compared to all-inclusive pricing. Also, if implementing add-on pricing results in a significant reduction in add-on usage (i.e., ρ is closer to 1), then the cost savings effect overwhelms the revenue effect of add-on pricing.

We now turn our attention to studying individual factors that affect the cost-saving effect. From Lemma 2 we know that add-on pricing leads to cost savings of $(1-\alpha)\rho c_a/2$. Therefore, a higher marginal cost of the add-on c_a or a larger proportion of consumptors among base consumers ρ (who will discontinue add-on usage in response to add-on pricing) will lead to higher cost savings. One might intuitively think that the benefit of add-on pricing should be increasing in each of these values. However, the following two propositions highlight the strategic consequences of these cost savings.

Proposition 2 *When either $t_a<t_b$ or $t_a>v_a-c_a+t_b$: the higher the marginal cost saved by add-on pricing (i.e., c_a), the lower the benefit of add-on pricing (i.e., $\pi_j^*-\hat{\pi}_j$) to firms.*

Proposition 2 shows that, surprisingly, higher cost savings from add-on pricing due to a higher marginal cost of add-on do not necessarily translate into higher firm profits. To understand this result, it is beneficial for us to first consider the impact of c_a on the cost-saving effect and the revenue-loss effect, respectively. From Lemma 2 we know that $\partial(\text{costs saved by add-on pricing})/\partial c_a=(1-\alpha)\rho/2$. Notice that the above marginal impact of c_a on the cost-saving effect is not related to price sensitivity. On the other hand, if $v_a-c_a>t_a-t_b$:

$$\partial(\text{revenue under add-on pricing - revenue under all-inclusive pricing})/\partial c_a$$
$$=\partial(p_{jb}^*/2+p_{ja}^*\alpha/2-\hat{p}_{jb}/2)/\partial c_a \quad =(\alpha-\partial\hat{p}_{jb}/\partial c_a)/2 \qquad \text{(From Lemma 1)}$$

Therefore, if a firm loses revenue under add-on pricing, the marginal impact of c_a on this revenue loss is linearly related to $\partial\hat{p}_{jb}/\partial c_a$: the higher $\partial\hat{p}_{jb}/\partial c_a$ is, the more significant is the marginal impact of c_a on this revenue loss associated with add-on pricing. Overall, Proposition 2 says that, when the non-base consumers are more price sensitive than the base ones (i.e., $t_b/t_a>1$), the marginal impact of c_a on the revenue-loss effect is strong due to competition-dampening under all-inclusive pricing, and it overwhelms the marginal impact of c_a on the cost-saving effect. Therefore the benefit of add-on pricing to a firm decreases in c_a.

We next consider how the benefit of add-on pricing changes with the number of consumptors (i.e. consumers who decide to stop add-on usage only when it carries a positive price).

Proposition 3 *When $t_a > t_b$: the more wasteful add-on usage prevented by add-on pricing (i.e., ρ), the lower the benefit of add-on pricing to firms.*

Proposition 3 shows that the effect on the benefit of add-on pricing of serving a greater number of *consumptors* who wastefully use the add-on when it is free can be in opposition to the effect of c_a. The revenue effect of increasing ρ can be expressed mathematically:

$$\partial(p^*_{jb}/2 + p^*_{ja}\alpha/2 - \hat{p}_{jb}/2)/\partial\rho$$
$$= -(\partial\hat{p}_{jb}/\partial\rho)/2 = -(1-\alpha)c_a t_a / 2(\alpha t_b + (1-\alpha)t_a) \qquad \text{(From Lemma 1)}$$

Note that this effect holds regardless of whether add-on prices are set at their maximum (i.e., $t_a > v_a - c_a + t_b$) or not. The issue becomes determining when $\partial\hat{p}_{jb}/\partial\rho$ will be large. Following a related logic as above, when t_b/t_a *decreases* non-base consumers are getting *less* price sensitive relative to base consumers and thus a price cut under all-inclusive pricing attracts increasingly more base consumers than non-base consumers. However, a higher ρ diminishes the value to a firm of attracting a base consumer because there is a ρ probability the base consumer will be a consumptor who carries an additional cost to serve. As a consequence more wasteful add-on usage dampens each firm's incentive to cut the all-inclusive price, and this dampening effect is strengthened when t_b/t_a decreases. When the non-base consumers are less price sensitive than the base consumers (i.e., $t_b/t_a < 1$), the profit benefit of add-on pricing is positive but is decreasing in ρ. Thus the competition dampening revenue effect of ρ drives the counter-intuitive finding of Proposition 3 that the benefit of add-on pricing can be decreasing in the number of consumers for whom it lowers the cost-to-serve .

Collectively, Propositions 2 and 3 demonstrate that the source of cost reduction plays a critical role in its effect on the profitability of add-on pricing and interacts with the nature of heterogeneity in consumer price sensitivity. In fact, it is possible for a greater reduction in the number of add-on users to diminish the benefit of add-on pricing in circumstances such that the marginal cost of the add-on has the opposite effect.

5. Concluding Remarks

From a practical standpoint, there is some truth to a 2009 statement made by Jeff Smisek, then CEO of Continental Airlines: "*Airlines are really depending on fees in order to stay in business. It's either that or higher fares, or airlines are going to go under.*" Our model shows that without add-on fees, the equilibrium posted prices will be higher with all-inclusive pricing than with the add-on fees. However, there are several market characteristics that can create a recipe for contradicting the CEO's statement regarding the value to firms of charging add-on fees. Our model shows that, while a firm can save costs by using add-on pricing to discourage some consumption, doing so does not automatically lead to improved profits because charging separately for add-ons can lead to strategic and negative consequences on the revenue side that might negate the benefit from cost savings. Our research thus provides theoretically new and managerially relevant insights on when and how add-on pricing leads to a stronger revenue-loss effect than the cost-savings effect, thereby actually resulting in diminished profitability.

(Proofs and references omitted due to page limit.)

Comparing Two Rating Mechanisms in Crowdsourcing Contests

Liang Chen
University of Kentucky
liang.chen@uky.edu

Pei Xu
University of Kentucky
pei.xu@uky.edu

De Liu
University of Kentucky
de.liu@uky.edu

Abstract

Designing crowdsourcing contests is a central issue for any crowdsourcing platform. This study highlights an important design element of crowd sourcing contests: the rating schemes. To our knowledge, no comparison has been made between crowd-rated contests and expert-rated contests. Using a unique data set that permits both kinds of rating mechanisms in the same contest, we are the first study to examine the differences between crowd-based and expert-based rating mechanisms from the perspective of winning probability. Based on our empirical analysis, we find that winning expert- and crowd-rated prizes requires very different past performance, entry timing, and social capital. Winning an expert prize decreases future participation while winning a crowd prize increases it. Finally, the probability of participation has a U-shaped relationship with the ratio of crowd prize sum. Increase in the crowd prize ratio discourages expert-prize winners.

Keywords: Crowdsourcing contest, Expert Rating, Popular Voting, Rating Mechanism

1. Introduction

Crowdsourcing platforms (CSPs) are Internet platforms that let firms, governments, and individuals acquire solutions, seek creative ideas, and complete projects using a "crowd" of Internet users (Yang et al. 2011). In recent years, CSPs have grown in popularity and importance in a number of areas such as idea generation, product development, marketing, R&D innovation, problem solving, free-lancer jobs, and application developments. Many CSPs, especially those using contests as an organizing mechanism, rate and choose entries using various rating mechanisms, including subjective expert ratings and crowd ratings, along with automated objective rating mechanisms (e.g. algorithm-based ratings for coding/programming contests). This paper aims to investigate differences between two subjective rating mechanisms, namely expert and crowd ratings, in the context of crowdsourcing contests for creative work. Rating mechanisms are central to the CSP design, as they pertain not only to the quality of winning (or promoted) entries, but also the participation into these contests and the strategies one may adopt in order to win (or get promoted). This issue, despite its critical importance, has so-far been overlooked in the literature (Bullinger and Möslein, 2011).

We address our research questions from two perspectives: (a) whether/how do expert and crowd ratings differ in terms of *winner determination*? (b) whether/how do expert and crowd-based rating mechanisms differ in terms of attracting *participation*? For the first perspective, we identify three key winner determinants based on the literature: past performance, submission timing, and social capital. We examine whether the effects of these determinants differ for winning expert-rated and crowd-rated prizes. For the second perspective, we compare the effects of expert- and crowd-rated prizes on a user's participation decision (award appeal) and the effects of past winning different types of prizes (award impact). By comparing expert and crowd rating mechanisms in award appeal, winner determination, and award impact, we develop a more holistic understanding of two rating mechanisms.

2. Hypotheses Development

2.1 Winner determination

Both experts and crowds have been used to judge the quality of contest entries. Prior accounts suggest that there are differences in expert and crowd judgments. However, because quality is usually latent, and we do not directly observe the decision criteria used by these raters, we choose to infer difference in judgment criteria by examining differences in winning determinants for expert- and crowd-rating prizes. Among the winning determinants examined in previous crowdsourcing contests, several seem most frequently used: contestants' past performance, submission timing, and contestants' social capital. We discuss them below in turn.

2.1.1 Past Performance

Past performance indicates how a contestant performs in the past. In our context, past performance can be represented by the ratio of the number of prizes a contestant has won under a particular rating scheme relative to the number of entries the contestant has submitted (i.e., *past experience*). Prior research has found that past performance is strongly correlated with the future winning probability in crowdsourcing contests. Due to the many differences in expert and crowd rating mechanisms, we consider that characteristics used by expert raters for quality assessments to generally differ from those by crowd raters. Though we do not directly observe these decision criteria, an entry that wins an expert prize may embody such characteristics favored by expert raters. This suggests that the submitter of the entry has or knows what it takes (e.g., skills, sophistication, tastes, strategies) to gain favorable evaluation of experts. This would afford the contestant a higher chance to win an expert prize in the future (Terwiesch and Yi, 2008). Similarly, contestants who win crowd prizes in the past may have a better chance of winning a crowd prize in the future. However, two kinds of raters generally use divergent decision criteria, so past performance embodies different decision rules, and one type of past performance may not carry over to the other. Therefore:

H1a-b*: An entry by a contestant with higher past performance in expert prizes is (a) more likely to win an expert prize; (b) NOT more likely to win a crowd prize.*

H1c-d*: An entry by a contestant with higher past performance in crowd prizes is (c) more likely to win a crowd prize; (d) NOT more likely to win an expert prize.*

2.1.2 Submission Timing

In a crowdsourcing contest, a contestant can strategically choose when to submit her entry as long as the contest is open. After the contest ends, expert raters typically rate all or most entries in a systemic fashion. Because expert raters focus more on the intrinsic quality of submissions, late entries give contestants the needed time to improve their ideas and thus may have better quality and win favorable expert evaluations than entries that were submitted early. Additionally, in contests where entries are posted publicly, a contestant can improve her entry by learning from good elements in other entries. In contrast, crowd raters vote any time before the deadline. They tend to arrive stochastically and may not make additional effort to come back again for the later entries. Thus, early entries get noticed by more crowd voters and have an advantage. Adding this advantage is herding effects – that is, online users also copy the decisions of those who proceed from them (Duan et al., 2009). Indeed, entries which have already received many votes may get more crowd attention and thus build up their votes more quickly than those who received few votes (Muchnik et al., 2013). One reason for such herding effect is that raters as peer users may be attracted to entries with more votes, and thus are more likely to vote on these entries. Thus:

***H2a-b**: A late entry is more (less) likely to win an expert prize (crowd prize).*

2.1.3 Social Capital

Social capital is defined as resources which are embedded in the crowdsourcing community and enable a community member to coordinate or to achieve desired goals via her connections or interactions with other members. A distinctive feature of our study context is that the crowd voters are connected by social networks. In particular, a contestant can follow/be followed by others as well as give and/or receive comments from others. When users are socially connected to potential raters, their likelihood of getting votes and the valence of the vote are likely affected by the social relation between users and crowd raters. A social connection or interaction increases the awareness of each other. This alone can increase the chance that a user's entry is rated by others. Social network research has shown that socially connected individuals are more likely to offer and return favor. A contestant may mobilize her social capital to increase her self-promotion performance, leading to a higher probability of getting crowd votes. These effects suggest that higher social capital would lead to a higher chance of winning crowd prizes. Another view holds that contestants with great social capital can also gain access to diverse information, which enables them to reduce fixation effects and generate new ideas favored by the crowd (Bayus, 2013). However, extensive exposure to the crowd may cause a contestant to deviate from her original design toward compromised alternatives favored by the crowd, and thus may not win favor from the experts (Hildebrand et al., 2013). Therefore, we hypothesize:

***H3a-b**: An entry by a contestant with greater social capital is (**a**) NOT more likely to win an expert prize; (**b**) more likely to win a crowd prize.*

2.2 Participation

2.2.1 The Effect of Past Performance under Two Rating Mechanisms

Winners of expert and crowd prizes may behave differently when faced a future contest. A contestant with higher past performance in expert prize is less likely to participate in a future contest for a few reasons. First of all, expert prizes may be considered more creditable signal of a contestant's talents. Once getting an expert panel's seal of approval, the marginal utility of getting another prize is substantially decreased, implying that a contestant is less likely to participate in a future contest. Indeed, in some cases, experts may offer expert-prize winners an outside opportunity and thus further reduce the chance of the contestant returning to future contestants. Even if expert-prize winners stay for future contests, they are likely more selective as the learning benefits for them are low and they may not submit an entry if they deem the winning chance to be low (Archak and Ghose, 2010). In contrast, contestants with higher past performance in crowd prizes are more likely to participate in a future contest because the more crowd prizes they win, the more identified they become with this community. As the popularity they have gained with this community is not portable, they are more likely stay with the community and become increasingly involved. Thus:

***H4a-b**: The higher past performance in expert prizes (crowd prizes), the less (more) likely to participate in a new contest.*

2.2.2 The Effect of Crowd Prize Ratio

Previous studies find that the amount of rewards can increase the number of contestants. When expert and crowd prizes coexist in a same contestant, they may appeal to a different crowd of participants. As we have argued earlier, expert and crowd ratings have differentiated winning determinants. In other words, they appeal to contestants with different skill levels, goals, and strategies. For example, those who are more likely to win a crowd prize may choose a contest that offers predominant crowd prizes and vice versa. When the ratio of crowd prizes (i.e., *crowd ratio*) is very low or even zero, potential contestants who have an advantage in crowd prize may avoid it because they are not competing to their advantage and they do not see this an opportunity of engaging the community. Likewise, when crowd ratio is very high or even one (the maximum value), potential contestants who are oriented in winning expert prizes may perceive this contest as a popularity game and refuse to participate. Thus, a contest providing either pure expert prizes or pure crowd prizes can cater to a particular kind of participants. On the other hand, a contest with an intermediate crowd ratio sends a mixed signal to potential contestants. Contestants are not sure whether the contest is suitable for them and unclear what strategies are needed to win the contest, resulting in a low participation probability. Thus:

H5: *There is a U-Shaped relationship between crowd ratio and participation probability.*

3. Methodology and Results

3.1 Data Collection and Model Selection

We collect contest data from Zooppa, a global social network of creative talents who help partner companies launch user-generated advertising campaigns. Zooppa was founded in Italy in 2007, and then launched in the U.S. in 2008. Zooppa not only permits multiple prizes for each contest, the client (i.e. sponsor of the contest) can also give both crowd and expert (client) prizes.

Variables in Table 1 are collected from Zooppa. Past performance in expert prizes (crowd prizes) is measured by *hitrateexpert* (*hitratecrowd*). Social capital is measured by two variables: *commentsgiven_before* and *commentsreceived_before*. Submission timing is measured by *gsubmitOrder*. We calculate past performance and social capital of a contestant before entry submission, and thus remove the possibility of reverse causality. All the three dependent variables (*winExpertPrize*, *winCrowdPrize*, and *Participation*) are binary (i.e., 1, 0). We also collect controlled variables such as *tenure*. Overall, our dataset includes all the contests between Dec. 10, 2007 and Feb. 8, 2013, resulting in a total of 110 contests and 7,407 video entries (we focus on video entries and do not collect information on other entries).

We estimate conditional logit models to test our hypotheses. For H1-H3, we estimate the probability of an entry winning a contest conditional on the total number of winning entries in this contest. Because time-invariant contest characteristics are common for all entries, they are canceled out in a conditional logit model. All contest level factors such as number of competitors are canceled out. Because entries are not random, above estimations are subject to selection bias caused by endogenous entries. We therefore use the Heckman correction to correct the likely selection bias. For H4-H5, we estimate the probability of participation conditional on the total number of participations by a contestant (i.e., a conditional logit model grouped by contestants).

3.2 Results

Our results are presented in Table 2 (H1-H3) and Table 3 (H4 & H5). In particular, H1a-d are strongly supported, indicating a positive effect of past performance on winning probability for the same rating mechanism and one type of past performance does not carry over to the other. Interestingly, past performance in crowd prizes is negatively associated with the probability of winning an expert prize, suggesting that the community may view an expert-prize winner as an outsider. H2a and H2b are strongly supported, indicating that different temporal strategy are required for winning two types of prizes. Whereas H3a is supported, H3b is partially supported. A contestant's outgoing, rather than incoming social interactions help her win a crowd prize.

H4a and H4b are strongly supported, demonstrating that two types of prizes vary greatly in promoting a contestant's future participation. H5 is supported as well. When we further examine the interaction between crowd ratio and past performance, we find that a high crowd ratio discourage many contestants with high performance in expert prizes (namely, talented contestants) – consistent with our finding that past performance in expert prizes negatively affects winning future crowd prizes. However, there is no significant interaction effect between crowd ratio and past performance in crowd prizes. In addition, consistent with previous studies, we find both total awards and the number of awards significantly increase participation.

4. Conclusion

Designing crowdsourcing contests is a central issue for crowd sourcing platforms. This study highlights an important design element of crowd sourcing contests: the rating schemes. Previous theoretical literature has distinguished different evaluation schemes but not tested them. Using a unique data set that permits both kinds of rating mechanisms in the same contest, we contribute to theory testing by examining the differences between crowd-based and expert-based rating mechanisms from the perspectives of winning probability and of contest participation. Furthermore, our empirical analysis suggests that expert and crowd rating mechanisms are very different in terms of award appeal, winning determination, and award impact. These findings will help a CSP choose between different mechanisms.

5. Acknowledgement

Due to space limitation, literature review and numerous references are omitted in this paper. Researchers may contact with any author for the available references.

References

1. Archak, N., and Ghose, A. "Learning-By-Doing and Project Choice: A Dynamic Structural Model of Crowdsourcing," *ICIS 2010 Proceedings*, St. Louis, MO, 2010.
2. Bayus, B.L. "Crowdsourcing new product ideas over time: an analysis of the Dell IdeaStorm community," *Management Science* (59:1), 2013, pp. 226-244.
3. Bullinger, A.C., and Moeslein, K. "Innovation Contests: Systematization of the Field and Future Research," *International Journal of Virtual Communities and Social Networking* (3:1), 2011, pp. 1-12.
4. Duan, W., Gu, B., and Whinston, A. "Informational cascades and software adoption on the internet: an empirical investigation," *MIS Quarterly* (33:1), 2009, pp. 23-48.
5. Hildebrand, C., Häubl, G., Herrmann, A., and Landwehr, J.R. "When social media can be bad for you: Community feedback stifles consumer creativity and reduces satisfaction with self-designed products," *Information Systems Research* (24:1), 2013, pp. 14-29.
6. Muchnik, L., Aral, S., and Taylor, S.J. "Social influence bias: A randomized experiment," *Science* (341:6146), 2013, pp. 647-651.
7. Terwiesch, C., and Yi, X. "Innovation Contests, Open Innovation, and Multiagent Problem Solving," *Management Science* (54:9) 2008, pp. 1529-1543.
8. Yang, Y., Chen, P.-y, and Banker, R. "Winner Determination of Open Innovation Contests in Online Markets," *ICIS 2011 Proceedings*, Shanghai, China, 2011.

Table 1. Variable Definitions

Variables	Definition
winExpertPrize	Whether or not an entry was selected as an expert prize (win, 1; otherwise, 0)
winCrowdPrize	Whether or not an entry was selected as a crowed prize (win, 1; otherwise, 0)
Participation	Whether or not an contestant submitted her entry to a contest (participate, 1 otherwise, 0)
hitrateexpert	The ratio of the number of expert prizes a contestant has won relative to the number of entries he submitted
hitratecrowd	The ratio of the number of crowd prizes a contestant has won relative to the number of entries he submitted
videoExperience	The number of video entries the contestant has submitted in the past.
scores	The score an entry gets from crowd raters.
tenure	The number of years from the time a contestant first started her activity (e.g. commenting or submitting)
commentsgiven_before	The number of comments posted by a contestant before a contest (a measure of social capital)
commentsreceived_before	The number of comments received by a contestant before a contest (a measure of social capital)
gsubmitOrder	Equals submitOrder (the order of entry when submitted) divided by total number of entries. contestant
crowdratio	The ratio of the amount of crowd prizes relative to the total amount of prizes in a contest

Table 2. Winner Determination

Variables	winExpertPrize (2SLS)	winCrowdPrize (Condit Logit)
hitrateexpert	0.108***	0.027
hitratecrowd	-0.117**	2.121***
gsubmitorder	0.108***	-1.615***
logvideoexperience	-0.002	-0.596*
tenure	-0.000	-0.000
logcommentsgiven_before	-0.005	0.393***
logcommentsreceived_before	0.009	0.291
lambda	-0.002	-0.244
logscores	0.071***	
Adjust R-squared	0.024^	0.139
N	6791	2871

* p<0.05, ** p<0.01, *** p<0.001; Other control variables are not shown in this Table.
^ denotes that R square comes from the OLS regression.

Table 3. Participation

Variables	Participation (Conditional Logit)_
hitrateexpert	-0.301**
hitratecrowd	0.544**
tenure	1.408***
$crowdratio^2$	11.314***
crowdratio	-11.302***
hitrateexpert × crowdratio	-2.653***
hitratecrowd × crowdratio	-0.138
totalawardsum	0.006***
numawards	0.034***
Adjust R-squared	0.227
N	218205

* p<0.05, ** p<0.01, *** p<0.001; Other control variables are not shown in this Table.

Consumer-Generated Reviews and Social Learning: Implications for Quality Perception of New Products

Jiao Feng
School of Economics and Management,
Beihang Univetsity
fengjiao198705@126.com

Zhong Yao
School of Economics and Management,
Beihang Univetsity
iszhyao@buaa.edu.cn

Abstract

We investigate the implications of consumer-generated reviews and social learning on consumers' purchase decisions and a monopolist firm's profit by consumers updating their quality perception for a new product. We consider two modes of social learning, for perfect social learning, we use Bayesian updating model to stimulate the process of consumers' social learning, but for imperfect social learning, we use another model to almost representation of consumer behavior in real-world settings. The results show that the rating of reviews has an opposite trend with the quantity of reviews in the long term, and firms have higher profit and potential consumers have higher quality perception of products under imperfect social learning than perfect social learning. The paper gives some effective implications and suggestions for online retailers to develop their marketing strategies.

Keywords: Social learning, Herding theory, Consumer-generated reviews, Self-selection, Market outcome

1. Research Problem

Consumer-generated reviews have been recognized as a major influence of customers' purchase decisions and product sales. When new and innovative products launch, whose quality is hard to assess before use, surveys suggest that online review websites play a greater important role in helping potential consumers objectively and comprehensively know the products than any other medium (DoubleClick 2004), such as advertising.

A large amount of scholars have shown that volume and valence positive associations with consumers' purchase intentions and product sales. Moreover, Dellarocas et al. (2004) proposed the valence (average numerical rating) of online reviews is a better forecaster of future movie revenues than other measures they considered. However, Duan et al. (2008) found that the number of online postings significantly influenced movies' box office revenues, while rating of online user reviews has no significant effect on box office sales. A small number of literatures focus on which one is more important than the other one affecting on product sales, and majority of them are using empirical methodology but not from theory perspective. More importantly, there has been considerably less research on the essential relationship between volume and valence of online reviews.

Consumers observe and learn reviews may help them adjust and evolve their estimates of product quality. Gaur and Park (2007) established a model of consumer learning and choice behavior in response to uncertain service in the marketplace. Li and Hitt (2008) consider self-selection biases and consumer learning that affect consumer purchase decisions and consumer surplus. All of them studies from consumer learning, individual learning from online reviews, not from social perceptive, in which individual not only learns from online reviews but also learns by observing the behaviors of other consumers (herding behavior). Few authors had analyzed social learning from online buyers' reviews. Celen et al. (2010) find that individual is willing to follow the advice given to them by their predecessor and that the presence of advice increases subjects' welfare. Ifrach et al. (2012) analyze the social learning mechanism from online reviews and others behaviors, and its influence on the seller's pricing decision. However, all of them not place emphasis on consumer-generated reviews and social learning impacting on market outcome (product sales and consumer surplus).

More importantly, in practice, we find an interesting phenomenon that rating of reviews has an opposite trend with quantity of reviews through investigated consumer-generated reviews of iphone from taobao's 576 online stores whose monthly sales exceed 20 before December 12^{th} 2013 (see figure 1). Another interesting thing is that majority of amount of reviews of online stores focus on below 1000, then we statistic online stores in different range of reviews. It shows that 91% online stores' quantity of reviews below 1000, average rating of reviews is 4.89 and hold 2% total sales volume (monthly), but just only 1% online stores' quantity of reviews surpass 5000, average rating of reviews is 4.83 and hold 35% total sales volume (monthly) (see figure 2). We intuitively think that online stores who have a great quantity of and low rating of reviews get more sales than those who have a small quantity of and high rating of reviews. They arouse us some questions. Whether the phenomenon is special or not? Why quantity and rating of consumer-generated reviews have negative relationship? And why numerous and low rating of reviews boost product sales? Whether they are contradict with original researches? In order to solve the problems, we had looked up literatures but not found answers, which stimulate us to explore the presence and implication of these problems by addressing several research questions:

- What is the relationship between quantity and rating of reviews, positive or negative?
- How do quantity and rating of reviews affect consumers' purchase decisions by social learning for launching a new product?
- How do quantity and rating of reviews affect market outcomes (sales and consumer surplus) and total welfare by social learning?

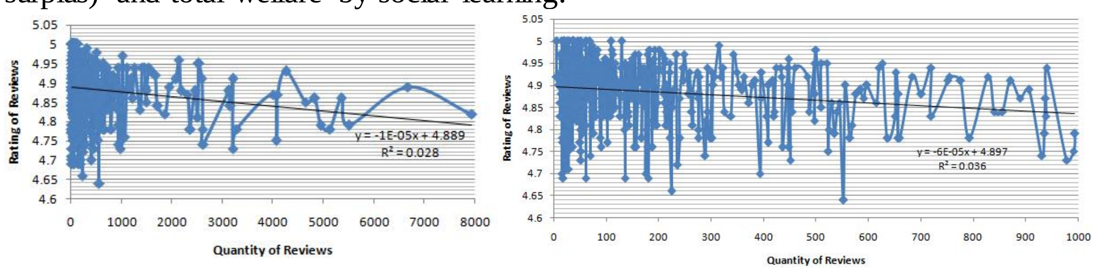


Figure 1 Quantity and rating of reviews of iphone 5 (data collected from Taobao.com via web crawler)

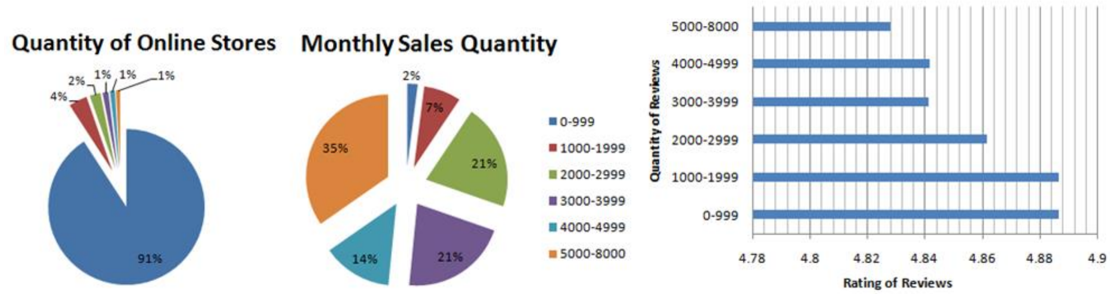


Figure 2 Online stores quantity and monthly sales quantity of iphone 5 in different regions of reviews (data collected from Taobao.com via web crawler)

2. Research Methodology

In this paper, we develop a mathematical model of social learning from consumer-generated reviews, and use it to analyze the role of consumer-generated reviews and the implications of social learning on consumers and firms, and then through making full use of social learning and consumer-generated reviews strategies improving product sales. We think of a profit-maximizing monopolist launching a new product's selling season, t ($t = 1, 2$), in which, in each period, consequential consumers, $N^{(T)}$ ($N^{(T)} = N^{(1)}, N^{(2)}$), arrive the market. Consumers have a common ex-ante quality perception $\hat{q}_i$ ($\hat{q}_i \sim N(u_q, \sigma_q^2)$), but heterogeneous preference $x_i^{(T)}$ ($x_i^{(T)} \sim N(u_X, \sigma_X^2)$) for the product. We assume that selling season of the product includes two periods. Demand function is, $D_t^{(T)}(p)$, in selling period t of consumers arrived at T-period.

2.1 First-Period

During the first period which is an introductory phase (short time), consumers who purchase the product and their utility function are,

$$E(U_{i1}^{(1)}) = u_{q1} + \rho \frac{\sigma_{q1}}{\sigma_X}(x_i^{(1)} - u_X) + u_\alpha u_X - p \ge 0 \tag{1}$$

Consumers purchased the product and choose whether provide their reviews (consumer-generated reviews) about the product, following binomial distribution $\Psi_i \sim B(D_1^{(1)}, \frac{1}{D_1^{(1)}} \frac{n_1}{D_1^{(1)}})$. We assume that the distribution of consumer-generated reviews is power-law distribution, we have, $f_{n_1}(D_1^{(1)}(p)) = \frac{b}{D_1^{(1)}(p)^r}, r > 0, b > 0$. In the first period, the quantity of consumer-generated reviews n_1 is,

$$n_1 = E(f_{n_1}(D_1^{(1)}(p))) = \frac{bD_1^{(1)}(p)^{2-r}}{2-r}, 1 \le r < 2, b > 0 \tag{2}$$

Moreover, the average quality of consumer-generated reviews R_1 is given by,

$$R_1 = E(\Psi_i \hat{q}_i \mid x_i^{(1)} \ge \varpi_1^{(1)}(p)) = (\frac{2-r}{b})^{-\frac{1}{2-r}} n_1^{\frac{1-r}{2-r}} [u_q + \rho \sigma_q \sigma_X h(\varpi_1^{(1)}(p))] \tag{3}$$

Let $h(\cdot)$ denote the hazard ratio $h(\cdot)=\dfrac{f(\cdot)}{1-F(\cdot)}$, $h(\varpi_1^{(1)}(p))=\dfrac{f(\varpi_1^{(1)}(p))}{1-F(\varpi_1^{(1)}(p))}$.

2.2 Second-Period

In the second period which is a large-scale selling phase, new arriving consumers and remained consumers who not purchase the product in the first period but still interest in the product and stayed in the market, all of them change or update their quality perception of the product by social learning from consumer-generated reviews and their heterogeneous preference. Whether consumers will purchase the product depend on their utility function, if they purchase the product, we have,

$$E(U_{i2}^{(2)})=u_{q2}+\rho\frac{\sigma_{q2}}{\sigma_x}(x_i^{(2)}-u_x)+u_\alpha u_x-p\geq 0 \tag{4}$$

The second period is different with the first one is that consumers learn consumer-generated reviews and change their quality perception from $q_{i1}^{(T)}\sim N(u_{q1},\sigma_{q1}^2)$ to $q_{i2}^{(T)}\sim N(u_{q2},\sigma_{q2}^2)$. In the presence of social learning, we provide two modes of updating processes (perfect and imperfect social learning) to make $q_{i1}^{(T)}$ to $q_{i2}^{(T)}$.

2.2.1 Perfect Social Learning

The first one is called perfect (Bayesian) social learning, which consumers in the second selling period know all information of buyers in the first selling period, therefore, we use Bayesian updating model to help consumers in the second selling period to update their quality perception by learning consumer-generated reviews from buyers in the first selling period. In the first period, the quantity of consumer-generated reviews is n_1, the quality of consumer-generated reviews is R_1, buyers' average preference $u_{x_i^{(1)}}=\frac{1}{n_1}\sum_{i\in n_1}x_i^{(1)}$. Under perfect social learning, potential consumer i ' updated quality perception of the product (i.e., $q_{i1}^{(T)}\sim N(u_{q1},\sigma_{q1}^2)->q_{i2}^{(T)}\sim N(u_{q2},\sigma_{q2}^2)$) is given by the Bayesian and Gaussian updating models.

$$\bar{u}_{q2}=u_{q1}+\frac{n_1\sigma_{q1}^2}{n_1\sigma_{q1}^2+\sigma_q^2(1-\rho^2)}([\frac{D_1^{(1)}R_1}{n_1}-\rho\frac{\sigma_q}{\sigma_x}(u_{x_i^{(1)}}-u_x)]-u_{q1}),\ \sigma_{q2}^2=\frac{\sigma_{q1}^2\sigma_q^2(1-\rho^2)}{\sigma_{q1}^2+\sigma_q^2(1-\rho^2)} \tag{5}$$

2.2.2 Imperfect Social Learning

As everyone knows that, it is hard for consumers in the second selling period know all information of buyers in the first selling period, therefore, for the second one we show a more accurate representation of consumer behavior in real-world settings. The second one is called imperfect social learning, in the second selling period, in which consumers not know all private information of buyers in the first period and treat consumer-generated reviews as a representative sample of ex-post quality perception. Updated quality perception of the product (i.e., $q_{i1}^{(T)}\sim N(u_{q1},\sigma_{q1}^2)->q_{i2}^{(T)}\sim N(u_{q2},\sigma_{q2}^2)$) is given by

$$\bar{u}_{q2} = u_{q1} + \frac{n_1\sigma_{q1}^2}{n_1\sigma_{q1}^2 + \sigma_q^2}\left[\left(\frac{2-r}{b}\right)^{\frac{1}{2-r}} n_1^{\frac{r-1}{2-r}} R_1 - u_{q1}\right] \quad , \quad \sigma_{q2}^2 = \frac{\sigma_{q1}^2\sigma_q^2}{\sigma_{q1}^2 + \sigma_q^2} \tag{6}$$

We model the processes of consumer-generated reviews and consumers' social learning. Therefore, we can use them to analysis the relationship between the quantity of reviews and rating of reviews. Then we compare perfect social learning to imperfect social learning from firm profit, consumer surplus and total welfare to analysis the role of social learning on consumers and firms. Some effective results got by the analysis, which are described below.

3. Research Results

We pay more attention on how social learning from consumer-generated reviews affects potential consumers' quality perceptions to directly determine consumers' purchase decisions and consumer surplus and indirectly act on firm profit and total welfare in the perfect and imperfect social learning. Our models show that the quantity of consumer-generated reviews and rating of consumer-generated reviews have a positive effect on consumer updated quality perception, which is consistent with original literatures. However, (1) the rating of reviews has an opposite trend with the quantity of reviews in the long term (consistent with the real-world examples); (2) in a time, the quantity of reviews plays an important role under a slight fluctuation of rating of reviews on product sales, which illustrates why online stores who have a great quantity of and low rating of reviews get more sales than those who have a small quantity of and high rating of reviews; (3) when quantity of reviews exceeds a threshold, it does not play a significant role on updating consumers' quality perception, because of information overload theory and the limits of human information processing capacity, so potential consumers not take enough time and effort to learn all reviews;

We emphasize the role of social learning from consumer-generated reviews on consumers and firms. Through comparable analysis, we find that (1) social learning from consumer-generated reviews has a positive effect on total welfare, firms' profit and consumer surplus (exceed a value) for high-quality product, which has a different effect on low-quality product in perfect social learning. Numerical experiments and real-world examples also illustrate the results (consistent with real-world firms' willingness to invest in online reviewing platforms). (2) potential consumers have higher quality perception under imperfect social learning than perfect social learning. (3) Compared to perfect social learning, under imperfect social learning the firm sometimes has higher optimal price in the same ex-post (similar as real) quality perception. Our models and numerical experiments suggest that imperfect social learning generally provide higher firm profit than perfect social learning because potential consumers cannot learn the reviews completely and deeply to update their quality perception too close to ex-post quality perception, especially for low-quality product.

4. Conclusion and Future Work

This paper considers the effects of quantity and rating of consumer-generated reviews and social learning on consumers' purchase decisions and product sales. The paper explains how quantity and rating of reviews affect consumers' quality perception of a new product by social learning, and it illustrates the negative relationship between quantity of reviews and rating of reviews and which one is more important than the other one effect on product sales. The results show that both of quantity and rating of reviews are positive with product sales, however, rating of reviews is decreasing in quantity of reviews, and it slowly reduces and not has dramatically declined when quantity of reviews reaches to some value, therefore, if quantity of reviews exceeds the value, which has more effect on product sales than rating of reviews (not has a significant effect) for high quality products because of social learning and herding theory.

The paper also focuses on the implication of social learning on the firm's optimal price and market outcomes. We examined the two modes of social learning from reviews and others' behaviors, namely, perfect and imperfect social learning. The main different of them is whether potential consumers in second-period know all information of buyers in first-period. When real quality of product exceeds consumers' ex-ante quality perceptions or below some value, optimal price and firm profit not have dramatically difference in perfect social learning and imperfect social learning. However, when real quality of product is larger than some value and smaller than consumers' ex-ante quality perceptions, firm profit is huge distinction in two modes of social learning. The reason is that consumers, in imperfect social learning, not completely understand or learn real quality of the product and improve product sales, but dislike consumers in perfect social learning reduce product sales.

Our research addresses the impact of consumer-generated reviews and social learning by focusing on the naturally relationship between them and effect on market outcome. However, our study has several limitations. In this section, we also provide some interesting avenues for future research. Firstly, we only investigated data consumer-generated reviews of iphone at Amazon.com and Taobao.com, while this choice minimized aspects of products and data size. Whether the results of this paper are suit for both of experience and search goods or not, which need to give further research. Secondly, we have assumed a monopolist launching a new product, but more realistically, we may assume existed a competitive or alternative product. Finally, we only divide selling season into two periods, short time first-period and long time second-period, however, from a more long-term perceptive of the implication of consumer-generated reviews and social learning, it is better to subdivide the second period into multiple subperiods.

5. References

1. Celen, B., Kariv, S., and Schotter, A. "An Experimental Test of Advice and Social Learning," *Management Science* (56:10), 2010, pp.1687-1701.
2. Dellarocas, C., Zhang, X., and Awad, N. F. "Exploring the Value of Online Product Reviews in Forecasting Sales: The Case of Motion Pictures," *Journal of Interactive Marketing* (21:4), 2004, pp. 23-45.
3. DoubleClick. "DoubleClick's touchpoints II: The changing purchase process," 2004.
4. Duan, W., Gu, B., and Whinston, A. "Do Online Reviews Matter?-An Empirical Investigation of Panel Data," *Decision Support Systems* (45:4), 2008, pp.1007-1016.
5. Gaur, V., and Park, Y. "Asymmetric Consumer Learning and Inventory Competition," *Management Science* (53:2), 2007, pp.227-240.
6. Ifrach, B., Maglaras, C., and Scarsini, M. "Monopoly Pricing in the Presence of Social Learning," 2012, Columbia University, Working Paper.
7. Li, X., and Hitt, L. M. "Self-Selection and Information Role of Online Product Reviews," *Information Systems Research* (19:4), 2008, pp.456-474.

Personalized recommendation based on overlapping communities using time-weighted association rules

Haoyuan Feng[1], Jin Tian[1], Harry Jiannan Wang[2], Minqiang Li[1], Fuzan Chen[1], Nan Feng[1]

[1] Tianjin University, Tianjin, 300072, P.R. China
[2] University of Delaware, Newark, DE, 19716, USA
jtian@tju.edu.cn

Abstract

Modeling users' ever-changing interests has been a critical topic in recommender system research. In this paper, we propose a new personalized recommendation framework by leveraging and enhancing overlapping community concepts from complex network analysis literature and developing a time-weighted association rule mining method. Experiment results show that our proposed approach outperforms several existing methods in recommendation precision and diversity.

Keywords: personalized recommendation; overlapping community; time-weighted association rules; user interests

1. Introduction

Recommender systems have been implemented by many commercial websites, such as Amazon and eBay, to help users discover products of their interests. High-quality recommender algorithms and strategies can greatly increase profits and improve user loyalty. One of the most important aspects in personalized recommendation is the user interest modeling. Most of the conventional user interest models are static models, such as the user-based collaborative filter model, assuming that the users' interests do not change over time. However, users' interests are rather dynamic, e.g., users may prefer different styles of clothes at different ages. Users' interests for products like music and movie are even easier to change. Therefore, it is critical to capture user interest change over time in order for the recommender system to predict users' preferences more accurately.

In this paper, we propose a novel recommendation model based on overlapping communities using time-weighted association rules (OCTW). Temporal factors of user interests are fully considered when generating both the overlapping communities and the association rules. The overlapping community method is adopted to generate the users' interest relationship network based on their ratings. The association rules, which have good scalability, are generated to represent user interests. Then for each community, the proposed model operates the time decay function on association rules, which satisfy the downward closure property. Experimental results show that the proposed model can predict the user interests change better and achieve a higher recommendation accuracy compared with some traditional methods.

The key contribution of this paper is two-fold: (1) presenting a temporal overlapping community method to generate the dynamic user-user interest relationship network over time and to depict the users' multi-interest characteristics, and (2) proposing a new time-weighted association rule algorithm to model user interests change over time. The rest of the paper is organized as follows: the details of the proposed OCTW model are elaborated in Section 2. Section 3 presents the experimental studies for verifying the proposed model. Finally, Section 4 summarizes the key points of the paper and concludes with remarks for the future research.

2. Brief Literature Review

When observed data is generated from a distribution that changes over time it is known as concept drift (Sahoo and Singh 2012b). Interest drift can be considered as the special term of concept drift in recommendation. Research on user interest drift can be mainly divided into two groups. The first group is to build user models on a limited number of subsets of the original dataset and then ensemble the recommendation results. Elwell proposes an ensemble of classifiers-based approach for incremental learning of concept drift by training one new classifier for each data batch (Elwell 2011b). Masud et al. classify users into several different subsets according to their interests and establish the interest models separately in different subsets' feature space (Masud et al. 2011a). Research in the second group tends to reduce the weights of outdated data for recommendations. Rafeh and Bahrehmand develop a time-adaptive collaborative filtering model to identify the user interest change by introducing a time-decay function (Rafeh and Bahrehmand 2012a). Huang et al. propose a two-stage recommender system based on time series to capture the main interests of users in supermarket purchase (Huang and Huang 2009a). Koren use the user ratings to represent user's changeable interest in a potential matrix factorization model (Koren 2010b).

People may have multiple interests at the same time period resulting in the diversity of the user interest. There has been some recent work on modeling users' multiple interests. Cantador and Castells describe a proposal to automatically extracting multilayered communities of interest from semantic user profiles and applied it to group modeling and hybrid recommendations (Cantador and Castells 2011b). Chen et al. proposed a generalized cross domain collaborative filtering framework to make recommendations from multiple domains (Chen et al. 2013). The overlapping community methods provide an alternative to represent the variability and diversity of the user interests. The users in the overlapping part belong to more than one community and thus have multiple interests. Research has shown that local fitness maximisation (LFM) algorithm, i.e., a local optimized overlapping-community algorithm, is better than the traditional global optimized ones (Lancichinetti et al. 2009a).

Association rules (AR) have been successfully used to represent user interests in some static recommendation model (Wang and Shao 2004b). Generally, user interests can be provided by AR in the rule form 'A→B'(A and B are the implicit/explicit user interest or items), which means that the users that have the interest 'A' are likely interested in 'B'. Each rule represents a specific association with a certain interest. As for dynamic interest models, some researches have adopted the calendar-based temporal association rules to show the date effect on the user interest (Li et al. 2003a), while others have conducted the time-interval rules to describe the purchase cycle (Joong and Nam 2012a). Time association rules with time-varying minimum support are also proposed, which changes the downward closure property (Cagliero 2013a).

The proposed model aims to capture the user's multiple interest in a temporal context with a modified LFM-based overlapping community method. And a time-weighted AR method is further precisely designed to predict the user interest that is changeable over time in each community.

3. OCTW Configuration

3.1 Framework of OCTW

The framework of proposed model has three parts in Figure 1. The first part is to generate the overlapping communities, which contains three steps: user-user network setting, overlapping community detection and community combination. The second part of the proposed model is to mine the association rules in those overlapping communities separately, which includes frequent itemset mining and time-weighted rules' generation. It's the core part of proposed model by fully considering the time effect and increasing the ability to follow user interest change. The third part is the personalized recommend part which filters association rules for users. As one person may belong to a few communities, the recommendation list could contain recommendations from different interest groups. So the third part combines the recommended items by membership to communities and gets Top-N items from list for each person. These make the model personalized.

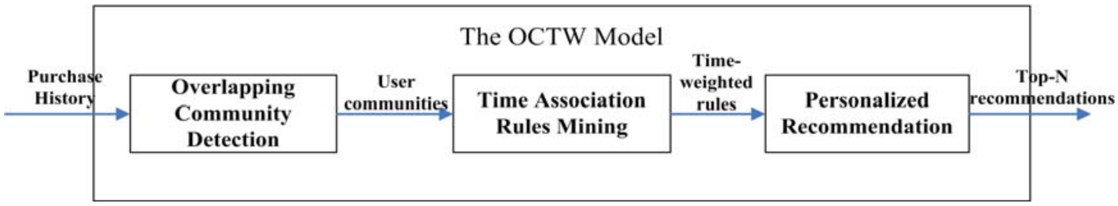


Figure 1. Framework of OCTW

3.2 Overlapping community generation

Split the training data into a pre-designed period of time (six months in this paper) as the timestamp of data. User purchase/rating history can be transformed into the user-item matrix $\boldsymbol{M} = [m_{ij}]_{nu \times ni}$ by a forgetting-curve-liked function as follows:

$$m_{ij} = \begin{cases} e^{\frac{-(TL-period_{ij})}{\theta}} & if\ period > 0 \\ 0 & if\ period = 0 \end{cases} \tag{1}$$

where *TL* is the duration of training data, $period_{ij}$ is the serial number of month when the user *i* bought/rated the item *j*, *θ* is the cycle of user interests change in a certain domain. *nu* is the number of the user and *ni* is the number of item. Thus user-user network is generated based on the matrix, by adding edges between users who "like" the same items. The weight of edge is decided by both the number of same items the two correlated users have bought/rated and the difference between the two users' buying/rating time. So one same item contributes a weight ranged from 0 to 1.

Then a modified LFM algorithm is adopted to detect the overlapping communities in user-user network. We use both the identical main process and the same fitness function of the original LFM method (Lancichinetti et al. 2009a). But the choosing principle of the start node in our modified LFM method is based on the degree ranking. An additional community combination process is executed after the communities could not expand and a hybrid community combination strategy is adopted in the proposed recommendation model.

For one thing, if the original communities obtained by LFM method have high combined degree (*CD*), it means there are too many points in the overlapping domain of two communities. Accordingly, we combine these communities if the combined degree is higher than a pre-designed threshold. The combined degree is calculated as follows (Wang 2011a):

$$CD_{pq} = \beta \times \frac{|C_p \cap C_q|}{|C_p \cup C_q|} + (1-\beta) \times \frac{|N(C_p) \cap N(C_q)|}{|N(C_p) \cup N(C_q)|} \tag{2}$$

where C_p, C_q are the overlapping communities. $N(C_p)$ is the neighbor set of C_p. Thus CD_{pq} is decided by both inside nodes and outside neighbors. $\beta \in [0,1]$ is a pre-designed parameter.

On the other side, if one of the two communities is much smaller than the other, that is, the proportion of the two communities is smaller than a pre-designed value, the *CD* is calculated as:

$$CD_{pq} = \frac{|C_p \cap C_q|}{|min(C_p, C_q)|} \tag{3}$$

This kind of *CD* definition aims to obtain *CD* values based on the proportion of the overlapping part to the smaller community.

3.3 Association rule mining

The frequent itemset mining is a time-consuming process with scanning the dataset many times. FP growth algorithm can speed up the process in parallel processing without the generation of the candidate itemsets. Based on FP model, we proposed a novel time-weighted rule mining method.

First, split the training set into *TL* periods, calculate the weight in each time period using expression (1), and normalize the weights. $Weight_s$ denotes the normalized weights and $Weight_s \in (0,1)$, *s*=1,…, *TL*. Then calculate the support of rules in different durations and get time conferences $Timeconf$:

$$Timeconf(A \to B) = \frac{1}{support(A)} \times \sum_{s=1}^{TL} \left(Weight_s(A) \times \sum_{t=1}^{TL} \left(Weight_t(B) \times support_{st}(A \cap B) \right) \right) \quad (4)$$

where $support_{st}(A \cap B)$ is the support of rule 'A→B' beginning from time period *s* to time period *t*.

3.4 Personalized recommendation

The proposed model needs to combine the recommendations from different communities by considering the target user's degree of membership in every community. The degree of membership is calculated by the percentage of the user node's edges in a certain community to all edges of this node in all communities. $Recom_{ij}$ is the global recommendation value of item *i* to user *j* and defined as:

$$Recom_{ij} = max\left\{ Degree(C_p, i) \times max\{Timeconf_{C_p}(x \to j)\} \right\} \quad (5)$$

where $Degree(C_p, i)$ is the membership degree of user *i* in community C_p, $p = 1, 2, \ldots$. If user *i* is not in C_p, $Degree(C_p, i) = 0$. $x \to j$ denotes any rule which have item *j* as the later part in community C_p. Thus the global recommendation value of item *j*, $Recom_{ij}$ is the maximum in the product values of the maximum of $Timeconf(x \to j)$ in each community C_p and the corresponding degree of membership in C_p. After ranking the recommendation values, a Top-N recommendation strategy is used to generate the recommendation list for the target user.

4. Experiments

Experiments were conducted to testify the recommendation performance of the proposed model in predicting the user multiple and variable interest. The MovieLens dataset was adopted which has 943 users and 1682 movies in 7 months. The user rating data in the first six month were used for training and the remaining data in the seventh month for testing. Four evaluation indices, *precision*, *recall*, *F*1 and *diversity*, were used. We calculate the average values of *precision*, *recall*, *F*1 of all active users who like at least one item in testing month. *Diversity* is denoted as the number of the unique recommendations (Adomavicius and Kwon 2012a). Since the data is sparse, there are many items, which are not scored by users. These items couldn't be judged whether people like them in the testing month. We ignore these items in the experiments.

Firstly, we assessed the proposed model with different community combination methods. Figure 2 shows the recommendation performance of different algorithms with Top-10 recommendation. The hybrid community combination method (shown as method iv in Figure 2) was compared with the models (i) with no community combination at all and (ii) only used the combination strategy mentioned in expression (2) or (iii) only used the combination strategy mentioned in expression (3). The recommendation model that used the hybrid combination strategy performed best on both precision and recall among the compared four models. The F1 score obtained by this model was also higher than that of the other three models. It means that communities with high CDs need to be combined to improve the result, and the proposed hybrid method really takes advantage of the two single strategies and keeps balance of them. Thus in the next experiments, the proposed algorithms OCTW adopted the hybrid combination strategy.

Secondly, we compared the proposed algorithm OCTW with some conventional methods, such as the traditional AR method, the user-based collaborative filtering method (UBCF), the item-based collaborative filtering method (ITCF), the time-weighted association rules model (TWAR), the association model with overlapping communities (OCAR), and one recent research, the two-stage recommender system (TSTS) (Huang and Huang 2009a). All the algorithms adopted Top-10 recommendations strategy. The average precision, the average recall, the average F1 score and the diversity are listed in Table 1.

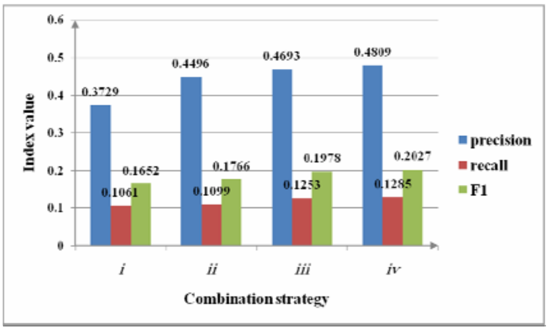


Figure 2. Recommendation performance with different community combination strategies

Table 1 Performance comparison of the seven algorithms (Top-10)

Algorithm	*precision*	*recall*	*F*1	*diversity*
OCTW	**0.4809**	**0.1285**	**0.2027**	139
OCAR	0.4221	0.0752	0.1276	173
TWAR	0.3169	0.0371	0.0664	65
AR	0.2747	0.0335	0.0597	65
TSTS	0.4673	0.0995	0.1641	59
UBCF	0.2490	0.0256	0.0465	239
ITCF	0.1721	0.0143	0.0264	**338**

Shown as Table 1, the proposed model performed best among the seven algorithms. Both the F1 score and the diversity obtained by OCTW are more than three times to that of AR. The OCTW increases the F1 scores 200.5% and 58.86% compared with TWAR and OCAR respectively. Moreover, OCTW achieves the highest diversity value among the first five algorithms except OCAR, which illuminates that the overlapping community method plays an important role in depicting a certain user's various interests so as to obtain diverse recommendations. Although the CF-based algorithms, UBCF and ITCF perform higher in diversity, they perform poorly in both precision and recall, as well as F1 score. TSTS performs similar to OCTW in precision, but much worse than OCTW in both F1 score and diversity.

Finally, we conducted experiment to testify the performance of the algorithms with different N values in Top-N recommendation strategy. Figure 3 depicted the performance curves of the algorithms with different Ns. Except UBCF, the OCTW's F1 is always higher than other models. And OCTW is also higher than UBCF when N is smaller than 50, which indicates OCTW can achieve higher recommendation accuracies when the required recommendation number is limited. For example, the screen could not show too many recommendations in a mobile recommendation context. Note that those algorithms that adopt association relationship can achieve high-quality rules as well as good recommendations when N is small. However, when N increases, there are not enough rules to extract for such big number of items with a fixed minimum support. Therefore, the performance curves of such algorithms are nearly flat as the value of N increases.

5. Conclusions and Future works

This paper proposes a dynamic model OCTW to predict user multiple and changeable interests. A modified overlapping community method is employed to describe users' relationships based on their buying/rating actions, in fully consideration of the time weights of the relationship edges. A novel time-weighted association rule method is then designed to capture the frequent itemsets.

Finally the recommendation result is generated with top-*N* recommendation strategy. Experimental results show that the OCTW is able to achieve a high recommendation accuracy and diversity than other conventional algorithms. We are planning to extend our research in a number of directions, such as adopting a self-tuned minimum support in the model and enhancing the model in an incremental learning frame to deal with new sequential data.

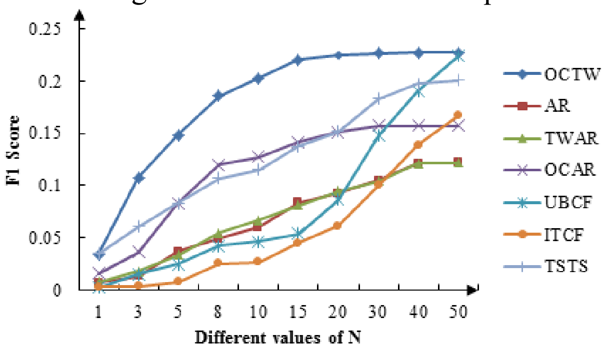


Figure 3. F1 scores of seven models with different Top-Ns

Acknowledgements

The work was supported by the National Science Fund for Distinguished Young Scholars of China (Grant No. 70925005) and the General Program of the National Science Foundation of China (Grant Nos. 71001076, 71101103, and 71271149).

References

1. Adomavicius, G., Kwon, Y. O. “Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques,” *IEEE Transactions on Knowledge and Data Engineering* (24:5), 2012a, pp. 896-911.
2. Cagliero, L. “Discovering Temporal Change Patterns in the Presence of Taxonomies,” *IEEE Transactions on Knowledge and Data Engineering* (25:3), 2013a, pp. 541-555.
3. Cantador, I., Castells, P. “Extracting Multilayered Communities of Interest from Semantic User Profiles: Application to Group Modeling and Hybrid Recommendations,” *Computers in Human Behavior* (27:4), 2011b, pp. 1321-1336.
4. Chen W., Hsu W., Lee M. L. “Making Recommendations from Multiple Domains,” ACM Proc KDD, 2013, pp. 892-900.
5. Elwell, R., and Polikar, R. “Incremental Learning of Concept Drift in Nonstationary Environments,” *IEEE Transactions on Neural Network* (22:10), 2011b, pp. 1517-1531.
6. Huang, C. L., and Huang, W. L. “Handling Sequential Pattern Decay: Developing a Two-Stage Collaborative Recommender System,” *Electronic Commerce Research and Applications* (8:3), 2009a, pp. 117-129.
7. Joong, H. C., Nam, H. P. “Comparative analysis of sequence weighting approaches for mining time-interval weighted sequential patterns,” *Expert Systems with Applications* (39:3), 2012a, pp. 863-873.
8. Koren, Y. “Collaborative Filtering with Temporal Dynamics,” *Communications of the ACM* (53:4), 2010b, pp. 89-97.
9. Lancichinetti, A., Fortunato, S., Kertesz, J. “Detecting the Overlapping and Hierarchical Community Structure in Complex Networks,” *New Journal of Physics* (11:3), 2009a, pp. 033015.
10. Li, Y. J., Ning, P., Wang, X. S., Jajodia, S. “Discovering Calendar-Based Temporal Association Rules,” *Data and Knowledge Engineering* (44:2), 2003a, pp. 193-218.
11. Masud, M. M., Gao, J., Khan, L., Han, J. W. “Classification and Novel Class Detection in Concept-Drifting Data Streams under Time Constraints,” *IEEE Transactions on Knowledge and Data Engineering* (23:6), 2011a, pp. 859-874.
12. Rafeh, R., and Bahrehmand, A. “An Adaptive Approach to Dealing with Unstable Behaviour of Users in Collaborative Filtering Systems,” Journal of Information Science (38:3), 2012a, pp. 205-221.
13. Sahoo, N., Singh, P. V., Mukhopadhyay, T. “A Hidden Markov Model for Collaborative Filtering,” *MIS Quarterly* (36:4), 2012b, pp. 1329-1356.
14. Wang, F. H., and Shao, H. M. “Effective Personalized Recommendation Based on Time-Framed Navigation Clustering and Association Mining,” *Expert Systems with Applications* (27:3), 2004b, pp. 365-377.
15. Wang, Y. P. (eds.). *Research on Overlapping Community Detection in Complex Networks*, Taiyuan University of Technology, Taiyuan, 2011a.

A Thematic Analysis of Analyst Information Discovery and Information Interpretation Roles

Allen Huang
Department of Accounting, Hong Kong University of Sci. and Tech.
allen.huang@ust.hk

Reuven Lehavy
Ross School of Business, University of Michigan
rlehavy@umich.edu

Amy Zang
Department of Accounting, Hong Kong University of Sci. and Tech.
amy.zang@ust.hk

Rong Zheng
Business Statistics and Operations Management, Hong Kong University of Sci. and Tech.
rzheng@ust.hk

Abstract

Extant literature on financial analysts has sought to understand their information discovery role and information interpretation role but offers inconclusive evidence. We employ a novel methodology based on topic modeling – an advanced technique in computer science, to compare and contrast the thematic content of the analyst reports to that in the conference calls. It allows us to separately identify and measure analyst information discovery and interpretation roles without referencing to stock price. We find that analysts discuss topics that were not referred to in conference calls, consistent with information discovery. We also find that analysts use different vocabularies from managers to discuss the conference call topics, suggesting information interpretation. Our further analyses indicate that analysts respond to investor demand for their services: they discovery more information when firms' proprietary cost is high and provides more interpretation when the processing cost of the conference call is high.

Keywords: Financial analyst, Textual analysis, Topic modeling

1. Introduction

Financial analysts serve an important information intermediary role in capital markets (Frankel et al. 2006; Ramnath et al. 2008). Through their research, they process and interpret public corporate disclosures, corporate events, and news for investors, and also provide investors new information they discover from their private research efforts.

Evidence in the extant academic literature on the relative importance of the analyst information interpretation and information discovery roles is mixed. For example, Francis et al. (2002) and Frankel et al. (2006) conclude that the informativeness of analyst research and earnings announcements complement each other, consistent with the analyst information interpretation role. In contrast, Ivkovic and Jegadeesh (2004) and Chen et al. (2010) interpret their evidence to suggest that the discovery of new information by analysts is more valuable to investors than their interpretation of the information in earnings announcements. Livnat and Zhang (2012) extend the analyses of Chen et al. (2010) by including other corporate disclosures; they reach the opposite conclusion – the analysts' information interpretation role after firms announce their earnings dominates their information discovery role. Recognizing the conflicting nature of these findings, a literature review by Ramnath et al. (2008) calls for more research on the distinction between the analyst interpretation and discovery roles.

While these studies employ different samples and research methodologies, they uniformly infer analyst information roles from the immediate equity market reaction to the issuance of analyst reports. This research design suffers from two potential limitations. First, it assumes that investors fully understand and instantaneously incorporate the information provided by analysts into stock prices. Second, because a vast number of analyst reports are issued immediately after corporate disclosures (46.5% of analyst revision reports in our sample are issued on the same day or one day after earnings conference calls), researchers are unable to disentangle the immediate market reaction to the information content of analyst reports and that of the earnings release or the conference call.

In this study, we compare and contrast the information content of sell side analyst reports issued on the day of or the day after conference calls (AR) to that in earnings conference calls (CC). Specifically, we ask (1) do the vast number of analyst reports issued promptly after one of the most important regular corporate disclosure events – earnings release and the adjoining conference call – provide incremental information to the discussions in these calls, and (2) when do analysts play an information discovery role and when do they play an information interpretation role? In stark contrast to prior work, we do not rely on the immediate stock price reaction around the earnings release as a measure of information content; rather, we introduce several novel measures of the thematic content of analyst reports and conference calls and use these measures to test the relative information content of analyst reports as well as their information discovery and interpretation roles.

To construct these measures, we rely on an advanced technique in information retrieval research called *Latent Dirichlet Allocation* (LDA). Developed in a seminal paper by Blei et al. (2003), LDA is a robust method that relies on statistical correlations among words in a large set of documents to discover and quantify their latent thematic structure (or, topics). LDA can be thought of as a dimensionality-reduction technique, similar to cluster analysis or principle components analysis, but designed specifically for use with text. The algorithmic procedure of LDA is able to handle a massive collection of documents impossible for human coders to process. Moreover, as an unsupervised statistical learning method, LDA does not require training data, annotation, or any prespecified topic labels from the researchers. These desirable features lead to a widespread application of LDA in a variety of fields.

2. Hypotheses development

2.1. Analyst information role immediately after earnings conference calls

In contrast to prior research that assigns analysts either an information interpretation or an information discovery role, a more realistic depiction is that analysts engage in a combination of these roles. First, theoretical models in Kim and Verrecchia (1994) suggest that public information disclosure triggers analysts to produce idiosyncratic information, because analysts combine their private research effort with the information disclosed during public corporate disclosure to produce "uniquely privately inferred information" about firms' future prospects (Mayew et al. 2013). Second, because of the importance of quarterly earnings announcements combined with the adjoining conference calls, analysts compete to issue reports quickly after this significant information event. Analysts likely take this opportunity to signal their superior ability to interpret the large amount of qualitative and quantitative information in the conference calls, as well as provide new, private information heretofore unknown to their clients and investors. Accordingly, we hypothesize that:

H1a*: Analysts serve the information discovery role in reports issued immediately after earnings conference calls;*

H1b*: Analysts serve the information interpretation role in reports issued immediately after earnings conference calls.*

2.2. Cross-sectional determinants of analyst information discovery role

We predict that the importance of the analyst information discovery role is increasing in firms' proprietary costs. An extensive theoretical literature on proprietary cost (see reviews in

Verrecchia 2001) demonstrates that proprietary costs represent a significant consequential disclosure cost that prevents managers from being forthcoming because disclosing proprietary information can damage the company's competitive position in the product market. When proprietary cost is high, the impact on the capital market due to proprietary information being withheld tends to be greater, implying a greater value of the withheld information to investors, as well as a higher investor demand for additional sources of information. Therefore, we hypothesize that:

H2a*: The importance of analyst information discovery role immediately after earnings conference calls increases with firms' proprietary cost.*

Implicit in our arguments motivating *H2a* is the assumption that analysts engage in greater information discovery for firms with high proprietary cost as a response to investors' demand for more information. We empirically test this assumption by examining whether investors value the analyst information discovery role, and, specifically, whether investors place a greater weight on analyst information discovery for firms with high proprietary cost. Thus, we hypothesize:

H2b*: Investors value the analyst information discovery role, and view this role as more valuable for firms with high proprietary cost.*

2.3. Cross-sectional determinants of analyst information interpretation role

Evidence in prior research suggests that the information processing cost of earnings conference calls is generally high because some of managers' spoken disclosure tends to be informal and unstructured, involves ambiguous language, subjective evaluation, and a significant amount of non-financial information (Frankel et al. 1999). Prior research also documents that corporate disclosures that involve high processing costs result in an increasing demand for analyst service and a greater collective effort by these analysts (Lehavy et al. 2011). Accordingly, we predict that:

H3a**:** *The importance of the analyst information interpretation role immediately after earnings conference calls increases with the costs of processing the information in these calls.*

Similar to our discussion above, *H3a* relies on the assumption that analysts increase their efforts to interpret the discussions in conference calls when these discussions are hard to interpret, in response to investors' demand for clearer information. Accordingly, our final prediction is that:

H3b*: Investors value the analyst information interpretation role, and view this role as more valuable when the information processing cost of the discussions in earnings conference calls is high.*

3. Measurement of information discovery and information interpretation

We use LDA output to (see Blei 2012 for a review) to construct four related measures for the information discovery and information interpretation roles played by analysts.

3.1 Measuring information discovery based on differences in the distribution of topics between analyst reports and conference calls

We use the following procedure to construct the topic vector (T_d) of a document d: first, we separate each sentence in a document into words; then, using the topic-word frequency matrix $\mathbf{\Phi}$, we construct a frequency vector for each word containing the number of times it appears in each of the K topics. For each sentence, we then sum the frequencies of the words in each topic and assign the sentence to the topic with the highest combined frequency. The fraction of document d that is dedicated to a discussion of topic k (S_{dk}) equals the number of sentences that are assigned to the topic k divided by the total number of sentences in document d. Formally,

$$Topic\ vector\ of\ document\ d = T_d = \ (S_{d1}, S_{d2}, \dots S_{d60}), \tag{1}$$

where S_{dk} represents the fraction of the discussion in a document devoted to topic *k.*

3.2 Measuring information interpretation based on differences in word usage between analyst reports and conference calls

Our empirical tests of the analyst interpretation role are based on a statistical comparison of the distribution of words used by analysts and managers to discuss the top ten topics in the presentation part of the conference call. To conduct this test, we extract the amount of discussion dedicated of each of the top ten CC topics in the AR and the CC and construct a vector of the word usage for each topic:

$$\begin{aligned} Word\ vector\ of\ topic\ k\ in\ CC &= W_{CC,k} = (v_{1k}, v_{2k}, \dots v_{Nk})\ ; \\ Word\ vector\ of\ topic\ k\ in\ AR &= W_{AR,k} = (w_{1k}, w_{2k}, \dots w_{Nk}); \end{aligned} \qquad (2)$$

where each element of these vectors (v_{wk}) is the frequency of word w in the discussion of topic k in the respective document (N is the total number of unique words in the corpus).

3.3 Measuring the amount of information discovery in analyst reports

Our empirical tests of the determinants of the analyst information discovery role require a summary measure of the amount of information discovery contained in analyst reports issued promptly after earnings conference calls. We define this measure as one minus the cosine similarity between the distribution of topics in a conference call and that in the adjoining analyst reports (i.e., one minus the cosine similarity between the topic vector of CC and AR in Eq. 1). Formally, we measure analyst information discovery as:

$$Discovery = 1 - \frac{\sum_{k=1}^{K}\left(S_{AR,k} \cdot S_{CC,k}\right)}{\sqrt{\sum_{k=1}^{K}\left(S_{AR,k}\right)^2} \cdot \sqrt{\sum_{k=1}^{K}\left(S_{CC,k}\right)^2}}. \qquad (3)$$

where $S_{AR,k}$ ($S_{CC,k}$) is the fraction of the discussion in AR (CC) devoted to topic k. Intuitively, information discovery occurs when analysts introduce topics that are not included in CC, or are less emphasized by managers in their CC.

3.4 Measuring the extent of information interpretation in analyst reports

Our empirical tests of the determinants of the analyst information interpretation role require a summary measure of the extent to which analyst reports provide interpretation of the information contained in CC. We define this measure as the average, over the top ten CC topics, of one minus the cosine similarity between the words used by the analysts to describe each of these topics and the words used by managers to discuss the same topics in their CC (i.e., the difference between W_{AR} and W_{CC} for each of the top ten CC topics.Formally, our interpretation measure is:

$$Interpret \ = \frac{1}{10}\sum_{k=1}^{10}\left(1 - \frac{\sum_{j=1}^{N}\left(w_{jk} \cdot v_{jk}\right)}{\sqrt{\sum_{j=1}^{N}\left(w_{jk}\right)^2} \cdot \sqrt{\sum_{j=1}^{N}\left(v_{jk}\right)^2}}\right). \qquad (4)$$

where, w_{1k} is word 1's frequency in the discussion of topic k in the AR; v_{1k} is word 1's frequency in the discussion of topic k in the CC; N is the total number of unique words in the corpus; k is one of the top ten topics discussed in the CC.

4. Empirical results

4.1. Tests of Analyst Information Discovery and Information Interpretation Roles

Our first hypothesis (H1a) is that analysts serve the information discovery role in reports issued immediately after earnings conference calls. To test this hypothesis, we compute the Pearson's chi-square statistics and test for the homogeneity of the distribution of topics discussed in each AR and CC pair (i.e., we test the null that $T_{CC} = T_{AR}$, see equation 1). The Pearson's chi-square test is a standard statistical test for the homogeneity of the frequency distribution of certain events (i.e., the frequency of the sentences in each of the 60 topics in our setting) observed in two or more samples.

The mean (median) value of the chi-square statistic across all 17,750 pairs of AR and CC is 103.1 (94.17), indicating that the homogeneity between the topic distribution in these documents is rejected 71.7% of the time (at the 10% level). That is, in 71.7% of prompt analyst reports, the set of topics discussed is statistically different than those discussed in the immediately preceding CC. This evidence supports the information discovery role in analyst reports.

Next, we empirically test the analyst information interpretation role (H1b). Our tests attempt to statistically distinguish between analyst reports that describe the key topics in CC using words that are similar to those used by managers to describe the same topics, from those reports that use a different set of words to describe these topics. The latter set of reports likely transformed and paraphrased the information contained in the public disclosure or provided a new perspective, consistent with a meaningful interpretation role.

Empirically, for each CC-AR pair, we use the Pearson's chi-square to test whether the words used by managers and analysts for a given CC topic are significantly different (i.e., we test the null that $W_{AR,k} = W_{CC,k}$ in equation 2 for each of the top ten CC topics). We focus on the top ten CC topics to avoid the noise introduced by economically less important topics. Out of a total of 167,544 top ten CC topics in our sample, the homogeneity between the distribution of words used to describe these topics in CC and AR is rejected (at the 10% level) for 49.4% of the sample. That is, in each set of reports issued promptly after the CC, analysts provide statistically significant interpretation for an average of five of the top ten CC topics.

4.2 Tests of the determinants of analyst information discovery and information interpretation roles

We follow literature (Li et al. 2013;Loughran and McDonald, 2013;Huang et al. 2014;Frankel et al. 2006) to measure firm's proprietary costs and to capture the cost of processing the information in earnings conference calls.

Untabulated regression results of tests of the cross-sectional determinants of the analyst information discovery role show that the coefficient estimates on the proprietary cost measures are positive and significant in all specifications. This evidence supports H2a that the amount of information analysts discover in reports issued promptly after conference calls is increasing in firms' proprietary cost. The coefficient estimates on the control variables indicate that the amount of information discovery in prompt analyst reports is also increasing in the sign and magnitude of the earnings news, but is decreasing in the length of the analyst reports.

The regression results of tests of the cross-sectional variation in the analyst information interpretation role suggests that, consistent with H3a, the amount of information interpretation analysts provide in reports issued promptly after earnings conference calls increases when the information contained in these calls is more difficult for investors to process. The coefficient estimates on the control variables indicate that analyst information interpretation role is greater for firms reporting negative earnings surprise, smaller firms, and firms with lower ratios of book-to-market (i.e., growth firms).

To test H2b and H3b, we regress investor reaction to the issuance of AR on our measures of analyst information discovery and information interpretation, as well as interaction terms of these measures with $Competition$ and $Uncertain$. The coefficients on $Discovery$ and $Interpret$ are positive and significant, indicating that the information discovery and interpretation provided by analysts trigger incremental market reactions. This result provides further validation that our information content measures based on the thematic topics generated by LDA likely capture the informativeness of these documents, as perceived by investors.

5. Conclusions

We analyze the information content of analyst research reports and the role they play in discovering and interpreting corporate financial disclosures to capital market participants. To do so, we introduce novel measures of the information content of textual data that rely on algorithmic analyses of the themes (or topics). Our study advances the understanding of the information role of analysts as well as the determinants of their information discovery and information interpretation roles, by explicitly quantifying the semantic content of analyst research reports and contrasting it with managers' discussions in earnings conference calls. Additionally, we introduce measures of the information content of textual disclosures that do not rely on equity market reactions to the release of these disclosures.

References

Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent Dirichlet allocation. *The Journal of Machine Learning Research* 3, 993-1022.

Blei, D., 2012. Probabilistic topic models. *Communications of ACM* 55(4), 77-84.

Chen, X., Cheng, Q., Lo, K., 2010. On the relationship between analyst reports and corporate disclosures: Exploring the roles of information discovery and interpretation. *Journal of Accounting & Economics* 49, 206-226.

Francis, J., Schipper, K., Vincent, L., 2002. Earnings announcements and competing information. *Journal of Accounting & Economics* 33, 313-342.

Frankel, R., Kothari, S.P., Weber, J., 2006. Determinants of the informativeness of analyst research. *Journal of Accounting & Economics* 41, 29-54.

Frankel, R., Johnson, M., Skinner, D.J., 1999. An empirical examination of conference calls as a voluntary disclosure medium. *Journal of Accounting Research* 133-150.

Huang, A., Zang, A., Zheng, R., 2014. Evidence on the information content of text in analyst reports. Working paper, Hong Kong University of Science and Technology

Ivković, Z., Jegadeesh, N., 2004. The timing and value of forecast and recommendation revisions. *Journal of Financial Economics* 73, 433-463.

Kim, O., Verrecchia, R.E., 1994. Market liquidity and volume around earnings announcements. *Journal of Accounting & Economics* 17, 41-67.

Lehavy, R., Li, F., Merkley, K., 2011. The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review* 86, 1087-1115.

Li, F., Lundholm, R., Minnis, M., 2013. A measure of competition based on 10-K filings. *Journal of Accounting Research* 51, 399-436.

Livnat, J., Zhang, Y., 2012. Information interpretation or information discovery: Which role of analysts do investors value more? *Review of Accounting Studies* 17, 612-641.

Loughran, T., McDonald, B., 2013. IPO first-day returns, offer price revisions, volatility, and form S-1 language. *Journal of Financial Economics* 109, 307-326.

Mayew, W.J., Sharp, N.Y., Venkatachalam, M., 2013. Using earnings conference calls to identify analysts with superior private information. *Review of Accounting Studies* 18 (2), 386-413.

Ramnath, S., Rock, S., Shane, P., 2008. The financial analyst forecast literature: A taxonomy with suggestions for future research. *International Journal of Forecasting* 24, 34-75.

Verrecchia, R.E., 2001. Essays on disclosure. *Journal of Accounting & Economics* 32, 97-180.

Two-sided Peer Influence on Content Creation in Social Media Platforms

Bin Zhang
University of Arizona
binzhang@arizona.edu

Anjana Susarla
Michigan State University
asusarla@board.msu.edu

Ramayya Krishnan
Carnegie Mellon University
rk2x@cmu.edu

Abstract

While prior research has studied the motivations of individuals to consume content on social media platforms, limited work exists on how contributors are motivated to create content in an environment of limited attention. We examine the role of peer production on YouTube, where content creators are competing for attention. We face a novel set of statistical challenges in estimating peer influence. Identification of social influence in large-scale social networks such as YouTube is difficult due to interdependent decision of users, correlations between the video's observable and unobservable characteristics, and attributes over time. We design a new method – Network Auto-probit model with Fixed Effects (NAFE), to identify peer influence among YouTube users. Implications for research and practice are also discussed.

1 Introduction

Tremendous advances in business analytics and capabilities to leverage large data have completely transformed digital content creation (Manyika et al., 2011). Social media is providing a wealth of opportunity for content creators to express themselves and for individuals to share their preferences with others, ultimately promising to transform how consumers engage with each other at a global level. This paper is motivated by the growth and popularity of YouTube. On average, over 4 billion videos are viewed every day; over 800 million unique users visit YouTube each month; Given the volume of search that occurs on YouTube, it has been considered to be the second largest search engine (Rosenbaum, 2010), next to Google. We investigate the role of peer influence in content creation by examining a content creator's decision to upload a new video. We assume that content creation efforts are driven not only by the personal preferences of individual content creators but also by the content creation decisions focal node's network neighbors. The inter-linkages between content creators and the scarcity of attention suggest that neighborhood effects in content creation could take two forms: competition and substitution. The first effect is when content creators simultaneously try to compete for the audience's attention or awareness. It is possible that content creators compare themselves with the uploading patterns of other content creators. For example if a channel uploads a video, the neighboring channels also need to upload videos in order to direct attention towards their offerings. Alternately, another scenario would be that each user acts as a content contributor for the entire neighborhood. Since YouTube showcases related videos and directs viewers' attention towards content produced by neighboring channels, spillovers in attention from neighboring channels could make it less likely for a content creator to upload a video when other channels in the neighborhood post videos on related topics.

2 Context

On YouTube users can create their own page or channel and subscribe to or accept subscriptions from other channels. A channel is personalized to each user and allows for users to display content that they uploaded, videos from other members, videos favorited by them and channels that they have subscribed to, and channels that subscribe to their content. As a YouTube user, one can establish a relationship with other users either as a friend or a subscriber. When a new video is posted, all the subscribers and friends of a channel are alerted through email or RSS feeds. A friend relationship is initiated by an invitation from one person to another, and for the relationship to be confirmed, it requires an agreement from the other person. Thus, it is more likely that a friend relationship on YouTube represents the pre-existing social ties between individuals, either from friendship based on real life, or friendship developed online. Since the friend network we observe is the result of such mutual agreements, we characterize friendship networks on YouTube as an undirected network (e.g., Girvan and Newman, 2002). YouTube demonstrates a two-sided model of interaction where the value of the YouTube platform lies in bringing together the interaction between groups of content creators and content consumers. Channels can build subscriptions and friend relationships with other channels. Ratings and comments on videos are all user driven as well. Thus, content creators have networked ties to other content creators as well as content consumers. Further, the YouTube model blurs the distinction between content creators and content consumers since users on YouTube are identified by their channel, encouraging closer links between consumers and content creators.

3 Data Collection

Given the purpose of analyzing the patterns of content creation of YouTube, we implemented a PHP program to access YouTube channel data through YouTube Data API. But before collecting data , we face another challenging problem: determining an ideal subset of the YouTube friend network. The YouTube friend network is large – starting from a random seed channel, the average number of distinct channels embedded in three-hop friend networks, the lower bound of the diameter of a network ensuring sufficient condition for identification (Bramoulle et al., 2009), is at about 1.5 million with over 20 million of edges on average. With this wealth of interconnectedness and access to data, though, come several challenges. The first is estimation consistency. Starting from any random channel, it is almost certain that one or more extremely popular channel (e.g., a channel with more than a million subscribers) will be included in a four-hop network, because of small-world effect. In the local network of these popular channel, with high probability there exists more equally popular channels. Even we have enough computing power, any analyses of such network would be biased towards high-degree nodes in terms of preferential attachment; thus network influence will be overestimated. The second challenge is computational cost. Most of the social network analysis packages do not scale well. That is, for some class of questions and analytic routines, standardized desktop systems are not able to complete analysis of large networks in a realistic time frame. Memory-wise, systems do not have enough resources to accommodate the structure of a large network. For example, a standard PC will not be able to allocate enough memory for an adjacency matrix with a size of 100,000$\times$100,000. Third, since the global YouTube friend network is a sparse network and the real strength of connections among internet users is unobservable. A randomly extracted network is highly likely to be a sparsely connected network. This kind of network is less likely to have peer influence from within, because the connections among neighbors are weak. So in a context of studying peer influence, a dense network is favored. A densely inter-connected subnetwork indicates that a group of individuals have strong relations, making peer influence more likely to happen.

A legitimate and efficient method to build one (or more) subnetwork(s) in this situation is snowball method (Coleman, 1958; Goodman, 1961). We extract three-hop networks starting from different random channels, until 100 subnetworks are extracted. With network having distance larger than or equal to three exist, identification of social influence can be archieved Bramoulle et al., 2009.

During the extraction, we only keep new subnetworks that are not overlapping with any subnetworks that are retained. We then remove all the users with more than 10,000 friends, together with the edges connecting to them in the subnetworks. After the abnormal users are removed, we calculate *I*-*E* ratio for each subnetwork. We keep the subnetwork with the highest *I*-*E* ratio at 0.15 (focal network), resulting in a network with 4600 nodes (channels). Once the focal subnetwork is determined, we want to construct panel dataset for channels in the network. Every two weeks we run the data collection program, The information we collect through Google YouTube API for each channel are: (1) Channel ID (2) The number of subscribers for the channel (3) The friends that the channel has (4) The number of videos uploaded for the channel.

4 Model

As explained earlier, we are interested in the role of peer influence on content creators propensity to introduce new content on YouTube. We assume that a content creator has a greater propensity to introduce new content in a particular period depends upon the following factors. First, we consider the role of observable characteristics of users (on YouTube every user is referred to as a channel), such as a channel's centrality in the network of subscribers and its betweenness centrality in the friend network. Second, we consider the positive dependence on attention by examining the impact of views of prior content created by the channel on the propensity to create content in the present period. Since there could be a positive shock to views in a period that could also be correlated with the propensity to introduce content, we consider a method of moments approach where we take the lagged difference of video to account for dynamic effects (e.g., Arellano and Bond, 1991). Given no existing model supports binary response variable, fixed effects, and panel data, we design a new model, network auto-probit with fixed effects (NAFE), to study network influence for the same group of channels on YouTube across time. NAFE is implemented using Bayesian methods and the estimates are generated through Markov chain Monte Carlo routine. The matrix notation of the model is specified as below.

$$
\begin{aligned}
\mathbf{y}_t &= \mathbb{I}(\mathbf{z}_t > 0) \\
\mathbf{z}_t &= \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\theta}_t + \rho_2\mathbf{W}_{t-1}\mathbf{y}_{t-1} + \boldsymbol{\eta} + \boldsymbol{\epsilon}_t \\
\boldsymbol{\theta}_t &= \rho_1\mathbf{W}_t\boldsymbol{\theta}_t + \mathbf{u}_t \\
\boldsymbol{\epsilon}_t &\sim \text{Normal}(0, I); \\
\mathbf{u}_t &\sim \text{Normal}(0, \sigma^2 I); \\
\mathbf{X}_t &= \boldsymbol{\kappa} + \mathbf{e}_t \\
(\boldsymbol{\eta}, \mathbf{e}_t) &\sim \text{Normal}(0, \boldsymbol{\Omega}); \\
\boldsymbol{\Omega} &= \begin{pmatrix} \Omega_{e,e} & \Omega_{e,\eta} \\ \Omega_{\eta,e} & \Omega_{\eta,\eta} \end{pmatrix}
\end{aligned}
$$

$\mathbf{y}_t$ is the vector of observed binary choices, whether a YouTube user uploaded new videos in time period t. It is an indicator function of the latent preference of users, $\mathbf{z}_t$. For any channel i, if z_{ti} is larger than a threshold 0, channel i choose y_{ti} as 1, uploading new videos; if z_{ti} is smaller than 0, then i would choose y_{ti} as 0, not to upload new videos at time period t.

$\mathbf{z}_t$ could be represented as a function of both covariates $\mathbf{X}_t$, autocorrelation term $\boldsymbol{\theta}_t$, and fixed effects $\boldsymbol{\eta}$. $\boldsymbol{\epsilon}_t$ is $\mathbf{z}_t$'s error term. The covariates matrix $\mathbf{X}$ included are observable characteristics of users, and can be represented as $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \triangle\mathbf{v}_{t-1}\}$. The first covariate, $\mathbf{x}_1$, is the number of subscribers a channel has; $\mathbf{x}_2$ is the influence of a channel in the friend network, measured by betweenness centrality. $\triangle \mathbf{v}_{t-1}$, is the change of video consumption information of each channel in $t-1$, i.e. the change of total views of all videos of a focal channel in time period from $t-2$ to $t-1$. Its coefficient $\boldsymbol{\alpha}$ is included in the vector of covariate coefficient $\boldsymbol{\beta}$. $\mathbf{X}_t$ is modeled as two parts, time-invariant individual effect κ, and error term $\mathbf{e}_t$ ($\mathbf{e}_t = \{e_{tij}\}$).

$\boldsymbol{\theta}_t$ is the autocorrelation term, which is responsible for those nonzero covariances in the $\mathbf{z}_t$. $\mathbf{u}_t$ is the error term for $\boldsymbol{\theta}_t$. $\boldsymbol{\theta}_t$ can be described as the network structure $\mathbf{W}_t$ and coefficient ρ_1. $\mathbf{W}_t$ is the network structure describing whether two YouTube channels have friend relationships in time period t. If channel j is a friend of channel i then element $W_{tij} = 1$ otherwise $W_{tij} = 0$. Network influence could also show in observed decision. We thus include $\mathbf{y}_{t-1}$, the lagged decision in the last time period made by the focal channel in the model. $\mathbf{W}_{t-1}$ is the structure of friend network in time period $t-1$, ρ_2 is the correspondent coefficient of $\mathbf{W}_{t-1}\mathbf{y}_{t-1}$. $\boldsymbol{\beta'} = \{\boldsymbol{\beta}, \alpha, \rho_2\}$. $\boldsymbol{\eta}$ is the fixed effects to account for video level unobservables. Such effect accounts for intrinsic attributes that would affect video consumption such as whether it follows fashion trend (Yoganarasimhan, 2012). $\boldsymbol{\eta} = \{\eta_1, ..., \eta_n\}$ is correlated with $\mathbf{X}_t$ errors $\mathbf{e}_t$. The variance-covariance between η and X_t is described by matrix $\boldsymbol{\Omega}$, whose entry $\Omega_{\eta,\eta}$ represents the variance of the $\boldsymbol{\eta}$ errors; entry $\Omega_{e,e}$ represents the variance of the $\mathbf{X}_t$ errors; and entry $\Omega_{e,\eta}$ represents the covariance of the $\mathbf{X}_t$ errors with $\boldsymbol{\eta}$, capturing the correlation between fixed effects η_t and attributes $\mathbf{X}_t$.

5 Results

The subnetwork extracted by using snowball method has a node size of 4600, with an I-E ratio at 0.15, which gives us a network has both much less heterogeneity and can be computed much easier. We then use NAFE to estimate the network effect on users decision of uploading new video, by using this network in a time span of six weeks. The analysis results of NAFE model are presented in the Table 1 below.

We find peer influence is significantly associated with users' decision on content creation on YouTube platform, $\rho_1 = 0.46, p < 0.01$. For a focal YouTube channel, the more friends in the network upload new videos in the current time period, the more likely that he/she also uploads new videos in the same period. This finding shows, as a social media, YouTube provides different motivation for users to create content. Furthermore, network influence takes two forms in this context – both in intrinsic preference and observed decision. ρ_1, the coefficient of interdependent preference $W\theta$, could be looked as the measure of intrinsic value, while ρ_2, the coefficient of observed creation decision in the lagged time period $\mathbf{y}_{t-1}$, is a measure of interdependent observed decision. The influence coming from friends' decision of creation is negatively significant at $-0.26, p < 0.10$. Using betweenness centrality as the measure of strength of influence in network, we found it is also positively significant, ($\beta = 1.54 \times 10^{-6}, p < 0.05$), meaning influential users in the friend network are more likely to upload new videos. It suggests that an opinion leader in the network is more likely to create new content, possibly to solidify and enhance he/her influence, and not losing the leadership position in the network.

Channels that have higher increase in the change of views for all the videos provided by the channel are more likely to upload new videos in time period t. We use 1500 as a threshold, because the threshold separate channels into two groups at similar size. For group of channels that have higher increase in the number of video views, on average, they are 16% ($p < 0.05$) more likely to upload new videos than the group of channels that have lower increase in the number of video views). Channels that are more influential are more likely to create new content in period t. The influence

Table 1: Results of YouTube Content Creation Peer Effect Analysis

Parameter	Estimate	Lower	Upper
Subscriber	4.14×10^{-6}**	1.61×10^{-6}	6.66×10^{-6}
	(1.29×10^{-6})		
Betweenness	1.54×10^{-6}**	1.32×10^{-6}	1.76×10^{-6}
	(1.13×10^{-7})		
$\triangle \mathbf{V}_{t-1}$	8.03×10^{-7}*	3.15×10^{-8}	1.57×10^{-6}
	(3.93×10^{-7})		
Peer effect (ρ_1)	0.46**	0.35	0.57
	(0.057)		
$\mathbf{y}_{t-1}$ (ρ_2)	−0.26*	−0.49	−0.037
	(0.12)		

†: $p < 0.10$; *: $p < 0.05$; **: $p < 0.01$;

of each channel in the friend network is measured by using centrality. We use 200 as the threshold to categorize high centrality and low centrality. Channels that are more influential are 18% more likely to create new content than less influential channels ($p < 0.05$). On average, likelihood of content creation increase with the number of subscribers. Channels with more subscribers are more likely to create new content in current time period t. Channels with more 300 subscribers are 17% higher likelihood than those who have less than 300 subscribers ($p < 0.05$).

6 Discussion

In the context of social media and crowdsourcing platforms that have millions or even billions of nodes, data gathered from networks is usually a subnetwork of a much larger network. We consider a method of estimating peer influence that is robust to network size and that is computationally tractable. Our method of network collection also preserves the network properties and structure of the larger YouTube network. Further, an analysis of influence maximization in online settings should consider the clusters of influence and the sub-networks within which individual content creators are embedded in. Focusing on an individual content creator might overstate or understate influence. Our research offers several directions to incorporate subnetwork extraction strategy to study business analytics. Since we consider both content creation (uploading behavior) as well as social ties such as friend ties, our research suggests the need to incorporate large-scale network user interaction data for a meaningful analysis of growth and value of social media platforms. The methodological issues involved in identifying individuals' propensity for content creation could also apply to other settings beyond social media.

Our results suggest that peer effects on YouTube significantly increase competition between content creators. Given the scarcity of attention on YouTube, channels on YouTube seek to actively distinguish themselves from neighboring channels in their content provision strategies. YouTube has recently tried to implement a partnership based channel strategy (Seabrook, 2012) to enable channels to better monetize content. However, a model that effectively changes content provision with voluntary content creation to one where payment strategies are used to compensate content creators could be less effective when peer effects primarily drive content creation. A focus on individual channels and neglecting the community features of YouTube wherein channels are actively connected to other content creators as well as content users might be detrimental. Further, a decision by YouTube to partner with a handful of channels could end up merely reinforcing the dominant position of the top few content creators, similar to results from search engines where pages ranked at the bottom are almost never viewed (Halvey et al., 2006). Such a situation could end up reducing novelty and variety for a casual viewer, and lower the loyalty towards YouTube. The results could have implications in other crowdsourcing and user generated content phenomenon where sites face a tradeoff between content consumer engagement and monetization.

The key to the growth and monetization of crowd sourcing platforms such as YouTube is to motivate and incentivize content creators and simultaneously promoting engagement from content consumers. Recently YouTube has revised the guidelines for commenting and user participation in hopes of increasing an individual user's engagement with the channels he/ she is likely to visit. However, our results suggest that content consumers may be more engaged by novel content rather than encouraging loyalty towards a channel. Indeed we find that peer effects in content creation outweigh the impact of having a large base of subscribers or that of a high number of subscribers in the friend network, suggesting that the scarcity of attention amongst competing channels is the characteristic feature of content creation on YouTube.

References

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies*, 58:277–297.

Bramoulle, Y., Djebbari, H., and Fortin, B. (2009). Identication of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55.

Coleman, J. S. (1958). Relational analysis: The study of social organization with survey methods. *Human Organization*, 17(4):28–36.

Girvan, M. and Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826.

Goodman, L. A. (1961). Snowball sampling. *The Annals of Mathematical Statistics*, 32(1):148–170.

Halvey, M., Keane, M. T., and Smyth, M. (2006). Mobile web surfing is the same as web surfing. *Communications of the ACM - Self managed systems*, 49(3):76–81.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A. H. (2011). *McKinsey Global Institute*.

Rosenbaum, S. (2010). *Meet The Father Of The Second Largest Search Engine In The World*. Business Insider. http://articles.businessinsider.com/2010-11-02/entertainment/29985901_1_web-video-jawed-karim-video-views.

Seabrook, J. (2012). *Streaming Dreams*. New Yorker. http://www.newyorker.com/reporting/2012/01/16/120116fa_fact_seabrook?currentPage=all.

Yoganarasimhan, H. (2012). Impact of social network structure on content propagation: A study using youtube data. *Quantitative Marketing and Economics*, 10:111–150.

An Empirical Analysis of Rational Addiction to Mobile Social Apps

Hyeokkoo Eric Kwon
Korea Advanced Institute of Science and Technology
hkkwon7@business.kaist.ac.kr

Hyunji So
Korea Advanced Institute of Science and Technology
namjjong@business.kaist.ac.kr

Sang Pil Han
City University of Hong Kong
sangphan@cityu.edu.hk

Wonseok Oh
Korea Advanced Institute of Science and Technology
wonseok.oh@kaist.ac.kr

Abstract

As the social technology advances, users become increasingly empowered to promote and maintain their social solidarity through mobile social applications, such as social network services, instant messaging, and social games. However, heavy and uncontrolled use of these apps foster habit-forming activities, and such behaviors can develop into an addiction, just like consuming alcohol, cigarettes, and drugs. In this paper, we extend the rational addiction model of Becker and Murphy (1988) to three types of mobile social apps that are offered at a free of charge. Furthermore, using unique panel data set consisting of the mobile app time-use histories of three mobile social apps, we investigate their rational addictiveness. The initial findings suggest that all three mobile social apps are addictive. However, only users of game apps are found to be rational, while such rational behaviors are not evident among users of social networks service and those of instant messaging. Finally, the subsample analysis reveals that a rational addiction behavior is more pronounced for users with relatively low levels of education and income.

Keywords: Mobile Apps, Rational Addiction, Econometrics

1. Introduction

The emergence of mobile platforms as a social catalyst has been a blessing and a curse. Users are empowered to promote and maintain their social relationships and solidarity through mobile social applications, such as social network services (e.g., Facebook), instant messaging (e.g., KakaoTalk), and social games (e.g., Anipang), which have become more and more ingrained into their everyday lives. Because it is very easy to carry and use the mobile devices anytime and anywhere, users are never out of touch and they are constantly connected to people in their social circles. However, the portability and convenience furnished by the mobile platform also have adverse ramifications on human nature. Heavy and uncontrolled use of mobile social networking and gaming apps foster habit-forming activities, and such behaviors can develop into an addiction, just like consuming alcohol, cigarettes, and drugs. While having breakfast, attending meetings, crossing the street, or even driving, people compulsively check their mobile devices in a desire not to miss a satisfying "social update." As we move more deeply into the mobile era, the apparent addictive preoccupation and impaired dependency with mobile platforms have become vexing social challenges of our times.

In academic circles, the "nature versus nurture" controversy surrounding addictive behaviors continues to interest scholars and medical professionals. In fact, scholars who support the "nature" perspective have long taken it for granted that addiction is an acute type of irrational behavior and a chronic disease that requires psychological and medical treatment to overcome. However, in his seminal article, “Theory of Rational Addiction”, the Nobel Laureate economist Gary Becker and his colleague defied conventional knowledge by claiming that an addiction to substances (e.g., cigarettes and alcohol) can be also explained by a form of rational and utility-maximizing human behavior that can be “nurtured” (e.g., controlled) by economic factors (e.g., price). More specifically, in the rational addiction model, an addict is assumed to not only account for the dependence of current addictive consumption on past consumption, but to anticipate the expected future consequences when making current consumption decisions. This is in contrast to prior “myopic” theories of habit formation where one’s current utility depends only on past and current consumption of the addictive good (Pollak 1970; 1976). The myopic consumer ignores the possibility that his or her current choices would affect future consumption. In the current study, we seek to explore the mobile “app-diction” phenomenon (i.e., addictions to apps) from a perspective of rational addiction frameworks.

In Becker and Murphy’s (1988) theory, the price manipulations of addictive substances can affect the consumption propensity of those types of goods. For example, to discourage cigarette smoking, governments can make a pre-announcement of an increased taxation of the addictive goods (Becker et al. 1994). However, in the context of mobile apps users do not directly incur monetary costs to consume them after they have downloaded and installed the apps. Instead, the number of active users and players within a network (hereinafter, a ‘platform size’) may influence addictive behaviors. This is because when all else is equal, the utility from the use of mobile platform apps, such as social networking and social gaming apps, grows commensurate with the number of active participants. Thus, we extend Becker and Murphy’s rational-addiction model by assuming that the number of active users influences addictive behaviors towards mobile platform apps, and users face a time constraint rather than a budget constraint.

In this study, we use a unique panel data set containing weekly individual-level mobile app usage of thousands of active users to examine whether users’ addictive behaviors for mobile platform apps exhibit ‘socially’ rational patterns. In particular, we address the following research objectives:

(i) To estimate the rational-addiction model for mobile platform apps, we ask:
   a. Do users exhibit socially addictive behaviors for mobile platform apps?
   b. How does the degree of socially rational addiction vary by different mobile platform apps (for example, social networking apps vs. social gaming apps)?
   c. To what extent does user heterogeneity, such as age, gender, income, and education, account for the degree of socially rational addition behaviors?

(ii) To quantify the socio-/economic impact of mobile app addiction, we ask:
   a. What are the short-run and long-run ‘platform size’ elasticities of different mobile platform apps?
   b. How can the preventable cause of addiction to mobile platform apps be characterized?

2. Empirical Context and Data Description

We provide a brief overview of the empirical data. We have gathered a large-scale panel data set consisting of the mobile app time-use histories of three mobile social apps including a popular social networking app (Facebook), a social messaging app (KakaoTalk), and a social gaming app (Anipang). KakaoTalk is a free mobile messenger application for smartphones with free text and free call features.[1] Users can share diverse content and information from photos, videos, voice

[1] http://en.wikipedia.org/wiki/KakaoTalk

messages, and URL links to contact information. Both one-on-one and group chats are available, and there are no limits to the number of friends who can join in group chat. KakaoTalk automatically synchronizes the user's contact list on their smartphones with the contact list on KakaoTalk and finds friends who are on the service. Anipang is a smartphone-based social puzzle game. There are 6 animals consisting of vertical and horizontal lines. Users have to match any three or more animals to get a 'combo,' and when a combo meter fills up, users can get a 'bomb' that destroys a vertical and horizontal row. The basic structure of the game is similar to Candy Crush. One game only last for one minute. After completing a round, players can check their ranking among their friends in KakaoTalk who play the game and compete with them. Further, to participate in this game, a player needs hearts, which are in-app currency that let the player continue to play the game. In addition to just waiting another 8 minutes, there are two ways to get hearts. First, one can get hearts from friends for gifts. Second, by inviting one's KakaoTalk friends to join the game. This mechanism encourages users to continue to communicate with other users.

The data is provided by Nielsen KoreanClick, a Korean market research company that specializes in online and mobile internet audience measurement. Nielsen KoreanClick maintains a panel of mobile app users, ranging between the ages of 10 and 70, selected on the basis of stratified sampling in Korea. After an individual agrees to be a panel member, he or she downloads and installs a Nielsen Mobile App on his or her mobile devices. This app collects data on user's consumption of mobile apps for the purpose of market research. We collected the data between October 1, 2012 and October 31, 2013 (13 months). The data include 5,382 panel members who used the aforementioned three mobile platform apps throughout the sampling period. Moreover, our data incorporate individual-level, weekly information on the app title, date, and duration of the mobile apps used. We discuss our modeling approach in detail below.

3. Statistical Challenges and the Modeling Approach

One of the key statistical challenges we face here is to identify the impact that past and future consumption of a particular mobile app has on the user's current assumption. The rational-addiction model poses endogeneity issues due to the presence of leads and lags of the dependent variable and a potential serial correlation among the disturbances. We consider the following econometric model (see the Appendix for details):

$$C_{i,t} = \delta_0 + \delta_1 C_{i,t-1} + \delta_2 C_{i,t+1} + \delta_3 N_t + \mu_i + \lambda_t + u_{i,t}$$

where the subscript i denotes the i-th user and the subscript t denotes the t-th week. $C_{i,t}$ is the consumption of a given mobile platform app (measured in seconds) by user i and time t. N_t is the number of active users in the network during time t. μ_i is a user-specific effect, λ_t is a week-specific effect, and u_{it} is a remainder disturbance.

In the usual dynamic panel-data models (for example, Arellano and Bond 1991), only $C_{i,t-1}$ appears and not $C_{i,t+1}$. Hence, if the $u_{i,t}$s are serially correlated, we cannot use the usual prescribed instruments, lagged $C_{i,t}$'s such as $C_{i,t-2}$, $C_{i,t-3}$, and so forth, for our model. The Becker and Murphy model suggests that we use both a lagged and a future number of active users [$N_{i,t-1}$, $N_{i,t-2}$, $N_{i,t+1}$] among the set of instruments. However, we found that these instruments may be invalid in our empirical context. Therefore, to address the endogeneity issue, we propose to use a consumption for a *different* mobile app $C'_{i,t}$ as instruments. The rationale for using these instruments is that the current consumption $C_{i,t}$ for one app is less likely to be correlated with the consumption for another app. We found that these proposed instruments are valid in our data and they conform to the requirements necessary for analyzing our panel data.

Table 1. GMM Estimates of Rational Addiction Model for Three Mobile Social Apps

	Anipang[a]	Facebook[b]	KakaoTalk[c]
C(t-1)	0.242***	0.253**	0.0501**
	(0.0575)	(0.122)	(0.0218)
C(t+1)	0.222***	0.133	0.00939
	(0.0653)	(0.143)	(0.0236)
N(t)	10.51***	8.793***	2.948***
	(2.571)	(3.213)	(1.114)
Observations	4,737	4,900	27,652
Number of id	493	522	1,820

NOTE: Robust standard errors are in parentheses. All GMM estimators pass the Arellano-Bond test for AR(1) at the 5% level and pass the same test for AR(2) test, Hansen test of over-identification, and Difference-in-Hansen tests of exogeneity of instrument subset test at the 10% level.

[a]We use consumption levels for Anipang app [$C_{i,t-3}$, …, $C_{i,t-11}$] and Facebook app $C'_{i,t}$ as instruments.

[b]We use consumption levels for Facebook app [$C_{i,t-3}$, …, $C_{i,t-11}$] and Anipang app $C'_{i,t}$ as instruments.

[c]We use consumption levels for KakaoTalk app [$C_{i,t+14}$, …, $C_{i,t-16}$] and Anipang app $C'_{i,t}$ as instruments.

Table 2. Result of Sub-sample Analyses for Anipang

	Education		Income (over the age of 30)		Lifestyle			
	High School Graduated	University Graduated	below \$3000	above \$3000	Conspicuous Consumer	Rational Familist	Sociable Activist	Trend Setter/Leader
C(t-1)	0.382***	0.263***	0.355***	0.267***	0.383***	0.226***	0.221***	0.291***
	(0.0556)	(0.0582)	(0.0373)	(0.0622)	(0.0631)	(0.0708)	(0.0561)	(0.0758)
C(t+1)	0.325***	0.230***	0.360***	0.207***	0.369***	0.120**	0.210**	0.300***
	(0.0612)	(0.0703)	(0.0480)	(0.0686)	(0.0694)	(0.0548)	(0.0844)	(0.0760)
N(t)	12.62***	8.154***	4.725	11.90***	2.979	17.30***	17.00***	4.568
	(4.067)	(2.679)	(2.937)	(3.271)	(1.970)	(4.633)	(4.774)	(3.151)
Observations	473	3,618	662	3,159	932	1,039	1,177	1,476
Number of id	48	357	74	320	84	97	131	164

NOTE: Robust standard errors are in parentheses. We use consumption levels for Anipang app [$C_{i,t-3}$, …, $C_{i,t-11}$] and Facebook app $C'_{i,t}$ as instruments. All GMM estimators pass the Arellano-Bond test for AR(1) at the 5% level and pass the same test for AR(2) test, Hansen test of over-identification, and Difference-in-Hansen tests of exogeneity of instrument subset test at the 10% level

4. Key Results

Table 1 shows GMM estimates of the rational addiction model for three mobile social apps. We find that past consumption and platform size have positive impacts on current consumption in all three apps. However, future consumption has a positive impact on current consumption only in Anipang. This result implies that all three mobile social apps are addictive but, as opposed to conventional wisdom, people addicted to social gaming app are only rational. That is, social gamers are forward-looking and they rationally adjust their current consumption amounts for social game apps based on their prediction of future consumption. In contrast, social networking and social messaging addicts do not rationally decide their current consumption levels. Table 2 shows the result of sub-sample analyses according to user demographic profiles for social gaming app Anipang. Results demonstrate that the rational addiction behavior is more pronounced for social gaming users who have lower levels of education and income than users with higher levels of education and income. In addition, regarding the impact of the users' lifestyles on rational addiction, users characterized as rational familists and social activists exhibit more rationally addictive behaviors than trend-setters and conspicuous consumers.

5. Contributions

Rational addictive behaviors have been documented in the consumption of a variety of products, such as cigarettes (Chaloupka 1991), alcohol (Grossman et al. 1998), caffeine (Olekalns and Bardsley 1996), cocaine (Grossman and Chaloupka 1998), illicit drugs (Saffer and Chaloupka 1999), opium (Liu et al. 1999), and other substances. This paper contributes to the body of literature on addictive behaviors as one of the first studies that examines the rational addiction phenomenon for mobile platform apps. Through the lens of rational addiction theory, the findings of our study provide policy makers and app constituencies with initial implications regarding the preventable cures for IT addiction problems in general and mobile app addiction problems in particular. In addition, we will also discuss the future research directions that our paper heralds for IS and other inter-disciplinary researchers.

References

1. Arellano, M., and Bond, S. 1991. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," The review of economic studies (58:2), pp 277-297.
2. Becker, G., Grossman, M., and Murphy, K. 1994. "An empirical analysis of cigarette addiction," The American economic review (USA)).
3. Becker, G. S., and Murphy, K. M. 1988. "A theory of rational addiction," The journal of political economy), pp 675-700.
4. Chaloupka, F. 1991. "Rational Addictive Behavior and Cigarette Smoking," Journal of Political Economy (99:4), pp 722-742.
5. Grossman, M., and Chaloupka, F. J. 1998. "The demand for cocaine by young adults: a rational addiction approach," Journal of Health Economics (17:4), pp 427-474.
6. Grossman, M., Chaloupka, F. J., and Sirtalan, I. 1998. "An empirical analysis of alcohol addiction: Results from the Monitoring the Future panels," Economic Inquiry (36:1), pp 39-48.
7. Liu, J.-L., Liu, J.-T., Hammitt, J. K., and Chou, S.-Y. 1999. "The price elasticity of opium in Taiwan, 1914–1942," Journal of Health Economics (18:6), pp 795-810.
8. Olekalns, N., and Bardsley, P. 1996. "Rational addiction to caffeine: an analysis of coffee consumption," Journal of Political Economy (104:5), pp 1100-1104.

9. Pollak, R. A. 1970. "Habit formation and dynamic demand functions," Journal of Political Economy (78:4), pp 745-763.
10. Pollak, R. A. 1976. "Habit formation and long-run utility functions," Journal of Economic Theory (13:2), pp 272-297.
11. Saffer, H., and Chaloupka, F. 1999. "The demand for illicit drugs," Economic inquiry (37:3), pp 401-411.

Appendix.
Assume a representative individual *i* with the instantaneous utility function:

(A1) $\mathrm{U_{i,t}} = \mathrm{u}\left[\mathrm{C_{i,t}}, \mathrm{A_{i,t}}, \mathrm{Y_{i,t}}, N_t\right]$

where $\mathrm{C_{i,t}}$ is time spent on an addictive app of *i* at *t*, $\mathrm{A_{i,t}}$ is 'addictive stock' capturing cumulative past consumption of addictive app of *i* at *t*, $\mathrm{Y_{i,t}}$ is time spent on any other non-addictive activities of *i* at *t*, and $\mathrm{N_t}$ is the platform size of addictive app at *t*. Further, we assume that the utility function is concave with negative second derivatives, and that the addictive goods consumption does not affect the marginal utility of non-addictive good consumption.

(A2) $\mathrm{U_{CC}} < 0, \mathrm{U_{AA}} < 0, \mathrm{U_{YY}} < 0, \mathrm{U_{CY}}, \mathrm{U_{AY}} = 0$

There are three characteristics of addictive consumptions—withdrawal, tolerance, and reinforcement— that are represented in mathematical forms below.

(A3) $\mathrm{U_C} > 0, \mathrm{U_A} < 0, \mathrm{U_{CA}} > 0$

Because time is finite, each individual should allotted his (her) time for playing an addictive application within his (her) time budget.

(A4) $\mathrm{C_{i,t}} + \mathrm{Y_{i,t}} = W$

where W is the length of time period t. Due to this time constraint, $\mathrm{Y_{i,t}}$ can be expressed as a function of $\mathrm{C_{i,t}}$. In addition, following prior work on rational addiction (Becker et al. 1994), we assume that addictive stock is equal to the consumption of previous period.

(A5) $\mathrm{A_{i,t}} = C_{i,t-1}$

As a result, we can rewrite the utility function as a function of $\mathrm{C_{i,t}}$, $\mathrm{C_{i,t-1}}$, and $\mathrm{N_t}$.

(A6) $\mathrm{U_{i,t}} = \mathrm{u}\left[\mathrm{C_{i,t}}, \mathrm{C_{i,t-1}}, \mathrm{N_t}\right]$

Then the individual i's problem is to choose $\mathrm{C_{i,t}}$ to maximize the sum of lifetime utility discounted at the rate r.

(A7) $\max_{C} \sum_{t=1}^{\infty} (1+\mathrm{r})^{-\mathrm{t}}\, \mathrm{u}\left[\mathrm{C_{i,t}}, \mathrm{C_{i,t-1}}, \mathrm{N_t}\right]$

where r is constant discount rate. We assume a quadratic functional form for the utility function in $\mathrm{C_{i,t}}$, $\mathrm{C_{i,t-1}}$, and $\mathrm{N_t}$ to obtain the following empirical demand function.

(A8) $$\mathrm{U_{i,t}} = a_1\mathrm{C_{i,t}} + a_2\mathrm{C_{i,t-1}} + a_3\mathrm{N_t} + \frac{1}{2}u_{cc}\mathrm{C_{i,t}^2} + \frac{1}{2}u_{AA}\mathrm{C_{i,t-1}^2} + \frac{1}{2}u_{NN}\mathrm{N_t}$$
$$+u_{CA}\mathrm{C_{i,t}}\mathrm{C_{i,t-1}} + u_{CN}\mathrm{C_{i,t}}\mathrm{N_t} + u_{NA}\mathrm{N_t}\mathrm{C_{i,t-1}}$$

Finally, substitute this equation into the equation (A7), then maximize the equation (A7) to obtain the empirical demand function of an addictive application C(t):

(A9) $C_{i,t} = \delta_0 + \delta_1 C_{i,t-1} + \delta_2 C_{i,t+1} + \delta_3 N_t$

where $\delta_0 = -\frac{a_1+\frac{a_2}{1+r}}{u_{CC}+\frac{u_{AA}}{1+r}}$, $\delta_1 = -\frac{u_{CA}}{u_{cc}+\frac{u_{AA}}{1+r}}$, $\delta_2 = \frac{\delta_1}{1+r}$, and $\delta_3 = -\frac{u_{NC}+\frac{u_{NA}}{1+r}}{u_{cc}+\frac{u_{AA}}{1+r}}$.

Based on (A2) and (A3), positive value of δ_1 indicates that a good is addictive, and positive value of δ_2 indicates that an addiction is result of rational forward-looking behavior.

An Empirical Analysis of Mobile App Time Use

Sang Pil Han
City University of Hong Kong
sangpil78@gmail.com

Sungho Park
Arizona State University
Sungho.Park.1@asu.edu

Wonseok Oh
Korea Advanced Institute of Science and Technology
wonseok.oh@kaist.ac.kr

Abstract

The rapid adoption of smartphones and tablets, as well as the widespread use of mobile applications, has fueled the growth of the mobile app economy. In this study, we use a unique panel data set detailing individual-level mobile app time-use and build a utility theory-based model for multiple discrete/continuous choice of app use. We quantify the baseline utility and satiation levels of different mobile app categories and examine how these vary with user demographics. To allow for category (dis)similarity in unobserved attributes, we employ a factor analytic structure in our multiple discrete/continuous model. The result shows that in terms of use interdependence among different categories of apps, we find that people who spend a great deal of time on social networking apps (e.g., Facebook) also spend a lot of time on multimedia apps (e.g., YouTube), indicating that social networking apps and multimedia apps are economic complements. We discuss the implication of our findings in the context of optimal media planning in the mobile app economy.

Keywords: Mobile Apps, Time-Use Modeling, Interdependence, Econometrics

1. Introduction

The rapid adoption of smartphones and tablets, as well as the widespread use of mobile applications, has fueled the growth of the mobile app economy. According to a recent industry report, more than 220 million people in the U.S. have already become active users of mobile apps, thereby surpassing the number of Internet users on laptops and desktops. The app economy, in which transactions for products and services are conducted through mobile apps, is expanding at an incredible pace; it generated revenue of $72 billion in 2013, and it is projected that the amount of revenue will reach $151 billion by 2017 (APPNATION 2013). Despite the pervasiveness of mobile apps in our everyday economic and social exchanges, however, not all apps are created equal. Some (e.g., communication messengers, games, news/stock feeds) are utilized with great frequency, while a torrent of apps are used only once or twice in their lifetimes. App time-use also depends on audience heterogeneity. User demographics and preferences affect users' intrinsic preference for an app, as well as the marginal utility of an app, as consumption quantity changes. Moreover, some apps are utilized together, while others substitute for each other in their use frequency and amount.

In this study, we use a unique panel data set detailing individual-level mobile app time-use and build a utility theory-based model for multiple discrete/continuous choice of app use that assumes diminishing marginal utility as the level of consumption of any particular app increases (i.e., satiation). In particular, we address the following research objectives:

(i) To quantify the baseline utility and satiation levels of different mobile app categories, we ask:
   a. What is the baseline marginal utility of a particular app category when none is consumed?

b. To what extent does the marginal return (i.e., satiation) diminish as the consumption of a particular category of apps increases?
c. How do the baseline utility and satiation levels vary by different user demographics?

(ii) To examine the app time-use interdependence among different categories of apps, we ask:
a. Are Facebook and YouTube app time-use economic substitutes or complements?
b. How do we estimate app category (dis)similarity in unobserved attributes when the number of app categories is large (e.g., dimensionality issue)?
c. How can optimal media planning in the app economy be characterized?

2. Data Description

We provide a brief overview of the empirical background for our data. We have gathered large-scale panel data comprising mobile app and web time-use histories provided by Nielsen KoreanClick, a Korean market research company that specializes in online and mobile internet audience measurement. Nielsen KoreanClick maintains a panel of mobile app users, aged between 10 and 70, selected on the basis of stratified sampling in Korea. After an individual agrees to be a panel member, he or she downloads and installs a Nielsen Mobile App on mobile devices. This collects data on his or her use of mobile apps and the mobile web for the purpose of market research. The system automatically uploads the encrypted log files from the meter application on panels' mobile devices to a server via secure cellular connection or Wi-Fi.

We collected the data between March 5 and April 30 2012 (8 weeks). These data include 1,425 panel members who used mobile apps and the web throughout the sampling period. Moreover, our data incorporate individual-level, weekly information on the type, name, date and duration of mobile apps used and mobile websites visited. The mobile content types in our data were classified into seven categories, specifically games, multimedia, lifestyle, utilities, communication, social media and other apps, and combined mobile web activities. Table 1 shows that users spend the most time (24%) with communication apps such as mobile messaging apps, followed by multimedia apps (17%), mobile web (17%), and game apps (12%). Table 2 demonstrates that approximately 98% of users in our data used more than five categories of apps during a given week. This descriptive finding of the joint use of multiple app categories lends support to the validity of our econometric model in which we incorporate the multiple-discrete choice into the continuous time-use decision. We discuss our modeling approach in detail below.

Table 1. Time Use According to App Categories

Social Networking Services	6%
Game	12%
Multimedia	17%
Life	8%
Utility	8%
Communication	24%
Others	8%
Mobile Web	17%

Table 2. Number of Apps Jointly Used

Number of Apps Jointly Used	Percent (%)
2	0.5
3	0.4
4	0.8
5	2.7
6	13.2
7	38.3
8	44.2

3. Statistical Challenges and the Modeling Approach

A key statistical challenge we face here is to identify the baseline utility and satiation levels of different categories of apps while allowing for user heterogeneity and cross-app use interdependence even when the number of app categories is large. We also seek to observe which app categories are used and how much time is spent on each selected app category in our data. Accordingly, we describe mobile app usage behavior using a multiple discrete/continuous process. Compared to discrete choice models (i.e. logit or probit models) or continuous dependent variable models, multiple discrete/ continuous models can better handle both the discrete and continuous nature of observed data within a single utility-based framework. Because of this advantage, multiple discrete/continuous models have been successfully applied in several academic fields, including marketing, transportation, and economics (Kim et al 2002; Hendel 1999; Bhat 2008). To allow for category (dis)similarity in unobserved attributes, we employ a factor analytic structure in our multiple discrete/continuous model. The factor model introduces correlation in the latent baseline utilities and satiation parameters across app categories with relatively few parameters. We can thus reduce the number of parameters required to estimate a full covariance matrix while remaining highly flexible. Because of these advantages, researchers have applied a factor analytic structure in random coefficient logit or probit models (Elrod and Keane 1995; Singh et al. 2005; Hansen et al 2006). One methodological contribution of our proposed model is that it extends factor analytic structure to multiple discrete/continuous models (see the Appendix for details).

Table 3. Estimates for Baseline Utility and Satiation Parameters

		Constant	Demographic Variables						
			Age 30's	Age 40's & over	Female	Income mid-class	Income upper-class	Education mid-class	Education upper-class
Baseline Utility	Category	$\bar{\beta}$	Π_β						
	Others	**-2.12**	0.00	0.07	-0.01	0.00	0.03	**-0.16**	0.00
	Game	**-2.80**	0.09	-0.04	**-0.09**	**-0.10**	**-0.25**	0.01	**-0.11**
	Multimedia	**-1.61**	-0.05	**-0.13**	**0.17**	**-0.08**	-0.04	**-0.28**	**-0.29**
	Lifestyle	**-1.64**	-0.05	0.04	**0.12**	0.01	**0.09**	0.06	0.04
	Utility	**-1.52**	**-0.14**	-0.09	0.00	-0.03	-0.04	-0.08	**-0.16**
	Communication	**-1.15**	-0.08	-0.08	**0.28**	0.05	**0.15**	**-0.17**	**-0.21**
	Social Network	**-2.42**	**-0.47**	**-0.63**	**0.08**	-0.05	0.07	-0.11	**-0.12**
	Web	**-1.45**	0.07	-0.01	**-0.08**	-0.07	-0.02	**-0.29**	**-0.21**
Satiation	Category	$\bar{\lambda}$	Π_λ						
	Others	**3.02**	-0.02	0.01	-0.03	**-0.06**	**-0.16**	-0.04	**0.09**
	Game	**1.97**	**0.08**	**0.34**	**-0.15**	**0.11**	**0.17**	-0.06	**-0.13**
	Multimedia	**1.89**	**0.39**	**0.49**	**0.14**	**-0.06**	**0.07**	-0.02	**0.10**
	Lifestyle	**2.92**	**-0.10**	**-0.23**	**0.13**	0.03	**0.11**	**0.20**	0.00
	Utility	**3.26**	**0.17**	**0.17**	**0.16**	**-0.09**	**0.12**	**0.08**	-0.03
	Communication	**1.84**	**0.32**	**0.35**	**-0.19**	-0.02	-0.01	**0.25**	**0.19**
	Social Network	**2.60**	0.04	**0.22**	**-0.33**	**-0.12**	**0.12**	**0.30**	**0.38**
	Web	**2.44**	**0.12**	**0.37**	**-0.11**	**-0.14**	0.02	-0.05	**-0.12**

(Bold: significant at the .05 level)

4. Key Results

Our results in Table 3 show that mobile users' baseline utility for communication apps is the highest and their baseline utility for game apps is the lowest among the mobile web and app categories. Note that as $\overline{\beta}$ increases, the baseline utility increases. Among several demographic variables, gender and education explain substantial heterogeneity in baseline utilities across mobile users. For example, women exhibit a higher intrinsic preference for communication and multimedia apps and a lower intrinsic preference for game apps and the mobile web compared to men. In addition, users with high education levels show significantly lower baseline utilities in most categories, indicating relatively higher utility of outside options (i.e. time used for other activities). The satiation level is the highest in the utility app category and lowest in the communication app category. Note that as $\overline{\lambda}$ increases, the satiation effects also increases. We find that gender and age explain substantial user heterogeneity in terms of the satiation level. As age increases, satiation with communication and multimedia apps increases, while that with lifestyle apps simultaneously decreases. In addition, women show significantly lower satiation levels regarding social media apps than men. Moreover, unobserved factors play an important role in explaining variations in baseline utility and satiation level across mobile users. In our preferred two-factor specification, about 50% of random variation has been explained by the factors. Although we do not report the results due to brevity, in terms of use interdependence among different categories of apps, we find that people who spend a great deal of time on social networking apps (e.g., Facebook) also spend a lot of time on multimedia apps (e.g., YouTube), indicating that social networking apps and multimedia apps are economic complements.

5. Contributions

This paper contributes to the field as the first study to quantify the baseline utility and satiation levels of mobile app categories; moreover, it examines the use interdependence among app categories, which is also a new stream of research. The proposed method also offers a methodological contribution to empirical frameworks concerning multiple discrete/ continuous extreme value choices (Bhat 2005) by reducing the number of parameters required to estimate a full covariance matrix using the factor analytic structure. This allows for high flexibility in estimation even when the number of alternatives is large, which is increasingly common in big data analyses. Further, we will discuss the implications for optimal media planning in the mobile app economy by performing media selection analyses (e.g., Danaher et al. 2010) using our data and results. Thus, this paper contributes to an emerging stream of literature on the economics of mobile internet (e.g., Ghose and Han 2011, Ghose et al. 2013, and Ghose and Han 2014) and mobile marketing (e.g., Bart et al. 2014, Luo et al. 2014). We will also discuss the future research possibilities that our paper heralds for IS and other inter-disciplinary researchers.

References

1. APPNATION. 2013. State of the App Economy. http://appnationconference.com/main/research/.
2. Bart, Y., Stephen, A., and Sarvary, M. "Which Products Are Best Suited To Mobile Advertising? A Field Study of Mobile Display Advertising Effects on Consumer Attitudes and Intentions," *Journal of Marketing Research*. Forthcoming, 2014.
3. Bhat, C. "A Multiple Discrete-Continuous Extreme Value Model: Formulation and Application to Discretionary Time-Use Decision," *Transportation Research Part B*. (39). 2005, pp. 679-707.
4. Bhat, C. "The Multiple Discrete-Continuous Extreme Value (MDCEV) Model: Role of Utility Function Parameters, Identification Considerations, and Model Extensions," *Transportation Research Part B*. (42), 2008, pp. 274-303.

5. Danaher, P., Lee, J., Kerbache, L. "Optimal Internet Media Selection," *Marketing Science*. (29: 2), 2010, pp. 336-347.
6. Elrod, T. and Keane, M.P. "A Factor-Analytic Probit Model for Representing the Market Structure in Panel Data," *Journal of Marketing Research*, (32: February), 1995, pp. 1-16.
7. Ghose, A. and S. Han. "An Empirical Analysis of User Content Generation and Usage Behavior on the Mobile Internet," *Management Science,* (57: 9), 2011, pp. 1671-1691.
8. Ghose, A., A. Goldfarb, S. Han. "How is the Mobile Internet Different? Search Costs and Local Activities," *Information Systems Research*, (24: 3), 2013, pp. 613-631.
9. Ghose, A. and S. Han. "Estimating Demand for Mobile Applications in the New Economy," *Management Science*. Forthcoming, 2014.
10. Hansen, K., Singh, V., Chintagunta, P. "Understanding Store-Brand Purchase Behavior across Categories," *Marketing Science,* (25: 1), 2006, pp. 75–90.
11. Hendel, I. "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns," *Review of Economic Studies*. (66), 1999, pp. 423–446.
12. Kim, J., Allenby, G.M., Rossi, P.E. "Modeling Consumer Demand for Variety," *Marketing Science*, (21: 3), 2002, 229–250.
13. Singh, V., Hansen, K., and Gupta, S. "Modeling Preferences for Common Attributes in Multicategory Brand Choice," *Journal of Marketing Research*, (42: 2), 2005, 195–209.
14. Luo, X., Andrews, M., Fang, Z., and Phang, Z. "Mobile Targeting," *Management Science*. Forthcoming, 2014.

Appendix.

We specify the following latent utility of mobile app use-time:

$$\begin{aligned} U_{ht} &= \frac{1}{\alpha_0} \cdot \exp(\varepsilon_{h0t}) \cdot q_{h0t}^{\alpha_0} + \sum_{j=1}^{J} \frac{1}{\alpha_{hj}} \cdot \mu_{hjt} \cdot \{(q_{hjt}+1)^{\alpha_{hj}} - 1\}, \\ \mu_{hjt} &= \exp(\beta_{hj} + \varepsilon_{hjt}), \text{for } j = 1,...,J, \\ \alpha_{hj} &= 1 - \exp(\lambda_{hj}), \text{for } j = 1,...,J, \end{aligned} \tag{1}$$

where $h = 1,...,H$ denotes individual mobile users, $j = 0$ and $j = 1,...,J$ denote an outside option (activities other than mobile phone use) and mobile web and mobile apps categories, respectively, and $t = 1,...,T$ denotes time period (weeks). q_{hjt} $(j = 0,...,J)$ is time allocated to *j* by user *h* in time period *t*. This specification is referred to as "alpha-profile" in the literature. μ_{hjt} represents the "baseline marginal utility" or the marginal utility when none is consumed. When a user decides which category to use first, categories with large value of μ_{hjt} have higher probabilities of being selected compared to those with small μ_{hjt}. Also it can be interpreted as a measure of "perceived quality" because higher values of μ_{hjt} mean that the alternative confers higher levels of utility from any level of consumption, all else the same. α_{hj} is interpreted as a satiation parameter in that it determines how the marginal utility of alternative changes as consumption quantity increases. If $\alpha_{hj} = 1$, then the marginal utility of j is constant for user *h*, resulting in a linear indifference curve. In this case, user allocates all her time to the category with the highest perceived quality. As α_{hj} decreases, satiation rises, the utility function in alternative *j* becomes more concave, and satiation occurs at a lower value of q_{hjt}. Due to this diminishing marginal utility, choice options are very close substitutes for each other, several options will be selected instead of only one selected option. It can be shown that U_{ht} becomes a proper utility function when $\mu_{hjt} > 0$ and $\alpha_{hj} < 1$ (Bhat 2008).

Note that we specify μ_{hjt} and α_{hj} to satisfy these conditions in (1). ε_{h0t} and ε_{hjt} represent unobservable idiosyncratic part in utility and are assumed to follow Type I Extreme Value distribution.

For user- and category-specific β_{hj}, we specify the following factor analytic structure:

$$\beta_h = \bar{\beta} + \Pi_\beta D_h + \Gamma_\beta \psi_h + \Lambda_\beta \upsilon_h, \tag{2}$$

where a $(J \times 1)$ vector $\beta_h = [\beta_{h1}, \beta_{h2}, ..., \beta_{hJ}]'$ and $\bar{\beta}$ is a $(J \times 1)$ constant vector. D_h is the $(K \times 1)$ vector of observed demographic variables of user *h* and Π_β is a $(J \times K)$ coefficient matrix. Γ_β is a $(J \times F)$ factor loading matrix and ψ_h is a $(F \times 1)$ vector of orthogonal Gaussian factors ($\psi_h \sim N(0, I_F)$). Λ_β is a $(J \times J)$ diagonal matrix and υ_h is a $(J \times 1)$ vector of independent unit-variance Gaussian random variables ($\upsilon_h \sim N(0, I_J)$). The proposed specification decomposes individual-level heterogeneity in mobile usage utility into three parts. The first part is what the observed demographic variables explain. The second part is explained by parsimonious factors. In factor analysis literature, ψ_h is referred to as "common factor". *F* elements in ψ_h influence all *J* elements in β_h. We can understand main characteristics of the factors by interpreting the factor loading matrix Γ_β. Also note that, along with $\Pi_\beta D_h$, this common factors generate correlations in β_h. The remaining variation in β_h is explained by the last term in (2), $\Lambda_\beta \upsilon_h$, which is referred to as "specific factor." Unlike ψ_h, *j*-th element in υ_h influences β_{hj} only.

From (2), we can derive the following covariance matrix of β_h:

$$Cov(\beta_h) = \Pi_\beta \Omega_D \Pi'_\beta + \Gamma_\beta \Gamma'_\beta + \Lambda_\beta \Lambda'_\beta, \tag{3}$$

where Ω_D is a covariance matrix of D_h. The variance decomposition of (3) allows us to quantify the relative contribution of each part. For example, the proportion of variation in β_{hj} explained by observed demographic variable is (*j*-th diagonal element of $\Pi_\beta \Omega_D \Pi'_\beta$)/(*j*-th diagonal element of $Cov(\beta_h)$).

The value of user- and activity-specific satiation parameter α_{hj} is determined by λ_{hj}. Similar to (2), we specify the following factor analytic model structure to λ_{hj}:

$$\lambda_h = \bar{\lambda} + \Pi_\lambda D_h + \Gamma_\lambda \varphi_h + \Lambda_\lambda \nu_h, \tag{4}$$

where $(J \times 1)$ vector is $\lambda_h = [\lambda_{h1}, \lambda_{h2}, ..., \lambda_{hJ}]'$, and $\bar{\lambda}$ is a $(J \times 1)$ constant vector. Π_λ is a $(J \times K)$ coefficient matrix. Γ_λ is a $(J \times F)$ factor loading matrix and φ_h is a $(F \times 1)$ vector of orthogonal Gaussian factors ($\varphi_h \sim N(0, I_F)$). Λ_λ is a $(J \times J)$ diagonal matrix and ν_h is a $(J \times 1)$ vector of independent Gaussian random variables ($\nu_h \sim N(0, I_J)$). The covariance matrix of λ_h can be decomposed as follows:

$$Cov(\lambda_h) = \Pi_\lambda \Omega_D \Pi'_\lambda + \Gamma_\lambda \Gamma'_\lambda + \Lambda_\lambda \Lambda'_\lambda. \tag{5}$$

For the estimation of the proposed model, we first solve the Kuhn-Tucker conditions derived from (1) for constrained utility maximization and then obtain multiple discrete/continuous demand functions. We derive the probability that observed bundles are chosen from these demand functions. Model parameters are estimated by maximizing the probability within the maximum likelihood framework.

The Effect of Interruptions and Page Switching Mode On Mobile Shopping Behavior

Yong Wang
Nanjing University, China
wangyong7994@163.com

Xue Yang
Nanjing University, China
yangxue@nju.edu.cn

Quansheng Wang
Nanjing University, China
wangqs@nju.edu.cn

Abstract

In mobile shopping applications, different page switching modes designed to display product information have been applied to facilitate effectively consumer experiences. However, consumers may encounter interruptions such as short messages and phone calls are frequently in the mobile shopping context with may affect the shopping process and outcomes. By leveraging on Distraction-Conflict Theory and Event Segmentation Theory as the theoretical basis, this study investigates the influence of interruptions on consumers' mobile shopping behavior on the conditions of two types of page switching modes designed in mobile shopping applications. Results from a laboratory experiment indicate that interruptions enhance shopping efficiency in terms of reducing decision time under vertical page switching mode but inhibit shopping efficiency under horizontal page switching mode. The shopping efficiency further influences consumers' information recall and attitude to the mobile shopping application.

Keywords: Mobile Commerce, Mobile Shopping, Interruptions, Page Switching Mode, Lab Experiment, Distraction-Conflict Theory, Event Segmentation Theory

1. Introduction

Mobile shopping is proposed to be the next driver to the technology-enabled business development and has been evolving into a critical part of the online shopping market. For example, it has reached a 9.2% penetration ratio in the whole online shopping market in China in 2013, 4.4% higher than 2012 and expected to reach 24.2% in 2017 (iResearch 2014). Mobile shopping basically supports consumers to enjoy the pleasure and convenience of online shopping through various mobile applications anytime and anywhere.

In order to attract consumers' attentions and ease their shopping effort, mobile shopping application designers pay much attention to the interface design, while page switching mode of product list is one important feature among many others. Since the screen size of mobile devices is much more limited than the PC screens, much fewer amount of product information can be displayed on each one page in the mobile shopping applications. Consumers need to lightly swipe a finger on the touch screen of mobile devices in order to view more products contained in different pages. In existing mobile shopping applications, there are two major types of page switching mode: vertical page switching and horizontal page switching. The former mode is to scroll page up or down to access more information, while the latter is to swipe page forward or backward like reading a book. Previous research has found that page switching mode may influence cognitive effort and task performance

(Bernard et al. 2002, Robertson et al. 1998). For example, in mobile web browings, vertical scrolling allows for faster interface switching with low cognitive load, while horizontal swiping increases the task time and cognitive load (Warr and Chi 2013). However, there is no consensus on which type of page switching mode can motivate more effective mobile shopping task processing and more positive decision outcomes. More importantly, it remains unclear whether the effects of page switching mode on mobile shopping behavior and outcomes differ for consumers in different context, such as interrupted vs. uninterrupted environment.

Interruptions are a frequent occurrence in mobile shopping scenarios. For example, due to the rich functionality of mobile phone, consumers are often interrupted by phone calls, pop-up messages or pushed advertisements when shopping online with mobile phones. Prior research shows that interruptions may result in the loss of working memory content and increased cognitive load (Norman and Bobrow 1975), as a result, affecting decision making and user behavior, such as decision time, accuracy and satisfaction (Wendy 2008, Xia and Sudharshan 2002). In addition, interruptions nay interact with task complexity to improve decision-making performance on simple tasks with low cognitive load, versus lower performance on complex tasks with high cognitive load (Baron 1986, Speier et al. 1999, Speier et al. 2003). Considering different levels of cognitive load incurred under the two types of page switching mode, there is a need to investigate whether and how interruptions make a difference to consumers with different page switching modes when shopping online with mobile applications.

In this research, by lying on Distraction-Conflict Theory and Event Segmentation Theory as the theoretical basis, we focus on the common issues of interruptions during mobile shopping and page switching mode design for mobile shopping applications, and investigate their joint effects on consumers' mobile shopping behavior and outcomes. A lab experiment is designed to test the research model and hypothesis. Theoretical contributions to the existing theories and literature and practical implications to mobile shopping application designers as well as consumers are discussed.

2. Theoretical Background and Hypothesis

2.1 Page Switching Mode

Previous research has compared two types of page switching mode, vertical page switching and horizontal page switching, and discussed the effects of them on consumers' cognition and behavior. An experimental research showed that, when browsing web in mobile devices, vertical scrolling allows for faster interface switching and is less frustrating, while horizontal swiping increases the number of interactions and time to switch as well as users' cognitive load (Warr and Chi 2013). Literature suggests that, in general, vertical page switching facilitates quick scanning od the search results and allows for skipping irrelevant information effectively (Bernard et al. 2002). On the contrary, horizontal page switching inhibits users' memory of information location and consequently makes it difficult to find the target information, resulting in increased cognitive load (Robertson et al. 1998). Some other research indicates that when information is presented in horizontal page switching mode, users are better able to relocate information and remember more details of the information than when the same information is presented in vertical page switching mode, since the latter one exacerbates the cognitive load on users (Piolat et al. 1997, Sanchez and Wiley 2009). The present study will introduce the Event Segmentation Theory to analyze the different effects caused by the two modes further.

2.2 Interruptions

An interruption is “an externally generated, randomly occurring, discrete event that breaks continuity of cognitive focus on a primary task” (Coraggio 1990, pp.12) and typically “requires immediate attention” and “insists on action” (Covey 1989, pp.150-152). Interruptions create interference which increases the overall cognitive processing load and forces an individual to focus or narrow his or her attention on one task at the expense of another (Speier et al. 2003). It may result in forgetting some of the information needed for processing the primary task and consequently, some information cues for decision-making are lost or never enter working memory (Kahneman 1973). After dealing with an interruption, a decision maker returns to the primary work when a recovery period is needed to reprocess information that was forgotten while attending to the interruption or lost from working memory (Edwards and Gronlund 1998). In such cases, decision efficiency is decreased (Dawis et al. 1989). Interruptions have been mainly studied in the context of organizational settings in which the employees' main tasks were interrupted by secondary tasks such as instant messages, phone calls or visitors (Froehle and White 2013, Mansi and Levy 2013). Although there is little research on interruptions in the context of mobile commerce, we propose that they may be more salient in the mobile shopping environment. The reason is that the mobile shoppingvenvironment is very information saturated with enormous amount of product information placed on the screen (Bakos 1997). Therefore, interruptions such as short messages and phone calls requiring immediate response could be more attention catching and consequently impose a more salient influence on consumers’ mobile shopping behavior and outcomes. The study will base on Distraction-Conflict Theory to explain the influence of interruptions on consumers’ mobile shopping behavior with two types of page switching modes designed in mobile shopping applications.

2.3 Research Model and Hypothesis

Event Segmentation Theory (EST) is adopted to discuss the differences between two types of page switching mode in terms of the cognitive resources and attention required. People make sense of continuous streams of activities in daily life partly by segmenting them into events (Kurby and Zacks 2008). According to Zacks and Tversky (2001), an event is defined as a segment of time at a given location that is conceived by an observer to have a beginning and an end. The information entered into sensory is segmented spontaneously into events, and “what is happening now”, is separated from “what just happened”. EST proposes that the segmentation of activities into events happens on an ongoing basis and plays a role in cognitive control (Zacks et al. 2007). When an event model, namely the mental representation of “what is happening now” is stable, effective perceptual predictions about “what will happen” arise, which requires little additional conscious attention and conserves cognitive resources. But when event model is instable, an event boundary is perceived as perceptual predictions cease to be effective. The event boundary is like a "gate" through which the working memory storing "what is happening now" is updated, and more perceptual processing occurs to reset the event model with more cognitive resources and increased attention. Here is an example of how the mechanism in EST might work when one observes an everyday activity: imagine watching a man scrape dishes. He scraps one plate and then a second plate. At this point, it can be effectively deduced from previous observations that he will probably continue to scrape the plates. For the duration of the plate-scraping activity, event models are stable and affairs could be predicted with little cognitive resources. However, when the man scraped the last plate, things would become less predictable, as the inference of his goal to scrape all of the plates would no longer have predictive value for following activities. At this point, event models become instable and perceptual prediction would decline, leading to updating of the event model and consume more cognitive resources. On these grounds, the vertical page switching mode which supports continuous and smooth transition of information, can guarantee the stability of mental representation of product information and consequently a stable event model. As a result, it requires fewer additional attention and cognitive resources when the page is transited to the next. In contrast, the horizontal page switching mode makes the information change completely from one page to the next, similar to going through a "gate" and updating working memory, leading to the instability of mental representation and the event model. Users have to reset event model every time when entering an absolutely new page, resulting in consumption of more cognitive resources and attention.

As to interruptions, existing research on the influence of interruptions on decision performance is guided by Distraction-Conflict Theory (DCT), which was originally proposed to exam distractions. Distractions and interruptions both occur while a decision maker is performing a primary task. The difference between them lies in the sensory channels that detect interruptions/distractions. Distractions are detected by a different sensory channel from those of the primary task and may be ignored or processed concurrently with a primary task (Cohen 1980, Groff et al. 1983), while interruptions use the same sensory channel for both the interruptions and the primary task and cannot be ignored (Kahneman 1973). In short, interruptions are more intrusive than distractions (Speier et al. 2003). For example, when a consumer is shopping online with mobile phone, background music is distraction while phone calls and pop-up short messages are interruptions. However, the cognitive processing and resulting effects of them on decision performance will be similar as both disrupt, and potentially overload, the finite cognitive capabilities of the decision maker (Speier et al. 2003). Therefore, DCT can be adopted in present research. According to DCT, Interruptions facilitate performance on simple tasks and inhibit performance on complex tasks (Baron 1986, Nicholson et al. 2005). The task complexity lies in the number of information cues that must be processed (Wood 1986). Simple tasks require processing fewer cues and consequently fewer cognitive resources than complex tasks (Baron 1986). Interruptions during simple tasks generate increasing stress and narrowed attention, resulting in the possible dismissal or exclusion of part of information cues, thus facilitating decision performance. However, the complex tasks require significantly more mental attention and cognitive resources and a decision maker's excess cognitive capacity decreases when attending to complex tasks. The narrowing of attention when interruptions occur in complex tasks likely results in a decision maker processing fewer information cues, some of which may be relevant to completing the task efficiently, thus resulting in deteriorated performance. In other words, interruptions may facilitate performance on tasks requiring fewer cognitive resources and inhibit performance on tasks requiring more cognitive resources.

Based on the above discussion, EST suggests that vertical page switching mode consumes fewer cognitive resources than horizontal switching mode, and DCT suggests that interruptions will facilitate performance on tasks requiring fewer cognitive resources while inhibit performance on tasks requiring more cognitive resources. Therefore, we believe that, in the context of mobile shopping, there is a joint effect of interruptions and page switching mode on consumers' performance. In particular, when interruptions occur when users are using mobile shopping application with vertical page switching mode, task performance may be heightened because fewer cognitive resources are needed. In contrast, interruptions on users whose mobile shopping application is designed with vertical page switching mode can increase their cognitive load and thus impair their mobile shopping performance. Among various performance measures, decision time is often used as an indicator of performance in decision-making tasks involving interruptions (Mansi and Levy 2013, Speier et al. 2003), page switching mode (Stuart and Austin 1986, Warr and Chi 2013), and online shopping tasks (Hong et al. 2004). Less time spent indicates higher efficiency of decision-making. In this study, we focus on decision time as a performance measure and define it as the time that consumers spend on browsing and comparing products and buying successfully through mobile shopping applications, subtracting time spend on interruptions if any. Therefore, we propose that in the context of mobile shopping:

H1a: When the product display in a mobile shopping application applies vertical page switching mode, decision time will be shorter when the shopping task is interrupted than when the shopping task is not interrupted.

H1b: When the product display in a mobile application applies horizontal page switching mode, decision time will be longer when the shopping task is interrupted than when the shopping task is not interrupted.

Information recall is an indicator of the degree to which the media is successful in conveying information to the consumers (Hong et al. 2004). Previous literature indicates that the exposure time will positively affect information recall (Helsen and Schmittlein 1993, Lim et al. 2011). For example, studies have shown that the longer a person is attending to an advertising, the more advertising content they tend to remember (Krugman et al. 1995). Research also indicates the longer a person stays on a particular web page the more likely they are to remember a banner ad on that page (Danaher and Mullarkey 2003). When consumers shop online through mobile shopping applications, information of different products of the same category is exposed to them in the whole process of decision making. Thus decision time can be treated as similar to information exposure time. Therefore, we propose that in the context of mobile shopping:

H2: Decision time in mobile shopping positively influences information recall, that is, *information recall will be higher when decision time is longer.*

Attitude toward a behavior refers to a person's judgment that performing the behavior is favorable (Hong et al. 2004). Attitude is a reliable indicator of behavior intention, which subsequently affects actual behavior (Ajzen and Fishbein 1980). Consumers adopt electronic commerce or mobile commerce partly because of its time-saving and efficiency (Buellingen and Woerter 2004, Pedersen 2005). Task efficiency will influence attitude toward mobile commerce (Buranatrived and Vickers 2002, Lin 2013). Therefore, as decision time is an indicator of task efficiency, we propose that in the context of mobile shopping:

H3: Decision time in mobile shopping positively influences attitude to using mobile shopping application, that is, attitude toward using a mobile shopping application will be more negative when decision time is longer.

The research model is presented in Figure 1.

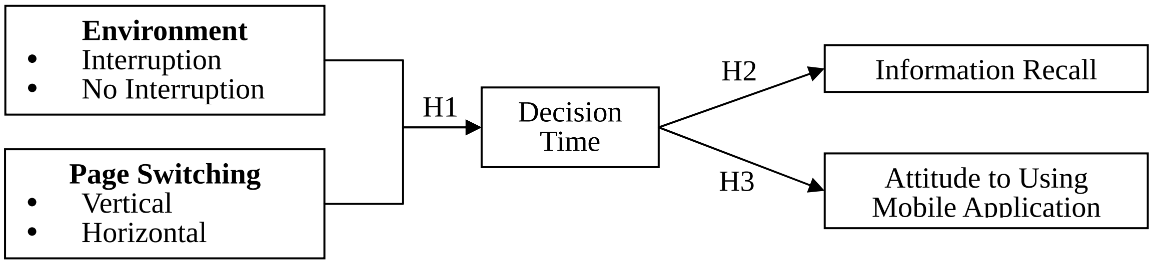


Figure 1. Research Model

3. Research Method

A 2×2 between-subject full-factorial experimental design was employed to test the hypotheses. A pretest was conducted to test the experimental application for mobile shopping. A total of 209 graduate and undergraduate students were recruited to the experiment and randomly assigned to each of the four experimental conditions. Participants' average age is 23.89 years, and 64 percent are female. Average Internet experiences are above 5 years. In addition, 86.6% participants have experience of online shopping and 77% of them ever purchase through mobile phone.

A mobile shopping application was developed for this study. It was installed on a Samsung Galaxy S4 mobile phone with Android system version4.3. Three categories of product, each containing 20 options, were chosen as target for mobile shopping. All the brand names in the product descriptions and pictures were removed. Two types of page switching mode, vertical scrolling and horizontal swipe, were developed for the application. Product information, font size, image size, and color scheme were held constant in both types of page switching. As to interruptions, a simulated short message unrelated with shopping was sent to the experimental mobile phone randomly. Each participant took part in three shopping tasks and was required to buy one product in each product category. Before the formal shopping tasks, a trial product shopping task was provided for participants to familiarize with shopping application.

Decision time was measured in seconds by the average shopping time across the three product categories and was calculated from the mobile log files. Recall of product information was assessed by two measures: recall of price range and recall of attributes contained in product introduction of each product category. Demographic information and attitude toward using mobile shopping application were measured by a questionnaire. Following the guidelines by Ajzen and Fishbein (1980), attitude was measured by three items using ten-point Likert scales.

4. Results Analysis

We performed control checks on gender and participants' experience with online shopping and mobile shopping. There is no significant difference between females and males in their decision time ($t=-1.057$, $p=0.292$), information recall ($t=-2.408$, $p=0.117$), and attitude ($t=-0.606$, $p=0.546$). Statistics tests shows that there are no significant correlations between control variables, including experience with online shopping and mobile shopping, and the intermediary variable or the dependent variables. Hence, none of the control variables have an effect on the intermediary and dependent variables under investigation.

We used SPSS-based Two-Way ANOVA test to analyze the data. The result shows a significant interaction effect between the interruption and the page switching on decision time ($F=3.968$, $p<0.05$), supporting H1a and H1b. Subjects shopping with vertical page switching spend less time when interrupted than uninterrupted, while those shopping with horizontal page switching spend more time when interrupted than uninterrupted. Figure 2 displays the mean levels of decision time in the four experimental conditions. Correlation analysis was conducted to test the other hypotheses. Decision time has a positive correlation with information recall ($R=0.278$, $p<0.01$) and a negative correlation with attitude toward mobile shopping application ($R=-0.297$, $p<0.01$), supporting H2 and H3.

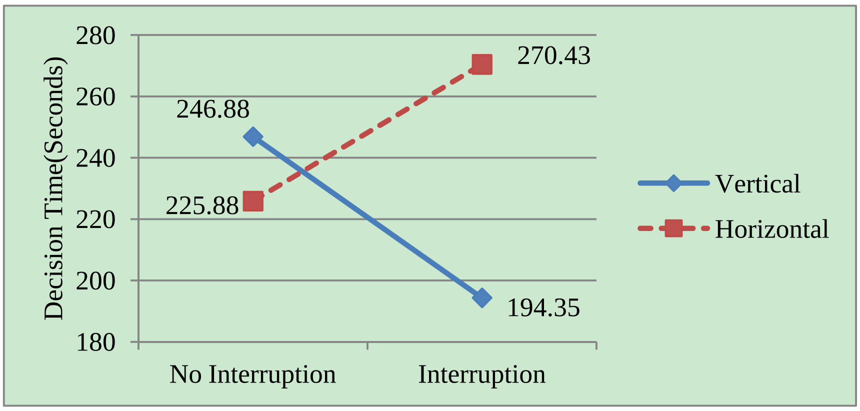


Figure 2. Mean Levels of Decision Time

5. Conclusions

The present study has examined the role of page switching mode and interruptions in influencing consumers' mobile shopping behavior. Utilizing Distraction-Conflict Theory and Event Segmentation Theory as the theoretical basis, our analysis suggests that interruptions may not always negatively influence consumers' decision making. Instead, they interact with page switching mode of the mobile shopping application and result in different effects on shopping behavior and outcomes.

The contributions of this research are twofold. From a research perspective, it applies psychological theories on the issues of interface design in the mobile commerce domain and provide evidence of the importance of page switching mode and interruptions in this new context. From a practical perspective, the findings will have implications for the designers of mobile shopping applications on how to best display product information under interrupted or uninterrupted environment.

Time saving is a indicator of shopping efficiency and affects consumers' attitude to the mobile application. However, besides facilitating consumers' shopping, online retailers hope to retain consumers as long as possible on the application in order to have them more exposed to product information. Future research should pay more attention to the application design that promotes information recall and attitude to mobile application without sacrificing shopping efficiency.

References

Ajzen, I. and Fishbein, M. *Understanding attitudes and predicting social behavior*, Prentice Hall, Englewood Cliffs, NJ, 1980.

Bakos, J. Y. "Reducing Buyer Search Costs: Implications for Electronic Marketplaces," *Management Science*(43:12), 1997, pp. 1676-92.

Baron, R. S. 1986. 'Distraction-conflict theory: Progress and problems.' *Advances In Experimental Social Psychology:* 1-39.

Bernard, M., Baker, R. and Fernandez, M. "Paging vs. Scrolling: Looking for the Best Way to Present Search Results", [online at http://usabilitynews.org/paging-vs-scrolling-looking-for-the-best-way-to-present-search-results/].

Buellingen, F. and Woerter, M. "Development perspectives, firm strategies and applications in mobile commerce," *Journal of Business Research*(57:12), 2004, pp. 1402-08.

Buranatrived, J. and Vickers, P. 2002. 'An investigation of the impact of mobile phone and PDA interfaces on the usability of mobile-commerce applications.' Paper presented at Networked Appliances, 2002. Liverpool. Proceedings. 2002 IEEE 5th International Workshop on. 30-31 Oct. 2002.

Cohen, S. "Aftereffects of stress on human performance and social behavior: A review of research and theory," *Psychological Bulletin*(88:1), 1980, pp. 82-108.

Coraggio, L. 1990. 'Deleterious effects of intermittent interruptions on the task performance of knowledge workers: A laboratory investigation.' The University of Arizona.

Covey, S. R. *The seven habits of highly effective people*, Simon and Schuster, New York, 1989.

Danaher, P. J. and Mullarkey, G. W. "Factors Affecting Online Advertising Recall: A Study of Students," *Journal of Advertising Research*(43:03), 2003, pp. 252-67.

Dawis, R. V., Fruehling, R. T. and Oldham, N. B. *Psychology: Human Relations and Work Adjustment*, McGraw-Hill, New York, 1989.

Edwards, M. and Gronlund, S. D. "Task Interruption and its Effects on Memory," *Memory*(6:6), 1998, pp. 665-87.

Froehle, C. M. and White, D. L. "Interruption and Forgetting in Knowledge-Intensive Service

Environments," *Production and Operations Management*, 2013, pp. n/a-n/a.
Groff, B. D., Baron, R. S. and Moore, D. L. "Distraction, attentional conflict, and drivelike behavior," *Journal of Experimental Social Psychology*(19:4), 1983, pp. 359-80.
Helsen, K. and Schmittlein, D. C. "Analyzing Duration Times in Marketing: Evidence for the Effectiveness of Hazard Rate Models," *Marketing Science*(12:4), 1993, pp. 395-414.
Hong, W., Thong, J. Y. L. and Tam, K. Y. "The Effects of Information Format and Shopping Task on Consumers' Online Shopping Behavior: A Cognitive Fit Perspective," *Journal of Management Information Systems*(21:3), 2004, pp. 149-84.
iResearch. "China Online Shopping Market Slows Down While Mobile Buying Outbursts", 2014, [online at http://www.iresearchchina.com/views/5548.html].
Kahneman, D. *Attention and effort*, Prentice-Hall, Englewood Cliffs, NJ, 1973.
Krugman, D. M., Cameron, G. T. and White, C. M. "Visual Attention to Programming and Commercials: The Use of In-home Observations," *Journal of Advertising*(24:1), 1995, pp. 1-12.
Kurby, C. A. and Zacks, J. M. "Segmentation in the perception and memory of events," *Trends in Cognitive Sciences*(12:2), 2008, pp. 72-79.
Lim, Y.-M., Yap, C.-S. and Lau, T.-C. "The Effectiveness of Online Advertising in Purchase Decision: Liking, Recall and Click," *Journal of Applied Sciences Research*(7:9), 2011, pp. 1517.
Lin, H.-F. "Determining the relative importance of mobile banking quality factors," *Computer Standards & Interfaces*(35:2), 2013, pp. 195-204.
Mansi, G. and Levy, Y. "Do instant messaging interruptions help or hinder knowledge workers' task performance?," *International Journal of Information Management*(33:3), 2013, pp. 591-96.
Nicholson, D. B., Parboteeah, D. V., Nicholson, J. A. and Valacich, J. S. 2005. 'Using Distraction-Conflict Theory to Measure the Effects of Distractions on Individual Task Performance in a Wireless Mobile Environment.' Paper presented at System Sciences, 2005. HICSS '05. Proceedings of the 38th Annual Hawaii International Conference on. 03-06 Jan. 2005.
Norman, D. A. and Bobrow, D. G. "On data-limited and resource-limited processes," *Cognitive Psychology*(7:1), 1975, pp. 44-64.
Pedersen, P. E. "Adoption of Mobile Internet Services: An Exploratory Study of Mobile Commerce Early Adopters," *Journal of Organizational Computing and Electronic Commerce*(15:3), 2005, pp. 203-22.
Piolat, A., Roussey, J.-Y. and Thunin, O. "Effects of screen presentation on text reading and revising," *International Journal of Human-Computer Studies*(47:4), 1997, pp. 565-89.
Robertson, G., Czerwinski, M., Larson, K., Robbins, D. C., Thiel, D. and Dantzich, M. v. 1998. 'Data mountain: using spatial memory for document management.' *Proceedings of the 11th annual ACM symposium on User interface software and technology:* 153-62. San Francisco, California, USA: ACM.
Sanchez, C. A. and Wiley, J. "To Scroll or Not to Scroll: Scrolling, Working Memory Capacity, and Comprehending Complex Texts," *Human Factors: The Journal of the*

Human Factors and Ergonomics Society(51:5), 2009, pp. 730-38.

Speier, C., Valacich, J. S. and Vessey, I. "The Influence of Task Interruption on Individual Decision Making: An Information Overload Perspective," *Decision Sciences*(30:2), 1999, pp. 337-60.

Speier, C., Vessey, I. and Valacich, J. S. "The Effects of Interruptions, Task Complexity, and Information Presentation on Computer-Supported Decision-Making Performance," *Decision Sciences*(34:4), 2003, pp. 771-97.

Stuart, K., Card and Austin, H., Jr. "A multiple, virtual-workspace interface to support user task switching," *SIGCHI Bull.*(17:SI), 1986, pp. 53-59.

Warr, A. and Chi, E. H. 2013. 'Swipe vs. scroll: web page switching on mobile browsers.' *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems:* 2171-74. Paris, France: ACM.

Wendy, L. "Focusing on Desirability: The Effect of Decision Interruption and Suspension on Preferences," *Journal of Consumer Research*(35:4), 2008, pp. 640-52.

Wood, R. E. "Task complexity: Definition of the construct," *Organizational Behavior and Human Decision Processes*(37:1), 1986, pp. 60-82.

Xia, L. and Sudharshan, D. "Effects of Interruptions on Consumer Online Decision Processes," *Journal of Consumer Psychology*(12:3), 2002, pp. 265-80.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S. and Reynolds, J. R. "Event perception: A mind-brain perspective," *Psychological Bulletin*(133:2), 2007, pp. 273-93.

Zacks, J. M. and Tversky, B. "Event structure in perception and conception," *Psychological Bulletin*(127:1), 2001, pp. 3-21.

Inside the Minds of Software Pirates: A Comparison Study of American and Chinese Pirates

Xiaorui Hu
Saint Louis University
hux2@slu.edu

Bin Lu
Zhejiang University of Finance and Economics
lu1bin@yahoo.com.cn

Wenbo Kuang
Renmin University of China
wbkuang@vip.sina.com

Guohua Wu
California State University at Fullerton
mwu@exchange.fullerton.edu

Abstract

U.S. and China are the top two world's economies with the highest commercial value of pirated software. This paper aims to enter the minds of the "pirates" in these countries to study their pirating behavior and seek ways to reduce the piracy rates. We extend the software piracy model of Peace et al. (2003) by adding two paths from punishment severity to subjective norm and punishment certainty to subjective norm. Hypotheses on the relationships among attitude toward software piracy, perceived punishment, subjective norms, perceived behavioral control and piracy intention are developed and tested. With a student sample from the U.S. and Chinese universities, SEM analyses validate the extended software piracy model. The key factors that influence the piracy intention are identified. Our findings provide insights to the decision-making process of the pirates, and help seek effective ways to improve government involvement and legal enforcement to combat the worldwide software piracy.

Keywords: Software Piracy, Intellectual Property, Ethics, Deterrence Theory, Expected Utility Theory, Theory of Planned Behavior

1. INTRODUCTION

Software piracy has been a major problem for decades and costs billions of dollars worldwide annually. Despite the significant amount of effort made by many governments, organizations and manufacturers around the world, the worldwide retail value loss of the unlicensed software still reached $63.4 billion in 2011 (BSA 2012). As the United States and China have suffered the most from software piracy, with losses far exceed other countries, this research is dedicated to study the minds inside the "pirates" in these two countries.

Academic researchers have been paying attention to the software piracy issue in the past few decades since Richard Mason (Mason, 1986). Recently, studies on cross-cultural settings began to emerge (Gopal & Sanders 1998; Moores & Dhillon 2000; Walls & Harvey 2006; Yang et al. 2009), and several studies have been done specifically on Chinese software piracy issues (Davison et al. 2009; Martinsons & Ma 2009). This paper employs the theory of planned behavior (Ajzen 1991) as a framework, and extends an existing software piracy model (Peace et al. 2003) by adding two additional paths to the subjective norms.

2. LITERATURE REVIEW

Software piracy has been identified as an important research stream by many (Gopal & Sanders 1997, 2000; Moores & Dhillon 2000). Researchers have attempted to either offer explanations of individual pirating behavior (Gopal & Sanders 1997; Banerjee et al. 1998; Wagner & Sanders 2001; Lau 2003; Peace et al. 2003; Siponen, Vance & Willison 2012; Walls & Harvey 2006), or to establish models to predict individual behaviors (Ajzen 1991; Banerjee et al. 1998; Cronan & Al-Rafee 2008; Leonard et al. 2004; Peace et al. 2003). In this paper, we will focus on both streams -- (1) developing an extended software piracy model and (2) validating and testing the model using data from both the United States and China.

Tan (2002) developed an ethical decision making model to examine consumers' purchase intention of pirated software. He found that a consumer's purchase intention was influenced by certain aspects of their perceived moral intensity (magnitude of consequence, temporal immediacy and social consensus), perceived risks (financial, prosecution and social risks), and moral judgment (cognitive moral development and moral reasoning). Cronan and Al-Rafee (2008) used TPB and added "moral obligation" and "past piracy behavioral" as the antecedents to intention, and found those two factors had significant impact on piracy intention. Gopal and Sanders (1998) present a general model of ethical behavioral related to the impact of behavioral and cultural factors on software piracy, and ethical predisposition to deontology and teleology, age and gender were presented as antecedents to ethical intention in the model.

This research is to test a well-established software piracy model in the Chinese cultural setting, and we hope to provide a set of constructive feedbacks on the model and provide insights to both the American and the Chinese software piracy issues.

3. RESEARCH MODEL

In this paper, we adopted the framework of the Theory of Planned Behavior (Ajzen 1991) and extended the software piracy model from Peace, Galletta and Thong (2003) to analyze and compare the different factors between American and Chinese students. Peace et al.'s software piracy model (2003) has integrated the theory of reasoned action (Ajzen & Fishbein 1977), the theory of planned behavior (Ajzen 1991), the expected utility theory, the deterrence theory (Tittle 1980), and added the attitude dimension that includes an ethical behavior concern. We extend the Peace et al.'s model by adding two additional paths -- from punishment certainty to subjective norm and from punishment severity to subjective norm. The extended software piracy model is shown in Figure 1.

3.1 Antecedents of Piracy Intention

According to Ajzen 1991, intentions are assumed to capture the motivational factors that influence a behavior. They are indications of how hard a person is willing to try and how much effort he or she plans to exert to perform a certain behavior. Since attitude is the most significant predictor of intention, we propose:

H1: Attitude toward software piracy is positively related to software piracy intent.

Subjective norm refers to the perceived peer or social pressure toward a certain behavior. If the important referents around an individual approve of certain behavior, the individual will feel

more at ease while performing it. Previous researchers offer diverse findings on the effect of subjective norms. Therefore, we propose:
H2: Subjective norm is positively related to software piracy intention.

In the same fashion, perceived behavioral control refers to an individual's perceived ease or difficulty of performing a specific behavior of interest (Ajzen 1991). The previous studies have shown that a person's behavior was strongly influenced by his or her confidence in the ability to perform it. Therefore, we propose:
H3: Perceived behavioral control is positively related to software piracy intention.

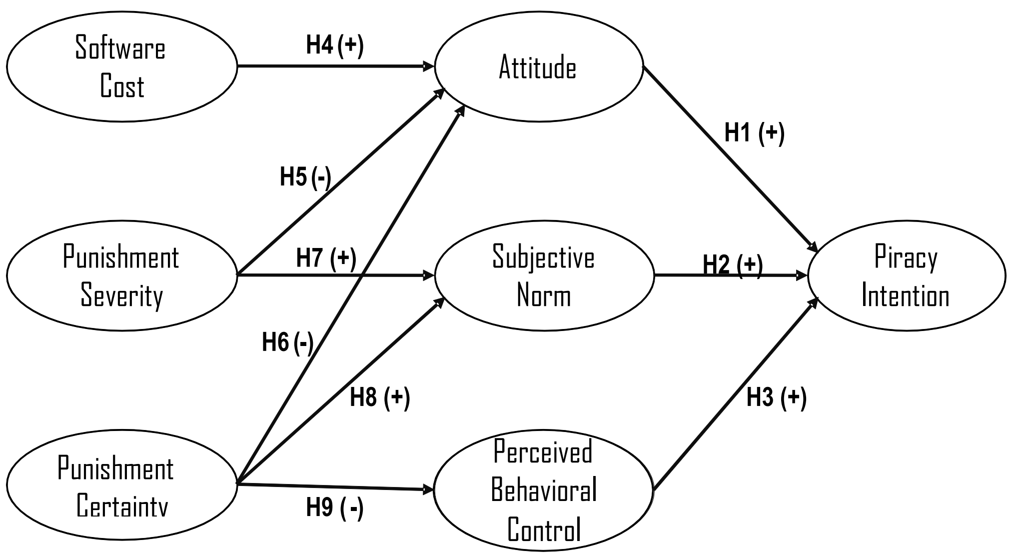


Figure 1: Extended Software Piracy Model

3.2 Antecedents of Attitude

According to TPB model (Ajzen 1991), three types of salient beliefs are distinguished, and behavioral belief is assumed to influence attitude toward the behavior. We apply the expected utility theory and deterrence theory to offer three salient behavioral beliefs to explain the mechanisms behind the pirates' attitude-forming process toward software piracy: software cost, punishment certainty and punishment severity.

We believe that an individual favors a behavior if he or she believes desirable consequences are presented in that behavior, and forms unfavorable attitude toward the behavior if he or she perceives mostly undesirable consequences and results with that behavior. Therefore, we propose:
H4: Perceived software cost is positively related to attitude towards software piracy.

Deterrence theory suggests that the perceived punishment certainty and punishment severity have impact on a person's attitude toward adverse behavior (Tittle 1980). Obviously, both the punishment certainty and severity serve as factors that affect feelings toward the pirating behavior. Therefore, we hypothesize:
H5: Perceived punishment severity is negatively related to attitude toward software piracy.

H6: Perceived punishment certainty is negatively related to attitude toward software piracy.

3.3 Antecedents of Subjective Norm

Subjective norm is a social factor that refers to an individual's perceived social pressure to perform or not to perform a specific action. The perceived opinion from those important people around an individual will comprise the individual's subjective norm. Normative beliefs are believed to constitute the underlying determinants of subjective norms (Ajzen 1991). We believe that the two deterrence factors contributed to subjective norms are punishment certainty and punishment severity. If the action of software piracy has a higher likelihood of being caught, the potential public disgrace and shame will elevate one's perceived disapproval from the important referents. Thus, we hypothesize:

H7: Perceived punishment severity is positively related to subjective norm.

H8: Perceived punishment certainty is positively related to subjective norm.

3.4 Antecedents of Perceived Behavioral Control

Perceived Behavioral Control (PBC) refers to the ease or difficulty of performing the behavior and it is assumed to reflect past experience as well as anticipated impediments and obstacles (Ajzen 1991). We believe that the more resources and opportunities an individual believes he or she owns, the lesser the hindrance he or she expects, the greater his or her perceived behavioral control over the behavior might be. Therefore, we propose:

H9: Perceived punishment certainty is negatively related to perceived behavioral control.

4. RESEARCH METHODOLOGY

4.1 Questionnaire Development

The survey questionnaire consists of two parts. Part I provides general demographic information, such as gender, year of college, and major. Part II contains questions that measure the constructs in the research model, including items for attitude toward software piracy, perceived punishment certainty, perceived punishment severity, subjective norms, software costs, perceived behavioral control and piracy intention. All scale items in part II are adopted from Peace et al. (2003).

4.2 Sample

Instructors at four different universities (two Chinese and two American universities) distributed paper-based survey questionnaires to their students in class. A total of 674 valid survey questionnaires were collected (364 Chinese surveys vs. 310 American surveys).

5. RESULTS

82.7% of the Chinese respondents (vs. 51.6% of the American respondents) reported prior software piracy behavior. We analyzed data by adopting a two-stage process – (1) the measurement model assessment and (2) the structural model assessment.

5.1 Measurement Model Assessment

To evaluate the measurement model, we conduct a confirmatory factor analysis (CFA) for the American subsample and the Chinese subsample, respectively. AMOS 17.0 is used with maximum likelihood method (Byrne 2001). For the American subsample, the CFA generates a significant chi-square ($\chi2$ = 212.98, p < .001), and the fit indexes also indicate an acceptable fit: comparative fit index (CFI = .97), incremental fit index (ILI = .97), and root mean square error of approximation (RMSEA = .045). The composite reliability (CR) and the average variance extracted (AVE) for each construct are then calculated using the formula proposed by Hair et al. (1998). Similar results are found in Chinese sub-sample.

5.2 Structural Model Assessment

We conduct structural equation modeling using AMOS 17.0. Table 1 summarizes the path estimates and goodness-of-fit indexes for the American subsample (Chi-square = 280.95, *df* = 139, p < .001; IFI (Incremental Fit Index) = .95; TLI (Tucker Lewis Index) = .92; CFI (Comparative Fit Index) = .95; and RMSEA = .057). Similarly, we found Chinese subsample also indicates the model fits the data reasonably well.

Structural Path	**Standardized Estimate (S.E.)**	**t-value**	**P**
H4 Software Cost → Attitude toward Piracy	.06 (.073)	1.06	n.s.
H5 Punishment Severity → Attitude toward Piracy	-.27 (.049)	-3.77	***
H6 Punishment Certainty → Attitude toward Piracy	-.57 (.122)	-6.35	***
H7 Punishment Certainty → Subjective Norms	.98 (.176)	7.44	***
H8 Punishment Severity → Subjective Norms	.22 (.047)	3.17	***
H9 Punishment Certainty → Perceived Behavioral Control	-.60 (.161)	-6.28	***
H1 Attitude toward Piracy → Piracy Intention	.45 (.115)	5.54	***
H2 Subjective Norms → Piracy Intention	.01 (.144)	.11	n.s.
H3 Perceived Behavioral Control → Piracy Intention	.57 (.109)	5.89	***
Model fit indexes:			
Chi-square = 280.95 *df* = 139, p < .001; IFI = .95 TLI = .92 CFI = .95 RMSEA = .057			

* p < .05 ** p < .01 *** p < .001 n.s. = not significant

Notes: standard errors are in parentheses.

Table 1: Model Fit and Tests of Proposed Relationships for the U.S. Sub-sample

6. CONCLUSION

The theoretical contributions of this research are three folds. First, it validates and challenges the Peace et al. (2003) model, and empirically tests it in both the U.S. and the Chinese data. Five paths in the original model are empirically validated by both countries' data, yet two insignificant paths from the original model are also found and strongly supported by the data from both countries. Second, the extended two additional paths proposed by this paper have been proven to be significant. The two additional paths from punishment severity and punishment certainty to subjective norm are supported by both the U.S. and the Chinese subsamples. These two factors are established as effective antecedents impacting subjective norm. Third, this paper provides insights to future researchers in terms of developing more effective software piracy models to better suit the young generations.

Our results also provide some useful insights for combating software piracy. The first insight the data offers is that among the three antecedents to software piracy intention only two are significant: attitude toward software piracy and perceived behavioral control. Subjective norm sadly exits from the significant list, posing a "scare" to the software industry that peer pressure and important referents' disapproval no longer matters to the young pirates' pirating intention. The second insight of the paper is that we reveal that the software cost is not significant in terms of deterring the favorable attitude toward piracy behavior. The third insight is that punishment certainty and punishment severity stand out among all antecedents. Though they are indirect impact factors to intention, yet they drive to the heart of deterring the piracy intention. The fourth insight is derived from the fact that perceived behavioral control has more explanation power to piracy intention than the attitude. One final insight of the paper is a call to make the subjective norm work. It is bothersome for us to realize that the important referents' disapproval is no longer effective on deterring the young generation to do something both illegal and unethical. Researches are needed to identify ways to put the subjective norm back to the game. Enhancing their moral standard and reducing individualism in the "pirates" might be some starting points.

References

1. Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211.
2. Cronan, T.P. and S. Al-Rafee (2008). Factors that influence the intention to pirate software and media. *Journal of Business Ethics*, 78(4), 527-545.
3. Davison, R.M., Drummond, D., Martinsons, M.G., Li, Y., Ou, C.X., Lo, H.W.H., and Murata, K. (2009). The ethics of IT professionals in Japan and China. *Journal of the Association for Information Systems*, 10(11), 835-859.
4. Gopal, R.D., and Sanders, G.L. (1998). International software piracy: Analysis of key issues and impacts. *Information Systems Research*, 9(4), 380-397.
5. Hair, J.F., Anderson, R.E., Tatham, R.L., and Black, W.C. (1998). *Multivariate data analysis*. (4th ed). Upper Saddle River, NJ: Prentice Hall.
6. Martinsons, M.G., and Ma, D. (2009). Sub-Cultural differences in information ethics across China: focus on Chinese management generation gaps. *Journal of the Association for Information Systems*, 10(11), 817-833.
7. Moores, T.T., Alexander, N., Rothenberger, M.A. (2009). Knowledge of software piracy as an antecedent to reducing pirating behavior. *Journal of Computer Information Systems*, 50(1), 82-89.
8. Peace, A.G., Galletta, D.F., and Thong, J.Y.L. (2003). Software piracy in the workplace: A model and empirical test. *Journal of Management Information Systems*, 20(1), 153-177.
9. Shadow Market, 2011 BSA global software piracy study, Ninth edition, Published May 2012, retrieved March 11, 2014, from http://globalstudy.bsa.org/2011/downloads/study_pdf/2011_BSA_Piracy_Study-Standard.pdf
10. Tan, B. (2002). Understanding Consumer Ethical Decision Making with respect to the Purchase of Pirated Software. *The Journal of Consumer Marketing*, 19(2/3), 96-111.
11. Tittle, C.R. Sanctions and Social Deterrence. New York: Praeger, 1980.
12. Yang, D. Sonmez, M. Bosworth, D., and Fryxell, G. (2009). Global Software Piracy: Searching for Further Explanations. Journal of Business Ethics, 87, 269-283.
13. Siponen, M., Vance, A., & Willison, R. (2012). New insights into the problem of software piracy: The effects of neutralization, shame, and moral beliefs. *Information & Management*, 49(7), 334-341.
14. Walls, W., and Harvey, P. (2006). Digital pirates in practice: analysis of market transactions in Hong Kong's Pirate Software Arcades. International Journal of Management, 23(2), 207-215.

Is Attribute-based Customization always better? An Extended Attribute-based online Customization system proposed and its advantage

Ning Wang
Renmin University of China
piao_nn@163.com

Kanliang Wang
Renmin University of China
kanliang.wang@gmail.com

Abstract

Attribute-based customization system is now commonly used for online product customization and it is validated better than Alternative-based by researchers. However, both practice and studies have neglected the disadvantages that it force consumer to make sequential choice and cannot support simultaneous comparison which do bad for choice confidence. For this, this paper first proposed an "Extended Attribute-based OCS"supporting simultaneous comparison. Then based on the underpinning of "Sequential choice "and "Simultaneous choice", the paper compared Extended Attribute-based customization system vs. Attribute-based customization system, illustrated their different effect on the decision confidence. Two moderators: task complexity and two types of consumers (Maximizer/Satisficer) were explored. It was found: Extended Attribute-based decision leads to more perceived confidence about the final choice. But this advantage is relatively large only if the task complexity is moderate; Maximizer matches with the Extended Att-based OCS better, while Satisficer is more suitable for Attribute-based OCS.

Keywords: Online customization system, Attribute-based customization, perceived choice confidence, Sequential choice, Simultaneous choice

1. Introduction

Online customization is now being a fashionable trend in varies industries. Firms like Adidas, Nike, and Dell and so on are examples of pioneers applying online customization platform allowing customers to configure products and services. Online Customization System (abbreviated as OCS) which is also called configurator, choice board, design system, user toolkits, online decision support system, online customization support system, customization interface, and so on (Franke and Piller 2003;kamis ,et al. 2008; Valenzuela, et al. 2009), is an indispensible bridge between the user and the manufacturer. Various studies showed that OCS can significantly affect customer decision making and behavior (Kamis,et al. 2008; Bharati and Chaudhury 2004; Haubl and Trifts 2000).

At present, Attribute-based customization (Att-based for short) vs. Alternative-based customization (Alt-based for short) are two kinds of OCS most widely used in many online customization websites. Att-based is more popular in practice, and it is also one of the focuses of OCS studies (Kamis,et al. 2008; Valenzuela, et al. 2009; Iyengar, et al. 2010;Randall, et al. 2007). Att-based OCS which presents users with all of the product attributes that can be customized along with all of the possible options for each attribute. Users then customize their product by selecting an option for each customizable attribute while the Att-based OCS continuously updates the image of the customized product. Alt-based OCS simply displays all of the possible product alternatives to the users who can then pick out the product alternatives they prefer. Users do not actively customize the product themselves while multiple alternatives are shown and customers may compare and evaluate multiple alternatives before settling on a single alternative (Kamis et al., 2008).

Researchers found that Att-based customization was superior to Alt-based in the aspects of decision difficulty, perceived ease of use, decision satisfaction, intention to use aid and intention to purchase the product (Kamis,et al.2008; Valenzuela, et al.2009).

However, there is an observation from practices that: by an Att-based OCS, only the last configuration of product can be presented in front of customers, the configuration customer did before is replaced and updated by recent choices, thus can't be seen and traced. Consumers have to decide to accept or reject each option as it is presented, with the former configuration unavailable. By this sequential presentation way of alternatives, consumers can't contrast their configurations together, they may not sure the last one is the best. Nevertheless, the perceive value of choice is largely from comparing different option (Huber, et al. 1982; Simonson and Tversky 1992). On the other hand, in an Alt-based OCS, multiple alternatives are shown and customers may compare and evaluate multiple alternatives simultaneously before settling on a single alternative (Kamis et al. 2008), but it does not provide change of unfavorite attributes and the choice set are too much, too complex to compare, bring about cognitive overload and thus the decision confidence is much lower (Iyengar, et al.2000). Therefore, these two systems both have advantages and disadvantages.

Furthermore, there is also newest evidence about consumer decision: customers may be less satisfied and certain with the choices they make if their options are presented one at a time rather than all at once (Mogilner et al.2012).while in the Att-based OCS, consumers are actually presented one configuration at one time.Unsatisfied with the present one or the expectation for unknown one, customers can move on and choose one by one sequentially, which is very similar to "sequential choices". So, Att-based OCS may not overall perfect in any perspective. It may have potential defects neglected by prior studies in customization field.

This paper is motivated by the above observation from practice and studies: How can the Att-based OCS gain the advantage of simultaneous choice of various configurations which is obsessed by Alt-based? Following this consideration, firstly we propose a new OCS, named Extended Attributed-based OCS, absorbing the advantages of Alternative-based OCS.Then we compare it with widespread Att-based OCS,and demonstrate: whether the Extended Att-based OCS can lead to more confident with final choice as compared to Att-based OCS? Which one is better for consumers' decision and under what conditions can the new proposed OCS be suitable for use?

2. An Extended Attribute-based OCS proposed

Whether there is a balance between Att-based and Alt-based emphasizing visual comparison of choices or configurations selected by customers? These directions have some explorations now. Christian et.al (2013) proposed a two-step customization model: (1) choosing a "starting solution" from a subsample of pre-specified alternatives first, and (2) refining that starting solution to create the final self-designed product, they revealed that this two-step customization model was superior to previously proposed customization models, leading to less choice complexity, increased preference certainty and product satisfaction. However, this research is more like a combination of Alt first and Att later, the need of configuration comparison can also not be realized.

Here we proposed a new OCS, named Extended Att-based OCS, which is a combination of

Att-based and Alt-based. Its interface inherited Att-based OCS, with the following differences: Att-based OCS only allow **single choice in** choosing options for a customizable attribute, while Extended Att-based OCS allow **multiple choices** in choosing options for a customizable attribute. In Extended Att-based OCS, customer can select more than one preferred options in a customizable attribute and then alternatives are automatically generated by randomly combination of the chosen options, thus customers can inspect and compare all their favorite options visually and simultaneously. It successfully solves the problem that alternatives passed-up cannot be unavailable later. In addition ,it enable the consumers to get all combination possibilities of their preferred options simultaneously ,without the need to manfully compose different attributes together and worry about the omission of potential better combination .An illustration of the interface is shown in figure 1.

Figure 1 Att-based OCS vs. Extended Att-based OCS

3. Theoretical Foundations and Hypotheses Development

The theoretical model for the study is presented in Figure 2. We develop the rationale for these relationships below.

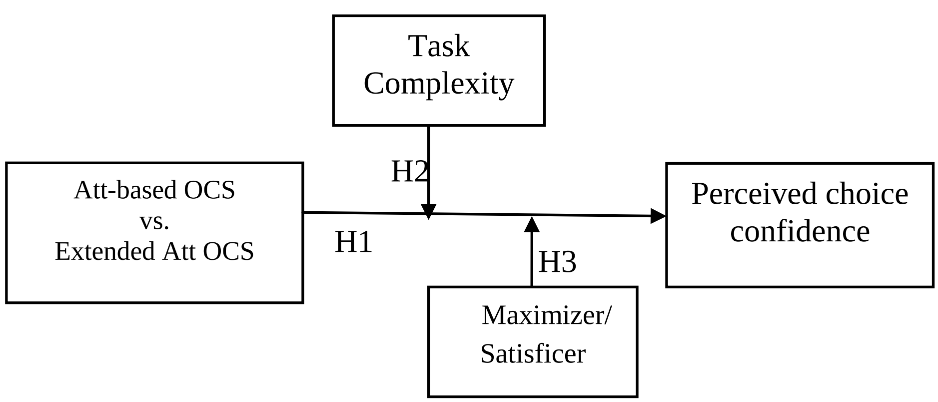


Figure 2 Research Model

Simultaneous Choice vs. Sequential Choice

The theoretical basis for our examination of the difference between an Extended Attribute--based and an Attribute-based OCS comes from two choice presentation modes: Simultaneous Choice vs. Sequential Choice.

Simultaneous Choice means all options of choices are presented at one time, while one option is presented at a time in Sequential Choice. The method of option presentation affects what choosers compare their selected outcome to when determining their satisfaction. Sequential choices force choosers to evaluate that option relative to a reference in their mind (Farquhar and Pratkanis 1993). Prior research has shown that when individuals evaluate an option separately, they compare that option's list of attributes against those of a typical option (Hsee and Leclerc 1998). Notably, however, when (a) choosers are presented with options that vary along an unspecified list of attributes, (b) a natural reference is not available, (c) choosers know that a series of alternatives will become available in the future, This increases the likelihood that choosers will conjure a more ideal reference option, rather than a prototypical one (Griffin and Broniarczyk 2010). Thus, sequential choosers are likely to compare each option presented against an imagined better option. By contrast, simultaneous choosers will remain focused on the fixed set of currently presented options (Fischer et al. 1999; Hsee, Dube, and Zhang 2008; Hsee and Leclerc 1998; Shiv and Huber 2000; Tversky and Kahneman 1973, 1974; Tversky, Sattah, and Slovic 1988).

Extended Att-based strategy is related to the Simultaneous Choice, with many generated alternatives implicitly presented before consumers. And traditional Att-based procedure is alike Sequential Choice, consumers and construct one alternative at one time, then decide whether to switch to next one or to accept the current alternative separately. Unlike simultaneously presented choice sets where choosers remain focused on the known set of alternatives, sequentially presented choice sets lead choosers to imagine and hope for a better option(Mogilner et al.2012),Sequential choosers are less likely to stick with their selected option, when given the opportunity to switch (Hafner et al.2011).

Perceived choice confidence

In deciding whether or not to select a particular option, people commonly compare it with other alternatives that are currently available as well as with relevant alternatives that have been encountered in the past experience. Consumer's perceived value of choice derives from comparison to a large extent (Huber, et al., 1982; Simonson and Tversky, 1992).Almost all purchase decisions consumers make involve a comparison between two or more alternatives. Consumers may feel more confident if brand evaluations are based on a comparison with other alterative (the comparison is perceived as credible) than if a brand is evaluated in isolation (Dhar and Simonson, 1992).

Multiple alternatives presentation provides the opportunity for customers to compare and evaluate multiple alternatives before settling on a single alternative Kamis, et.al, 2008). Under Extended Att-based OCS, consumers are in the mode of simultaneous choice, while under Att-based OCS, consumers become Sequential chooser. Simultaneous choosers will remain focused on the fixed set of currently presented alternatives. In contrast, sequential choosers are likely to compare each option presented against an imagined better option. Likely, when consumers customize under Att-based OCS, every time they click, there will be a different alternative emerging. Since one product has more than one attribute, the composition of different attributes seems miracles, consumers can never think out what will it like until its combination out. Even consumers know their preference well, they cannot imagine all the possibilities of their preferences, especially in case of huge amount of the attributes and options in each attribute.

They may think that there is a possibility that an upcoming option will be better, but it is also possible that the current option is the best they are going to get (Monglner et al 2012).So Att-based OCS make consumer's customization choice full of uncertainty. However, with the support of Extended Att-based OCS, all possible alternatives are automatically presented implicitly through the random composition of selected options, consumers just need to evaluate and choose one from multiples. In prior research, when alternatives are presented simultaneously, and the presence of alternatives makes the decision fairly certain: choosers can feel confident that they are selecting the best option from among those available. Sequentially presented choice sets, in contrast, resign choosers to greater uncertainty.

Based upon the above discussion, we propose that

H1: with the help of extended attribute-based decision support function, customers tend to perceive more confidence about the final customized product.

Task complexity

Task complexity has been studied extensively in decision-making research, where it has been operationalzed as the number of alternatives and attributes (Olshavsky 1979; Payne 1976; Payne et al. 1988; Kamis et al. 2008). When making a choice, cognitive capabilities are used to retain each potential option and the individual's preferences. Therefore, when there are a large number of alternatives and/or attributes, the decision task is more complex than one with fewer alternatives and/or attributes. Hence, the number of alternatives and attributes directly affect the complexity of the choice making task (Swait and Adamowicz 2001) and, thus, in this study we use choice set size (i.e., the number of product versions available) as a proxy for task complexity. While customizing the product, the consumer needs to make (creative) choices (e.g., choosing the design and colors of the product's appearance), which requires mental effort (Ruth, et.al 2009).

The sequential presentation of options tends to lead choosers to imagine and hope for a better option, but this process is contingent on having cognitive resources available (Monglner et al. 2012). Because the bigger the choice set size, the more cognitive resource needed to deplete, the less cognitive resources available. By this way of swamping choosers' cognitive resources keeps them from hoping for better possibilities, thereby allowing them to enjoy more satisfaction from the decisions they make. So, the more complexity the task is, the smaller the difference of perceived choice confidence between Extended Att-based OCS and Att-based OCS is. That is, task complexity moderates the effect of the two customization strategy on perceived choice confidence.

Based on the above arguments, we propose that

H2: the higher the task complexity, the smaller the gap of perceived choice confidence between Extended Att-based OCS and Att-based OCS.

Maximizer vs. Satisficer

There are two categories of consumers: Maximizers and satisficer. Maximizers are defined as those people who refuse to settle on a choice until they find only the very best. Satisficers weigh their options but are content with settling for the option that is good enough (Schwartz et al. 2004).According to this definition, like the sequential presentation of options, maximizer would have the tendency to wonder about future possibilities hoping for a better option to become available (Schwartz et al. 2002).So, under Att-based OCS,which is in the mode of sequential choice, compared with Satisficer, Maximizer may be more inclined to expect for better one thus the lower perceived confidence of his final choice. Moreover, Maximizer/Satisficers may be moderators, strengthen or attenuate e the difference between under Extended Att-based OCS and Att-based OCS.

H3a：in Att-based OCS, Maximizer's perceived choice confidence is lower than Satisficer；in Extended Att-based OCS, Maximizer's perceived choice confidence is higher than Satisficer.

H3b: Maximizers strengthen the gap between Extended Att-based OCS and Att-based OCS, while Satisficer attenuate the gap between Extended Att-based strategy and Att-based OCS

4. Research Method

This study will employ laboratory experiment. We are planning to employ a 2 × 2 × 3 between subjects design. There are two versions of the website (Extended attribute-based OCS, Attribute-based OCS) of one product (mobile phone shell) that will be used for two types of consumers (Maximizer VS.Satisficer) using three choice set sizes (8, 54, and 150 product versions available).To ensure the involvement of participants, the context is set to customize a mobile phone shell for the participants themselves. And the final customized mobile phone shells will be given to them.

On the measure of dependent variable-perceived choice confidence, a questionnaire which is adapted from Dhar and Simonson, 1992 is designed, All items called for ratings on a Likert scale ranging from 1(not at all) to 7 (extremely).

5. Expected Conclusion and Implication

The Extended attribute-based decision leads to more consumer confidence about the final customized product. But the gap of decision confidence between the two OCS is not significant when the task complexity is too large. That is, the advantage of Extended Attribute-based OCS is relatively large only if the task complexity is moderate. In the aspect of different type of consumers, Maximizer's perceived choice confidence is lower than Satisficer in Att-based OCS,while Maximizer's perceived choice confidence is higher than Satisficer In Extended Att-based OCS.Maximizer matches with the Extended Att-based OCS better ,while Satisficer is more suitable for Att-based OCS.

The findings of the study will help us enrich the theories related to human-computer interactions and online consumer customization behaviors. Firstly, we improve the traditional and currently widespread applied Att-based OCS this study proves that Att-based OCS may not be good at all aspects, and makes up the prior studies on the negative effects of pure Att-based OCS ignored before. Secondly, through proposing an Extended Attribute-based OCS, this study provides operation implications and guidance for system developers on the design of online customization procedure or systems, such as the conditions and context to implement these two OCS.

However, there is only one dependent variable investigated in this paper, other effects on the subjective perception in customization process as well as the objective behavior such as the variety of final choice, the novelty or uniqueness of consumers' final design can be explored in the future.

References

1. Bharati, P., and Chaudhury, A. 2004. "An Empirical Investigation of Decision-Making Satisfaction in Web-Based Decision Support Systems," Decision Support Systems (37:2), pp. 187-197.

2. Dhar,R., and Simonson,I. 1992. “The Effect of the Focus of Comparison on Consumer Preferences,”Journal of Marketing Research (XXIX: November), pp. 430-40.
3. Franke,N., and Piller ,F. T.2003 "Key research issues in user interaction with user toolkits in a mass customization system.”International Journal of Technology Management (26:5-6), pp.578-599.
4. Fischer, G.W., Carmon.Z., Ariely.D., and Zauberman,G.1999.“Goal-Based Construction of Preferences: Task Goals and the Prominence Effect,” Management Science (45: August), pp.1057–75.
5. Hildebrand, C., Herrmann, A., and Häubl, G. (2013). “Product Customization via Starting Solutions. In Theory and Practice in Marketing Conference on Marketing Impact, pp.5.
6. Häubl, G., and Trifts, V.2000. “Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids,” Marketing Science (19:1), pp. 4-21.
7. Huber, J., Payne, J.W., and Christopher, P.1982."Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis, " Journal of Consumer Research (9: June), pp. 90-95.
8. Hsee, C.K., and Leclerc,Franc. 1998. “Will Products Look More Attractive When Presented Separately or Together?” Journal of Consumer Research (25: September), pp.175–86.
9. Hsee, C., Dube, J.P., and Yan, Z.2008. “The Prominence Effect in Shanghai Real Estate Prices,” Journal of Marketing Research (45: April), pp.1–22.
10. Hafner, R.J., Matthew P. W, and Simon J. H. 2011, “Spoilt for Choice: The Role of Counterfactual Thinking in the Excess Choice and Reversibility Paradoxes,” Journal of Experimental Social Psychology (47:4), pp.629–37.
11. Iyengar, S., and Lepper, M. R. 2000. “When Choice Is Demotivation: Can One Desire Too Much of a Good.
12. Griffin, J., and Broniarczyk,S. 2010.“The Slippery Slope: The Impact of Feature Alignability on Search and Satisfaction,”Journal of Marketing Research (42:2), pp. 323–34.
13. Kamis, A., Koufaris, M., and Stern, T. 2008.”Using an Attribute-Based Decision Support System for User-Customized Products Online: An Experimental Investigation." MIS Quarterly (32:1), pp.159-177.
14. Mogilner,C.,Shiv,B., and Iyengar,S.2013.“Eternal Quest for the Best: Sequential (vs. Simultaneous) Option Presentation Undermines Choice Commitment.” Journal of Consumer Research, DOI: 10.1086/668534.
15. Olshavsky, R. W. 1979. “Task Complexity and Contingent Processing in Decision Making: A Replication and Extension,” Organizational Behavior and Human Performance (24:3), pp.300-316.
16. Payne, J. 1976. “Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis,” Organizational Behavior and Human Performance (16:2), pp. 366-387.
17. Peter, F., and Pratkanis,A. 1993.“Decision Structuring with Phantom Alternatives,” Management Science (39), PP.1214–26.
18. Randall, T., C. Terwiesch, and Ulrich, K.T. 2007. "User Design of Customized Products." Marketing Science (26:2), pp. 268-80.
19. Ruth, M., Jan P., Schoormans, L.,and Hendrik,N. Schifferstein,J.2009. “Emotional bonding with personalized products,” Journal of Engineering Design (20:5), pp. 467–476.

20. Shiv, B., and Huber,J. 2000, “The Impact of Anticipating Satisfaction on Consumer Choice,” Journal of Consumer Research(27 :September),pp.202–16.
21. Swait, J., and Adamowicz, W. 2001. “The Influence of Task Complexity on Consumer Choice: A Latent Class Model of Decision Strategy Switching,” Journal of Consumer Research (28:1), pp. 135-149.
22. Schwartz.B., Ward.A., Monterosso,J., Lyubomirsky,S., Katherine.W., and Darrin. L.2002. “Maximizing versus Satisfying: Happiness Is a Matter of Choice,” Journal of Personality and Social Psychology (83:5), pp.1178–97.
23. Schwartz, B., Ward. A., 2004. “Doing better but feeling worse: The paradox of choice “Positive psychology in practice.
24. Simonson, I., and Tversky,A. 1992."Choice in Context: Tradeoff Contrast and Extremeness Aversion," Journal of Marketing Research (29: August), pp. 281-96.
25. Tversky, A., and Kahneman, D. 1973, “Availability: A Heuristic for Judging Frequency and Probability,” Cognitive Psychology, (5 :September), pp.207–32.
26. Tversky, A., Sattah, S., and Slovic.P. 1988. “Contingent Weighting in Judgment and Choice,” Psychological Review, (95: July), pp. 371–84.
27. Taylor Randall, Christian. T., and Karl T. U. 2005. “Principles for User Design of Customized Products,” California Management Review (47: 4), pp. 68-85.
28. Valenzuela,A., Dhar, R., and Zettelmeyer,F.2009."Contingent Response to Self-Customization Procedures: Implications for Decision Satisfaction and Choice." Journal of Marketing Research (46: 6), pp.754-63.

An Exploration of the Risk Sources of Social Networking Site Usage

Rui Chen and Sushil K. Sharma
Miller College of Business, Ball State University, Muncie, IN 47306, U.S.A

Abstract
Social networking sites have flourished within the recent past. Despite the values one may derive from site usage, perceived risks post a constant threat to the individuals who engage in these popular online platforms. In this paper, we explore two major sources of influence on the risks that are perceived by social networking site members. We discuss the underlying mechanisms through which these individual sources affect the overall perceived risks. Additionally, we identify a set of important antecedents to these risk sources by drawing upon the pertinent literature. The current paper helps in developing a foundation for further research on social networking site risks.
Keywords: Social networking sites, Risk perception, Sources of influence.

1. Background
Social networking sites (SNS) have proliferated since the last several years. Leading SNS websites are such as Facebook, Tumblr, Twitter, and Pinterest, which collectively account for about 90% of all social network visits (Laudon and Traver 2014). The latest statistics shows that Facebook has 1.23 billion monthly active users, 757 million daily active users, and 945 million monthly active mobile users (Facebook 2014). When registered at a SNS, members create their profile pages to share information about events, activities, interests, and histories. They may also receive constant, personalized news feeds that update the status and stories from friends, pages, groups, and events. In addition, members may search for people, places, photos, and other information that is shared by others. Chen and Sharma found that engagement at SNS may bring members a variety of benefits, including improved network management, enhanced social image, learning about the external world, and enjoyment (Chen and Sharma 2013). Collin et al. recognized improvements on media literacy, creativity, individual identity and self-expression, belonging and collective identity, and civic and political participation as other benefits that members are likely to derive from their daily use of SNSs (Collin et al. 2012).

The use of SNSs is, however, not without costs as members perceive risks in their engagement on these popular platforms. Various factors may result in the perceived risks in SNS usage (Lehrman 2010; Maier et al. 2012). From the perspective of information assurance, in this paper, we explore the pertinent antecedents that influence the perceived risks held by the individual SNS users. Our study focuses on the research questions of (1) "*What are the major sources of influence that cultivate a member's overall perceived risk of SNS usage?*", and (2) "*What are the major predictors to these aforementioned risk sources?*" The rest of the paper is organized as follows: the subsequent section reviews the literature in social networking sites. Next, we present the research model along with a set of research hypotheses. The last section concludes the paper.

2. Brief Literature Review
The risk literature has pointed out that perceived risks held by individual decision makers may originate from a variety of sources (Cox and Rich 1967; Cunningham 1967). As a consequence,

it is important to recognize the imperative, distinct sources that foster the overall risk perceptions within the context of SNS. Unless the major sources of influence are properly recognized, we won't become able to develop an appropriate nomological network that recognizes (1) the impacts, and their relative strengths, of individual risk sources on the development of an overall risk perception, and (2) the likely antecedents to the various, major risk sources.

Unfortunately, prior studies of SNS have largely missed these research issues. In general, the existing studies fall into two main streams of research when perceived risks are concerned. The first stream of research has focused on selective sources of perceived risks and has examined the potential connections between specific risk sources and SNS members' cognition and affection. Posey et al. found that privacy risk beliefs significantly discourage members in self-disclosure (Posey et al. 2010). Shin tested the effects of security concerns and privacy concerns on SNS member trust and attitude and found empirical evidence (Shin 2010). Unfortunately, Shin's measurements of security concern assess privacy (unauthorized access and misuse) but not security risks. Lorenzo-Romero et al. examined the impacts of risk perception on perceived usefulness and on intention to use SNS and found marginal support on the latter (Lorenzo-Romero et al. 2013). The second stream of research has examined the overall perceived risks of SNS along with their antecedents and/or consequences. However, it has failed to recognize the important risk sources. Loh tested the likely impacts of perceived risks on trust and intention to purchase on SNSs and found support on the former (Loh 2011). What have been missing from the literature are (1) the studies that associate individual risk sources with the development of an overall perceived risks to understand their relative effects and (2) the studies that pinpoint specific predictors that relate to each, unique risk sources. As a result, there exists a significant gap in the risk research of SNS. Future studies are warranted to refine the nomological network that better explain perceived risks, their sources, and pertinent antecedents.

3. Research Model and Hypotheses

In this section, we develop a conceptual research model to underscore two important sources that shape SNS members' overall perceived risks. Drawing upon the literature, we identify the two sources of influence as cyber-attack concerns and privacy concerns. We further posit that cyber-attack concerns are affected by perceived Internet risks as well as perceived member trust. Moreover, we contend that website trust may affect privacy concerns. Finally, we project that perceived risks of SNS usage may affect members' actual site use behavior.

Members of SNSs constantly weight the benefits and costs in forming their decisions about site use behaviors. When they perceive a high level of risks, they will become reluctant to participate in the social networking activities due to an increased cost. At a low level of perceived risks, on the contrary, individuals are inclined towards site use as the cost is acceptable. We, therefore, expect that perceived risks of SNS may deter members from their use of social networking sites.
Proposition 1: perceived risks in using SNSs negatively associate to site use.

SNSs have grown to become a popular online platform where a large crowd gathers for interactions. As a result, SNSs have become a preferred platform where cyber criminals launch attacks. Attacks on SNS platforms are frequently documented. Facebook revealed that, every day, it receives about 600,000 login attempts from impostors who are interested in private member information (Ashford 2011). In November 2013, Trustwave researchers discovered compromised credentials for 300,000 Facebook accounts, over 20,000 Twitter accounts, and 8,000 LinkedIn accounts, which were stolen due to the use of keylogger malware created by hackers (Pagliery

2013). The anti-virus firm Sophos discovered that 40% of SNS users have encountered malicious attacks (McMillan 2011). When a member falls for cyber-attacks, he or she will become victimized and suffer from a series of potential losses and damages on finance, reputation, emotion, and computer systems. In this paper, we define cyber-attack concerns as one's uneasiness about using a given social networking site considering the likely exposure to cybercrimes. When the cyber-attack concerns increase, members will expect a greater odd of adverse consequences, which results in a greater amount of perceived risks. We, therefore, expect that cyber-attack concerns may increase the overall perceived risks of SNS.

Proposition 2: cyber-attack concerns positively associate to perceived risks.

Members' cyber-attack concerns may be susceptible to own assessments of the general computing environment. In this paper, we define perceived Internet risk as one's uneasiness about using the Internet (Wang et al. 2009). As the largest network in the world, the Internet connects over 2 billion users across the globe (InternetWorldStats 2012). Due to the extrasensory complexity and difficulty in its governance, Internet renders as a "*Wild West*" as malicious attackers prey on innocent internet users on a 24/7 basis. According to the 2013 Norton Report by Symantec, the global loss of consumer cybercrime is expected to be $113 billion and the average cost per victim is $298 (Merritt and Haley 2013). Perceived Internet risk is, therefore, indicative of an individual's concern about the Internet, as a result of his or her past unsuccessful use experience (e.g., victimization by prior attacks). When the Internet risk is perceived to be high, a member may exaggerate the uncertainties in SNS use, exhibit unjustified confidence in their judgment, and consequently overrate possible cyber-attacks in using SNSs. To its contrary, a member with low Internet risk perception may underestimate cyber-attacks in SNS usage. Therefore we expect:

Proposition 3: perceived Internet risk positively associates to cyber-attack concerns.

An average SNS member interacts with a large number of others. Among these friends, a small portion represents close ties while the rest represent loose ties such as brand new acquaintance. Due to information asymmetry, it remains hard for one to verify the other contacts. Anyone can create a bogus account on SNSs with false information. Studies have found that stalkers, spammers, and scammers are frequently found on SNSs and they attack unsuspected users for fraudulent purposes such as distribution of malware and identity theft (Grayer 2011). Trust concerns ability, benevolence, and integrity of a trustee (Colquitt et al. 2007). In this paper, we focus on the integrity dimension as it directly relates to our research focus. Integrity is defined as the extent to which a trustee is believed to follow sound moral and ethical principles (Mayer et al. 1995). The trust literature has suggested that perceived trust facilitates the reduction of perceived risks (Jarvenpaa et al. 2000; Kim et al. 2008). When an individual evaluates the other members on the same SNS high on trust, he or she expects a reduced chance of cyber-attacks from those around him or her on a SNS. As a result, we expect:

Proposition 4: Perceived member trust negatively associates to cyber-attack concerns.

At its essence, a SNS offers a social environment where members may engage into interpersonal interactions. Interactions may take the forms such as profile creation and browsing, sharing of photos and videos, exchange of comments, participation in groups and events, and ramifications such as likes and retweets. Over the time, members reveal a large volume of private information into the SNSs. Example private information includes, but not limited to, name, birth date, contacts such as phone numbers and mailing addresses, interests, hobbies, marital status, and political opinions, etc. The sheer amount of privacy that an individual shares on a SNS is much

greater than what is disclosed on traditional online platforms such as online retailing stores. In this paper, we define privacy concerns as concerns about one's inability to control the personal information submitted to a SNS (Dinev and Hart 2006). When members perceive a high level of privacy concerns, they expect more chances of personal data being exploited. That is, they perceive a great extent of perceived risks of SNS usage. As a result, we expect:
Proposition 5: privacy concerns negatively associate to perceived risks.

The private data of members is at the possession of SNS providers. While one may safeguard privacy from other members by configuring the privacy settings, one's private data is collected by the service providers who may analyze the data for business intelligence and may share the data with affiliated business entities for secondary uses. In this paper, we define website trust as the perceived integrity of the service providers of a SNS. Past studies have questioned the trust of the leading SNS vendors in protecting user privacy. In 2010, Facebook and MySpace were found sending data, which could be used to find members' name and other personal details, to advertising companies without receiving consents from their members (Steel and Vascellaro 2010). Recently, Facebook were sued over allegations that it methodically intercepts private user messages and shares the data with advertisers and marketers for profits (Gullo 2014). When members rate a SNS high on trust, they expect a small chance of website mishandling and/or misusing of their shared privacy. That is, members will expect a low level of privacy concerns. Therefore, we project:
Proposition 6: perceived website trust negatively associates to privacy concerns.

4. Conclusion

SNSs have mushroomed over the Internet in the recent years and the use of SNSs has grown to become the No. 1 online activity for the U.S. (Richter 2013). While users reap benefits from the use of SNSs, they are also constantly challenged by the risks that may be involved. In this paper, we theorize two major sources of the perceived risks in SNS usage as cyber-attack concerns and privacy concerns. We elaborate on their likely associations with the overall perceived risks of SNS and we further identify their potential antecedents. The current study assists future research works in developing a conceptual framework that may be extended and validated. Future research involves the collection of empirical survey data from SNS users as well as the model validation through statistical approaches such as structural equation modeling.

References

Ashford, W. 2011. "Facebook Admits to 600,000 Cyber Attacks a Day," in: *ComputerWeekly.com*.

Chen, R., and Sharma, S. K. 2013. "Understanding Member Use of Social Networking Sites: A Value Analysis," *Communications of the Association for Information Systems* (33:7), pp. 97-114.

Collin, P., Rahilly, K., Richardson, I., and Third, A. 2012. "The Benefits of Social Networking Services," Cooperative Research Centre for Young People, Technology and Wellbeing.

Colquitt, J. A., Scott, B. A., and LePine, J. A. 2007. "Trust, Trustworthiness, and Trust Propensity: A Meta-Analytic Test of Their Unique Relationships with Risk Taking and Job Performance," *Journal of Applied Psychology* (92:4), pp. 909-927.

Cox, D. F., and Rich, S. U. 1967. "Perceived Risk and Consumer Decision Making: The Case of Telephone Shopping," in *Risk Taking and Information Handling in Consumer Behavior*, D.F. Cox (ed.). Boston, MA: Harvard University Press, pp. 487-506.

Cunningham, S. M. 1967. "The Major Dimensions of Perceived Risk," in *Risk Taking and Information Handling in Consumer Behavior,* D.F. Cox (ed.). Boston, MA: Harvard University Press, pp. 82-108.

Dinev, T., and Hart, P. 2006. "An Extended Privacy Calculus Model for E-Commerce Transactions," *Information Systems Research* (17:1), pp. 61-80.

Facebook. 2014. "Company Info," in: *Facebook.com*.

Grayer, M. 2011. "I Accepted a Fake Facebook Friend Request, Should I Be Afraid?," in: *Naked Security*. Burlington, MA: Sophos Ltd.

Gullo, K. 2014. "Facebook Sued over Alleged Scanning of Private Messages," in: *Bloomberg.com*. New York, NY: Bloomberg I.P.

InternetWorldStats. 2012. "Internet Users in the World: Distribution by World Regions - 2012 Q2." Miniwatts Marketing Group.

Jarvenpaa, S. L., Tractinsky, N., and Vitale, M. 2000. "Consumer Trust in an Internet Store," *Information Technology Management* (1:12), pp. 45-71.

Kim, D. J., Ferrin, D. L., and Rao, H. R. 2008. "A Trust-Based Consumer Decision-Making Model in Electronic Commerce: The Role of Trust, Perceived Risk, and Their Antecedents," *Decision Support Systems* (44:2), pp. 544-564.

Laudon, K., and Traver, C. 2014. *E-Commerce 2014: Business, Technology, and Society*, (10 ed.). Prentice Hall.

Lehrman, Y. 2010. "The Weakest Link: The Risks Associated with Social Networking Websites," *Journal of Strategic Security* (3:2), pp. 63-72.

Loh, A. 2011. "A Study on Influence of Trust, Social Identity, Perceived Risk and Ewom on Consumer Decision-Making Process in the Context of Social Network Sites." Karlskrona, Sweden: Blekinge College of Technology.

Lorenzo-Romero, C., Chiappa, G. D., and Alarcon-del-Amo, M. 2013. "The Users Adoption and Usage of Social Network Sites: An Empirical Investigation in the Context of Italy," in: *International Marketing Trends Conference*. Paris, France.

Maier, C., Laumer, S., Eckhardt, A., and Weitzel, T. 2012. "When Social Networking Turns to Social Overload: Explaining the Stress, Emotional Exhuastion, and Quitting Behavior from Social Network Sites' Users," *Europen Conference on Information Systems*, Spain.

Mayer, R. C., Davis, J. H., and Schoorman, F. D. 1995. "An Integrative Model of Organizational Trust " *Academy of Management Review* (20:3), pp. 709-734.

McMillan, G. 2011. "40% of Social Network Users Attacked by Malware," in: *Time, Inc.* New York, NY.

Merritt, M., and Haley, K. 2013. "2013 Norton Report," Symantec Corporation, Mountain View, CA.

Pagliery, J. 2013. "2 Million Facebook, Gmail and Twitter Passwords Stolen in Massive Hack," in: *CNN Money*.

Posey, C., Lowry, P. B., Roberts, T. L., and Ellis, T. S. 2010. "Proposing the Online Community Self-Disclosure Model," *European Journal of Information Systems* (19), pp. 181-195.

Richter, F. 2013. "Social Networking Is the No. 1 Online Activity in the U.S.." New York, NY: Statista, Inc.

Shin, D. H. 2010. "The Effects of Trust, Security and Privacy in Social Networking: A Security Based Approach to Understand the Pattern of Adoption," *Interacting with Computers* (22:5), pp. 428-438.

Steel, E., and Vascellaro, J. E. 2010. "Facebook, Myspace Confront Privacy Loophole," in: *The Wall Street Journal*. New York, NY: Dow Jones & Company, Inc.

Wang, J., Chen, R., Herath, T., and Rao, H. R. 2009. "Visual E-Mail Authentication and Identification Services: An Investigation of the Effects on E-Mail Use," *Decision Support Systems* (48), pp. 92-102.

Rational and Irrational Herding in P2P lending

Seongmin Jeon
Gachon University
smjeon@gachon.ac.kr

Byungjoon Yoo
Seoul National University
byoo@snu.ac.kr

Abstract

Peer-to-peer (P2P) lending has recently become a popular IS research field due to information asymmetry issues that exist and result in investors' herding behavior influenced by their peers. Rational herding occurs when lenders make decisions based on observational learning while irrational herding happens when lenders mimic the choices of others by referring to their decisions and following well-funded listings. We use a unique panel data set that tracks the funding dynamics of borrower listings on Popfunding.com, the largest P2P lending site in Korea. We find evidence of herding among lenders, which echoes the results of previous research on P2P lending. However, our results show that the existence of rational herding is ambiguous. The moderating effects of the total amount on herding are identified significant while the effects of the interest rate are not consistent.

Keywords: rational herding, observational learning, peer-to-peer lending, online marketplace, information asymmetry, econometric analyses, wisdom of crowds

1. Introduction

Peer-to-peer (P2P) lending provides a new platform where borrowers are able to raise funding from multiple individual lenders online. P2P lending has information asymmetry issues in that investors have a limited amount of information with which to distinguish between good and bad borrowers. Peer influence is expected to be a significant driver of lending and consequently, investors are likely to herd. The question remains as to whether this kind of herding will be rational or irrational. We use a unique panel data set that tracks the funding dynamics of borrower listings on Popfunding.com, the largest online P2P lending site in Korea. We find evidence of herding among lenders, however, the existence of rational herding is ambiguous, contrary to the results of Zhang and Liu (2012).

P2P lending represents an open marketplace for loans provided not by a bank, but instead by individuals online taking advantage of P2P architecture. In P2P lending, financial transactions are facilitated directly between individuals ("peers"). Potential borrowers create and post listings with an overview of their need for a loan; while potential lenders place bids on listings they would be interested in funding. A borrower is then provided a loan only in the case that his or her listing attracts enough bids to exceed a predefined amount or to fulfill a loan request by a number of lenders. One of the biggest P2P lending markets, Prosper.com, has come to have 200,000 listings seeking to raise $1 billion in funding despite that this market did not exist as recently as 2005. According to a market study by the Open Data Institute, the P2P lending market will be worth £1billion by 2016 if it continues to grow at its current rate (Financial Times, 2013).

P2P lending is different from traditional financial markets. A borrower will raise fund from multiple lenders who share some portion of the loan. P2P lending markets have a number of advantages over traditional financial institutions. The most widely known benefit of P2P lending is that borrowers are able to make loans at a lower interest rate with no security or collateral, while investors are able to get a higher return on their investments with little or no involvement by financial intermediaries. The value that P2P lending provides to borrowers is that borrowers who were not able to obtain loans now have a chance to raise funding thanks to the implementation of Internet and social network technology-driven microfinance platforms which rely upon social collateral. Peer influence is expected to be a significant driver of P2P lending. There are several previous studies on herding behavior in P2P lending (Puro et al. 2011; Shen et al. 2010; Zhang and Liu, 2012; Herzenstein et al., 2011). Zhang and Liu (2012) document evidence of the behavioral mechanism underlying herding among lenders in P2P lending. They discriminate rational herding and irrational herding using the data set from Proper.com, one of the biggest P2P lending markets in the U.S. Irrational herding happens when lenders mimic others' choices, referring to others' decisions and following well-funded listings (Croson and Shang, 2008; Simonsohn and Ariely, 2008), whereas rational herding occurs when lenders make decisions based on observational learning (Banerjee 1992; Bikhchandani et al. 1992).

This study identifies rational and irrational herding empirically mainly through two steps presented by Zhang and Liu (2012). First, we control herding factors for listings' unobserved heterogeneity and externalities among lenders. That is, the unobserved characteristics of listings may influence the bid participation of investors, and investors are likely to prefer a well-funded listing because the listing has high chances to get financed. Second, rational and irrational herding will be differentiated through the identification of the moderating effects of observational learning.

2. Literature Review

Investors in P2P lending are most likely to encounter three sources of risk, namely that regarding borrowers, the lending products and the online environment (Yum et al. 2012). P2P lending platforms attract broader customer segments of borrowers beyond those of traditional financial institutions and deal with anonymous unknown borrowers. More often than not, borrowers do not have a credit history and have had trouble obtaining transactions from traditional financial institutions. This Internet's anonymous environment keeps trust from being established among participants and fraud issues are rising as not a small number of users tend to have fictitious user names (Greiner and Wang, 2007). Consequently, the fundamental issue with P2P lending is the default which results when strangers do not pay back the money at all. Moreover, an individual investor is vulnerable to the risks brought on by loan default, while financial institutions manage risk through portfolio management.

Information asymmetry models developed in the area of information economics assume that one party to a transaction has, at the bare minimum, related information while the other party does not (Akerlof, 1970). Investors in P2P lending appear to be aware of risks involving the borrowers' opportunistic behavior through the social networks on P2P lending. It is reported that only a small portion of the requests listed on Prosper are successfully funded (Freedman and Jin, 2011). Berger and Gleisner (2009) find that investors are able to screen potential borrowers and monitor the repayment through group leaders. Group leaders are expected to reduce information asymmetry, particularly when there are riskier borrowers.

3. Hypotheses

Herding may help attract bids and subsequently improve a loan's chances of getting financed. Like herding in the traditional financial market, herding is found to be a factor in non-diagnostic decisions in the context of online auctions (Simonsohn and Ariely, 2008). A number of investors have the incentive to follow other investors' decisions as they learn from their peers by imitating their behaviors in situations of imperfect information. Thus, it can be inferred that lenders show the behavior of herding by following the other investors behavior. In this context, the following statement is presented:

H1: Investors' choice of participation in bidding is significantly affected by the participation of other investors.

If investors exhibit herding behavior, they just mimic others' behavior or undertake observational learning to review the requests' attributes. Like what Zhang and Liu (2012) claim, investors in P2P lending are expected to show the behavior of rational herding. They are likely to identify which requests they would bid on by looking at how financed the request is, that is, how many peers have already bid. Then they will review the details of the identified requests to make their own decisions to bid. These processes can be understood as a part of social learning. To clarify the effects of social learning and rational herding, we present the hypothesis on investors' decisions relating to our investigation:

H2: Investors in P2P lending show the behavior of rational herding by undertaking their own reviews and making decisions after having the behavior of herding.

4. Data and Model

4.1. Data

Starting to provide its services to the public in May 2007, Popfunding.com has grown rapidly into one of the largest online microloan markets in Korea. Early in 2010, Popfunding.com had mediated more than $1 million in funding. The Korea Institute of Finance estimates the size of Korea's private money lending market at 18 trillion Korean won (KRW), approximately equivalent to $19.9 billion. The FSS reported that more than 1 in 3 Koreans have borrowed money at least once from consumer loan providers (The Korean Herald, 2007).

We track a sample of listings posted on Popfunding.com from July 2009 to December 2009. We have listings with duration of 15 days, the typical duration on Popfunding.com. Our sample contains 903 listings. Each listing has a set of its attributes and contains panel information such as bids and funding status. The minimum amount for a bid is 1,000 KRW, equivalent to $1. In this sample, listings request between 500,000 KRW or $500 and 5,000,000 KRW or $50,000, with a mean of $2105 or about $2,000. Borrowers propose interest rates between 3% and 30%, with an average of 29.72%.

4.2. Econometric Model

We implement the econometric model for panel data in order to analyze if herding influences lending decisions. In addition, if so, we come to have the model to identify if such herding is irrational or rational. The intermediary publicly provides the funding status such as the percent funded in real time, aiming to help the decisions of subsequent investors. This is likely to lead to peer influence. We focus on the bid amount and the number of bids as the measure for how a request attracts an investor. We assume that investors optimally allocate their investment money among many choices of listings on Popfunding.com. At the same time, the number of bids represents the number of people who recognize the listing as being deserving to be financed.

A simple test of herding may find serial correlations in lending. Both the amount funded on the day and the number of bids attracted on the day, y_{it}, are likely to be correlated positively with the lagged cumulative amount and the cumulated number of bids, y_{it-1}.

$$y_t = \beta_0 + y_{t-1}\alpha + X_t \beta_1 + Z_i\beta_2 + u_i + v_t \qquad (1)$$

where X_{it} represents time-varying request attributes and Z_i describes time-invariant request attributes. The time-varying request attributes X_{it} include *Lag Percent Needed*, the percentage of the amount requested by listing i, *Lag Total Bids*, the cumulative number of bids listing i has attracted by the day t-1. To control the influences on certain days of the week or certain days during a request's duration, we include Day-of-Week Fixed Effects and t th Day-of-Period Fixed Effects in X_{it}. Time-invariant request attributes Z_i include Amount Requested and Borrower Rate. u_i represents a fixed effect control term for unobserved request characteristic. The assumption here is that unobservable characteristics of requests are time-invariant. We may accept the assumption in that the requests' characteristics would not change during the funding period as the conditions for the requests are likely to remain unchanged.

It is crucial to discriminate the irrational herding from rational herding. Irrational herding as well could have serial correlation patterns. That is, investors simply mimic other investors by participating in the bids according to the popularity of listings such as leaderboard. Zhang and Liu (2012) have the independent variable of *Lag Percent Needed* to explain how much investors decide to bid based on the social comparison theory, referring the work of Croson and Shang (2008). We assume that investors are irrationally herding as long as they simply mimic other investors' decision. Meanwhile, if investors are rational, their decisions are likely to be moderated through reviewing the requests' attributes. To discriminate the rational herding from irrational herding, we supplement the equation (1) with the interaction terms between the lagged total amount and observable listing attributes.

$$y_t = \beta_0 + y_{t-1}\alpha + X_t \beta_1 + Z_i\beta_2 + Y_{t-1}Z_i\beta_3 + u_i + v_t \qquad (2)$$

If rational herding exists, the coefficient of β_3 should have the opposite signs of request attributes' influence on funding amount. The moderating effect of a request's number of bids on its cumulative funding amount will vary by whether investors are rational and irrational. Rational investors are likely to review the number of investors who already participated in bidding, not alone the total percent funded. Therefore, we add the variable of interaction term between lag total amount and lag total bids.

5. Results and Discussions

Column (1) in Table 4 reports the estimation results for Equation (1). Having the fixed effects controlled for the heterogeneous characteristics of each request, the result of R-squared increases from 44% to 71%. The effect of the total amount that a request attracted on the previous day is positive and statistically significant. The total number of bids that a request attracts on the previous day has a significant and positive effect as well. This can be interpreted as they have herding—a request's past bid attraction is effective in attracting more subsequent funding. This finding confirms the research of Herzenstein et al. (2011) and Zhang and Liu (2012) on Prosper. Additionally, we track the effect of the externality which is negative counter-intuitively. The interaction term of *Lag Total Amount X Lag Percent Needed* is significant and negative. Investors are not likely to take parts in having the requests financed by bidding on the requests that need more amount to get financed.

Column (2) of Table 1 presents the estimation results of Equation (2). The statistical significance of the interaction term between and Lag Total Amount and time-invariant variables are somehow complex. Displaying greater total amount does not have significant effects on funding. After controlling the effects of revealed information to herding investors, the total attracted amount does not play effective role to raise the subsequent funding.

Borrower rate has weakly positive and significant effect on funding. The interaction term with *Lag Total Amount* is significant as well. We will be able to say that the borrowing conditions the borrowers present become important issues after a review by herding investors.

The interaction term between the *Total Amount* and *Lag Percent Needed* is negative and significant. Moreover, the coefficient decreases compared with the result of (2). In the study by Zhang and Liu (2012) takes this variable as a proof of rational herding if it is positive. When we look at the coefficient for *Lag Percent Needed* is positive and significant. This is also interesting because a higher percentage needed will normally discourage funding. This is generally contrary to the result of the strategic behavior of investors who would participate in the bids in the listings that are likely to get funded easily. However, when we interact with the total funded amount with the percent needed, we find that investors behave strategically to fund the listings that need a small amount to get financed. This echoes the findings of Herzenstein et al. (2011).

Table 1. Results for Herding and Rational Herding

	(1) Herding	(2) Rational Herding
L.Total Amount	0.442*** (0.041)	-1.481 (0.940)
L.Percent Needed	345,087*** (54,323)	548,145*** (79,535)
L.Total Bids	5,479*** (884.2)	16,393*** (841.5)
Lag Total Amount X Percent Needed	-1.573*** (0.034)	-2.358*** (0.039)
Lag Total Amount X Amount Requested		2.58e-07*** (1.68e-08)
Lag Total Amount X Borrower Rate		0.0689** (0.031)
Lag Total Amount X Lag Bids		-0.013*** (0.000)
Listing fixed effects	Yes	Yes
Observations	3,160	3,160
R-squared	0.71	0.79
Number of listings	662	662

The dependent variable is the amount of funding that a request attracts on a day. Panel regression with standard errors clustered by listing and reported in parentheses under parameter estimates. *p <0.1;**p <0005;***p<0.001.

Moreover, to investigate the moderating effects of time invariable variables, we find that Amount Requested statistically significantly moderates the herding effects while the effects of Interest Rate are ambiguous.

We confirm evidence of herding analyzing a unique panel data set from Popfunding.com. Imitation and mimicry may work even in the traditional financial markets. Herding is found in fashion and fads, not only in a simple decision but also in investment decisions. Investors are frequently influenced by the decisions of others. Herding happens through a coordination mechanism. We investigate the mechanism behind the presence of the herding phenomenon, following the approaches in the previous research that concludes the existence of rational herding. However, our results show that the existence of rational herding is ambiguous, somewhat different from the results of Zhang and Liu (2012). The moderating effects of the total amount on herding are identified significant while the effects of the interest rate are not consistent. Most presumably, the investors in Popfunding.com have limited information on the borrowers in that the credit grades are mostly low and thus it is hard to tell who will be good borrowers. Information such as whether the borrower is a homeowner, has an endorsement and his or her debt-to-income ratio, are not provided to the investors as well. In addition, the investors join the bid with minimum bid of $1 and are likely to think of their participation as charity rather than a form of investment.

References

1. Akerlof, G. The Market for Lemons: Qualitative Uncertainty and the Market Mechanism. Quarterly Journal of Economics. 84(3), 1970, pp. 488-500.
2. Banerjee, A. A Simple Model of Herd Behavior. Quarterly Journal of Economics. 107(3), 1992.
3. Berger, S. C. and Gleisner, F. Emergence of financial intermediaries on electronic markets: the case of online P2P lending. Journal of Verband der Hochschullehrer für Betriebswirtschaft, 2, 1, 2009, 39–65.
4. Bikhchandani, S., Hirshleifer, D. and Welch, I. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. Journal of Political Economy. 100(5), 1992.
5. Croson, R. and Shang, J. The impact of downward social information on contribution ecisions. Experiment. Econom. 11(3) 2008, 221–233.
6. Financial Times, Digital finance lending set to hit £1bn., 15 Jan. 2013.
7. Freedman, S. and Jin, Z. Learning by Doing with Asymmetric Information: Evidence from Prosper.com. NBER Working Paper. 16855. 2011.
8. Greiner, M. and Wang, H. Building consumer-to-consumer trust in e-finance marketplaces. In Proceedings of the 13th Americas Conference on Information Systems, 2007.
9. Herzenstein, Michal, Utpal M. Dholakia, and Rick L. Andrews. "Strategic herding behavior in peer-to-peer loan auctions." Journal of Interactive Marketing 25.1 (2011): 27-36.
10. Puro, L., Teich, J.E., Wallenius, H. and Wallenius, J. Bidding strategies for real-life small loan auctions. Decision Support Systems. 51, 2011, pp. 31-41.
11. Simonsohn, U. and Ariely, D. When Rational Sellers Face Non-Rational Buyers: Evidence from Herding on eBay. Management Science. 54(9), 2008, pp. 1624–1637.
12. The Korean Herald, Subprime lenders under tight scrutiny, 2007
13. Yum, H., Lee, B. and Chae, M. From the wisdom of crowds to my own judgment in microfinance through online peer-to-peer lending platforms. Electronic Commerce Research and Applications 11, 2012, 469–483
14. Zhang and Liu, Rational Herding in Microloan Markets, Management Science 58(5), 2012.

Motivating User Contributions to Online Communities: A Structural Modeling Approach

Wei Chen
Rady School of Management,
University of California,
San Diego
Wei.chen@rady.ucsd.edu

Xiahua (Anny) Wei
School of Business,
University of Washington,
Bothell
xwei@u.washington.edu

Kevin Zhu
Rady School of Management,
University of California,
San Diego
kxzhu@ucsd.edu

Abstract

User contribution determines the rise and fall of online communities. This paper models the dynamics of user contribution to an online community using a hidden Markov model in a public goods framework. Our model characterizes the fluctuations of user contributions by two latent activity states (active vs. inactive), and their effects on participation behavior. We estimate the model with a Bayesian approach using a user-level panel dataset from an online community. We find that social recognition and reciprocity are effective motivating mechanisms that drive users to switch to active state from inactive state. Further, a user respond differently to stimulates such as demand of knowledge and group size is in different states, which highlights the importance of modeling user states to understand contribution behavior. Our model is useful for platform designers to motivate online user contributions, and to build sustainable online communities.

Keywords: online community, user contribution, behavior dynamics, public goods, structural model

1. Introduction

In recent years, online communities have emerged as an important way to share knowledge, to forge collaboration and to foster innovation. This phenomenon has become significant for the knowledge economy in the Internet age where users are also producers. Despite the prosperity of these online communities, many others failed to achieve critical mass (e.g., Google Answers) due to the lack of sustainable participation of users (Ransbotham and Kane 2011). Consequently, it becomes a crucial issue to understand the dynamics of user participation, so as to design sustainable online communities and promote sustainable innovation.

Since the contents are often freely available to all users, contributions to online communities are public goods in nature. As a result, they may suffer from the traditional dilemma of public goods (Dietz et al. 2003), where rational individual decision-making may lead to free-riding and under-provision of public goods. The literature shows that voluntary cooperation is inherently fragile, even if most people are not free riders but conditional cooperators (Fischbacher and Gächter 2010). Then an important question follows: how to design effective mechanisms to encourage user contributions in online communities where knowledge is shared very differently and at large scale?

Users contribute to online communities for various reasons (Lerner and Tirole 2005). Several mechanisms (e.g. social comparison) have been proposed to encourage user participation (Chen et al. 2010). However, current models are inadequate to capture the fluctuations of contributions by individual users. A user may contribute actively for some time, and remains less active for other times. Such dynamic fluctuations are common in online communities. Inactive users and free riders can dry up the online communities (Bayus 2013). We need to capture this important feature in order to understand user contributions and further to design sustainable online communities. To narrow the gap in the literature, we seek to study two research questions: (1) How should we *model* the dynamics of user contributions in online communities? (2) What would drive users to be more active participants?

To understand these questions, we propose a structural approach that integrates hidden Markov model (HMM) with the public goods framework. This structural approach characterizes user contributions under different activity states (*active* and *inactive*) and the transition between states. We then examine the effect of specific motivating mechanisms on user contributions. We empirically evaluate our structural model by Bayesian estimation with data collected from a representative online community StackExchange. We find that (1) motivating mechanisms work differently in active and inactive states; (2) community interactions, especially social recognitions and reciprocity, are important channels to promote users to active state; and (3) community size and the demand of knowledge on the site motivates user contribution. Compared with the traditional approach without incorporating contribution states, our approach better characterizes the fluctuation of user behaviors, and thus has greater predictive power. Our findings provide managerial implications for designing mechanisms to motivate user contributions and sustain online communities.

2. Model Development: Structural Modelling of User Behavior

In this section, we propose an individual level model of user contribution behavior, where each user interacts repeatedly with the community and decides the contribution level. To account for the fluctuation of user contributions, we use a hidden Markov model (HMM) to capture the dynamics of a user's valuation of one's own contributions.

2.1 A Static Public Goods Model of User Contributions

We model user contributions to an online community in a public goods framework. Each self-interested user chooses the extent of contribution to the community. A user's utility consists three parts: (1) her valuation of the accumulated knowledge, (2) her valuation of her own contribution, and (3) a cost of contribution. By incorporating the above three parts, we can specify the utility of user *i* at time *t* and derive the equilibrium contribution:

$$Y_{it}^{*} = \alpha_i + X_{it}\beta_{s_{it}} + \varepsilon_{it}, \quad \varepsilon_{it} \mid X_{it}, \alpha_i, s_{it} \sim N(0, \sigma^2) \tag{1}$$

where α_i stands for the mean valuation of user *i* to her own contribution, X_{it} is a vector of user *i*'s interactions in the community in period *t*. We assume that the error term ε_{it} follows a normal distribution with mean zero and variance σ^2. Our goal is to estimate the coefficient vector $\beta_{s_{it}}$, which captures the influence of community and individual characteristics on the user's contribution. Note that this vector depends on user *i*'s contribution state s_{it}, which is a feature of our model, to be explained in the next section.

2.2 Contribution States in HMM

To model the fluctuation of user contribution, we consider it as two stochastic processes: a process of observed contributions, and an underlying hidden process of the user's states. The hidden state captures the time-dependent feature of a user's valuation of her own contribution. The observed contribution could be seen as a noisy signal of the hidden state process. The pair of processes together form a hidden Markov chain.

In our proposed HMM, a user could have two different hidden states: *active* or *inactive*. We denote $s_{it} \in \{0,1\}$ the state of user *i* in period *t* (0 for inactive and 1 for inactive). A user does not always stay in one state, but could transit between the two states. We then model the vector-valued stochastic process (Y_{it}, s_{it}) as a hidden Markov chain. The transition kernel for the hidden Markov chain is then defined as $P((Y_{it}, s_{it}) | (Y_{i,t-1}, s_{i,t-1})) = P(Y_{it} | s_{it}) p(s_{i,t-1}, s_{it})$, where $p(s_{i,t-1}, s_{it})$ is the transition probability from state $s_{i,t-1}$ to state s_{it}, and $P(Y_{it} | s_{it})$ is the emission probability which describe the state-dependent contributions. We explain these two probabilities in the next two sections respectively.

2.3 Transition Probabilities in HMM

As a Markov chain, the hidden contribution state could transit between different states. Let $p(j,1)$ be the transition probability from state j ($j \in \{0,1\}$) to the active state ($s_{it} = 1$), and $\sum_j p(j,0) = 1$. We assume that the transition probability $p(j,1)$ is affected by the user's interactions with the community, which may create certain social or personal norms (Bénabou and Tirole 2006) for the user. The user may then evaluate her own contributions differently based on the norms. Mathematically, we model the nonhomogeneous transition probabilities with a probit model (Wooldridge 2010). We assume that the states are determined by a latent propensity for transition L_{it}:

$$L_{it} = W_{i,t-1}\xi_{s_{i,t-1}} + u_{it}, \quad u_{it} | W_{i,t-1}, s_{i,t-1} \sim N(0, \sigma_u^2) \tag{2}$$

such that $s_{it} = 1$ if $L_{it} > 0$ and $s_{it} = 0$ if $L_{it} \leq 0$. In equation (5), $W_{i,t-1}$ is a vector of lagged variables related to the user's interactions with the community, $\xi_{s_{i,t-1}}$ is a vector of coefficients that captures the effects of these interactions, and u_{it} is a norm error term from the probit model. Note that we allow $\xi_{s_{i,t-1}}$ to depend on the state to capture the different effects of interactions under different states. Then we can obtain the transition probability.

2.4 State-dependent Contributions

Given the contribution states, we now return to state-dependent user contributions in equation (4). We need to derive the emission probability $P(Y_{it} | s_{it})$. Since our observations of user contributions are non-negative, we adopt the standard Tobit model (Wooldridge 2010). Then we can derive the emission probability.

3. Estimation and Results

We estimate our HMM using a Bayesian procedure developed by Kim and Nelson (1999). The Bayesian estimation algorithm treats both the parameters $\theta = \{\beta, \xi, \sigma, \sigma_u\}$ and the state space $S = \{s_{it}\}_{t=1,\ldots,T;\, i=1,\ldots N_t}$ as random variables with prior distributions. The algorithm then updates the distributions $\pi(\theta, S | Y, X, W)$ using Gibbs sampling (Albert and Chib 1993), which gives the posterior distribution of the parameters by incorporating the information we obtain from our observations.

We apply the proposed structural model to a dataset from a large online community to test the ability of the model empirically. The goal of the empirical analysis is to show that our model could explain and predict the dynamics of community contribution in the individual level, and help us understand the effects of different mechanisms on this dynamic process.

3.1 Data Description

We compiled individual level panel data from SuperUser (superuser.com), the third largest site among all StackExchange sites by number of contributions in 2014. We collected detailed behavior and interaction data of each user on each date from July 12th 2009 to March 1st 2012. To focus on the vital contributors, we include only users who contributed at least 10 answers during the sample period. Our full sample contains 2147 users who have contributed 127,360 (80.9%) out of the 157,375 answers in our data. We then construct a user-date panel data for these users. We use the first 100 days of the data to estimate our model. Our estimation sample has 628 unique users, and 47,517 user-date observations.

3.2 Variables

In our study, the dependent variable is the quantity of answers measured by $Answers_{it}$, i.e., the number of answers provided by user *i* on date *t*. The definitions and summary statistics of our independent variables are presented in Table 1. We adopt two sets of explanatory variables that affect transition probabilities (*W*) and conditional contributions (*X*), respectively. The variables in vector *W* captures the social interactions that could have an enduring effect on the user's valuation of her own contributions. These variables come from categories such as reciprocity, social recognition, and badges.

Table 1 Model Variables and Descriptive Statistics

Variable	Description	Mean	Stdev
$Answers_{it}$	Number of answers provided by user *i* on date *t*	0.29	1.17
$New\_questions_t$	Number of new questions asked on the site on date *t*	103.45	27.21
$Group\_size_t$	Number of participated users on the site on date *t*	139.15	30.94
$Membership_{it}$	Number of days since user *i* registered on the site	43.14	27.27
$Answers\_received_{i,t-1}$	Number of new answers to user *i*'s past questions on date *t*-1	0.12	0.69
$Upvotes\_answer_{i,t-1}$	Total number of new up-votes to user *i*'s past answers on date *t*-1	0.49	2.09
$Accepted\_answers_{i,t-1}$	Number of answers of user *i* that are accepted on date *t*-1	0.06	0.32
$Badges_{i,t-1}$	Number of badges earned by user *i* on date *t*-1	0.11	0.47
Observations	47,517		

3.3 Estimation Results

Table 3 reports the posterior means and standard deviations of the structural model based on Bayesian estimation. The coefficients in vector β and ξ vary across states, which indicates that a change in states could lead to a change in the contribution behavior. The initial probabilities of being in inactive and active states are 0.856 and 0.144, respectively. Hence, a user is more likely to be in an inactive state when he first becomes a new member of the site. This confirms the importance to study how to energize users to active state.

Table 2 Model Variables and Descriptive Statistics

	Contribution States	
Variable Name	State 0 (Inactive)	State 1 (Active)
X_{it}	β - Mean (Std)	
C_1	-1.235*** (0.038)	2.221*** (0.270)
$New\_questions_t$	0.0002 (0.0005)	0.007*** (0.002)
$Group\_size_t$	0.004*** (0.000)	0.025*** (0.002)
$Membership_{it}$	-0.006*** (0.000)	-0.013*** (0.002)
σ^2	1.392*** (0.024)	
$W_{i,t-1}$	ξ - Mean (Std)	
C_2	-2.555*** (0.026)	-0.248*** (0.098)
$Answers\_received_{i,t-1}$	0.040*** (0.018)	0.003 (0.028)
$Upvotes\_answer_{i,t-1}$	0.095*** (0.018)	0.048* (0.030)
$Accepted\_answers_{i,t-1}$	0.574*** (0.055)	0.125*** (0.035)
$Badges_{i,t-1}$	0.169*** (0.033)	-0.076*** (0.030)
σ_u^2	1.000*** (0.010)	
Initial Probability	0.856*** (0.016)	0.144*** (0.016)

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

We investigate the state-dependent contributions first. The coefficients of $New\_questions_t$ indicate the response of users to the demand of knowledge. We can see that users respond very differently under different states. An *active* user may contribute more (0.007, significant at 1% level), while an *inactive* user may not respond to the new demand much (0.0002, insignificant). Hence, active users are more responsive to the demand of knowledge on the site, and they supply more knowledge when the need arises. Regarding $Group\_size_t$, we find that a user may contribute more when she feels that the community is bigger, which confirms the result in Zhang and Zhu (2011). Meanwhile, the effect of $Group\_size_t$ is much bigger when a user is in active state (coefficient is 0.025 rather than 0.004, both significant at 1% level).

We now turn to how community interactions (e.g., reciprocity and image rewards) and the badge system would affect the state transition probabilities. For users in an inactive state, receiving more answers on their previous questions tends to boost them to an active state (coefficient 0.040, significant at 1% level). In contrast, the coefficient is insignificant for users in an active state. Hence, the *reciprocity* of providing answers can be more useful to stimulate inactive users.

The coefficients on $Upvotes\_answer_{i,t-1}$ and $Accepted\_answers_{i,t-1}$ are positive and highly significant (all at 1% level across states). We interpret these as the result of the reputational motivation, or "image rewards" as in Benabou and Tirole (2006). When a user receives more "up" up-votes or has more accepted answers, her reputational motivation is satisfied and her inference about the value of her own contribution thus increases, which can boost a user to the active state. Further, this "image reward" is more prominent in the inactive state. Together, these results highlight the effectiveness of *image reward* as a motivating scheme, especially for inactive users to switch to the active state.

Earning more badges on answers seems to motivate an inactive user to become active (coefficient 0.169, significant at 1% level). However, the effect of badge is even negative (coefficient -0.076, significant at 1% level) for active users, which means that having past badges may not help to keep users in the active state. This may be due to the licensing effect in prosocial behavior (Gneezy et al. 2011). If so, using badge system to motivate user contributions should be gauged carefully. This could be a good topic for future research.

4. Conclusion

Motivating voluntary user contributions has become a key issue for the sustainable development of online communities in the knowledge economy. Toward this end, we propose a structural approach that embeds hidden Markov model in the framework of public goods. We estimate the model using actual data from a representative online community, StackExchange, and find that motivating mechanisms work differently depending on whether a user is in the active or inactive state. Social recognition and reciprocity are effective to turn inactive users into active contributors, especially for those who are inactive. And users respond more to group size and demand of knowledge when they are in active states.

Our study enhances the understanding of voluntary contributions/public goods problems in online settings. Our structural model combines the economic theory of public goods, behavioral aspect of user motivations, and stochastic process of the HHM model. This framework is applicable to modeling other online communities, and more broadly, other settings of open innovation communities. As more and more companies and organizations are leveraging "wisdom of the crowds" through online community for open innovation, our results provide guidance to design effective mechanisms and sustainable platforms.

References

1. Albert, J. H., and Chib, S. 1993. "Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts," *Journal of Business & Economic Statistics* (11:1), pp. 1–15.
2. Bénabou, R., and Tirole, J. 2006. "Incentives and Prosocial Behavior," *American Economic Review* (96:5), pp. 1652–1678.
3. Chen, Y., Harper, F. M., Konstan, J., and Xin Li, S. 2010. "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens," *The American Economic Review* (100:4), pp. 1358–1398.
4. Dietz, T., Ostrom, E., and Stern, P. C. 2003. "The Struggle to Govern the Commons," *Science* (302:5652), pp. 1907–1912.
5. Fischbacher, U., and Gächter, S. 2010. "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments," *American Economic Review* (100:1), pp. 541–556.
6. Kim, C.-J., and Nelson, C. R. 1999. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, The MIT Press.
7. Lerner, J., and Tirole, J. 2005. "The Scope of Open Source Licensing," *J Law Econ Organ* (21:1), pp. 20–56.

8. Ransbotham, S., and Kane, G. C. 2011. "Membership Turnover and Collaboration Success in Online Communities: Explaining Rises," *Management Information Systems Quarterly* (35:3), pp. 613–627.
9. Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*, (second edition.) The MIT Press.
10. Zhang, X., and Zhu, F. 2011. "Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia," *The American Economic Review* (101:4), pp. 1601–1615.

A Feature-based Approach to Sentiment Analysis of Online Chinese Product Reviews

Meiming Xing
Beijing Institute of Technology
minghaizai@163.com

Zhijun Yan
Beijing Institute of Technology
yanzhijun@bit.edu.cn

Dongsong Zhang
University of Maryland, Baltimore
zhangd@umbc.edu

Baizhang Ma
Beijing Institute of Technology
gufeiyan_295259083@qq.com

Wenyang Liu
Beijing Institute of Technology
1281909767@qq.com

Abstract

Online product reviews are a main channel for online shoppers to obtain product information. With the rapid accumulation of online product reviews, it has become impossible for users to go through them one by one. The sentiment analysis of online English product reviews has been explored extensively. But the approaches couldn't be applied to product reviews in Chinese directly due to differences between the two languages. In this paper, we propose a new approach to sentiment analysis of online Chinese product reviews. Based on existing feature extraction results, some of the implicit product features and domain-dependent sentiment terms were recognized. A sentiment analysis based on product features was conducted according to syntactic dependency relationships. The empirical results show the good performance of the proposed method.

Key words: sentiment analysis, feature scoring, text mining, online product reviews

1. Introduction

In the past decade, online shopping has become a dominant trend. Many e-commerce websites have emerged in China, including Taobao, Jingdong, Vancl, etc. Consumers tend to obtain product information through online product reviews before making their purchase decisions. However, it is time-consuming and ineffective to navigate reviews one by one in order to obtain useful information. Sentiment analysis of online product reviews can help consumers gain a better understanding of others' positive and negative opinions on a certain product and enable them to make a more informed purchase decision (Ding et al. 2008).

There have been increasing studies on online product reviews. However, most of existing literature focuses on sentiment analysis of product reviews in English. More importantly, different consumers are often interested in different product features. It is difficult to improve the buy conversion rate if a consumer can't quickly identify individual reviews that have commented on certain product features of his/her interest. Although a few e-Business companies such as taobao.com have integrated some methods to classify reviews based on product features, further effort is still necessary to ensure a more precise review analysis. At the same time, lots of sentiment words in a review are context-dependent, and there lacks an efficient way to incorporate context-dependent words into a sentiment term lexicon that can improve the sentiment analysis performance considerably.

Aiming at improving the accuracy of feature-level sentiment analysis, this paper proposes a sentiment analysis approach that integrates a domain-dependent sentiment term lexicon. The rest of the paper is organized as follows. Section 2 introduces the related research on online product reviews. Then, Section 3 presents the proposed sentiment analysis method. The experiment result is given and discussed in section 4. The final section gives the conclusions.

2. Related research

Currently, there are two categories of sentiment analysis methods. The first type of methods relies on existing lexicons like sentiment and polarity lexicons (Kim et al. 2004; Rao et al. 2009; Williams et al. 2009). The second one is based on supervised machine learning method (Boiy et al. 2009; Jiang et al. 2011), in which a classification model for predicting sentiments (i.e., positive/neutral/negative) is constructed with a training dataset.

Sentiment lexicons are major tools for lexicon-based sentiment analysis. There are two major ways to build sentiment lexicons, namely lexicon-based and document set-based. With lexicon-based building methods, a set of sentiment terms called the seed set is obtained from standardized sentiment lexicon generating tools. The seed set refers to sentiment terms extracted from those tools with polarity labels (positive/negative) on them. Then a complete set of sentiment terms for a certain product can be obtained iteratively(Hassan et al. 2011; Kamps et al. 2004). With document set-based building methods, the polarity of sentiment terms are obtained based on the relationships among different words in the same sentence, or different sentences in a document set(Kaji et al. 2007; Kanayama et al. 2006). Studies concerning lexicon-building for sentiment terms also exist (Wu et al. 2010).

Currently, many studies on sentiment analysis have adopted supervised learning models. As sentiment analysis shares the same principles at a sentence level and at a word level, learning models adopted in sentence-level analysis are also applicable to word-level sentiment analysis. Dependency parsers are widely used to analyze review structure, extract dependency relations and other information relative to sentiment properties(Ganapathibhotla et al. 2008; Jiang et al. 2011).

In this research, the lexicon-based sentiment analysis method is applied to online Chinese product reviews. Context-dependent sentiment terms, which were not addressed before, will be incorporated in the sentiment lexicon and help to improve the analysis precision.

3. Sentiment analysis based on dependency relationships

In our method, the sentiment lexicon is firstly built, and the sentiment analysis based on dependency relationship is conducted to determine the sentiment polarity of product features and derive the sentiment polarity of the review. Moreover, the sentiment lexicon can be expanded automatically according to the product feature polarity.

3.1 Lexicon building

A complete lexicon is the core of lexicon-based sentiment analysis. In this research, our lexicons consist of a basic polarity lexicon, a polarity modifying lexicon, and a domain-dependent lexicon.

A basic polarity lexicon includes terms that are widely adopted and with a relatively fixed sentiment polarity. We use HowNet as the basic polarity lexicon in this study, which consists of 219 degree-level terms, 3,116 negative evaluation terms, 1,254 negative sentiment terms, 3,730 positive evaluation terms, 836 positive sentiment terms, and 38 proposition terms.

Although the basic polarity lexicon covers the polarity of a large number of sentiment terms, it is not sufficient to deal with all possible sentiments in product reviews. We take negation, strengthening and weakening of basic polarity terms into consideration. And the polarity modifying lexicon is constructed, which includes a negation lexicon and an emphasis lexicon. The emphasis lexicon is mainly originated from degree-level terms in HowNet, and the negation lexicon is built by analyzing the online product review. For each level of adverbs, a degree number is assigned.

Polarity terms and modifiers can define the sentiment that a term conveys. However, some adjectives don't have an explicit sentiment polarity, or only display a positive/negative polarity under certain contexts. These adjectives are called domain-dependent terms, and their polarity is generally determined according to linguistic rules(Yan 2010). Linguistic rules include rules of conjunctions within sentences, conjunctions between sentences and non-adversative conjunctions. The first two rules determine sentiment polarity through adversative and parity conjunctions, such as "and"(和, 而且) and "but"(但是). Rules for non-adversative conjunctions indicate that if there are no adversative conjunctions within sentences or between sentences, the polarity stays the same.

We also determine the polarity of a domain-dependent term by occurrence frequency, precisely, by calculating the difference in the frequency in which the term displays a positive polarity and a negative polarity. We then judge its polarity by comparing the difference to a pre-specified threshold F_0 (e.g., 5). Once a domain-dependent term displays a positive polarity F_0 times more than when it displays a negative polarity within review text, its final polarity will be determined to be positive.

3.2 Extraction of dependency pairs

The polarities of product features and review sentences are determined by analyzing the dependency relationship between feature terms and sentiment terms, as well as the adversative relationship between sentences.

Before the analysis of dependency relationship, some pre-processing needs to be done on the review corpus, including word segmentation, POS tagging, noun phrase extraction, and sentence segmentation. The ICTCLAS (Institute of Computing Technology, Chinese Lexical Analysis System) system is used for word segmentation and POS tagging. We adopt an existing approach to noun phrase extraction(Ma et al. 2013).

Based on the result of ICTCLAS, the sentence segmentation is conducted to shorten the length of review sentences by dividing review text into sentences that end with periods, exclamation marks, or ellipsis. The Stanford parser is used to perform dependency parsing which generates dependency trees and the set of dependency pairs. Each review will generate a dependency tree. Then we identify dependencies that possibly contain both product features and sentiment terms from all dependency pairs.

By analyzing reviews, we find that dependencies containing product features usually exist in subject-predicate relationships (nsubj), verb-object relationships (dobj) or structures with "的" (rcmod). Therefore we search product features mainly by detecting those 3 structures in the reviews.

3.3 Sentiment analysis

After extracting dependency pairs, product feature is derived by adapted PageRank algorithm (Ma 2014). And each feature term in a review, as well as each review, will be assigned a sentiment score during the sentiment analysis process. Additionally, the lexicon of sentiment terms of product reviews is expanded automatically. The whole process of sentiment analysis includes clause segmentation, basic scoring, context-dependent scoring, and sentence scoring. Cues to clause segmentation include periods, commas, and ellipsis, with a variety of punctuation marks taken into account. For example, the review "屏幕(screen)/n 大(big)/a ，/w 电池(battery)/n 比较(pretty)/d 耐用(durable)/a 。/w" could be segmented into two clauses: "屏幕/n 大/a" ("the screen is big") and "电池/n 比较/d 耐用/a".

For each clause, if a dependency pair contains a product feature and a sentiment term, a basic scorer is first used to detect the polarity of the sentiment term. The function of the basic scorer is to search the sentiment lexicon for the terms and their modifiers and obtain a basic score. However, there may be negations and degree modifiers that change the intensity or even the polarity of sentiment terms. There are also occasions where the polarity terms and sentiment terms don't correspond one-to-one. Therefore the basic scorer needs to deal with these special circumstances.

(1) If there is a negation before a sentiment term, the basic scorer sets the polarity as the opposite of the original polarity of the term.
(2) If there is a modifier before a sentiment term, the basic scorer sets the polarity as the polarity of the term multiplied by the degree of the modifier.
(3) If there are both a negation and a modifier before a sentiment term, the basic scorer sets the polarity as the combined result of (1) and (2).
(4) If a feature term corresponds to more than one polarity term, the feature's sentiment polarity is the sum of all polarities of the sentiment terms attaching to it.

If the basic scorer successfully performs scoring, the proposed approach continues scoring the polarity at a sentence level. Otherwise, a clause contains domain-dependent sentiment terms, which indicates that context-dependent scoring is further required. For example, the clause "屏幕(screen)/n 大(big)/a" has a dependency of type nsubj (according to Stanford parser, where "屏幕"-1, "大"-2). "屏幕(screen)" is a product feature, "大(big)" is a sentiment term that is domain-dependent. Therefore, the polarity of "大(big)" couldn't be determined by the basic scorer, and should be forwarded to the context-dependent scorer.

The context-dependent scorer judges the polarity of a sentiment term based on review context in which a term resides after the basic scorer fails to produce a result. The context mainly refers to relationships implied by conjunctions, including adversatives and juxtapositions. The context-dependent scorer adds a term and its polarity to a self-expanding sentiment lexicon if it successfully assigns a score to the term, and the lexicon in turn helps the scorer when it reaches a certain scale. So if an adversative or juxtaposition conjunction is located, we will search the basic polarity lexicon for the polarity of the neighboring sentiment terms, and use the result as the terms' polarities. If the neighboring terms cannot be found in the basic polarity lexicon, the proposed approach will set the terms' polarities to neutral as default and terminate the scoring process.

The whole polarity of a review is determined by aggregating the polarities of all its clauses and that's the end of the sentiment analysis process.

4. Evaluation and analysis of results

4.1 Evaluation indices and data

Precision has been the most frequently used measure for evaluating approaches to sentiment analysis. The definition of precision is shown in equation (1) as follows:

$$Precision = \frac{number\ of\ reviews\ whose\ polarities\ are\ correctly\ labeled}{total\ number\ of\ reviews} \quad (1)$$

A web crawler written in Java was utilized to collect online consumer reviews on a camera and a mobile phone generated between December 2012 and August 2013 from jd.com. The dataset includes product name, consumer id, consumer level, review scoring (by the consumer), review content and review date. The number of camera reviews is 11,291, and the number of phone reviews is 8,901.

4.2 Results

The evaluation result is shown in Figure.1, which illustrates that the proposed approach achieves high precision when dealing with positive reviews, and the overall precision is also satisfactory. However, the precision for neutral and negative reviews is relatively low. Possible reasons are: (1) The absolute number of neutral and negative reviews are small (Samsung Note II has 65 negative reviews, constituting 0.73% of all reviews; Canon EOS 600D has 293 negative reviews, constituting 2.6% of all reviews), thus there are inevitably high fluctuations resulted from imprecise polarity labeling;(2) Neutral reviews are generally ambiguous, so it's difficult for the model to generate results that are consistent with manual labeling, leading to low precision;(3) The basic polarity lexicon is far from perfect (i.e., some items are labeled differently from manual labeling), causing low precision.

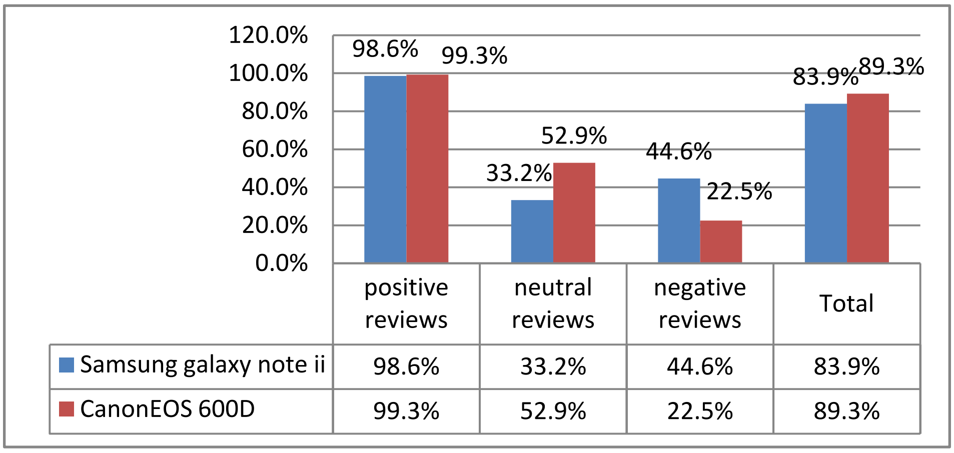


Figure.1 The result of sentiment analysis

5. Conclusion

This paper proposes a sentiment analysis method based on dependency relationships. Our research covered the construction process of lexicons, including basic polarity lexicon, polarity modifying lexicon, and domain-dependent lexicon. We explored a self-expanding domain-dependent lexicon based on occurrence frequencies by considering semantic structures. By considering both the dependency relationships between product feature terms and sentiment terms and the connections between sentiment terms and between sentences, the proposed approach analyzes sentiment polarities of review sentences and obtains sentiment labels for reviews. Finally, we conducted an experiment using the dataset crawled from jd.com, and evaluated the actual performance of the proposed method. In future research, we will compare this method with other existing methods, and try to improve the precision of the sentiment labeling process.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (Award #: 71128003, 70972006, 71272057) and Program for New Century Excellent Talents in University (Award #. NCET-11-0792).

Reference

Boiy, E., and Moens, M.-F. 2009. "A machine learning approach to sentiment analysis in multilingual Web texts," *Information retrieval* (12:5), pp 526-558.

Ding, X., Liu, B., and Yu, P. S. Year. "A holistic lexicon-based approach to opinion mining," Proceedings of the international conference on Web search and web data mining, ACM2008, pp. 231-240.

Ganapathibhotla, M., and Liu, B. Year. "Mining opinions in comparative sentences," Proceedings of the 22nd International Conference on Computational Linguistics, ACM, Manchester, 2008, pp. 241-248.

Hassan, A., Abu-Jbara, A., Jha, R., and Radev, D. R. Year. "Identifying the Semantic Orientation of Foreign Words," ACL 2011, pp. 592-597.

Jiang, L., Yu, M., Zhou, M., Liu, X., and Zhao, T. Year. "Target-dependent Twitter Sentiment Classification," ACL2011, pp. 151-160.

Kaji, N., and Kitsuregawa, M. Year. "Building Lexicon for Sentiment Analysis from Massive Collection of HTML Documents," EMNLP-CoNLL2007, pp. 1075-1083.

Kamps, J., Marx, M., Mokken, R. J., and De Rijke, M. 2004. "Using wordnet to measure semantic orientations of adjectives," The International Conference on Language Resources and Evaluation Conference2004, pp. 1115–1118.

Kanayama, H., and Nasukawa, T. Year. "Fully automatic lexicon expansion for domain-oriented sentiment analysis," Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Sydney, 2006, pp. 355-363.

Kim, S.-M., and Hovy, E. Year. "Determining the sentiment of opinions,"The 20th international conference on Computational Linguistics, Association for Computational Linguistics2004, pp. 1367-1373.

Ma, B. 2014. *E-commerce consumer online reviews mining research*, Beijing Institute of Technology.

Ma, B., Zhang, D., Yan, Z., and Kim, T. 2013. "An Lda and Synonym Lexicon Based Approach to Product Feature Extraction from Online Consumer Product Reviews," *Journal of Electronic Commerce Research* (14:4), pp 304-314.

Rao, D., and Ravichandran, D. Year. "Semi-supervised polarity lexicon induction," Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics, Association for Computational Linguistics2009, pp. 675-682.

Williams, G. K., and Anand, S. S. Year. "Predicting the Polarity Strength of Adjectives Using WordNet," ICWSM2009, pp. 346-349.

Wu, Y., and Wen, M. Year. "Disambiguating dynamic sentiment ambiguous adjectives," Proceedings of the 23rd International Conference on Computational Linguistics, Association for Computational Linguistics2010, pp. 1191-1199.

Yan, S. 2010. *Research on Opinion Mining of Product Reviews in Chinese*, Beijing Jiaotong University.

Complimentary Return Freight Insurance as Signals: An Innovation in E-commerce Return Policy

Shidao Geng
Dalian University of Technology
shidao@mail.dlut.edu.cn

Wenli Li
Dalian University of Technology
wlli@dlut.edu.cn

Abstract

With the increasing number of online commodities and retailers，to distinguish the quality of commodities and retailers in e-commerce is more complicated. E-commerce is virtual, anonymous and remote. Therefore, these characteristics lead information asymmetry to be serious more than ever before. Product and online retailer quality are communicated through website signals. Using signaling theory, we focus on an innovative return policy, the online retailers offer the complimentary return freight insurance to the consumers, and show the condition of separating equilibrium on an important variable, online retailer's return rate. Comparing with the old return policy, the new policy that consumers get free return freight insurance is a little more rigorous to the good retailers and much more helpful for the new retailers whose return rate is not fixed. We also propose a series of implication for the online platform to encourage this innovation.

Keywords: Return policy, Return freight insurance, Signaling theory, Innovation

1. Introduction

With the continuous development of e-commerce, the number and types of online commodities are constantly increasing, so consumers have more choice for online purchase. However, series of new problem have emerged, such as how to select the one that can obtain the best shopping experience in the similar products, or how to search the better one in the imperfect information online system. Retailers who are aware of that they could provide better commodities and services are more eager to share this fact with the online consumers. Though online reputation, commodities description and reviews could transmit the information to the buyers as signals (Mavlanova et al. 2012), which can be easily interfered by fake reviews or opinion spam(Lau et al. 2012; Liu 2010) and make consumers confused.

Return freight insurance was created by the insurance company and online platform Alibaba Corporation in 2010. If insured, it will compensate the freight fee by insurance company in case consumers return goods after they purchase. With the low premium, easy-insured and convenient-claim features, return freight insurance provides the psychological and financial security for online consumers. Meanwhile, it is an innovation in return policy that online retailers buy this insurance and then send it to consumers for free, which will convey the information that their product or service has high quality.

We consider that loose return policy such as complimentary return freight insurance is familiar with apology, because they both would give online retailers another opportunity to repair their image(Ho 2011). But the two signals appear in different time. An apology is made after the 1st stage and before the next cooperation, while the free insurance is offered as soon as consumers decide to buy.

2. Literature Review

2.1 Return Policy Influence

Return is actually a good opportunity to please consumers and a recognized tool to create more loyal customers (Griffis et al. 2012; Mukhopadhyay et al. 2005). This is even truer for increasingly popular Internet sales where the opportunity to physically examine the product is absent. Return policy was still an essential part even in the context of the financial crisis and loose return policy was more profitable than harsh return policy (Petersen et al. 2010). The purchase decision is more likely to be framed as two separate decisions: consumers' decisions to order and, upon receipt, their decisions to keep or return the item, and the endowment effect suggests some surprising benefits of return policy leniency to the retailer (Wood 2001). From manufacturer's point of view, following a policy of modularization and offering a generous return policy would increase revenue, but also increase the cost due to increased likelihood of return and increased cost of design (Mukhopadhyay et al. 2005).

2.2 Signaling In E-commerce

Using signaling theory, a three-dimensional framework was developed to classify website signals and found that low-quality sellers were likely to avoid costly and easy-to-verify signals and used fewer signals than did high-quality sellers, who used costly and difficult-to-verify signals and displayed more signals (Mavlanova et al. 2012). Using a signaling model, conditions were obtained when Internet retailers (e-retailers) use price to manage their customers' service expectations. In contrast to extant theory, it is possible for both low and high service e-retailers to use price in signaling their service levels (Mitra et al. 2010). Investigate three possible signals to distinguish between ''trustworthy'' and ''untrustworthy'' Web merchants in the case of B2C Internet commerce (Lee et al. 2005). A firm may signal the unobservable quality of its products through several marketing-mix variables. The authors develop a typology that classifies signals and discuss the available empirical evidence on the signaling properties of several marketing variables (Kirmani et al. 2000).

2.3 Apologize As Signals

Apologies are more frequent in long relationships, early in relationships, and between better-matched partners (Ho 2011). Apologies without cost have no value at all, so the return policy only is effective with purchase.

3. Base Model

The basic context is that the consumers make decision to buy or not in many merchants sell similar products driven by the existed demand, after the purchase, if there are problems in the product whether to return or not.
The two-stage actions should be considered, and the signal depends on the all payoffs.

3.1 Assumption

Assumption1. The consumer has no shopping experience in the specific online shop, which means he or she has no idea of the retailer type before.
Assumption2. The retailer's information such as commodity introduction, reputation, and online review is general acceptance. The consumer evaluation of the different retailer's product is approximately the same.
Assumption3. The old return policy commands the retailer to offer the freight fee for the product defect.
Assumption4. This paper is concerned about the payoff of retailer and consumer only calculates during return service, so the cost of commodity is considered as sunk cost.
Assumption5. The good retailer has a stronger integrate ability, whose return rate is lower.

3.2 Base Model

Set freight fee as f , premium of return freight insurance $p_i(r)$. When the consumer is a fixed group, the premium is increasing by the return rate of retailer.

Let compensation of return freight insurance as $C_i \approx f$, opportunity cost of return c_r , consumer product evaluation e , product quality q , consumer use value $v(e,q)$, and use value for defect goods v_d

Set commodity price P , p , $P \succ p$, rate of return $r \in (r_1, r_2)$, $r_1 \prec r_2$, cost of goods C .

The signaling model of return policy in E-commerce is shown in figure1. The arrow shows the different payoff between old return policy and the innovative one.
The 1st stage: without return policy
If the consumer decide to buy, then
$E_r(1-r)v + E_r(r)v_d - p \succ 0$ (1),
which means without the signaling the highest offered price is $E_r(1-r)v + E_r(r)v_d$, sum of the use value expectation and the use value expectation of the damage goods.
From the perspective of retailer,

$p \succ E_r(1-r)C$ (2), combined with (1)
$E(1-r)v + E(r)v_d \succ E(1-r)C$ (3)
As the assumption for the return rate of good and poor retailer, $r_1 \prec r_2$,
$E(1-r_1) \succ E(1-r_2)$ (4)
the good retailer's profit margins is low.
If the price $p \prec E(1-r)v + E(r)v_d \prec E(1-r_1)C$, the good retailer is squeezed out of the market, which entails they are more willing to send signals to consumers.

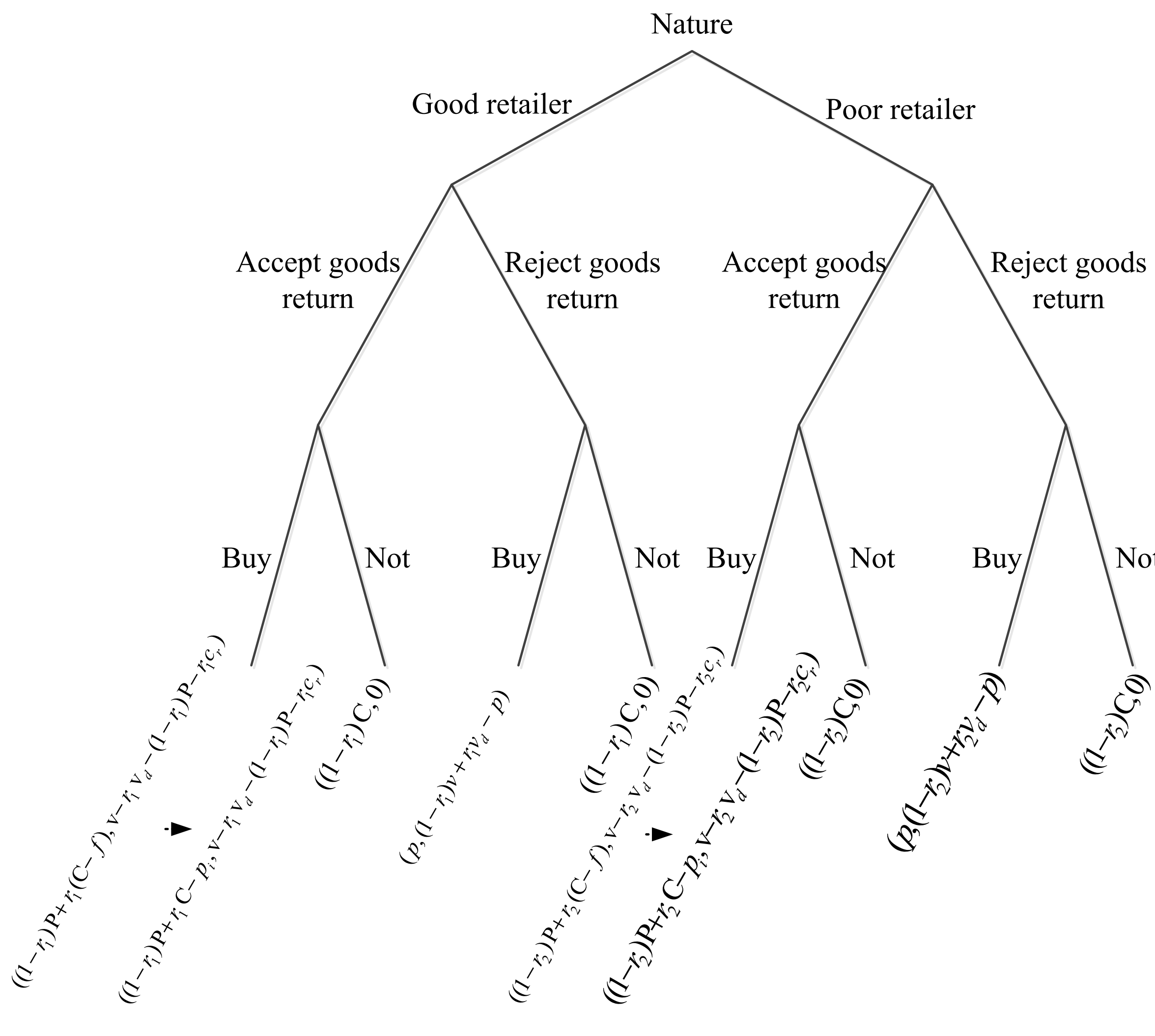


Figure1: Signaling model of return policy in E-commerce

The 2nd Stage:
1.Return policy as signals
$p \prec E(1-r)v + E(r)v_d \prec E(1-r_1)C$ (5)
the good retailer take the acceptance of goods return as the signal, which consumer will offer a higher price P, and without the return policy the price is still p.
The condition of consumer return goods is

$v_d \prec p - c_r$ (6),

indicate that the damage goods use value is smaller than the difference between the price and opportunity cost of return.

For the good retailer, accept goods return, the payoff is $(1-r_1)P + r_1(C-f)$, and reject return the payoff is still p; for the poor retailer, accept goods return, the payoff is $(1-r_2)P + r_2(C-f)$, and reject return the payoff is still p.

The good retailers want to accept goods return, while the poor always reject. If the signals work (Spence 1973), the separating equilibrium is as followed

$r_1 f \le E_r(1-r)P + E_r rC - p \le r_2 f$ (7)

and because the cost of goods C is sunk,

turns out $r_1 f \le E_r(1-r)P - p \le r_2 f$ (8)

it entails that the good retailer's difference of two expectation prices should be between the maximum and minimum return freight fee.

2.return freight insurance as signals

For the good retailer, offer the insurance, the payoff is $(1-r_1)P + r_1 C - p_i$, and reject return the payoff is still p; for the poor retailer, offer the insurance, the payoff is $(1-r_2)P + r_2 C - p_i$, and reject return the payoff is still p. The payoff of consumer is still the same.

The separating equilibrium for return freight insurance is $p_i(r_1) \le E_r(1-r)P - p \le p_i(r_2)$ (9)

The good retailer's difference of two expectation prices is between the two type premiums of return freight insurance.

And for the insurance company sake, there are $p_i(r_1) \succ r_1 f$, $p_i(r_2) \succ r_2 f$. The policy separating equilibrium of two return intervals get changed.

4. Result And Implication

Result1

Deliver the return freight insurance to the consumer for free is an easy return policy indeed, and comparing the return policy has almost the same form as signals.

Result2

We find out a new separating equilibrium for the good seller in E-commerce and show the applicable conditions.

Implication1

For the insurance company sake, it's more rigorous to distinguish the good retailer to use the return freight insurance as a signal.

Implication2

It gives the poor retailer an opportunity to adjust their return rate to satisfy the consumers. Especially for the new retailers, whose return rate is not fixed, the return freight insurance provides protection for their growth.

Implication3

Online platform could design the mechanism that set the complimentary return freight insurance as a Third-Party Assurance seal to make sure send this signal efficiently.

5. Future Research

The replacement is more common than goods return in online commerce, therefore, we consider to take the replacement as the 3rd stage for the future study. In the replacement circumstance, old return policy as signals may not applicable any longer, because the retailer return rate may not fixed in this situation for the different consumer risk preferences. Complimentary return freight insurance remains shining.

Reference

1. Griffis, S. E., Rao, S., Goldsby, T. J., and Niranjan, T. T. "The customer consequences of returns in online retailing: An empirical analysis," *Journal of Operations Management* (30:4) 2012, pp 282-294.
2. Ho, B. "Apologies as Signals: With Evidence from a Trust Game," *Management science* (58:1) 2011, pp 141-158.
3. Kirmani, A., and Rao, A. R. "No Pain, No Gain: A Critical Review of the Literature on Signaling Unobservable Product Quality," *Journal of Marketing* (64:2) 2000, pp 66-79.
4. Lau, R. Y. K., Liao, S. Y., Kwok, R. C.-W., Xu, K., Xia, Y., and Li, Y. "Text mining and probabilistic language modeling for online review spam detection," *ACM Trans. Manage. Inf. Syst.* (2:4) 2012, pp 1-30.
5. Lee, B.-C., Ang, L., and Dubelaar, C. "Lemons on the Web: A signalling approach to the problem of trust in Internet commerce," *Journal of Economic Psychology* (26:5) 2005, pp 607-623.
6. Liu, B. "Sentiment analysis: A multi-faceted problem," *IEEE Intelligent Systems* (25:3) 2010, pp 76-80.
7. Mavlanova, T., Benbunan-Fich, R., and Koufaris, M. "Signaling theory and information asymmetry in online commerce," *Information & Management* (49:5) 2012, pp 240-247.
8. Mitra, D., and Fay, S. "Managing Service Expectations in Online Markets: A Signaling Theory of E-tailer Pricing and Empirical Tests," *Journal of Retailing* (86:2) 2010, pp 184-199.
9. Mukhopadhyay, S. K., and Setoputro, R. "Optimal return policy and modular design for build-to-order products," *Journal of Operations Management* (23:5) 2005, pp 496-506.
10. Petersen, J. A., and Kumar, V. "Can Product Returns Make You Money?," *MIT Sloan Management Review* (51:3) 2010, pp 85-89.
11. Spence, M. "Job Market Signaling," *The Quarterly Journal of Economics* (87:3) 1973, pp 355-374.
12. Wood, S. L. "Remote Purchase Environments: The Influence of Return Policy Leniency on Two-Stage Decision Processes," *Journal of marketing research* (38:2) 2001, pp 157-169.

Administering and Capitalizing on Product sampling In an online context

Xianghua Lu	Chee Wei Phang	Xinlin Yao	Sulin Ba
Fudan University	Fudan University	Fudan University	University of Connecticut
lxhua@fudan.edu.cn	phangcw@fudan.edu.cn	12210690033@fudan.edu.cn	Sulin.Ba@business.uconn.edu

Abstract

Product sampling, a promotional tactic long employed by brands looking to enter a new market, release a new product, or increase existing sales, has recently been applied in the online context. By analyzing product sampling campaign data from a leading e-commerce store in China, we show that administering product sampling online can lead to increased sales, but only for experience products and not search products. Additionally, publishing product consumption experience information from product sampling (i.e., product sampling reviews), alongside regular consumer reviews, can stimulate sales of experience products directly and indirectly through enhancing subsequent consumer reviews. However, it is interesting to note that when product sampling reviews were overwhelmingly positive, they instead lowered subsequent consumer reviews. This implies that firms hoping to influence consumer reviews through positive product sampling reviews need to be wary of their negative repercussion.

Keywords: Product sampling, electronic commerce, consumer reviews, word-of-mouth (WOM).

1 INTRODUCTION

Product sampling, which offers a free amount or a trial of a product for consumers, is an old but effective way to enter a new market, promote a new product, or increase existing sales (Jain et al. 1995; Marks and Kamins 1988). By distributing samples to target customers to try out for free, this promotional tactic allows firms to demonstrate their product quality while taking the risk off the consumers. Consequently, product sampling may help create brand awareness, enhance brand perception, induce purchases, and potentially lead to brand loyalty (Bawa and Showmaker 2004; Gedenk and Neslin 1999; Heiman et al. 2001; Lammers 1991). Furthermore, it may be used as a means to poach customers from competitors (Jain et al. 1995).

Recently, this promotional tactic has been extended to the online context. For instance, Walmart.com has a section that offers free product samples provided by various brands to customers. In China, major e-commerce portals such as taobao.com and yihaodian.com also allow brands to offer free samples through their platform. The online context allows firms to reach a wide audience (Rowley 1996), which makes it easier to identify target consumer segments for administering product sampling. In addition, compared to offline product sampling campaigns where consumer responses are difficult to track, in the online context firms are able to measure the outcomes of this strategy in an accurate manner.

Firms may also acquire product consumption experience information from the consumers who are given product samples, and use it as word-of-mouth (WOM) to further stimulate sales. The online context enables firms to obtain such information from the consumers (e.g., by asking them to write a product sampling review) and disseminate the information to a wide audience to influence sales. These advantages of administering product sampling online has led to a prediction that it will be the “next big thing” (Costa 2010). Yet surprisingly, there has been very scant research on product sampling in the online context.

Given the lack of research and empirical evidence, firms would be interested to know whether and to what extent it can lead to increased sales to justify their investment in the online channel.

To more fully capitalize on administering product sampling online, firms would also be keen to make use of the product consumption experience information generated to further stimulate sales. However, it is not clear whether and how this can be achieved. To investigate these issues, we collaborate with a leading e-commerce store that administers product sampling online for data collection. In the following, we discuss the research hypotheses of this study.

2 HYPOTHESES development

We expect product sampling applied in the online context will lead to increased sales. Previous research on product sampling has highlighted three theoretical perspectives that may explain the positive effect of this promotional measure on consumer purchase, namely learning theory, self-perception theory, and attribution theory (Lammers 1991). First, sampling may lead to an effect akin to "shaping" in learning theory (Nord and Peter 1980). Consumers who sampled a product may be shaped to think of themselves as users of the product, thus increasing their probability of purchasing the product in future (Lammers 1991). Second, by sampling a product, consumers may go through a process of forming self-perceptions about their behavior (Bem 1972; Lammers 1991), e.g., as someone who is willing to try and use the products sampled. This leads them to be more willing to purchase the product when opportunity arises. Third, through the perspective of attribution theory (Sawyer and Dickson 1984), a product sample may heighten or accentuate the cues associated with the consumption of the product, and increase the consumers' tendency to purchase.

Beyond the consumers who are given a product sample, product sampling when applied in an online context may also non-trivially influence other consumers at a broader level. The online channel increases the visibility of a product sampling campaign by expanding the audience base it can reach (e.g., visitors of an e-commerce store). However, when examined at a deeper level, we expect the positive effect is only salient for experience products but not for search products, particularly in the online context. Search products are goods or services for which the most essential attributes can easily be evaluated prior to purchase (Franke et al. 2004; Huang et al. 2009). In contrast, experience products are goods and services for which the evaluation of their essential attributes incurs high cost in terms of time, money, cognitive effort, or other resources (Franke et al. 2004; Huang et al. 2009). Thus, we hypothesize:

H1. Administering product sampling online will promote sales, but only for (a) experience products; and not for (b) search products.

To maximize the impact of product sampling, it is hoped that consumers who sampled the product will share WOM information with others (Heiman et al. 2001; Jain et al. 1995), e.g., their friends and family members. In the online context, firms may more systematically leverage on this measure to amplify its effect. For instance, firms may require consumers to register to receive product samples, follow up with them afterward to assess their evaluation (e.g., via email), and incentivize them to submit product sampling review online. The product sampling reviews may then be published on the product webpage, along with regular consumer reviews (e.g., on top of the latter reviews), to stimulate greater sales. As previous research indicates, e-WOMs, in particular those of positive nature, are an effective driver of sales (Chen and Xie 2008; Dellarocas et al. 2007; Zhu and Zhang 2010). Intuitively, product sampling reviews tend to be more positive in nature because consumers who received a product sample are likely to perceive favorably about the brand out of goodwill (Heiman et al. 2001). For the same reason, these consumers may also be willing to expend effort in providing greater details about the product in writing its review, which can help reduce uncertainty for other consumers. Hence, we expect the incorporation of product sampling reviews into the pool of regular consumer reviews would serve to bring about greater sales. However, for the similar rationales provided for H1 above, we expect the positive effect is only salient for experience products but not search products. Specifically, the marginal increment of useful information due to the incorporation of sampling reviews of search products is likely to be minimal, given that information about their essential attributes can be obtained online, and might already be available in existing consumer reviews. In contrast, experience products are likely to benefit non-trivially from the greater details that consumers willingly put into describing their experience with such products. Such details might not have been available under regular circumstances (i.e., consumers write reviews after purchasing and trying an experience product). Hence, we hypothesize:

H2. Publishing product sampling reviews alongside regular consumer reviews (within the same section) will promote sales, but only for (a) experience products; and not for (b) search products.

3 DATA DESCRIPTION

We collaborated with a leading e-commerce store in China to obtain data of product sampling campaigns and their effects. A product sampling platform was set up in Oct. 2012, which allows brands to offer product samples to the store customers. Customers may apply for free product samples on the platform, from which recipients of the samples are randomly selected. A product sampling campaign is typically run for one week, during which customers were invited to apply for sampling the product offered. The samples would be mailed to the recipients' address right after the end of sampling campaign, and they are requested to write a product sampling report detailing their experience of using the product sample within one month. In September 6, 2013, the product sampling platform decided to incorporate the product sampling reviews gathered for a product into the regular consumer reviews of the product to assess if this could promote sales (see Figure 2). The selected product sampling reviews are placed on top of regular consumer reviews to increase their visibility. Customers know this review is not a regular one as website put an indicator to show this is a product sampling review.

Through the platform, we collected two panel datasets, Dataset I (DAS-I) and Dataset II (DAS-II). DAS-I was employed to assess the effects of administering product sampling campaigns on sales (H1(a) and H1(b)). The dataset contained 147 products that ran sampling campaigns during June and July, 2013 on the platform and their sales data. In this dataset, we also included the sales data of 86 products that did not run sampling campaign as a control. This product subset was purposely selected such that they consisted of similar products as those that ran sampling campaigns for comparison purpose. The whole period of this dataset was from May 1 to August 31, 2013 (to capture the sales data before and after the campaigns). The reason for this separate dataset was to have a clean assessment of the effects of administering product sampling, without the potential influences from incorporating product sampling reviews into the regular reviews on the platform in September 6, 2013.

DAS-II was employed to examine the effects of incorporating product sampling reviews into regular consumer reviews (H2(a), H2(b)). It included the sales data of 217 products which conducted product sampling campaigns during the period from October 2012 to July, 2013. We captured their sales data from August 6, 2013 to January 6, 2014, to examine the change in sales before and after the incorporation of product sampling reviews on September 6, 2013. Table 1 is the descriptive statistics of main variables in this study.

Table 1. Descriptive statistics of the datasets

Source	Variable	Obs.	Mean	Std. Dev.	Min.	Max.
DAS-I	*dailysales*	21760	17.21	78.88	0	2964
	product_price	21760	238.67	333.43	4	2280
	business_unit	21760	0.45	0.50	0	1
	experience_goods	21760	0.76	0.42	0	1
DAS-II	*dailysales*	33418	23.34	202.37	0	16923

	ccumulative_rating	30436	4.75	0.21	3	5
	latest_rating	30437	4.66	0.68	1	5
	product_price	33418	199.94	344.72	11.9	2790
	dailypost	33418	0.87	3.78	0	145
	impression	33418	67.80	343.55	0	36264
	experience_goods	33418	0.75	0.43	0	1
	sample_rating	21256	4.98	0.15	4	5
	photos	21256	6.76	3.88	0	12
	length	21256	275.36	224.82	50	1270
	trial_word	21256	0.53	0.50	0	1
	positive_points	21256	10.00	6.96	1	54
	negative_points	21256	0.21	0.53	0	3

4 MODELS

4.1 Study I: Effects of Administering Product Sampling on Sales

For this study, our focal dependent variable was sales (quantity of products sold); and we controlled for other factors such as price, product type and date. We employ a difference-in-difference (DiD) model to examine the effects of administering product sampling online on sales. The DiD approach is a research design for estimating causal effects. It is popular in empirical economics, e.g., to estimate the effects of certain policy interventions and policy changes that do not affect everybody at the same time and in the same way. It is used in other social sciences as well and in many applications, where “time” is an important variable that may hamper the assessment of the true effect of a treatment. In this study, we set the products which conducted sampling campaign as treatment group and those (similar products) that did not conduct such a campaign during the same time period as control group. Thus, we could compare the differences of means of sales between the two groups to examine effects of administering product sampling online.

Using a difference in deference model, we specify the regression model as follows:

$$(1)\log(dailysales)_{it} = \gamma_0 + \gamma_1 before\_after_{it} + \gamma_2 treatment_i + \gamma_3 treatment_i Xafter_{it} + \gamma_4 \log(dailysales)_{i,t-1}$$
$$+\gamma_5 \log(product\_price)_i + \gamma_6 business\_unit_i + \gamma_7 holiday_t + \gamma_8 weekend_t + \mu_i + \varepsilon_{it}$$

4.2 Study II: Effects of Incorporating Product Sampling Reviews

In this study, we would investigate the effects of incorporating Product Sampling Reviews. We use *Time*=1 to indicate that the date is after September 6, when product sampling reviews was incorporated into regular consumer reviews section. In our dataset, some products do not have product sampling reviews incorporated in our observation period, either because there is no appropriate sampling review available or brands choose delay the combination. For these products, consumers could not read sampling reviews in review section even when *time*=1. This phenomenon naturally create a control group for this study. We then are able to examine the effects of product sampling reviews incorporation with a DiD model. We use dummy variable *sampling_review* to distinguish the treatment group and control group.

In addition, Duan et al (2008) found there is a dynamic relationship between sales and regular review volume. In our case, we also need to considerate this relationship when we investigate the effects of sampling review on sales. Because if the sampling review could increase the sales, then more regular reviews would be posted. The posted review inversely would impact the sales. Thus, based on DiD specification, we employ a simultaneous equations model as follows to capture their dynamic relationship (Duan et al 2008).

$$(2)\log(dailysales)_{it} = \alpha_0 + \alpha_1 time_t + \alpha_2 sampling\_review_{it} + \alpha_3 timeXsampling\_review_{it} + \alpha_4 \log(price)_i$$
$$+\alpha_5 \log(impression)_{it} + \alpha_6 \log(dailypost)_{it} + \alpha_7 \log(dailypost)_{i,t-1} + \alpha_8 \log(dailypost)_{i,t-2} + \alpha_9 \log(cumulative\_rating)_{it}$$
$$+\alpha_{10} \log(latest\_rating)_{it} + \alpha_{11} weekend_t + \alpha_{12} holiday_t + \alpha_{13} promotion_{it} + \delta_i + u_{it}$$

$$(3)\log(dailypost)_{it} = \beta_0 + \beta_1 time_t + \beta_2 sampling\_review_{it} + \beta_3 timeXsampling\_review_{it} + \beta_4 \log(price)_i$$
$$+\beta_5 \log(dailysales)_{i,t} + \beta_6 \log(dailysales)_{i,t-1} + \beta_7 \log(dailysales)_{i,t-2} + \beta_8 \log(cumulative\_rating)_{it}$$
$$+\beta_9 \log(latest\_rating)_{it} + \beta_{10} weekend_t + \beta_{11} holiday_t + \beta_{12} promotion_{it} + v_i + \eta_{it}$$

5 RESULTS AND DISCUSSION

5.1 Study I: Effects of Administering Product Sampling on Sales

The estimation result is shown on table 2. Model 1 includes all products. Model 2 and model 3 investigate the search products and experience products respectively.

Table 2. Estimated Results of First Model

Variable	(a) All products	(b) Search products	(c) Experience products
before_after	0.025 (0.027)	0.044 (0.060)	0.022 (0.031)
treatment	-0.614*** (0.090)	-0.975*** (0.107)	-0.184 (0.123)
treatmentXafter	0.342*** (0.106)	0.084 (0.126)	0.410** (0.121)
log(*dailysales*)$_{i,t-1}$	0.545*** (0.037)	0.451*** (0.070)	0.562*** (0.040)
log(*product_price*)$_i$	-0.252*** (0.017)	0.122*** (0.017)	-0.083*** (0.007)
bussiness_unit	0.222*** (0.063)	0.296*** (0.042)	1.171*** (0.138)
holiday	-0.180*** (0.040)	-0.154** (0.078)	-0.194*** (0.047)
weekend	-0.114*** (0.023)	-0.133** (0.030)	-0.112** (0.027)
product_id	Included	Included	Included
Obs.	7299	1458	5841
Note: *: p <0.10, **: p<0.05, ***: p <0.01; robust standard error in parentheses			

From Table 1, we learn the coefficient of the interaction term for all product is positive, suggesting overall the sampling campaign works. If we separately examine this effect for experience product and search product, the coefficient is positive and significant for experience products but not for search products, we can conclude that the administration of product sampling online significantly influenced the sales of experience products but not for search products (i.e., H1(a) and H1(b) were supported). The result of online product sampling is consistent with traditional sampling literature that sampling could significantly improve the short term sales (Bawa et al, 2004).

5.2 Study II: Effects of Incorporating Product Sampling Reviews

The simultaneous equation (2) and (3) satisfy both order condition and rank condition, indicating this model is identified (Wooldridge, 2010). We employ three-stage least squares method (3SLS) to estimate this simultaneous model.

The estimated results of simultaneous equations are reported in Table 3. In column (a), the coefficient of *timeXsampling_review* is significant for sale_num, which implies that the incorporation of product sampling reviews could significantly increase the product sales. As for review number, we can observe that *timeXsampling_review* is significant but negative, which implies incorporating product sampling reviews on the platform decreases the posted review number. It could due to the content of sampling review is very rich, covers all the perspectives of purchase. The buyer who read this sampling review may hardly have any new opinion need to be complement, so they are less likely to post their review.

Table 2. Estimated Results of Simultaneous Equations

	(a) All products		(b) Search products		© Experience products	
Variables	Eq. (2)	Eq. (3)	Eq. (2)	Eq. (3)	Eq. (2)	Eq. (3)
time	0.171 (0.285)	-0.073 (0.064)	0.685 (0.509)	-0.138 (0.098)	0.104 (0.347)	-0.097 (0.084)
sampling_review	0.739 (0.550)	-0.214* (0.125)	-1.033 (3.647)	0.702 (0.497)	0.406 (0.547)	-0.159 (0.132)
timeXsampling_review	0.623* (0.350)	-0.238*** (0.078)	-1.267 (0.799)	-0.344** (0.149)	0.777* (0.401)	-0.302*** (0.097)
log(*price*)	4.177 (3.093)	-2.171*** (0.650)	-13.845 (10.251)	4.346** (1.769)	4.705 (3.522)	-3.159*** (0.725)
log(*cumulative_rating*)	-14.046 (9.926)	6.638*** (2.144)	43.560 (30.949)	-13.645*** (5.139)	-16.112 (11.089)	9.783*** (2.392)
log(*latest_rating*)	-0.222 (0.681)	0.336** (0.136)	-1.015 (1.258)	0.229 (0.257)	0.121 (0.767)	0.314** (0.159)
promotion	-0.179 (0.226)	0.141*** (0.046)	0.154 (0.535)	-0.049 (0.108)	-0.190 (0.247)	0.174*** (0.051)
product_id	Included	Included	Included	Included	Included	Included
Control variables	Omitted	Omitted	Omitted	Omitted	Omitted	Omitted
Obs.	2177	2177	527	527	1650	1650

We categorized products from DAS-II into search products (e.g., mobile phones, electrical appliances) and experience products (e.g., food items, cosmetic products) and separately examined the effects of incorporating product sampling reviews. From column (b) and (c), we can see that this intervention has different effects on sales across search products and experience products. For experience products, the incorporation of product sampling reviews positively affected the sales ratings at $p<0.01$ level. However, this effect vanished in the case of search products (i.e., H2(a) and H2(b) were supported).

6 DISCUSSION and CONCLUSION

In this study, we attempt to provide a better understanding of applying product sampling in the online context, which seems a viable strategy given the Internet’s capabilities to reach a wide audience, track consumer responses, and acquire and exploit information generated from product sampling. Indeed our empirical findings suggest that administering product sampling online can promote sales for both search and experience products. Additionally, the incorporation of product sampling reviews into the regular consumer reviews can stimulate accumulated ratings, but only for experience products. This insight has important implications for managers contemplating to administer product sampling online- it would be more worthwhile to adopt this promotional tactic when their products are of experience type. Also our findings extend the extant research on product sampling in the offline context that does not emphasize product differences in terms of search vs. experience (e.g., Bawa and Showmaker 2004; Gedenk and Neslin 1999; Heiman et al. 2001; Lammers 1991).

References

Omitted

User-Generated Content and Product Design of Competing Firms

Young Kwark
University of Florida
young.kwark@warrington.ufl.edu

Jianqing Chen
University of Texas at Dallas
chenjq@utdallas.edu

Srinivasan Raghunathan
University of Texas at Dallas
sraghu@utdallas.edu

Abstract

This paper studies the effect of user-generated content on competing firms' product design decisions when they face uncertain consumer taste. User-generated content makes firms better informed about consumer taste because of the additional information revealed, and it also makes their information more correlated because user-generated content is common and accessible by all firms. We build a game theoretical model in which two competing firms design and sell horizontally differentiated products to a group of consumers with unknown taste. Each firm observes a signal regarding consumer taste. We find that the variance-reducing effect embedded in user-generated content hurts the firms, and the correlation-increasing effect benefits the firms. The former is because uncertainty in the consumer taste makes firms unsure about where the demand is and thus can soften the competition between them. The reduced uncertainty by user-generated content intensifies the competition. The latter is because more correlated information can reduce the instances in which intended differentiation realizes similar product designs because the blindness of each other's information, and thus more correlated information can help the firms better "coordinate" their product positioning. Whether firms lose or benefit because of user-generated content depends on the interaction between the two effects. When the correlation-increasing effect dominates, the firms benefit from the user-generated content. In this case, the expected product differentiation is larger, the expected prices are higher, but consumer surplus is lower. When the two effects are comparable, a "win-win" result might arise in which both firms and consumers benefit from user-generated content.

Keywords: User-Generated Content, Product Design, Competition

1. Introduction

With the help of different Web 2.0 applications (e.g., blogs, online product reviews, and online crowdsourcing platforms), user-generated content has been fast-growing and has become a significant part of the online world. Users discuss and comment on current products or services, and also express their wish in new products for the future. Many firms now follow users' conversation on the Internet, listen to their opinions, and try to incorporate users' ideas to their business. For example, many firms (e.g., Nike) have their Facebook company pages to listen to and interact with their fans. Firms (e.g., Southwest Airline) have also been following the discussion across different social media platforms such as Twitter. More directly, in the past few years firms launched different crowdsourcing platforms to seek for consumer input, ideas, and advice on how to improve their products or services. Examples include Dell's IdeaStorm, Starbucks' myStarbucksIdea, Salesforce's customer community, and BestBuy's Idea Exchange. Starbucks founded the open environment, myStarbucksIdea, to hear consumers' ideas for the new product development in 2008 and several recent Starbucks products, such as Mocha Coconut Frappuccino, are the examples. On IdeaStorm.com, there have been "over 18,971 ideas submitted" and Dell has made "535+ ideas implemented" since its launch in 2007. Evidently, many firms listen to online users and are interested in getting ideas from user-generated content for their business, including their product design. On the one hand, the massive amount of user-generated content provides more information about what consumers really want, and thus it makes firms better informed about consumer taste. On the other hand, user-generated content makes firms' information about consumer taste more correlated, because user-generated content is common and accessible by all firms. This paper studies the effect of user-generated content on competing firms' product design decisions when they face uncertain consumer taste.

2. Model

We consider two firms, each selling one product. The products are imperfect substitutes, and the two firms compete with each other. We call the firm that sells product i firm i, $i \in \{1,2\}$. The marginal production cost for each product is assumed to be zero. We use Hotelling's horizontal differentiation framework to capture consumer preference. Consumers are uniformly distributed along the line $[M-1/2, M+1/2]$. Product i is located at x_i. The distance between a consumer and a product measures the degree of misfit of the product to the consumer: the longer the distance, the greater the degree of misft. We denote V as the surplus enjoyed by each consumer when he consumes an ideal product with perfect fit. In general, the products have some degrees of misfit to the consumers. For the consumer located at x, we assume that the misfit cost is $t(x-x_i)^2$ when consuming product i. Denoting the price as p_i, we can formulate the utility of the consumer located at x from consuming product i as $U_i = V - t(x-x_i)^2 - p_i$. Based on this formulation, we find the indifferent consumer to be located at $\xi = [p_2 - p_1 + (x_2^2 - x_1^2)t]/[2(x_2 - x_1)t]$ who derive the same utility from consuming either product and thus is indifferent between purchasing product 1 or 2.

We assume that each consumer has a unit demand. If we assume $x_1 < x_2$ and $\xi \in [M-1/2, M+1/2]$, we can formulate the demand for each product as $D_1 = 1/2 + (\xi - M)$ and $D_2 = 1/2 - (\xi - M)$. We also assume that the surplus V is large, relative to the misfit costs and price, such that the market is fully covered.

Different from the standard horizontal differentiation models, firms are uncertain about the consumer space, M. Traditionally (in the absence of user-generated content), firms conduct their own market research regarding the demand before their product design decisions. We model the information that a firm has about the consumer space as the firm observes a signal regarding M. User-generated content provides additional and public information for both firms. In the presence of user-generated content, first, each firm has a better understanding of the consumer space because of the additional information revealed by user-generated content, and thus the signal each firm receives has a better precision. Second, because this information revealed is public and common to both firms, user-generated content increases the correlation between the two firms' signals. Formally, we assume that firm i observes a signal s_i, with mean M and variance σ^2, where M is a unknown constant. We assume that two signals are jointly normal, as commonly used in the literature. In the absence of user-generated content, we assume the two signals are independent and each is with variance σ_{wo}^2. In the presence of user-generated content, the variance for each is σ_w^2, $\sigma_w^2 < \sigma_{wo}^2$, and the correlation coefficient between them is ρ, $\rho \geq 0$. Technically, we assume that $\sigma^2 \leq [18 - 3\sqrt{36 - 2(1+\rho)}]/[2(1+\rho)]$ such that in equilibrium $x_1 < x_2$ and $\xi \in [M - 1/2, M + 1/2]$.

The sequence of events is as follows. In stage 1, the firms determine the positions of their products x_i, based on the information they have about consumer space. In stage 2, the demand uncertainty is resolved, and then the firms set prices p_i simultaneously. In stage 3, consumers make their purchase decisions and demand is realized. We consider two scenarios: one without user-generated content and the other with user-generated content. We use the first scenario as the benchmark to analyze the effect of user-generated content on the product design decision. Firms' own signals about demand and consumers' misfits are their private information. All other model parameters are common knowledge. All players are risk neutral.

3. Effect of User Generated Content on Product Design

In this section, we first explain how we derive the Bayesian Nash equilibrium, following the approach of backward induction. We then analyze the effects of user-generated content on the competing firms' product design decisions and payoffs, by comparing their equilibrium outcome with those without user-generated content.

In stage 2 of the game, after observing the consumer space, the firms maximize their profits by choosing the optimal prices for their products. By the first-order conditions, we can derive the firms' optimal prices, which are functions of the product design choices x_i. Substituting the optimal prices to the firms' profit functions, we obtain the firms' profit as functions of the location choices x_i and M. In stage 1 of the game, the firms do not know M and face uncertainty regarding the demand. Firm i observes his own private signal about consumer space, s_i, but does not observe his competitor's signal. We conjecture that each firm's equilibrium product design strategy is to choose a location linear in his signal regarding M; that is, $x_i = s_i + b_i$, where b_i is a constant to be determined. Based on this conjecture, we can derive the firms' optimal product design strategies and verify such a conjecture is correct. Substituting the firms' optimal locations back to their prices and profits, we can derive the equilibrium prices and profits. Ex ante, we do not know what signals that the firms will receive. Considering all possible signals that firms might receive, we can derive the firms' expected equilibrium locations, prices, and profits, as well as consumer surplus and social welfare.

We next examine the effect of uncertainty regarding the demand on firms' product design decisions and firms' profits. For the comparison purpose, we next use the regular notations (e.g., π_i) for the scenario without uncertainty and use the notations with hats (e.g., $\hat{\pi}_i$) for the scenario uncertainty.

Proposition 1 *Compared to the case without uncertainty, in the presence of uncertainty:*
(a) Product differentiation is higher; that is, $\mathbb{E}(x_2 - x_1) < \mathbb{E}(\hat{x}_2 - \hat{x}_1)$*;*
(b) Prices are higher; that is, $\mathbb{E}(p_i) < \mathbb{E}(\hat{p}_i)$*;*

(c) Firm profits are higher; that is, $\mathbb{E}(\pi_i) < \mathbb{E}(\hat{\pi}_i)$ *;*

(d) Consumer surplus is lower; that is, $\mathbb{E}(CS) > \mathbb{E}(\hat{C}S)$ *;*

(e) Social welfare is lower; that is, $\mathbb{E}(SW) > \mathbb{E}(\hat{S}W)$.

As in the standard "location-then-price" games, in choosing their locations, the firms tradeoff two effects: the demand effect and the strategic effect. In the case without uncertainty (such that the firms know exactly the consumer space), on the one hand, each firm wants to move toward the center M to increase his demand, which is called the demand effect. On the other hand, moving toward the center could increase the competition with his rival and decrease the prices, which is called the strategic effect. In balancing the two effects, in equilibrium one firm choose to locate at $M - 3/4$ and the other at $M + 3/4$. In the presence of uncertainty, the two effects continue to exist in the firms' location decision making process. However, the impact of the uncertainty on the two effects is asymmetric. The strategic effect works similarly as in the case without uncertainty, because, in stage 2, the uncertainty is resolved and differentiating from each other softens the price competition in each realization. The demand effect, in contrast, is weakened by the uncertainty, because, in the presence of the uncertainty, the firms are not sure about the location of M and the realizations in which M turns out to be close to a firm's location itself could give the firm a great demand. Because of the weakened demand effect in the tradeoff, in the presence of uncertainty, the firms choose to differentiate further from each other, compared to the case without uncertainty.

Because of the increased differentiation, the competition is softened and the firms charge higher prices and earn higher profits in equilibrium. On average consumers are hurt because of the increased prices and increased misfit costs resulting from the increased differentiation. Social welfare is also hurt by the uncertainty. Notice even in the case without uncertainty, the competition between the two firms leads to excess differentiation. To minimize the total misfit costs, a social planner would choose to locate the two firms equidistantly on either side of the middle of the segment; that is, one at $M - 1/4$ and the other at $M + 1/4$. But, the competition equilibrium leads to the structure with one firm at $M - 3/4$ and the other at $M + 3/4$. So the equilibrium differentiation in the case without uncertainty is excess. In the presence of uncertainty, the firms are even further differentiated, which hurts the social welfare because of larger misfit costs.

In the absence of user-generated content, we assume that firms conduct their own research and each firm uses his own private information independent with the competitor's; that is, $\rho = 0$. As user-generated content become commonly available to firms, it provides the same additional information to both firms. As a result, firms' estimate about consumer taste is more accurate, and meanwhile their estimates are more correlated; that is, $\sigma_w^2 < \sigma_{wo}^2$ and $\rho \geq 0$. Next we examine the effect of user-generated content on firms' product design decisions and profits. We use the notation with subscript "wo" and "w" for the case without and with user-generated content (e.g., $\hat{\pi}_{i,wo}$ and $\hat{\pi}_{i,w}$).

Proposition 2 *Compared to the case without user-generated content, in the presence of user-generated content (with* $\sigma^2 = \sigma_w^2 < \sigma_{wo}^2$ *and* $\rho > 0$*):*

(a) Product differentiation is higher (i.e., $\mathbb{E}(\hat{x}_{2,wo} - \hat{x}_{1,wo}) < \mathbb{E}(\hat{x}_{2,w} - \hat{x}_{1,w})$ *) if and only if* $\rho > (\sigma_{wo}^2 - \sigma_w^2)/\sigma_w^2$ *;*

(b) Prices are higher (i.e., $\mathbb{E}(\hat{p}_{i,wo}) < \mathbb{E}(\hat{p}_{i,w})$*) if and only if* $\rho > (\sigma_{wo}^2 - \sigma_w^2)/\sigma_w^2$ *;*

(c) Firm profits are higher (i.e., $\mathbb{E}(\hat{\pi}_{i,wo}) < \mathbb{E}(\hat{\pi}_{i,w})$*) if and only if* $\rho > (\sigma_{wo}^2 - \sigma_w^2)/\sigma_w^2$ *;*

(d) Consumer surplus is lower(i.e., $\mathbb{E}(\hat{CS}_{wo}) > \mathbb{E}(\hat{CS}_w)$*) if and only if*

$$\rho > \frac{\sqrt{2025\sigma_w^4 - 432\sigma_w^6 + 612\sigma_w^4\sigma_{wo}^2 + 4\sigma_w^4\sigma_{wo}^4}}{2\sigma_w^4} - \frac{45 + 2\sigma_w^2}{2\sigma_w^2}$$

(e) Social welfare is higher (i.e., $\mathbb{E}(\hat{SW}_{wo}) < \mathbb{E}(\hat{SW}_w)$*).*

The effects of user-generated content on firms' are two-fold. First, user-generated content provides additional information for each firm, and thus firms are better informed and their estimates of consumer taste are more accurate. We call this variance-reducing effect. Second, user-generated content is public and common, and thus it adds a common information resource to each firm's private information set. As a result, the firms' signals are more correlated because of user-generated content. We call this correlation-increasing effect.

The variance-reducing effect embedded in user-generated content tends to induce firms to less differentiate from each other and intensify the competition. As we explained for Proposition 1, with the uncertainty, because firms are not sure where the demand is, they has less incentive to move toward his rival to compete for the demand; in other words, the demand effect aforementioned is lessened. Because user-generated content reduces the uncertainty in the demand, it induces firms to compete more aggressively and position their products closer to each other, which hurts the firms.

The correlation-increasing effect, in contrast, can benefit the firms. When the firms' signals are positively correlated, each firm can infer the rival firm's signal, based on his own observed signal. The extreme case is that with perfect correlation each firm learns his competitor's signal exactly. When firms know each other's signal better, the strategic effect aforementioned works better and plays a more salient role in the tradeoff of positioning his product. Being blind to each other's signal (in the case with independent signals, as in the case without user-generated content), firms' intended differentiation could result in similar product designs in some realization, for example, in cases in which firm 1 receives a high s_1 and firm 2 receives a low s_2 (noticing firms' product design strategies: $x_1 = s_1 - 3/4 - (1+\rho)\sigma^2/6$ and $x_2 = s_2 + 3/4 + (1+\rho)\sigma^2/6$). Therefore, the more correlated information can help the firms better "coordinate" their product positioning.
Whether firms lose or benefit because of user-generated content depends on the interaction between the two effects. When the correlation-increasing effect dominates, the firms benefit from the user-generated content. In this case, the expected product differentiation is larger and the expected prices are higher. The effect of user-generated content on consumer surplus is similar.

The real benefit brought by user-generated content is realized in social welfare. Recall that the competition in general leads to excess differentiation. In the presence of user-generated content, the variance-reducing effect induces firms to compete aggressively and thus reduces the differentiation, which benefits social welfare. The correlation-increasing effect can reduces the variance of the difference between the two product designs, which also benefits social. As a result, both effects can help increase social welfare and the social welfare is always higher in the presence of user-generated content.

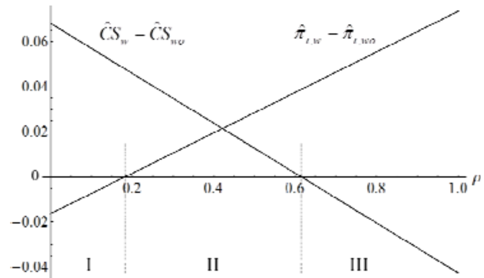


Figure 1: Firm Profit and Consumer Surplus
($\sigma_w = 0.23$, $\sigma_{wo} = 0.25$, $t = 5$, and $V = 20$)

We next delineate how the firm's profits and consumer surplus in the presence of user-generated content change with the correlation, compared with the case without user-generated content, as in Figure 1. In the first range (I) with sufficiently low correlation, firms lose while consumers benefit in the presence of user-generated content because the variance reduction in firm's signal plays a dominating role. In contrast, the third range of ρ (III) shows the opposite because the increased correlation plays a dominating role. In the second range (II), in which the firm signals are moderately correlated, both firms and consumers can be better off. We summarize this result as follows.

Corollary 1 *In the presence of user-generated content, both firms and consumers are better off when*

$$\rho \in \left[\frac{\sigma_{wo}^2 - \sigma_w^2}{\sigma_w^2}, \frac{\sqrt{2025\sigma_w^4 - 432\sigma_w^6 + 612\sigma_w^4\sigma_{wo}^2 + 4\sigma_w^4\sigma_{wo}^4}}{2\sigma_w^4} - \frac{45 + 2\sigma_w^2}{2\sigma_w^2} \right]$$

Corollary 1 suggests that user-generated content in our setting may lead to a win-win result in equilibrium; that is, both firms and consumers can be better off in the presence of user-generated content if the correlation between firms' signals is moderate. This win-win result is possible because the social welfare generated is higher or the total “pie” is bigger in the presence of user-generated content.

4. Conclusion

We examine the effect of online user-generated content on the competing firms' product design decisions when they sell substitutable products and face uncertain consumer taste. Because user-generated content provides additional information about consumers and is common to the firms, it reduces firms' uncertainties about consumers and increases the correlation of firms' information. We demonstrate that uncertainty reduction can hurt firms because reduced uncertainty intensifies the competition, while the increased correlation can benefit them because one firm can better infer the other's signal and thus better ``coordinate'' their product positioning.

Cardinality bundles with Constrained Prices

Karthik N. Kannan[*], Mohit Tawarmalani[*], and Jianqing "Fisher" Wu[*]

[*]Krannert School of Management, Purdue University

November 17, 2014

Abstract

We study the pricing of cardinality bundles (CB), where firms set prices that depend only on the size of the purchased bundle. The existing analytical framework lacks sub-additivity constraints on bundle pricing, which limits its application in reality. In this study, we solve the CB problem with additional constraints on bundle prices. We first study the CB problem with marginal decreasing prices and prove that it is a shortest-path problem. Second, we propose a dynamic programming algorithm to solve the CB problem with unit decreasing prices. Third, we analyze the CB problem with sub-additive prices and convert its MINLP formulation to a mixed-integer programming (MIP) one. Finally, we provide analytical and numerical analysis on the gaps between different CB models.

1 Introduction

This paper studies a bundling schemed called cardinality bundling (CB). in CB, sellers price for the number of goods and let consumers choose with specific products they what. Pricing for toppings of pizza is a simple example of CB. In many pizza stores, consumers are priced for the number of toppings regardless of the specific topping types. Similarly, Disney World uses CB to sell theme park tickets. Instead of selling tickets for each park separately, Disney World prices consumers for the number of visits to all its theme parks. More generally, information goods providers such as Netflix and Blockbuster, telecommunication service providers such as AT&T, and cable TV providers such as Eastlink, are also implementing CB in selling their products or services.

We next briefly review the literature on CB. Hitt and Chen (2005) develop the first cardinality bundling model. They discover some properties for the optimal solution of cardinality bundling problem with assuming that consumers' willingness-to-pays follow Spence-Mirrlees Single Crossing Property (SCP). Chu et al. (2011) show that the profitability of CB is more than component pricing and pure bundling, and is close to that of mixed bundling by using computational and empirical approaches. Wu et al. (2014) analytically studies the optimal pricing strategies for CB problems with SCP consumer valuations. They show that the optimal prices to the problem can be obtained, in strongly polynomial time, by solving a shortest-path problem. Based on the network structure underlying the shortest path formulation, they develop an algorithm to solve the quantity-discount problem proposed by Spence (1980). Lastly, they also study the characteristics of the underlying problem that lead to similar strongly polynomial time solution approaches.

The models in Hitt and Chen (2005) and Wu et al. (2014) assume that each consumer can only purchase no more than one bundle. This assumption is valid in some industries. For example, each home usually has no more than one cable TV connection and therefore is only able to purchase at most one cable TV bundle. Other examples include toppings of pizza and cellular data pricing plans. However, in some other industries, consumers are not restricted to only purchase one bundle. For example, consumers can easily purchase multiple bundles of songs at on-line music stores. As a result, the insights obtained by these works do not necessarily extend to situations when the consumers may purchase more than one bundle. In this paper, we relax the one bundle per consumer assumption. We introduce sub-additive constraints on bundle prices to ensure that the consumer incentive compatibility is not violated even if consumers are allowed to purchase more than one bundle of goods.

In reality, three main types of sub-additive price schemes are used in different industries. (1) Marginal decreasing prices (MDP) where the marginal price of each additional unit is weakly decreasing, which is also known as multiple-part tariff pricing (Wilson, 1993). (2) Unit decreasing prices where the unit price of each bundle is weakly decreasing. Since this type of price scheme is first introduced by Chu et al. (2011) as bundle-size pricing (BSP), we also call it BSP in this paper. (3) General form of sub-additive prices (CBSP) where the price of any bundle is no less than the total price of any two other bundles which can together form the previous one. In this paper, we study these various kinds of CB problems with different constraints on bundle prices. In order to get tractable and meaningful results, we additionally assume Spence-Mirrlees Single Crossing Property (SCP) on consumers' reservation price. We first develop a shortest-path solution approach for MDP. Second, we propose a dynamic programming algorithm to solve BSP. Third, we analyze the CB problem with sub-additive prices and convert its MINLP formulation to a mixed-integer programming (MIP) one. Finally, we provide analytical and numerical analysis on the gaps between different CB models.

2 Marginal Decreasing Prices (MDP)

2.1 Model

In this section, we consider the cardinality bundling problem with marginal decreasing prices. The model is built upon that in Wu et al. (2014) and we review it here for the sake of completeness. In this model, a seller implements cardinality bundling to sell his goods and seeks a optimal price scheme for each bundle size to maximize his profit. Each consumer makes her purchase decision to maximize her consumer surplus. If a consumer gets negative surplus from all the bundle sizes, she will purchase nothing.

In this section, the seller imposes marginal decreasing prices, or MDP, to insure that for each additional unit, the marginal price is no more than that of the previous unit. It is straightforward that if the marginal price for each additional unit is weakly decreasing, than the price of any bundle will always be weakly less than the total price of any other two smaller-sized bundles which can form the previous one. As a result, any rational consumer will never purchase more than one bundle.

Let $i \in \{1, \ldots, I\}$ denote consumer indexes and $j \in \{0, \ldots, J\}$ denote bundle size indexes. Notice, Bundle 0 is included in the model to represent consumer's choice of purchasing nothing. Let p_j and c_j denote the price and cost of Bundle j. Let $w_{ij} \geq 0$ be the willingness-to-pay of Consumer i for Bundle j. Let x_{ij} be a binary variable indicating whether Consumer i purchases Bundle j. Then, MDP can be formulated as follows:

$$
\begin{aligned}
\text{MDP1}: \quad \underset{x_{ij},p_j}{\text{Max}} \quad & \sum_{i=1}^{I}\sum_{j=0}^{J} x_{ij}(p_j - c_j) \\
\text{s.t.} \quad & \sum_{j'=0}^{J}(w_{ij'} - p_{j'})x_{ij'} \geq w_{ij} - p_j \;\forall i,\; \forall j && (1)\\
& p_0 = 0 && (2)\\
& p_j - p_{j-1} \leq p_{j-1} - p_{j-2} \;\forall j \geq 2 && (3)\\
& \sum_{j=0}^{J} x_{ij} = 1 \;\forall i && (4)\\
& x_{ij} \in \{0,1\} \;\forall i,\; \forall j. && (5)
\end{aligned}
$$

Following Hitt and Chen (2005) and Wu et al. (2014), we also assume that consumers' WTP follows the Spence-Mirrlees Single Crossing Property (SCP):

$$
\begin{aligned}
w_{ij} &\geq w_{i'j} \;\forall i > i', && (6)\\
w_{ij} - w_{ij'} &\geq w_{i'j} - w_{i'j'} \;\forall i > i',\; \forall j > j'. && (7)
\end{aligned}
$$

2.2 Properties of the Optimal Solution

First, we identify some properties of the optimal solution.[1] Let $w_{ij}^m = w_{ij} - w_{i,j-1} \;\forall i \;\forall j \geq 1$ and $w_{i0}^m = 0$ be the marginal WTP of each consumer i for each additional unit of goods j. Let p_j^m be the marginal price for each unit of goods j. Similarly, let c_j^m be the marginal cost for each unit of goods j.

Proposition 1 *There exists an optimal solution to* MDP1 *that satisfies:*

$$
\sum_{j'=j}^{J} x_{i+1j'} \geq \sum_{j'=j}^{J} x_{i,j'} \; i = 1, \ldots, I-1,\; \forall j. \qquad (8)
$$

[1] All the proofs are omitted.

Proposition 2 *There exists an optimal pricing scheme such that if two bundle sizes j and j' are bought by some consumers and $j' > j$ then $p_{j'} - c_{j'} > p_j - c_j$.*

Lemma 3 *Among the consumers purchasing a non-zero bundle size, the lowest indexed one is charged at her WTP in every optimal solution.*

Proposition 1, 2, and Lemma 3 are proved in Wu et al. (2014) when there is no constraints on bundle prices. In this paper, we show that in CB models such as MDP, BSP, or CBSP, where the bundle prices are constrained with various kinds of conditions, Proposition 1, 2, and Lemma 3 are still valid and therefore can be useful to derive solution approaches for these problems.

In $\mathrm{MDP1}$, we have an additional property that plays a critical roll to solve the problem.

Proposition 4 *In the optimal solution, p_j^{m*}, the optimal marginal price for any unit j, satisfies the following condition:*

$$p_1^{m*} \in \{w_{11}^m, \cdots, w_{I1}^m\},\ p_j^{m*} \in \{w_{1j}^m, \cdots, w_{i'j}^m, p_{j-1}^{m*}\} \forall j \geq 2,$$

where $i' = \arg\mathrm{Max}_i\{w_{ij}^m \leq p_{j-1}^{m*}\}$. That is, the optimal marginal price for any unit j is priced at the same marginal price as Bundle $j-1$, or at some consumer's marginal WTP on Bundle j that is no more than p_{j-1}^{m*}.

By Proposition 4, we know that $p_1^{m*} \in \{w_{11}^m, \cdots, w_{i'1}^m\}$, or, the marginal price of the first unit is priced at one consumer's marginal WTP for the first unit. Next, we can easily get $p_2^{m*} \in \{w_{11}^m, \cdots, w_{I1}^m, w_{12}^m, \cdots, w_{I2}^m\}$, or, the marginal price of the first unit is priced at one consumer's marginal WTP for the first two units. Recursively, we have $p_j^{m*} \in \{w_{11}^m, \cdots, w_{I1}^m, w_{12}^m, \cdots, w_{I2}^m, \cdots, w_{1j}^m, \cdots, w_{Ij}^m\}$.

Proposition 5 *Let $v_{jij'} = (I - i' + 1)(w_{ij'}^m - c_j^m)$ where $i' = \arg\mathrm{Min}_{i''}\{w_{i''j}^m \geq w_{ij}^m\}$. $v_{jij'}$ capture how the total profit will change if the marginal price of Bundle j is priced at the marginal WTP of Consumer i for Bundle j'.*

Next, we convert MDP1 to a shortest path problem. Let $x_{jij'}, j' \leq j$ be a binary variable to indicate whether the marginal price of Bundle j is priced at the marginal WTP of Consumer i for Bundle j'. Let $chi_{jij'}^{\tilde{i}\tilde{j}'}, j' \leq j, \tilde{j}' \leq j-1$ be a binary variable to indicate whether the marginal price of Bundle j is priced at $w_{ij'}^m$ while that of Bundle $j-1$ is priced at $w_{\tilde{i}\tilde{j}'}^m$. Let $v_{jij'}^{\tilde{i}\tilde{j}'}, j' \leq j, \tilde{j}' \leq j-1$ captures how the total profit will change if $\chi_{jij'}^{\tilde{i}\tilde{j}'} = 1$.

$$v_{jij'}^{\tilde{i}\tilde{j}'} = \begin{cases} 0 & \text{if } w_{ij'}^m = w_{\tilde{i}\tilde{j}'}^m \\ 0 & \text{if } w_{ij'}^m \leq w_{\tilde{i}\tilde{j}'}^m \text{ and } j' = j \\ -\sum_{i=1}^{I} w_{iJ} & \text{otherwise.} \end{cases} \tag{9}$$

Then we can reformulate MDP1 to the following shortest-path problem:

Theorem 6 MDP1 *is equivalent to the following shortest path problem on a graph which has* $(I+1)I(J+1)+2$ *nodes and no more than* $I^3(J+1)^2/2$ *edges:*

$$\text{MDP2}: \quad \underset{x_{jij'}, \chi_{jij'}^{\tilde{i}\tilde{j}'}}{Min} \quad -\sum_{j=0}^{J}\sum_{i=1}^{I}\sum_{j'=1}^{j} v_{jij'}x_{jij'} - \sum_{j=0}^{J}\sum_{\tilde{i}=1}^{I}\sum_{\tilde{j}'=1}^{j-1}\sum_{i=1}^{I}\sum_{j'=1}^{j} v_{jij'}^{\tilde{i}\tilde{j}'}\chi_{jij'}^{\tilde{i}\tilde{j}'}$$

$$s.t. \quad \sum_{i=1}^{I} x_{1i1} = 1 \tag{10}$$

$$\sum_{j'=1}^{J}\sum_{i=1}^{I} x_{Jij'} = 1 \tag{11}$$

$$\sum_{\tilde{i}=1}^{I}\sum_{\tilde{j}'=1}^{j-1} \chi_{jij'}^{\tilde{i}\tilde{j}'} = x_{jij'}, \ \forall i \ \forall j \ \forall j' \leq j \tag{12}$$

$$x_{j\tilde{i}\tilde{j}'} = \sum_{i=1}^{I}\sum_{j'=1}^{j} \chi_{jij'}^{\tilde{i}\tilde{j}'}, \ \forall \tilde{i} \ \forall j \ \forall \tilde{j}' \leq j-1 \tag{13}$$

$$\chi_{jij'}^{\tilde{i}\tilde{j}'} \in \{0,1\} \ \forall i \ \forall j \ \forall j' \leq j \ \forall \tilde{i} \ \forall \tilde{j}' \leq j-1. \tag{14}$$

3 Unit Decreasing Prices (BSP)

Instead of imposing the marginal decreasing prices, another way to solving the problem without the single bundle restriction on the consumer is to impose a non-increasing unit price constraint on the prices set by the vendor. In such a case, naturally, no consumer will have an incentive to buy more than one bundle to form the bundle of her desired size. Chu et al. (2011) considers such a restriction, $\frac{p_j}{j} \geq \frac{p_{j+1}}{(j+1)}\ \forall j \leq J-1$, in their cardinality bundling formulation and refer to it as the Bundle-Size Pricing (BSP). The vendor's decision problem is then

$$
\begin{aligned}
\text{BSP1}: \quad \underset{x_{ij},p_j}{\text{Max}} \quad & \sum_{i=1}^{I}\sum_{j=0}^{J} x_{ij}(p_j - c_j) \\
\text{s.t.} \quad & (1), (2), (4), (5), \\
& p_j/j \geq p_{j+1}/(j+1)\ \forall j \leq J-1 \qquad (15)
\end{aligned}
$$

The non-increasing unit price constraint is specified in Equation (15). Because of this constraint, the problem does not retain the structure of the shortest-path problem for MDP1. Notice, Proposition 1, 2, and Lemma 3 are still valid under BSP1 but Proposition 4 is not valid anymore. We next develop some new properties for BSP1. From now on, we relax the concavity assumption on consumers' WTP.

Proposition 7 *For a given price scheme, assume that Bundle j is purchased by some consumer(s). Also assume that $p_{j+1}, p_{j+2}, \cdots, p_J$ are all high enough so that no consumer purchase any bundle size greater than j. If we reduce p_{j+1} to a certain level such that some consumer change to purchase p_{j+1}, then this consumer purchases Bundle j before p_{j+1} is changed.*

Next, we develop a unit-price based dynamic programming algorithm for solving BSP_1 when the costs are separable in bundle sizes. In this algorithm, the unit prices can only take discrete values. The feasible set of unit prices correspond to a grid of length ϵ. There are total K points on the grid. K is determined by $K = W_{IJ}/\epsilon$, where W_{IJ} is last consumer's willingness-to-pay for the largest bundle size. We use the variable $k = 1$ to denote the individual grid points and u_k as the corresponding unit price. largest bundle size. We use $k = 0, 1, \ldots, K$ for grid step index and u_k for unit price on grid step k.

According to the definition of BSP, all bundle sizes are available in the market and the unit price of each bundle is no more than a smaller-sized bundle. Our algorithm start with finding out the maximum total profit when the unit price of Bundle 1 is priced at u_k and the unit price of any other larger-sized bundle is also priced at u_k. This situation is same as providing all bundles with the same unit price u_k. For each grid index k, we can easily find out which consumer i is the lowest type consumer starting to purchase and how many units she want to purchase according to her WTP. Similarly, we can also find out how many units each other higher type consumer purchases and then get the total profit by the vendor. We denote this profit value as Π_{i1k}. More generally, let Π_{ijk} be the maximum total profit if bundle size j is the first one to be provided at unit price u_k (i.e., the unit price of any smaller-sized bundle is greater than u_k) and consumer i is the first one to start purchasing this bundle.

We have already show how to calculate $\Pi_{i1k}\ \forall i,\ \forall k$. We can then calculate $\Pi_{i2k}\ \forall i,\ \forall k$, based on Π_{i1k} results. We use a function $\Delta(i, 2, k, i', 1, k')$ to calculate the change in total profit for a reducing in unit price. It basically calculates which consumers will switch from purchasing Bundle 1 to Bundle 2 because of the availability of Bundle 2 and how many units each of these consumers purchase with the new unit price. Therefore, we have $\Pi_{i2k} = \max \Pi_{i'1k'} + \Delta(i, 2, k, i', 1, k')$. By using the same recursive logic, we can continue to calculate Π_{ijk} for any larger bundle size j as well and can finally find the optimal solution for the BSP problem.

The pseudo-code for the algorithm is shown in Algorithm 1.

Theorem 8 *When the costs are separable in bundle sizes, for any given total error ϵ_t, let the grid step length parameter be $\epsilon = 2\epsilon_t/(J+1)JI$. Then the algorithm in Algorithm 1 guarantees that the gap between the*

for $i, j; i <= I, j <= J$ **do**
  $u_0 = w_{ij}/j$;
  $\pi_{ijK} = \Pi(i, j)$;
  for $i_1, j_1, k_1; i \leq i_1 \leq I, j \leq j_1 \leq J, k_1 \leq K$ **do**
    for $i_2, j_2, k_2; i_2 \leq i_1, j_2 \leq j_1, k_2 \leq k_1$ **do**
      $\Pi_{temp} = \Pi_{i_2 j_2 k_2} + \Delta(i_2, j_2, k_2, i_1, j_1, k_1)$
      if $\Pi_{temp} \geq \Pi_{i_1 j_1 k_1}$ **then**
        $\Pi_{i_1 j_1 k_1} = \Pi_{temp}$
      end if
    end for
  end for
  if $\max_k\{\Pi_{IJk}\} > \Pi_{max}$ **then**
    $\Pi_{max} = \max_k\{\Pi_{IJk}\}$;
  end if
end for

Algorithm 1: A Dynamic Programming Algorithm for BSP

optimal profit and the solution generated by the algorithm is no more than ϵ_t. Moreover, the computation complexity is $O(I^3 J^4 K^2)$, where $K = W_{IJ}/\epsilon$.

4 Sub-Additive Price (CBSP)

In the previous section, we remove consumers' incentives to purchase more than one bundle by imposing the non-decreasing unit price constraint. However, in some cases it may be a more strict constraint than necessary. The following example illustrates that the non-decreasing unit price constraint may reduce vendor's profits.

Example 9 *A music store can offer a single song for \$4 each, and a bundle size 10 for \$10. If someone wants 11 songs, she needs to pay \$14 to get a bundle and a single song, which has a higher unit price than that of bundle size 10. Imposing the non-decreasing unit price constraint in this scenario will reduce vendor's profits.*

To overcome this issue, we propose a CBSP model, cardinality bundling problems with sub-additive prices, in this section. Formulating the CBSP problem is similar to BSP1, except replacing the non-increasing unit price Constraints (15) with the following price sub-additivity constraints:

$$p_j \leq p_{j'} + p_{j-j'} \ \forall j \ \forall j' < \frac{1}{2}(j+1)$$

Proposition 10 *Solutions to CBP and BSP are respectively the lower and upper bounds for CBSP.*

CBP is the CB problem without any constraints on bundle prices. It is easy to understand the rationale behind this result. On one hand, CBP is the same problem as CBSP except that the price sub-additivity constraints are relaxed. One the other hand, price constraints in BSP are stricter constraints than sub-additivity constraints in CBSP, leading to an underestimation of CBSP.

When the costs are separable, it is possible to create an MIP formulation. Notice that the nonlinearity of the objective function in CBSP comes from $x_{ij}p_j$. Therefore, we introduce $q_{ij} = x_{ij}p_j$ to replace all the nonlinear items. By adding Constraints (20) - (23), we can reformulate CBSP as an MIP:

$$\begin{aligned}
\text{CBSP1}: \quad \underset{x_{ij},q_{ij}}{\text{Max}} \quad & \sum_{i=1}^{I}\sum_{j=0}^{J} q_{ij} - x_{ij}c_j \\
\text{s.t.} \quad & \sum_{j=0}^{J} x_{ij} = 1 \ \forall i && (16) \\
& \sum_{j'=0}^{J} (w_{ij'}x_{ij'} - q_{ij'}) \geq w_{ij} - p_j \ \forall i,\ \forall j && (17) \\
& p_j \leq p_{j'} + p_{j-j'} \ \forall j \ \forall j' < \frac{1}{2}(j+1) && (18) \\
& p_j \leq p_{j+1} \ \forall j \leq J-1 && (19) \\
& q_{ij} \geq x_{ij}p_j^L \ \forall j && (20) \\
& q_{ij} \leq x_{ij}p_j^U \ \forall j && (21) \\
& q_{ij} \geq x_{ij}p_j^U + p_j - p_j^U \ \forall j && (22) \\
& q_{ij} \leq x_{ij}p_j^L + p_j - p_j^L \ \forall j. && (23)
\end{aligned}$$

Table 1: Comparison of CBP, CBSP, and BSP

		Optimal profit			Gap	
Problem No.	Problem size(I,J)	CBP	CBSP	BSP	CBP	BSP
1	20,20	152.384	149.884	148.679	1.67%	-0.80%
2	20,20	0.8	0.78	0.775	2.56%	-0.64%
3	20,20	16.199	16.123	16.119	0.47%	-0.02%
4	20,20	22.536	22.536	22.504	0.00%	-0.14%
5	20,20	39.435	39.014	38.902	1.08%	-0.29%
Average		46.271	45.667	45.396	1.16%	-0.38%

Here, p_j^L and p_j^U are upper and lower bound for each p_j. Constraints (20) - (23) ensure that if $x_{ij} = 0$, then $q_{ij} = 0$, and if $x_{ij} = 1$, then $q_{ij} = p_j$. Therefore, MIP formulation $CBSP_1$ always has the same solution as the MINLP CBSP problem.

5 Gap Analyses

We also numerically evaluated how well the three mechanisms compare when the costs are zero. Table 1 shows five numerical examples with 20 consumers and 20 bundles sizes. All consumers' WTP is randomly generated according to SCP. In Column three to five, optimal profits for CBP, CBSP, and BSP are shown. We can see that for all the problems, CBP optimal value is (weakly) greater than that of CBSP which is (weakly) greater than that of BSP. We observe that the gaps can be large when using CBP compared to BSP. To investigate this issue further, we have also theoretically analyzed the gap between CBP and CBSP, and that between BSP and CBSP when the costs are separable in bundle sizes. Let Π^*_{CBP}, Π^*_{BSP}, and Π^*_{CBSP} be the optimal profits if the seller implements CBP, BSP, or CBSP respectively.

Proposition 11 *When the costs are separable in bundle sizes:*

- *The gap between the optimal profits of CBP and CBSP can be infinity.*

$$\max\left\{\frac{\Pi^*_{CBP}}{\Pi^*_{CBSP}}\right\} = \infty.$$

- *The gap between the optimal profits of CBSP and BSP is smaller than a factor of* 8.

$$\max\left\{\frac{\Pi^*_{CBSP}}{\Pi^*_{BSP}}\right\} \leq 8.$$

6 Conclusion

In this study, we first study the CB problem with marginal decreasing prices and prove that it is a shortest-path problem. Second, we propose a dynamic programming algorithm to solve the CB problem with unit decreasing prices. Third, we analyze the CB problem with sub-additive prices and convert its MINLP formulation to a mixed-integer programming (MIP) one. Finally, we provide analytical and numerical analysis on the gaps between different CB models. We reconcile the differences in the optimal solutions obtained via different formulations of cardinality bundling in the literature.

There are several ways to extend the current study. First, there is still room to improve the performance of proposed dynamic programming algorithm for the BSP problem by combining it with LP cuttings. Second, CBSP problem has only been converted to an MIP, which is still N-P hard. Third, the gap analysis between MDP to BSP is still missing. Last but not least, analyzing cardinality bundling problems without Spence-Mirrlees condition can provide a wider application of these pricing schemes in reality.

References

Chu, C., P. Leslie, and A. Sorensen (2011). Bundle-size pricing as an approximation to mixed bundling. *The American Economic Review 101*(1), 263–303.

Hitt, L. and P. Chen (2005). Bundling with customer self-selection: A simple approach to bundling low-marginal-cost goods. *Management Science 51*(10), 1481–1493.

Spence, M. (1980). Multi-product quantity-dependent prices and profitability constraints. *Rev. Econom. Stud. 47*(5), 821–841.

Wilson, R. B. (1993). *Nonlinear Pricing*. Oxford University Press.

Wu, J., M. Tawarmalani, and K. N. Kannan (2014). Cardinality bundling with spence-mirrlees reservation prices. working paper.

When Online Reviews Meets Market Share Signal: Firm's Pricing Under the Influence of User-Generated Information

Yang Liu
Dept of Information Management and E-Business Dept of Information System
Xi'an Jiaotong University City University of Hong Kong
jordanbull.23@163.com

Juan Feng
Department of Information System
City University of Hong Kong
juafeng@gapps.cityu.edu.hk

Xiuwu Liao
Department of Information Management and E-Business
Xi'an Jiaotong University
liaoxiuwu@mail.xjtu.edu.cn

Abstract

We study the interaction impact of market share signal and online review, on consumers' purchase decision as well as on firms' optimal pricing strategies. We build a two-period model in a duopoly context. We find that: (1) Once quality difference signal from the online reviews cannot offset the strongly unfavorable market share signal, the high quality firm gives up a penetration strategy; and low quality firm adopts a penetration strategy, unless the quality difference from online review is so large compared to the impact of the market share signal. (2) More information, such as the precision of market share information, is not always beneficial to the profits of both firms.

Keywords: Online reviews; market share signal; pricing strategies

1. Introduction

In recent years, the Internet commerce has become an essential component of commerce around the world, and consumers purchase decision is more and more influenced by online review due to its low cost, large scale and easy access (Dellarocas 2003). Furthermore, the development of mobile equipments such as smartphones and tablets enlarges the popularity of online reviews. Based on a Nielsen's report in 2014[1], reading reviews of recent/future purchases is the most common shopping activity among tablet owners (55%). After purchases, many mobile shoppers write reviews and comment on their purchases using social media (26% using smartphone and 23% using tablet). Online reviews help consumers learn more detail about the product they are interested in and reduce the information asymmetry between sellers and buyers (Chen et al. 2008; Kuksov et al. 2010). Also, from consumer reviews, firms can detect consumers' preference and segment market accordingly to avoid pricing unreasonably (Lin et al. 2011).

Besides online reviews, market share information may also influence consumers heavily (Chen 2008; Hanson and Putler 1996). It is observed that consumers tend to prefer products with a large market share (Caminal and Vives 1996). For example, when facing two restaurants without knowing their reputation, consumers are more likely to dine in the one with more diners. This phenomenon is known as herding or crowd effect. The development of IT facilitates the

[1] http://www.nielsen.com/us/en/newswire/2014/shopping-lists-how-mobile-helps-consumers-tick-all-the-boxes.html

communication of market share information, because it is convenient for sellers to record and disclose sales quantity information on the website.

It is not uncommon that firms utilize both online reviews and market share information in online market. However, these two kinds of information impact consumer differently. Online review is an active, direct and explicit revelation of products information in vertical (quality) or horizontal (taste) dimension (Kuksov and Xie 2010; Kwark and Chen 2012). From reviews, consumers learn details of products, such as the strengths and weakness of the product (vertically), why consumers like or dislike it (vertically and horizontally). On the contrary, market share seems to be a passive, indirect and implicit revelation of products information or even be misleading. Chen (2008) suggests that high sales volume will positively affect consumers' online choices. However, this herding effect cannot guarantee consumer to purchase a good product. For example, previous sales volume may be caused by randomness in consumer choices, brand fans or herding behaviors etc.

In this paper we examine the interaction of online reviews and market share signal in a duopoly context. We study: (1) how should firm optimally price its product based on the availability of information in the market? (2) Whether the more accurate information will be good for the firms? We build a two-period economic model to study the impact of both online review and market share information in a duopolistic setting. We find that: (1) the pricing strategies of firms depend on the interaction of market share information and online reviews. Once quality difference signal from the online reviews cannot offset the strongly unfavorable market share signal, the high quality firm gives up a penetration strategy; and low quality firm adopts a penetration strategy, unless the quality difference from online review is so large compared to the impact of the market share signal. (2) More information, the precision of market share difference, is not always beneficial to the profits of both firms.

2. Literature Reviews

This paper studies the impact of both online review information and the market share information on firms' pricing strategy. We discuss related literature in the following.

Online review

In the last decades literature on the effects of online review proliferates rapidly. A large volume of literature is developed studying the impact of online review on sales volume. Online reviews can be measured in three dimensions: valence, variance and volume (Dellarocas and Narayan 2006). Ye et al. (2009) find that variance has a negative impact on hotel bookings; Sun (2010) provides evidence that the impact of review variance depends on the average ratings valence; the positive impact of review valence on sales volume is also demonstrated in Chevalier and Mayzlin (2006) .

Recently, a growing stream of literature is developed to examine online reviews` impact on the marketing and pricing strategies. Chen and Xie (2008) show optimal timing and marketing communication strategy for sellers to respond to consumer reviews. Lin and Li (2011) suggest that consumers' rating attitude would affect the app price and app quality as well as the platform's revenue sharing policy. In the competition context, Kwark and Chen (2012) study the impact of online product reviews on an online retailer under a traditional wholesale-scheme and an emergent platform-scheme. Li and Hitt (2011) examine online reviews influences firms' pricing strategy in a context of repeat purchase and find that online reviews can alter consumers' propensity to switch among products.

Herding behavior

Herding is a common phenomenon in daily life, and the advances in IT intensify this power. In general, when a large sale volume is disclosed to consumers, it always attracts others to imitate. Salganik and Dodds (2006) find consumers are shown to influence each other when download music from a web application. Herding is also found in the online book sales (Chen, 2008). Chen, etc.(2011) consider herding effect from observational learning and the effect of online word of mouth, and find that the two factors have opposite effects: negative word-of-mouth is stronger than positive, whereas positive observational learning is stronger than negative. This commonly observed phenomenon may reflect a fact that consumers need to infer product quality based on the choices of other consumers. Caminal and Vives (1996) study the signal effect of market share information in a duopoly market, and find that consumers believe that a firm with a high market share is likely to offer a high quality good. Thus, viewed as a signal, information such as market share/sales volume may induce herding behavior and thus may signal product quality.

3. The Duopoly Model

We consider two firms H and L in an online market to sell two imperfectly substitutable products, h and l. The marginal cost of each product is assumed to be zero. Without loss of generality, we assume that firm H sells a higher quality product and the actual quality difference between H and L is q, which is unknown to consumers. However, consumers can infer the quality difference q by information they have.

Consumers are divided into two types: rational and irrational consumers. Rational consumers purchase only when they can obtain non-negative utilities, while irrational consumers purchase dose not[2]. We follow Hotelling model and assume that the taste of consumers are uniformly distributed in $[0,1]$, where product h located at 0 and product l located at 1. Each consumer demands only one unit of product. Let t represent the mis-fit cost per unit distance. Thus, if a rational consumer located at point x, the misfit cost incurred by product h is xt and by product l is $(1-x)t$.

Rational Consumers. They knows that the actual quality difference between the two products q is a random variable following normal distribution with mean m and precise τ_m.

That based on the signals of their own experiences, the known brand reputation and the products ads etc... The precise τ_m captures the consumer heterogeneous about either the products. Meanwhile, rational consumers are uncertain about their location on the market, which also means that they are not sure to what extent the products are unfit for them. So a consumer knows that he is at the position x with possibility 0.5 while knows that he may be at any point on the line equally with probability 0.5. So a rational consumer's prior belief about expected utility difference of each product i is:

$$m+\left(0.5-x\right)t-(p_1^h-p_1^l)$$

Where p_j^i is the price of product i in period j, $i=h,l$ and $j=1,2$. And the demands of both products are:

$$D_1^h=\frac{1}{2}+\frac{E(q)-(p_1^h-p_1^l)}{t}+u^h,\ D_1^l=\frac{1}{2}-\frac{E(q)-(p_1^h-p_1^l)}{t}+u^l$$

Where D_j^i is the demand of product i in period j. And u^i represents demand from irrational consumers, which is a random variable following normal distribution with mean 0 and precise τ_u which captures the irrational purchase behavior such as the fans of brand, random purchase and the herding behaviors etc…

[2] In the paper, we define irrational consumers as the one that purchase not depend on the utility, they may be the loyal consumers or they only care about the product quality without considering the price.

We consider a two-period model in which consumers only "live" for one period. In each period, the distribution of the rational consumers' prior belief about the quality difference q and the irrational consumers demand u^i are public information to both firms and consumers. Consumers also can observe the prices of both products in the current period, but not those in previous period.

Meanwhile, consumer can access two kinds of information at the beginning of the second periods. The first is the market share of the previous period. It is not uncommon for consumer to choose the product which has been bought by more consumers, although he does not know the true quality of the product. The second is the product reviews of the previous period. Consumer can read the review and learn more about both the quality difference and horizontal information.

The sequence of game is as follows: at the beginning of first period, both firms price their product simultaneously, and consumers make purchase decisions. They offer reviews to the product after their purchase. In the beginning of second period, both firms decide whether to publish the reviews [3] and then make the prices of their products accordingly. And the second period consumers make their purchase decisions based on the information available in the market.

3.1 Market Share Information Combined with Online Review

We assume that market share and online review information is available to consumers at the beginning of period 2. At the beginning of the second period, consumers have three information sources: consumers` prior belief about quality difference; and:

Market share information. The rational consumers who enter in the second period can learn the market share difference of the previous period: $\Delta x = \frac{2}{t}(m - \Delta p^e) + \Delta\tilde{u}$. Where $\Delta p^e = p_1^{he} - p_1^{le}$, $\Delta\tilde{u} = u^h - u^l$ and $\Delta\tilde{u} \sim N(0, \frac{2}{\tau_u})$. $p_j^{i\,e}$ is the rational consumers' expected about the previous period price p_j^i of product i . Consumers infer the public statistic $q_m{}^e = \Delta p^e + \frac{t}{2}\Delta x$, which is a normally distributed unbiased estimator of rational consumers` expected prior belief with variance $\frac{t^2}{2\tau_u}$ . Thus, the aggregated consumers` update belief about quality difference is: $\hat{\eta}^e = E(q \mid q_m{}^e) = \hat{x}_1 m + \hat{x}_3 q_m{}^e = \hat{x}_1 m + \hat{x}_3(\Delta p^e + \frac{t}{2}\Delta x)$.

Online reviews. Let *r* represent the information from online review, *r* is a random variable following normal distribution with mean q and precise τ_q . We assume that the expectation of reviews signal q reveals the real quality difference between two firms; and τ_q capture the consumer heterogeneous of quality difference.

Meanwhile, by reading reviews, rational consumers can update the uncertainty of their location in the market. So after reading the reviews, a consumer knows that he is at the position x with possibility $0.5(1+\gamma)$ while knows that he may be at any point on the line equally with probability $0.5(1-\gamma)$. Where γ represents the information signaled by product reviews. The more information from reviews, the more precise consumers know their location in the market.

Consumers` belief. In practice, if firms set the price to make the prices are p_1^h and p_1^l , it's easy to show that the expected quality difference signaled from market share difference is:

---

[3]. The decision of both firms is accordingly regarding whether to publish reviews or not. If not, then the behavior of publishing review itself is powerful information to signal the quality difference.

$\Delta x = \frac{2}{t}(m - \Delta p) + \Delta u$, where $\Delta u = u^h - u^l$ [4]. Assume without any additional information, consumers perceive the quality of the two products to be the same, so we have $m = 0$. Thus, based on the three kinds of information, and using Bayes update, the aggregated consumers` belief about the quality difference of these products in the second period is:

$$\eta^e == x_2 q + x_3(\Delta p^e - \Delta p + \frac{t}{2}\Delta u),$$

Where $x_2 = \frac{t^2 t_q}{t^2(t_m + t_q) + 2t_u}$ and $x_3 = \frac{2t_u}{t^2(t_m + t_q) + 2t_u}$. And in equilibrium, the firms set prices equal to the expected prices of consumers, that is $\Delta p^e = \Delta p$. So, $\eta = x_2 q + x_3\left(\frac{t}{2}\Delta u\right)$.

Assume that firms in the beginning of period 1 can conjecture what consumers will review. In practice, firms can learn consumers' preference and market segmentation through past experience, for example. We can obtain the equilibrium prices of both firms in the two periods from backward induction.

Proposition 1. (Equilibrium price).
With online reviews and market share information, firms' equilibrium prices in the two periods are:

For H firm: $p_j^h = \begin{cases} \frac{1}{2}t - \frac{1}{3}x_3\left(t + \frac{2x_2 q}{9(1+\gamma)}\right), & j = 1 \\ \frac{1}{2}t(1+\gamma) + \frac{1}{3}\eta, & j = 2 \end{cases}$, *for L firm:* $p_j^l = \begin{cases} \frac{1}{2}t - \frac{1}{3}x_3\left(t - \frac{2x_2 q}{9(1+\gamma)}\right), & j = 1 \\ \frac{1}{2}t(1+\gamma) - \frac{1}{3}\eta, & j = 2 \end{cases}$.

Where $\eta = x_2 q + x_3\left(\frac{t}{2}\Delta u\right)$, $j = 1, 2$ *represent the period 1 and 2.*

We are only interested in the equilibrium in which both firms' second-period prices are positve[5]. This is insured by the following two constraints: (1) $-u^* \le \Delta u \le u^*$ [6];(2) $0 \le q < z^*(\Delta u)t$. From now on we restrict our attention to the parameters that satisfy these two constraints.

Where $u^* = \frac{3(1+\gamma)}{x_3}$ and $z^*(\Delta u) = \frac{3(1+\gamma) - x_3\Delta u}{2x_2}$.

Proposition 1 shows the equilibrium prices depends on both the horizontal and vertical differentiation between the two products.

In period 2, both prices increase with the informativeness γ of online reviews. As γ increases, consumers learn more about the horizontal features of the products, as well as their own location are more precisely, which reduces the substitutability between the two products and in turn, reduces the price competition between the firms. Also, since θ represents the posterior belief about the quality difference between products, H firm always sets a higher price if $\eta > 0$.
In period 1, similar to Proposition 1, expecting that consumers will infer the quality difference between the two products from the market share difference, firms have incentive to reduce their first-period price to attract more consumers. The more consumers weigh the market share information (x_3), the higher incentive for firms to reduce the first-period price. Meanwhile, the

[4] . Since u^i is the realized demand from irrational consumer of firm i and $\Delta u = u^h - u^l$, so $\Delta u > 0$ signals the high quality has more market share than low quality firm has. In the following parts, we will refer $\Delta u > 0$ as "positive market share difference" and $\Delta u < 0$ as "negative market share difference"**.**
[5]. If not, the firm with negative price will exit the market since it can make no profit in the last period.
[6]. This constraint promise the quality difference q is non-negative. When $q < 0$, H firm become the low quality and L firm become the high quality. For the high quality firm, q is still non-negative.

high-quality firm knows that consumers can infer the quality difference from online review, which will lead to a higher price power, thus it can afford to lower its price in the first period to attract more consumers and enlarge its market share.

On the contrary, L firm will price higher in the first period to earn as much profit as possible from consumers from horizontal preferences, expecting that it will be in disadvantage in the second period due to unsatisfying review. The informativeness γ of online reviews in the second period will also affect the equilibrium price in first period. Firms expect that if consumers learn more horizontal information about products, they will rely more on horizontal information to make purchase decision. Thus, if the γ increase, H firm has less incentive to enlarge market share, and will set a higher price. And L firm knows that the quality disadvantage is offset by the horizontal differentiation, so it has the incentive to enlarge market share, because higher demands of first period will lead to more marginal profit. We compare the pricing strategies of the two firms in the two periods.

Proposition 2. (Firm Pricing Strategy)
Firms' pricing strategies are determined by the interaction of the quality difference revealed from the online review, and the impact of the market share signal.

1) *The H firm adopts a penetration strategy, unless the market share signal is strongly unfavorable, and the quality difference revealed from online review is not large enough to offset this unfavorable impact; that is, if $-u^* < \Delta u < -u_1$ and $q \le z_1(\Delta u)t$;*
2) *The L firm adopts a penetration strategy, unless the quality difference revealed from online review is so large compared to the impact of the market share signal; that is, if $q > max\{0, z_2(\Delta u)t\}$.*

Where $u_1 = \dfrac{3\gamma + 2x_3}{(1+x_2)x_3}$, $z_1(\Delta u) = -\dfrac{(3\gamma+2x_3)+x_3\Delta u}{2x_2(1+\varepsilon)}$, $z_2(\Delta u) = \dfrac{(3\gamma+2x_3)-x_3\Delta u}{2x_2(1+\varepsilon)}$ and $\varepsilon = \dfrac{2x_3}{9(1+\gamma)}$.

Proposition 2 shows that firms' equilibrium pricing strategies depend on the interaction between the impact of online reviews and market share signals. In general, expecting that consumers will be influenced by market share signals, both firms have incentive to reduce their first-period price to induce a larger market share, and set a higher price in the second period when the uncertainty of quality difference reduces and the market share signal realizes. The only exception is when such consumer-generated information is realized to be strongly unfavorable such that they are not able to set a higher price in the second period. More specifically, for the H firm, this is when the market share difference is extremely negative, while the quality difference indicated from online review is not large enough to overcome this negative impact; for the L firm, this is when the market share signal is extremely unfavorable, or when the quality difference indicated from online review is extremely unfavorable to offset any favorable market share signals.

3.2 The effect signal precision of market share difference on the equilibrium profits of both firms.

We use the simulation to analyze if the more precision of the information can benefit firms. Setting that $\tau_m = 1$, $\tau_q = 5$, $t = 0.5$, $r = 0.3$, and change q with different values, we get figure 1 and 2. From the figure 1, we can see that when market share difference is small ($\Delta u = 1$), the total profit of H firm decrease with the precision. When the market share difference become bigger ($\Delta u = 2$), the second period profit increase with precision when q is small ($q = 0.4$), and the increase rate excess the decrease rate in period 1, so the total profit is increase with precision. From the figure 2, we can see that when then market share difference support L firm ($\Delta u = -1$), the total profit increase with precision when the precision is not large. The increase is due to the precision`s positive effect on second period profit which is similar to that of H firm. When the precision become larger, the negative effect of the precision in period 1 is larger than that in period 2, so the total profit is decrease with the precision.

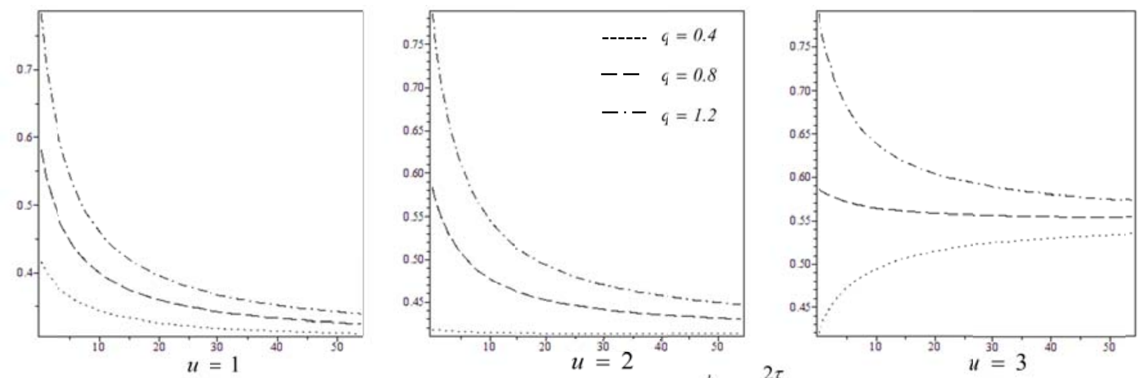


Figure1. The relationship between π^h and $\frac{2\tau_u}{t^2}$.

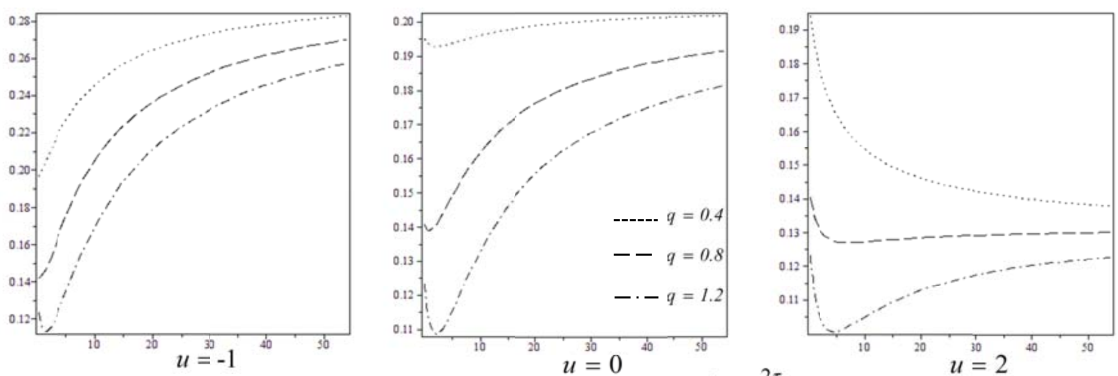


Figure2. The relationship between π^l and $\frac{2\tau_u}{t^2}$.

Combining the effect of precision on equilibrium prices in Proposition 1, and the figure 1 and 2, we learned that the effect of precision on total profits is the balance of that on the profit of each period. In the first period, the profits of both firms are decrease with precision. The large precision of market share difference suggests that it is a reliable signal. So in order to enlarge market share, both firms have more incentive to cut the first period prices and scarified part of the first period profit. In the second period, consumers will weigh the signals of online reviews and market share difference. If the precision of market share difference is large, the consumer will view infer the quality difference more from market share difference, and if the signal value of market share difference is larger than that of reviews, consumers` post belief about quality difference is increase, and the profit of $H$ firm will increase with precision, and the profit of $L$ firm will decrease with precision. If the signal value of market share difference is smaller than that of reviews, consumers` post belief about quality difference is decrease, the profit of H firm decreases with precision, and the profit of L firm increases with precision.

7. Conclusions and Future Research

Nowadays, consumer' online purchase decisions are heavily influenced by the information they obtain. Market share information and online reviews are two major kinds of information that are available to consumers. However, these two kinds of information influence consumers differently. This paper focuses on the interact effect of market share and online reviews information on firms optimal pricing strategy in a duopolistic context. We find that : (1) Once quality difference signal from the online reviews cannot offset the strongly unfavorable market share signal, the high quality firm gives up a penetration strategy; and low quality firm adopts a penetration strategy, unless the quality difference from online review is so large compared to the impact of the market share signal. (2) More information, such as the precision of market share information is not always beneficial to the profits of both firms.

From the theoretical perspective, this paper connects the signal of market share to that from online reviews and studies the interaction between these two effects. As a result, (1) the penetration pricing strategies is not always optimal to high quality firm; (2) The higher precision of information is not always detrimental to the low quality firm and beneficial to the high quality one.

This paper is not without limitation. Future extensions could include: First, we assume firms learnt the reviews distribution from past experience. This may not be always true. For example, when consumers entering in different periods are significantly different, or when a seller lacks experience. Secondly, it is worthwhile to conduct empirical study to test the validity of our theoretical findings.

Reference

[1] Caminal R, Vives X. 1996. Why Market Shares Matter: An Information-Based Theory. The RAND Journal of Economics. 27(2): 221-239.

[2] Chen Y F. 2008. Herd behavior in purchasing books online. Computers in Human Behavior. 2008, 24:1977–1992

[3] Chen Y, Xie J. Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix. Management Science, 2008, 54(3): 477-490.

[4] Chen Y, Wang Q, Xie J. 2011, Online Social Interactions: A Natural Experiment on Word of Mouth versus Observational Learning. Journal of Marketing Research, 48(2): 238–54.

[5] Chevalier J A, Mayzlin D. The Effect of Word of Mouth on Sales: Online Book Reviews. Journal of Marketing Research, 2006, 43(3): 345–354.

[6] Dellarocas C. The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms. Management Science, 2003, 49(10): 1407-1424.

[7] Dellarocas C, Narayan R. Tall heads vs. long tails: do consumer reviews increase the informational inequality between hit and niche products? Working Paper, 2007.

[8] Kwark Y, Chen J, Raghunathan S. Platform or Wholesale: Different Effects on Retailers of Online Product Reviews. ICIS 2012 proceedings, 2012:8.

[9] Li X, Hitt L M, Zhang Z J. Product Reviews and Competition in Markets for Repeat Purchase Products. Journal of Management Information System, 2011, 27(4): 9-41.

[10] Lin H, Li X, Tan Y, Xu J P. The Economic Role of Rating Behavior in Third-party Application Market. ICIS 2011 Proceedings, 2011: 26.

[11] Salganik, Matthew J, Peter Sheridan Dodds, and Duncan J. Watts. 2006. Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. Science,311(5762):854–6.

[12] Sun M. How Does Variance of Product Ratings Matter? Management Science, 2012, 58(4): 696-707.

[13] Ye Q, Law R, Gu B. The impact of online user reviews on hotel room sales. International Journal of Hospitality Management, 2009, 28(1): 180-182.

Can Health Information Sharing Reduce Duplicate Testing?

Sezgin Ayabakan[a], Zhiqiang (Eric) Zheng[a], Indranil Bardhan[a], and Kirk Kirksey[b]

[a] The University of Texas at Dallas, Richardson, Texas
[b] The University of Texas Southwestern Medical Center, Dallas, Texas

ABSTRACT

Recent healthcare reform has focused on reducing the excessive amount of waste and inefficiency in US healthcare spending, among which duplicate testing (or over testing) is one of the main culprits. We investigate the factors associated with duplication of radiology imaging tests when information sharing across healthcare providers is fragmented and patients may switch from one hospital to another. We hypothesize that patients' switching across hospitals will result in increased duplicate procedures due to lack of access to their medical history, and argue that implementation of intra- and inter-hospital information sharing technologies will help reduce the rate of duplicate tests. We utilize a comprehensive panel dataset consisting of 39,600 Congestive Heart Failure patient visits across outpatient clinics of 68 hospitals in North Texas from 2005 to 2012. Our results suggest that hospital switching leads to a higher incidence of duplicate testing and that inter-hospital information sharing technologies are associated with lower levels of duplication. Our results lend support to the argument for implementation of health information exchanges as a potential solution to reduce the incidences of duplicate tests and subsequently the overall cost of healthcare delivery.

Keywords: Duplicate testing, hospital switching, health information sharing

1. INTRODUCTION

It is estimated that 40-50% of U.S. healthcare spending amounts to waste, of which overuse of resources is a significant contributor (Bentley et al. 2008; Hillestad et al. 2005). The Congressional Budget Office estimates that around $700 billion per year, or 5 percent of US GDP, is spent on tests, and treatments that do not actually improve health outcomes (Orszag 2008). Waste due to inefficient use of resources can arise in situations like excessive antibiotic use for viral infections, avoidable hospitalizations for nursing home patients, unnecessary admissions of patients with chest pain, and overuse of screening and imaging procedures (Bentley et al. 2008). Recent enactment of the Affordable Care Act (ACA) aims to replace the current fee-for-service environment where providers are paid more for ordering more frequent tests and treatments with accountable care organizations that instead reward cost-effective care, as a way to coordinate and reduce avoidable costs (Beck 2013).

In this study, we specifically focus on the duplication and overuse of imaging procedures related to the diagnosis and treatment of congestive heart failure (CHF) outpatients. A likely cause of the excessive use of imaging tests is the lack of information sharing among disparate healthcare entities. Redundant medical procedures are likely to arise if patient medical data is not shared between different providers and healthcare systems (Bates et al. 1998; LaBorde et al. 2011). For example, Kripalani et al. (2007) report that only 3% - 20% of hospital attending physicians communicate with their patients' primary care providers. To make matters worse, it is estimated that between 33% and 63% of patient discharge summaries lack important information on diagnostic test results and other relevant information that may potentially cause readmission, dissatisfaction, delay in treatment, or other patient safety issues (Kripalani et al. 2007), thus exacerbating the possibility of incurring similar (i.e. duplicate procedures) in the future.

We argue that one of the main drivers of duplicate testing is the lack of information sharing among healthcare providers, a problem which is exacerbated when patients migrate across hospitals and health systems due to technological barriers to information sharing across organizations. Technological barriers are imposed when healthcare providers do not have access to IT platforms that allow for seamless sharing of patient data across health systems.

We empirically test our model using a comprehensive dataset of more than 39,600 CHF patient visits to outpatient clinics across 68 hospitals in North Texas. This dataset records information for each patient's visit tracked across a relatively long period from 2005 to 2012. Our results suggest that patients' switching behavior across hospitals has a significant association with the rate of duplicate imaging and tests among CHF patients. We show that inter-hospital information sharing technologies significantly reduce the rate of duplicate testing with respect to imaging and radiology tests conducted on CHF patients. Our study provides a foundation to estimate the avoidable costs attributed to duplication of outpatient tests that are incurred due to a lack of information sharing across healthcare providers. In the context of the current debate on healthcare reform and the need to reduce healthcare costs through reduction in redundant procedures, our study lends support to the possibility of reducing costs associated with patient diagnosis and treatment through implementation of health information exchanges (HIE).

2. BACKGROUND

There are many reasons for the high rate of duplicate or over testing (Dai et al. 2011). Some duplicate tests are necessary because a patient's condition may change from one visit to another and it is usually up to the physician to determine whether a test is necessary. Physicians have a tendency to adopt 'defensive medicine', a term reflecting the possibility of physicians' reducing their likelihood of future litigation if such tests lead to detection of significant findings (Currie and MacLeod 2013). Physicians also have other motives besides avoidance of lawsuits (Baicker et al. 2007; Currie and MacLeod 2008). Another cause of redundant testing is related to the current "fee-for-service" world of healthcare where providers' are paid based on fees for every service provided to patients, regardless of their necessity or impact on the quality of patient care (Gruber et al. 1999; Gruber and Owings 1994). Brill (2013) observes that there are significant variations in the prices charged to Medicare and non-Medicare patients, wherein self-pay and private insurance patients are charged higher prices for the same tests/procedures. Hence, the economic incentives behind resource overuse (i.e., duplicate testing) constitute one of the drivers of inefficiency and waste in the U.S. healthcare system (Dai et al. 2011).

A significant problem in any research on duplicate or over testing is the potential for patients to switch providers and visit different hospitals during an episode of a treatment due to personal preferences or medical reasons (LaBorde et al. 2011). When providers cannot easily access patient medical history, patients and/or their families are often asked to provide the relevant medical information prior to hospitalization (Kripalani et al. 2007). This information may include patients' prior medical history, diagnosis, allergies, and medication history. However, often times, patients may not be able to accurately remember crucial information related to previous hospital visits, such as diagnoses, medications, and procedures, due to a variety of factors including age, patient condition, and recall bias (Johnson et al. 2011). Such unavailability of information, especially when patients switch across different hospitals, can result in providers having to order repeat diagnostic tests and procedures which contribute to higher duplicate test rates. Hence, our focus in this study is to analyze the *impact of patient switching behavior (across hospital visits) on the incidence of duplicate testing*.

Another likely cause of the excessive use of diagnostic tests is the lack of information sharing among disparate entities. As health information technologies (HIT) facilitate information sharing within and across hospitals, we are interested in understanding the impact of intra- and inter-hospital image sharing technologies on the extent of duplicate imaging tests. Since these types of technologies are critical to build health information exchanges (HIE), we examine the potential benefits that may be accrued from implementation and usage of these technologies in order to develop a baseline understanding of the value of HIEs as an avenue to reduce the duplication rate associated with outpatient imaging tests in the U.S. healthcare system. Considering the fact that the US hospitals and providers have been slow to adopt HITs, and is at least four to thirteen years behind in implementing a national initiative compared to most European countries with more mature HIT systems (Shaver 1998), it is essential to inform policy makers about the implications of adopting HITs on healthcare costs based on empirical analysis of patient admission and diagnostic data. Such analysis may provide supporting evidence to better inform policy makers and other stakeholders who are considering funding the implementation and rollout of HIEs as a vehicle to reduce the inefficiencies related to costs of duplicate testing.

In this research, it is important to note that we focus on the determinants of duplicate testing and do not differentiate whether a particular duplicate test is necessary or truly redundant. Determining which test is necessary or redundant is a subjective exercise and can vary from patient to patient and one physician to another. We argue that studying over testing, regardless of whether it is redundant, is an interesting and significant problem in its own right since the U.S. healthcare system suffers from excessive overuse of tests. For instance, the number of procedures performed per patient in the U.S. is double that of the OECD average. Hence, developing a better understanding of the drivers of duplication is a first step to help healthcare policy makers, providers, insurance companies, and consumers determine the factors that contribute to the duplication rate, so that appropriate steps can be implemented to reduce the extent of duplication. Specifically, we focus on patients diagnosed with CHF since this represents one of two health conditions covered by the Department of Health and Human Services starting in 2012.

3. RESEARCH HYPOTHESES

We now develop our hypotheses to investigate the role of patients' hospital switching behavior, hospital health information sharing, patient payer type and admission status and their impact on the extent of duplicate testing during the course of patients' diagnosis and treatment over time.

3.1 Hospital Switching Behavior

A typical patient receives care from multiple, geographically disparate providers over the course of their treatment (Flanders 2009). Mobility of patients not only leads to more voluminous health data but also greater fragmentation and unavailability of information due to barriers to information sharing across providers (Flanders 2009). There exist significant organizational and technological barriers to sharing patient health information across disparate health providers in the present environment. In an ideal scenario, when a patient switches from one provider to another, the patient's prior health information must follow. However, patients may not always be able to accurately recall or communicate clinical information and treatment details with their prior care providers. Care providers may also lack incentives to share health information; they may be reluctant to retrieve clinical information because it is cumbersome and time consuming; and even if providers are willing to share, logistical barriers stemming from fragmented medical data across hospitals, laboratories and clinics may lead to inconsistencies and distortions in the patient's medical record (Johnson et al. 2011, Overhage et al. 2002)).

When medical data is unable to move between providers, diagnostic or treatment errors can arise as a result of current providers' inability to retrieve patients' prior health data (Bates et al. 1998; Flanders 2009; LaBorde et al. 2011). For example, LaBorde et al. (2011) contend that lack of healthcare IT (e.g EMR and HIE) integration across facilities can lead to more duplicate diagnostic laboratory tests. Hillestad et al. (2005) argue that EMRs can greatly facilitate information flow among providers and estimate that EMRs can save $7.9 billion annually by reducing the need for redundant laboratory and radiology tests. Understanding how duplicate testing occurs when patients switch hospitals is of importance to health policy makers, considering the Affordable Care Act's (ACA) initiative in reducing waste and abuse in the US healthcare system (http://www.hhs.gov/opa/affordable-care-act/index.html). We conjecture that patients' switching behavior across hospitals will result in an increase in duplicate testing.

H1. *The rate of duplicate testing when a patient is admitted to a different hospital within the same health system will be greater than when a patient is admitted to the same hospital.*

3.2 Health Information Sharing

Enabling access to patient health information sharing across various stakeholders can greatly reduce inefficiencies in the US healthcare industry. Health information technologies can facilitate actions related to capturing, storing, sharing and retrieving patient health information. Accessing related information about patient's medical and surgical history, allergies, current and past medication lists through health information technologies will allow providers to make better decisions regarding patient's diagnosis and treatment (Booth 2003). Electronic Health Records (EHRs) provide the technological foundation for sharing patient health information across organizational boundaries. EHRs are software platforms that enable interoperability among multiple stakeholders such as delivery and sharing of patient health information to physician offices and hospitals (Mishra et al. 2012). EHRs can enable intra-hospital information sharing through reduced barriers to information access such as digitized paper charts and digitized patient medical history on procedures and tests (HealthIT.gov 2013). In Canada, a national interoperable EHR (Canada Health Infoway) is currently being implemented in response to the problem of physicians having to re-order diagnostic imaging procedures due to patient referrals and transfers across various facilities (Mendelson et al. 2008) Yet, EHR systems are also criticized for not being able to communicate with each other, resulting in islands of fragmented information (Ozdemir et al. 2011). DesRoches et al. (2013) argue that exchanging patient clinical summaries, laboratory and diagnostic test results with outside entities were among the least likely adopted functionalities of an EHR. Although the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009 promotes adoption of EHRs through its 'meaningful use' incentive program, the first stage of meaningful use was fully met by only 12.2% of physicians in 2012 as estimated by Wright et al. (2013).

Raghupathi and Tan (2002) define a strategic IT framework to study health information integration in terms of internal and external integration. Internal integration refers to the degree to which systems and technologies are integrated with one another within an organization, whereas external integration refers to the degree to which technologies interface with outside organizations. Based on their theoretical framework, we examine the implications of HIT applications in terms of the impact of sharing patient health information within (intra) and across (inter) hospitals on duplication rates.

Computerized physician order entry (CPOE) and picture archiving and communication systems (PACS) are two types of health IT that can facilitate intra-hospital information sharing. Prior research has shown that CPOE implementation is associated with a reduction in hospital length of stay, nurses' administrative time, and drug and radiology usage in outpatient setting (Hillestad et al. 2005). For radiology, CPOE can benefit stakeholders through reduction in the rate of redundant or duplicate imaging tests, such as X-rays and CT scans (Bates et al. 1999; Chin and Wallace 1999)). Wang et al. 2003 estimate that decreased utilization of radiology services can save up to $86,700 per provider through more efficient sharing of imaging data across providers. As an imaging informatics tool, PACS enables distribution of radiology images to various medical units within a hospital (Branstetter 2007). The ability to obtain timely and accurately access to radiology images within hospitals, through EHRs that are integrated with PACS, reduces utilization of radiology imaging services and can provide efficiency gains through reduced repeat (duplicate) imaging examinations (Lu et al. 2012; Sodickson et al. 2011). Accordingly, we hypothesize:

H2a. *Intra-hospital information sharing of radiology images across departments is associated with a decrease in overall duplication rate.*

Information sharing across hospitals is crucial when patients are mobile and visit providers across disparate health systems. The fact that providers often do not have access to virtual private network (VPN) connection to disparate PACS has made inter-hospital information sharing uncommon or even impossible (Mendelson 2011). However, recent adoption of cloud-based systems and web-based personal health records (PHRs)have emerged as potential solutions to overcome barrier associated with inter-hospital information sharing (Shrestha 2011, Flanders 2009). Due to the distributed nature of digital images, radiologists need access to images not only at the point of care, but from multiple locations with full access to patient information. Cloud-based solutions have been developed to facilitate inter-hospital information sharing by offering the ability to share patient images and report data across multiple locations securely and conveniently. Such systems can allow access to patient medical information (including previously performed imaging tests and reports) from disparate locations-- resulting in efficiency gains through reduced costs and duplicate procedures (Shini et al. 2012). Likewise, Internet-enabled image transfer systems have been recently proposed as a solution to enable greater inter-hospital information sharing. For instance, Flanagan et al. (2012) report that implementation of an Internet-enabled image transfer system at Harborview Medical Center was associated with a reduction in duplication rate to 8.1%. In the light of possible benefits that can accrue from information sharing, we hypothesize that inter-hospital information sharing systems will increase information availability, reduce latency, and therefore reduce the extent of duplicate testing.

H2b. *Inter-hospital information sharing of radiology images across hospital locations will decrease duplication rate.*

4. RESEARCH METHODOLOGY

We obtain data from two sources: the Dallas-Fort Worth Hospital Council (DFWHC) Research Foundation and HIMSS Analytics. We use the DFWHC data to empirically test the hypotheses related to hospital switching (H1) while we utilize the HIMSS Analytics data along with DFWHC to test our hypotheses on the impact of information sharing (H2).

4.1. Data

We obtained a comprehensive dataset of 39,600 Congestive Heart Failure (CHF) patient visits across outpatient clinics of 68 non-Federal hospitals and 26 health systems in North Texas. Based on patient-level administrative claims data, each patient's visit history is tracked from 2005 to 2012 through a unique patient identifier number, the regional master patient index (REMPI) developed by the DFWHC Foundation (Bardhan et al. 2011). In this dataset we only include patients with CHF as the principal diagnosis, i.e., patient admissions with ICD-9 code of "428.xx". Focusing only on their principal diagnosis alleviates possible patient heterogeneity arising from treatment procedures and imaging tests that vary across different diagnoses.

Furthermore, we focus specifically on outpatient admissions because patients receive radiology imaging procedures, such as X-rays, CT scans, MRI scans, and ultrasound tests, primarily in an outpatient setting which also account for a majority of these tests (Lee et al. 2012).

Based on feedback provided by leading CHF physicians, we apply a conservative cutoff of 90 days to determine whether an imaging test can be considered as duplicate.[1] This is because the typical life span of a radiology imaging test is about 3 months. Similar cutoff windows have been used in the prior literature. For example, Lu et al. (2012) define repeat imaging as "that performed when a previous CT or MRI examination of the abdomen was followed by a second examination with the same modality, body part, and type in 4 months." In addition, Lee et al. (2007) define the time window of a repeat imaging test as seven months, but report that a majority of repeat radiological imaging tests happened in the first two months of initial examination. In our analysis, we exclude the index visit since the duplication rate and hospital switching event are calculated with respect to the prior visit information. Based on these criteria, we focus on the visit history of 4,038 CHF outpatients who exhibit at least two (or more) outpatient visits during our study period. Their histories comprise a total of 9,403 consecutive visits, where the consecutive visits occur within a ninety-day period (of the prior visit). Table 2 reports the descriptive statistics of our model variables. Table A1 in the Appendix provides an illustrative profile of the outpatient visits associated with a 62-year, old, non-white female CHF patient and the imaging tests that are performed on the patient during these visits.

We also collected hospital level information sharing data from the HIMSS Analytics database for these 68 hospitals. Inter-hospital information sharing data was available for the years between 2006 and 2012, whereas intra-hospital information sharing data was available for the years between 2009 and 2012 (Bardhan et al. 2011).

4.2 Variable Definitions

Clinical information about the outpatient procedures in our data is reported via the common procedure terminology (CPT) coding scheme. Since our focus is on measuring the duplication rate associated with outpatient imaging procedures, we use only CPT codes related to X-rays, computed tomography (CT scans), magnetic resonance imaging (MRI) and ultrasounds.[2] For each patient visit, we count the number of duplicate tests for each CPT code that appears in the current visit. Each CPT code is matched against the CPT codes recorded across all prior visits that occur within the 90 days prior to the current visit. If the CPT code appears in any of the prior relevant visits (i.e. $\leq$ 90 days), it is flagged as a duplicate procedure and counted towards the total number of duplicate procedures for the current visit. We then calculate the percentage of duplication as the ratio of the total number of duplicates to total number of all CPT procedures for the current visit. The visit-level averages for duplication count, procedure count and duplication percentage are 0.18, 0.40 and 15.35% respectively.

Our data contains three types of patient visits: Emergency/urgent, elective and other. We construct two new variables to index hospital switching events. These variables are termed: "Visit to different hospital within the same health system (*VisitDHSS*)" and "Visit to different health system (*VisitDS*)." For a current repeat visit, if the prior admission was to a different hospital within the same health system within the last 90 days, then variable *VisitDHSS* takes the value of one, and zero otherwise. However, if the prior visit was to a hospital in a different health system within 90 days, then variable *VisitDS* takes the value of one, and zero otherwise (Note

---

[1] We also ran sensitivity tests on our results using 60-day and 120-day cutoff windows in our calculation of duplicate testing rate, and observe that the results are qualitatively similar in terms of the sign and significance of the model coefficients.

[2] The list consists of 417 unique CPT codes in total.

that if a patient switches health systems, then he/she has also switched hospitals). Based on our data, 5.03% of all patient admissions are to different hospitals within the same health system, whereas 4.74% of all patient admissions are to different health systems. We observe that there are significant differences in the duplication rates between patients who are readmitted to the same hospital versus those who switch hospitals within or across health systems. T-tests of the means of the two distributions show statistically significant differences at $p < 0.001$.

For each patient visit, the type of health insurance is reported via the payer description information. We classify this information into six different insurance variables: Private, Medicaid, Medicare Part-A, Medicare Part-B, Self-pay and other.

In order to capture the extent of intra-hospital sharing of imaging data, we focus on five different measures of imaging distribution through PACS to different units/departments within a hospital. These five (dummy) variables capture the extent of intra-hospital distribution of images through health IT. Our data measures the extent to which images are distributed to the following departments within a hospital: (a) critical care unit ($ImageDistrCCU_{ht}$), (b) emergency room ($ImageDistrER_{ht}$), (c) intensive care unit ($ImageDistrICU_{ht}$), (d) operating room ($ImageDistrOR_{ht}$), and (e) over the Worldwide Web ($ImageDistrWEB_{ht}$). For each hospital year, we create a single variable ($ImageDistrCOMB_{ht}$) which measures the intensity of imaging distribution across various medical units. If a hospital has implemented all five imaging distribution systems, we assign a value of one to the combined imaging distribution variable (i.e. ImageDistrCOMB), and zero otherwise. Overall, our combined intra-hospital image distribution variable, *ImageDistrCOMB*, follows an increasing trend over time as more hospitals started implementing distribution of images to various departments over time.

In order to capture the extent of inter-hospital information sharing, we use the binary variable defined as "image access via the internet from outside locations ($ImageACCESS_{ht}$)" as provided by HIMSS Analytics. In other words, hospitals are assigned a value of one for $ImageACCESS_{ht}$ if they allow Internet-enabled access to images to providers and hospitals from external locations (beyond the boundaries of the hospital or health system). Table 5 reports the percentage of hospitals who implemented $ImageACCESS_{ht}$ between 2009 and 2012. We observe that hospitals have increased their access to images via the Web over time, as shown by the increase in value of $ImageACCESS_{ht}$ from 69.1% in 2009 to 86.8% in 2012.

Additionally, we track patient-specific demographic information on patient gender (female or male), age, race (white or non-white), and zip code. We also obtain hospital-specific information from CMS (Centers for Medicare and Medicaid Services). CMS classifies hospitals according to their teaching status and geographic locations (urban, rural), hospital case mix index (CMI), and hospital size (number of beds). Other variables include patient distance to hospital (measured in miles by using the distance between patient home and hospital zip codes), total visit charges, and an indicator for emergency department visit.

5. MODEL SPECIFICATION

5.1 Hospital Switching Behavior

In equation (1), we regress the dependent variable *DupPer* on hospital switching variables, insurance variables, patient visit type and admission source variables. We also control for other admission and hospital specific factors as well as year-specific fixed effects. Accordingly,

$$DupPer_{it} = \beta_0 + \beta_1 VisitDHSS_{it} + \beta_2 VisitDS_{it} + \beta_3 InsSelfPay_{it} + \beta_4 InsPrivate_{it} + \beta_5 InsOther_{it} + \beta_6 InsMedicareA_{it} + \beta_7 InsMedicaid_{it} + \beta_8 TypEmergency_{it} + \beta_9 TypOther_{it} + \beta_{10} SrcTrans_{it} + \beta_{11} SrcOther_{it} + Controls\ \beta_c + YearDummies\ \beta_y + \epsilon_{it} \quad (1)$$

where $Controls$ represents a vector of variables consisting of *PtFemale, PtWhite, PtAge, DaysSince, HsCMI, HsTeach, HsUrban, log(HsBeds).*[3] Note that we opt not to control for patient-specific fixed effects for two reasons: (a) we only observe one data point for many patients, and adding individual fixed-effects will disregard these observations; (b) we are also interested in estimating the effects of several time-invariant patient-specific variables such as sex and race. These variables have already absorbed a large portion of heterogeneity that individual fixed- effects aim to capture.

The specification in model (1) may be subject to potential endogeneity as arising from the hospital switching variables. That is, patient- or hospital-specific factors can drive hospital switching decisions of patients; at the same time, high-level of duplicate tests encountered in a prior admission may lead patients to switch hospitals. Since our switching variable is binary , we follow the two-step Heckman approach described in Bharadwaj et al. (2007), Mani et al. (2010) and Shaver (1998) to address endogeneity. We impose an exclusion restriction in the second stage equation by introducing three exogenous variables that are available in our data. These variables include the charges associated with the patient's previous visit ($log(TotCharge_{it-1})$), patient home to hospital distance in the previous visit ($log(PtHsDist_{it-1})$), and number of diagnoses in the previous visit ($numDiags_{it-1}$). After the first and second stage estimation, we correct the standard errors of the second stage with respect to the asymptotic variance-covariance matrix derived in Catsiapis and Robinson (1982). We report the OLS estimation results for Model (1), using the two-step Heckman correction approach, in Table 1.

5.2 Health Information Sharing

The second part of our econometric estimation focuses on the impact of health information sharing on duplicate testing. To examine if implementation of intra or inter-hospital information sharing systems has a bearing on hospitals' duplication rate, we deploy a difference-in-difference (DID) specification which is used in the literature extensively within natural and *quasi-natural experimental* settings (Kumar and Telang 2012; Meyer 1995). DID compares the *Treatment* group against the *Control* group, where the treatment effect is measured against a control group in the pre- and post-treatment periods. This setting allows us to handle potential confounding effects of unobserved factors and time-invariant features from treatment effects (Kumar and Telang 2012; Meyer 1995). We focus on two groups of hospitals with respect to the presence of imaging distribution systems. The first group of hospitals (control group) did not implement intra-hospital (inter-hospital) imaging distribution systems between 2006 and 2012, while the second group of hospitals (treatment group) started using imaging distribution services at some point between 2006 and 2012[4]. Since hospitals in the treatment group may start to implement intra-hospital information sharing technologies in different years, we match the

[3] Since patient insurance, visit type and admission source are categorical variables, their values are transformed into dummy variables.

[4] There were very few hospitals which always had intra-hospital (inter-hospital) information sharing systems throughout the years between 2006 and 2012. We excluded these few hospitals and thus our experimental setup focuses on examining two groups of hospitals: One control group without having any treatment factor and one treatment group having HIT implemented at some point in time.

hospitals from the control group to the treatment group using a propensity score matching approach (Rosenbaum and Rubin 1985). We assign the implementation year of the matched hospital (in the treatment group) to a hospital in the control group in order to calculate the $PostIntra_{it}$ variable (1 or 0) for visits happening to that hospital in the control group.

Our quasi-natural experiment involves comparison of the treatment group to the control group. Hence, we incorporate control variables into DID specification and estimate the following models for intra and inter-hospital sharing, respectively:

$$DupPer_{iht} = \alpha_0 + \alpha_1 TreatmentIntra_{iht} + \alpha_2 PostIntra_{iht} + \alpha_3 TreatmentIntra * PostIntra_{iht} + Controls\ \alpha_c + \vartheta_{iht} \quad (2)$$

$$DupPer_{iht} = \theta_0 + \theta_1 TreatmentInter_{iht} + \theta_2 PostInter_{iht} + \theta_3 TreatmentInter * PostInter_{iht} + Controls\ \theta_c + \tau_{iht} \quad (3)$$

where *i* denotes a patient, *h* denotes a hospital and *t* denotes admission time index. $TreatmentIntra_{iht}$ ($TreatmentInter_{iht}$) equals one if a patient *i* visits (at time *t*) a hospital *h* where intra-hospital (inter-hospital) information sharing has been implemented. $PostIntra_{iht}$ ($PostInter_{iht}$) equals one if patient *i*'s visit at time *t* is in the post-treatment time period of hospital *h* that has implemented intra-hospital (inter-hospital) information sharing technologies. The coefficient estimate of α_3 (θ_3) for $TreatmentIntra * PostIntra_{it}$ ($TreatmentInter * PostInter_{iht}$) is of interest since it captures the change in the duplication rate for hospitals which implement intra-hospital (inter-hospital) information sharing technologies relative to hospitals who do not. We also account for insurance type, admission source, visit type, patient age, gender, race and hospital characteristics in our DID estimation approach.

6. RESULTS

For hospital switching, our results indicate a positive association between switching across hospitals and health systems and the rate of duplicate tests. We find that the coefficients of *VisitDHSS* and *VisitDS* ($\beta_1 = 7.251, p < 0.10$ and $\beta_2 = 81.81, p < 0.01$) in Table 1 are statistically significant, supporting hypotheses H1. Accordingly, when a patient switches to a different hospital but stays within the same health system, her duplication rate increases by 7.25% (holding all other variables constant at their mean values) compared to the case if she had been readmitted to the same hospital. Furthermore, when a patient switches hospitals and moves to a different health system, her duplication rate increases by 81.81%. Hence, we find that patients, who switch to hospitals within or across health systems, are more likely to incur higher levels of duplicate tests, although the duplicate rate is much higher when patients move from one health system to another.

Next, we report the differences in duplication rates between the *Control* and *Treatment* groups of hospitals in Table 2. We report the results for the two cases of intra- and inter-hospital information. One may still argue that our $TreatmentIntra_{iht}$ and $TreatmentInter_{iht}$ variables might be subject to potential endogeneity. For example, hospitals with higher duplication rates may be more likely to implement intra- and inter-hospital information sharing technologies. To address this endogeneity problem, we applied an instrumental variable (IV) estimation approach. Potential IV candidates should explain the variation in our endogenous variables (i.e., $TreatmentIntra_{iht}$ and $TreatmentInter_{iht}$), while they should not be systematically determined by $DupPer_{iht}$. One possible IV is the age of a hospital in terms of the number of years that it has been in operation ($HsAge_{iht}$)[5]. We conjecture that relatively new hospitals would be more likely to implement new types of health information sharing technologies and older hospitals are usually slow adopters of such systems due to the difficulty of replacing legacy systems. At the same time, the age of a hospital may not be systematically co-determined with its imaging duplication rate. The Hausman test that compares the OLS and 2SLS estimates rejects the non-existence of endogenity in the intra-hospital information sharing ($p = 0.033$) and the inter-hospital information sharing models ($p = 0.085$).

Therefore, we apply two-stage least squares (2SLS) estimation to Models (2) and (3). To test the validity of $HsAge_{it}$ as an instrument for $TreatmentIntra_{iht}$, we first check the correlation between these two variables and report that it is significant at 0.01 level, i.e., Corr($TreatmentIntra_{iht}, HsAge_{iht}$) = 0.42 and Corr($TreatmentInter_{iht}, HsAge_{iht}$) = 0.48, and the first-stage regression is also significant. Next, we regress the residuals on $HsAge_{iht}$, $HsAge_{iht}^2$ and $HsAge_{iht}^3$, in which the residuals are obtained from regressing $DupPer_{iht}$ on the IVs and all other exogenous variables. The resulting coefficients of $HsAge_{iht}$, $HsAge_{iht}^2$ and $HsAge_{iht}^3$ were insignificant with p values of 0.49, 0.51 and 0.61, which suggests that endogeneity of these IVs is not a concern in our models

In Table 2, we present our 2SLS estimation results for the DID approach to study the role of health information sharing technologies on duplication rates. The 2SLS analysis results suggest that intra-hospital information sharing does not exert a significant impact on reduction in duplication tests with the coefficient of the interaction term, Treatment * Post, being statistically insignificant. Hence, our DID results do not support hypothesis H2a. ($\alpha_3 = -10.54$). However, we observe that the interaction term associated with H2b is negative and significant ($\theta_3 = -55.59, \mathrm{p} < 0.01$). This result suggests that hospitals with inter-hospital image sharing technologies exhibit a lower duplication (after implementation) rate compared hospitals without these technologies, lending further support to hypothesis H2b.

7. CONCLUSIONS

This paper examines the factors that are associated with duplicate imaging procedures and investigates the role of hospital switching and health information sharing technologies in terms of their impact on the extent of duplicate testing. Our results show that patients who switch hospitals but stay within the same health system are more likely to undergo higher duplicate tests than patients who do not. In addition, the increase in the rate of duplication is much higher for the cases where patients switch across health systems. This may be directly driven by the fact that hospitals are still unable to share medical information across health systems.

We also found differential impact of hospital information sharing on duplicate imaging procedures in terms of implementing intra- and inter-hospital image sharing technologies. According to our results, inter-hospital image sharing technologies (i.e., image access ability from external locations) is found to be associated with a decrease in the overall duplicate imaging procedures conducted on patients. As a result, providers who can access radiology images across hospitals exhibit a lower rate of duplicate imaging procedures. However, intra-hospital image sharing (i.e., image distribution to various medical departments within a hospital) is not significantly associated with a reduction in the overall rate of duplicate imaging. This result can be attributed to the possibility that EMR applications may already serve as a medium to share medical information of patients across various departments within a hospital. Providers

[5] We manually collected information about the age of hospitals. We searched hospital web pages, CEO letters, and also called some of these hospitals by phone to obtain their year of opening as a hospital provider.

may access necessary information (such as radiology imaging reports) through their EMR systems which may help reduce the rate of duplicate imaging procedures.

To the best of our knowledge, our study represents the first attempt to empirically explore the antecedents of duplicate tests using a large panel of patient data tracked across a relatively long period of time. In this research, we not only account for simultaneity between hospital switching and duplicate testing, but also include non-clinical, payer incentives as well as hospital admission and health status of patients. Our results indicate that imaging tests are overly utilized on Medicare compared to self-pay patients. Since duplicate tests represent a significant portion of the inefficiency in the U.S. healthcare system, balancing economic incentives among patients, providers and insurance companies may result in improved efficiencies through better allocation of resources, i.e., utilization of imaging equipment in our context. Emergency admissions and hospital switching also raise concerns regarding the inability to retrieve and share information across disparate stakeholders in the fragmented information environment of the current healthcare system in the US.

Selected References

Bardhan, I.R., J-h. Oh, Kirksey, K., and Zheng, Z. 2011. "A Predictive Model for Readmissions of Patients with Congestive Heart Failure: A multi-hospital study," Proceedings of the *Thirty-second International Conference on Information Systems*, Shanghai, December 2011

Bates, D.W., Boyle, D.L., Rittenberg, E., Kuperman, G.J., Ma'Luf, N., Menkin, V., Winkelman, J.W., and Tanasijevic, M.J. 1998. "What Proportion of Common Diagnostic Tests Appear Redundant?," *The American Journal of Medicine* (104:4), pp. 361-368.

Booth, N. 2003. "Sharing Patient Information Electronically Throughout the Nhs: Time for All Clinicians to Get Involved," *BMJ: British Medical Journal* (327:7407), p. 114.

Brill, S. 2013. "Bitter Pill." Time.

Flanders, A.E. 2009. "Medical Image and Data Sharing: Are We There Yet?," *RadioGraphics* (29:5), pp. 1247-1251.

Hendee, W.R., Becker, G.J., Borgstede, J.P., Bosma, J., Casarella, W.J., Erickson, B.A., Maynard, C.D., Thrall, J.H., and Wallner, P.E. 2010. "Addressing Overutilization in Medical Imaging1," *Radiology* (257:1), pp. 240-245.

Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R., and Taylor, R. 2005. "Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, and Costs," *Health Affairs* (24:5), pp. 1103-1117.

Johnson, K.B., Unertl, K.M., Chen, Q., Lorenzi, N.M., Nian, H., Bailey, J., and Frisse, M. 2011. "Health Information Exchange Usage in Emergency Departments and Clinics: The Who, What, and Why," *Journal of the American Medical Informatics Association* (18:5), pp. 690-697.

LaBorde, D.V., Griffin, J.A., Smalley, H.K., Keskinocak, P., and Mathew, G. 2011. "A Framework for Assessing Patient Crossover and Health Information Exchange Value," *Journal of the American Medical Informatics Association* (18:5), pp. 698-703.

Lu, M.T., Tellis, W.M., Fidelman, N., Qayyum, A., and Avrin, D.E. 2012. "Reducing the Rate of Repeat Imaging: Import of Outside Images to Pacs," *American Journal of Roentgenology* (198:3), pp. 628-634.

Table 1. Results of Two-step Heckman Approach for Duplication Percentage

	OLS[1] *DV: DupPer*		*Tobit*[2] *DV: DupPer*		*Tobit*[3] *DV: DupPer*	
Intercept	15.53**	(6.334)	-703.6***	(176.9)	-1027.67***	(189.97)
VisitDHSS	7.251*	(3.729)	195.1	(118.8)	228.90**	(114.29)
VisitDS	81.81***	(5.422)	2489.9***	(322.0)	2344.21***	(271.35)
InvMills_DHSS	-1.889	(1.950)	-55.46	(62.77)	-80.69	(62.39)
InvMills_DS	-46.35***	(2.593)	-1427.9***	(187.4)	-1366.98***	(150.19)
InsSelfPay	-4.005**	(1.884)	-174.4***	(57.43)	-114.47*	(62.02)
InsPrivate	1.955	(2.776)	104.0	(80.24)	-58.53	105.68
InsOther	3.964***	(1.382)	130.2***	(38.31)	30.26	(48.92)
InsMedicareA	-4.454***	(1.331)	-144.4***	(40.52)	-91.17**	(41.99)
InsMedicaid	9.408***	(2.100)	214.6***	(57.52)	192.46***	(65.16)
TypEmergency	27.04***	(1.276)	674.1***	(55.78)	962.07***	(78.75)
TypOther	11.47***	(1.246)	358.1***	(42.30)	699.69***	(70.91)
SrcTransfer	7.580	(7.363)	193.2	(198.3)	-938.00***	(2.47)
SrcOther	-17.24***	(1.861)	-432.8***	(63.37)	-572.49***	(131.09)
PtFemale	-0.928	(0.784)	-41.88*	(23.28)	-29.55	(23.09)
PtWhite	1.909**	(0.938)	90.42***	(29.17)	76.05***	(28.87)
PtAge	0.0713**	(0.0317)	2.614***	(0.963)	2.79***	(0.99)
DaysSince	-0.0387**	(0.0173)	-1.487***	(0.548)	-1.69***	(0.53)
HsCMI	0.204	(2.787)	-77.06	(82.74)	19.55	(77.82)
log(HsBeds)	-1.234*	(0.747)	-50.69**	(21.05)	-55.21***	(20.38)
HsTeach	-0.633	(1.720)	20.82	(50.46)	17.78	49.11
HsUrban	-0.975	(1.480)	57.16	(46.59)	-17.91	44.11
Sigma	-	-	660.7***	(40.04)	568.23***	(37.64)
R^2 (Psuedo for Tobit)	0.190		0.143		0.141	
Model p-value	0.000		0.000		0.000	
N	9403		9403		9403	

Std. error in parentheses, Time fixed effects included, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
[1]Asymptotic standard errors are reported for 2nd stage OLS estimation
[2]Robust standard errors are reported for 2nd stage Tobit estimation
[3]Heteroscedasticity consistent are reported for 2nd stage Tobit estimation

Table 2. Two-stage Least Squares Results for Difference-in-Difference Estimation

Parameters	**Intra-hospital Info. Sharing**	**Inter-hospital Info. Sharing**
Treatment^	28.76 (44.92)	-0.20 (6.86)
Post	10.68 (45.09)	29.28*** (10.29)
Treatment^ * Post	-10.54 (50.51)	-55.59*** (18.74)
Model	R^2 = 0.255 N = 8508	R^2 = 0.185 N = 1258

HsAge, HsAge2, HsAge3 are used as instruments for endogenous Treatment
Control variables included are insurance type, admission source, admission type, patient age, gender, race, and hospital characteristics

Online Care vs. Traditional Care: A Case Study of eVisit Usage

Changmi Jung
The H. John Heinz III College,
Carnegie Mellon University,
Pittsburgh PA, USA
Changmi@andrew.cmu.edu

Rema Padman
The H. John Heinz III College,
Carnegie Mellon University,
Pittsburgh PA, USA
rpadman@cmu.edu

Abstract

Ongoing digital transformation of healthcare delivery, particularly primary care, is partly being driven by the increasing gap between provider availability and patient demand for high quality, low cost, easily accessible care. Advances in information and communication technologies are enabling a transformation from traditional face-to-face encounters between clinicians and patients to online medical consultations, or eVisits, and empowering patients to self-manage their health by accessing their clinical information via portal sites. In this preliminary case study of a single primary care practice with multiple locations, we examine the potential impact of eVisit and portal usage on office visits by analyzing data from a major health system in Western Pennsylvania. Preliminary results indicate that 10 additional eVisits can reduce office visits by 2.6 and 100 more portal access is associated with 2.5 more office visits, providing new directions for investigating generalizable insights about digitized care delivery channels.

Keywords: Consumer health informatics, e-Health, patient portal, online medical consultation, online primary care delivery

1. Introduction

Digital transformation of medical care delivery, particularly primary care, is partly being driven by the increasing gap between provider availability and patient demand for high quality, low cost, easily accessible care [1, 2]. Online medical consultation is an increasingly accepted and adopted form of digitized care delivery, which can be used between patients and primary care physicians, or primary care physicians and specialty doctors. In this study, the term online medical consultation refers to the former use in which the initiator of the interaction is a patient with enquiries regarding non-urgent, acute symptoms. This is also called e-health, which is defined as health services and information delivered via internet and related technologies [3].

The growing interest in online medical consultation is attributed to several beneficial factors: immediacy of care, freedom from scheduling conflicts, affordable cost and convenience to patients, and ease of implementation and efficient/flexible time management options for physicians [2]. Many healthcare organizations are using this new service delivery strategy via internet patient portals to empower patients to access their clinical information and interact with their healthcare provider [4]. These patient portals enable users to practice self-service: check their own medical records, schedule the appointments, request prescription renewal and interact with their doctors electronically [5]. Thus, patients have an improved ability to actively participate in their health care and potentially achieve more favorable health outcomes [6, 7].

As patient portals and online medical consultations (eVisits, hereafter) become more widely available and increasingly sought, it is important to assess whether these technology-enabled medical services have brought intended changes to the traditional face-to-face encounters, particularly to provide an alternative for onsite office visits and phone-based care for non-urgent, acute conditions, and how those services affect clinic visits, specifically with regard to the volume of visits. Hence, we examine whether there have been changes in the traditional channel of clinical care, or more specifically, whether, increased eVisit use and portal usage have reduced clinic visits.

Thus, the purpose of this study is to analyze changes in office visit utilization of portal and eVisit users by specifically comparing pre-eVisit and post-eVisit periods. We analyze patients' utilization of office visits, eVisits, and patient portal for pre- and post- eVisit deployment at a primary care practice with three locations that is associated with a major health system in Western Pennsylvania. We examine the difference between eVisit users and non-users, and empirically test the effect of eVisit usage and patient portal utilization on the frequency of face-to-face encounters.

2. Patient Portal and eVisit

Similar to other patient portal services, our study site allows patients to take a more active role in their own health by providing secure and convenient online access to their electronic health information [2, 9]. Users can review clinical information such as health history, past visits and test results, as well as business services such as appointment scheduling, pre-registration, prescription renewal, and reminders for future appointment/health maintenance. The portal utilizes the underlying technical infrastructure and solutions offered by Epic Corporation via the EpicCare Electronic Medical Record (EMR) and MyChart patient portal. It has been in use for more than five years, has more than 100,000 current enrollees, and continues to grow along the two dimensions of users and services [9, 10].

eVisit was piloted as a new clinical service within the portal in August 2008, providing patients with an online consultation through a series of secure message exchanges with a physician [9, 10]. A limited set of eight acute, non-urgent health conditions were handled via this initial eVisit service, which included sinus/cold, cough, back pain, diarrhea, urinary symptoms (UTI), red eye, vaginal irritation, as well as an 'other' category (this category allows free text entry). The pilot system was deployed at a single practice with three locations where hundred percent of physician participation was achieved very quickly. Unlike free text messaging used in many online medical messaging services, eVisit uses structured templates for each condition which creates formatted documentation for the consultation. Thus, the template captures best practices and produces clearer evidence for communication by capturing necessary information, and is easy to use for both patients and physicians. It is integrated with practice workflow, and thus provides access to information that is stored in the Electronic Medical Record (EMR). A successfully completed eVisit is finally documented in the EMR as well [10].

In April 2009, eVisit was deployed in three additional practices and reimbursed by a few health plans. The system has been expanded to more practices since September 2010. The physicians and staff at the offices encouraged patients to sign up for the patient portal and to use eVisits for the treatment of the specified episodic illnesses. The use of the service was purely voluntary by patients and providers. The eVisit process starts when a patient, experiencing an acute, non-urgent health condition, chooses to complete and submit an eVisit on the patient portal. The completed message goes to a support staff pool that forwards the eVisit to a participating physician; if primary care provider is unavailable, an assigned on-call physician takes the responsibility to act on the submission. Once the physician reviews and responds with a diagnosis and treatment plan, the patient is alerted via their personal email to login for checking the response. The patient may choose to have further message exchanges before the physician closes the encounter. Once the encounter is closed, the support staff is notified and a claim is submitted for reimbursement. The details of the process flow are found in the literature [9].

3. Research Questions

Patient portal use may reduce the volume of patients' office visits since the services provided by the portal assist patients to self-service their healthcare needs, which could prevent unnecessary visits to clinics. The eVisit service is intended to resolve issues with non-emergent, acute symptoms, and thus theoretically, it is reasonable to expect that to some degree, eVisit substitutes for face-to-face office visit. Recent studies reported that increased web portal usage at a large health plan increased volume of phone-based customer service rather than act as a substitute [11], and users of online access to their health records have increased health care utilization [12].We hypothesize the following relationships accordingly:

H1: Patients with increased *portal usage* experience positive change in office visit frequency.

H2: Patients who utilize *eVisit* service have decreased office visits.

4. Data

The data used in this study is from the primary care practice where the pilot service was deployed. This practice is an appropriate study target since a pilot eVisit service was offered by the health system at this practice only, so there was approximately a 7 month lag until the reimbursement/copay structure was implemented, which allowed patients to explore the service during the free trial period. Additionally, all physicians in the practice participated in the eVisit service since deployment, and therefore, the confounding effects of individual physician's behavior/opinions on patients' choice regarding office visit vs. eVisit is minimized.

The study duration includes 14 months before and the same time period after eVisit deployment. In order to minimize seasonality effects, the same 14 months, but in different years, are selected as study period; i.e., pre-eVisit period is from April 1, 2007 to May 31, 2008 and post-eVisit is from April 1, 2009 to May 31, 2010. The study population includes all patients with office visit records in the pre-period to ensure that the target patients are members of the clinic. A possible confounding factor is that eVisit users may have better accessibility to internet or are more familiar with computer/internet use than non-eVisit users (hereafter, non-users). To minimize the potential bias due to differences in the comfort level of internet access and computer use to begin with among patients, we restrict our study population to the patients enrolled in the clinic's patient portal, and also control each patient's web preference. In summary, 4,794 portal users and their office visits, portal access logs, and eVisit records for the pre/post period are included in our analysis. Patient demographics, medical conditions, and payer information are included as well. Among the total 4,794 portal users, 290 patients sought physician consultation via the eVisit service during the 14 post-period months.

5. Method

We conduct descriptive analysis to first compare the two patient groups: eVisit users and non-users among the study population, followed by regression analysis to estimate the impact on office visits. Comparison factors include change in office visit volume/frequency and patient portal access for non-eVisit purpose (post-frequency minus pre-frequency), demographics, insurance coverage (whether the payer covers eVisit), and patients' health condition. The complexity of health condition is measured as the number of distinct comorbidities out of 19 conditions defined in Charlson Comorbidity Index (CCI). Demographic characteristics include gender, age, marital and employment status, ethnicity, and proximity to the clinic location. Simple two sample t-tests provide insights into the significance of factors that differentiate patient groups. For the hypotheses test, we use OLS (Ordinary Least Squares) regression to estimate the impact of patient portal and eVisit usage on face-to-face office visit encounters with physicians. The unit of analysis is a patient.

5.1 Specifications for testing H1: The dependent variable, the change in office visit frequency, is computed as the difference between total post-period office visits and total pre-period office visits. Explanatory variables are portal access frequency, eVisits usage, patients' health status, insurance, and demographic characteristics. The eVisit usage is the total number of eVisit submissions by each patient during the post period since the eVisit service was not available during the pre-period. *HealthStatus* is represented by the health complexity described earlier; it is computed as the total number of comorbid conditions if less than or equal to 2, and 3, if total number of comorbid conditions is more than 2. We control patients' web preferences (*WebPref*) by including a proxy value that compares the number of free, online medical advice functions that are utilized by each patient via the portal and the number of telephone calls made by the same patient (thus comparing online access vs. telephone access); this variable has a value of 1 for patients with more medical advice requests than telephone calls, 0.5 for patients who have the same volume of medical advice requests and calls, and zero for patients who have more telephone calls than medical advice or both volumes are zero. Insurance information is implied by two variables; *Insurance* is a binary variable with a value of 1 if the patient holds any type of health insurance, and *Cover* is a binary variable with value of 1 if the patient's insurance covers online consultation. Demographic characteristics include patients' gender, employment and marital status, and ethnicity. Since healthcare providers' readiness/attitude or recommendation influence patients' adoption of telemedicine technology [2], we add physician fixed effect with White cluster robust error for the physician. The regression model is as follows:

$$\Delta OfficeVisit_i = \alpha_0 + \alpha_1 Change\ in\ Portal_i + \alpha_2 eVisit_i + \alpha_3 HealthStatus_i + \alpha_4 WebPref_i + \alpha_5 Insurance_i + \alpha_6 Cover_i + \beta Demographic_i + \delta_1 Physician_i \quad (1)$$

(where i =patient, *Demographic* = vector of demographic variable values of patient i, β= vector of demographic variable coefficients)

5.2 Specifications for testing H2: In order to estimate the impact of availability of eVisit service on traditional face-to-face clinic visits, we apply Difference in Difference (DID) method. There was no other intervention during the study period other than eVisit implementation, and thus it is appropriate to apply DID to observe how the new online consultation service affects the volume of traditional encounters. The unit of analysis is pre or post-period visits in total for a patient, and thus each patient has two records of data, one for each time period. The response variable is the number of face-to-face clinic visits in pre/post-period, and explanatory variables are time dummy, patient type which is a binary variable with a value of 1 when the encountering patient is an eVisit user, 0 otherwise, and DID estimator which is a product of time dummy and patient type. Time dummy is a binary variable with a value of 1 for the post period. Patient characteristics are included as covariates in order to account for the heterogeneity of the patients due to the limitation of the study setting regarding self-selection issue. Instead of health status, the number of problems noted in pre/post periods, respectively, is included as time variant factor. The specification of model is as follows:

$$OfficeVisit_{it} = \alpha_0 + \alpha_1 Time\ dummy_{it} + \alpha_2 Patient\ type_i + \alpha_3 Time\ dummy_{it} \times Patient\ type_i + \alpha_4 Problems_{it} + \alpha_5 PortalAccess_{it} + \beta Demographics_i + \delta_1 Physician_i \quad (2)$$

(where i = patient, t = 1 if post-period, *Demographics* = vector of demographic variable values of patient i, and β= vector of demographic variable coefficients)

The outcome variable is discrete, thus one of the reasonable analysis methods is a Poisson regression. However, the outcome value's variance exceeds its mean value, which does not satisfy the assumption of Poisson regression in which those two values are equal, hence we use Negative Binomial regression which can be applied to over-dispersed data [12].

6. Results

A side-by-side comparison of the two groups of patients is shown in Table 1. Column 1 provides summary statistics on the patients who used eVisit during the post period, and the Column 2 presents the same on portal users who never used eVisit service. eVisit users are younger on average, predominantly female, and less likely to be retired, which has a strong correlation with age. Also, a higher percentage of eVisit adopters hold insurance that covers online consultation, which implies that cost concerns may affect their decision to utilize the service. This is mainly attributed to the fact that most patients in our study are covered by at least one health insurer. eVisit users appear to access the patient portal site more frequently and reside in zip codes different from the clinic. This may imply that patients who are geographically distant from the clinic see greater benefits from using eVisit service. The two groups are similar on the health complexity dimension.

The face-to-face encounters decrease for eVisit users with more eVisit encounters (Table 2). The negative difference between pre and post office visit volume show larger gap for patients who utilized eVisits more frequently. This implies that eVisit may be substituting for traditional clinic visits, supporting hypothesis 2. However, the opposite trend is found in the change in office visit volume for patient portal users who do not use online consultation (Table 3). As the portal access frequency increases in post-period relative to pre-period, the negative difference in office visit volume becomes smaller, and finally changes to positive difference (0.08). Simple comparison tables (Table 2 and Table 3) indicate potential support of both hypotheses 1 and 2.

The estimated effects of patient portal usage and eVisit use on the change in office visit volume in Table 4 show that both factors have significant influence on the patient's physical office visit, but in opposite directions. As indicated in Table 3, higher portal access is associated with increasing face-to-face encounters; if a patient's portal access increases by 100, the patient is likely to have 2.5 more office visits and vice versa. This supports our first hypothesis. Based on the result in Table 2, the eVisit service may have substitution power for office visit; 10 additional eVisits can reduce office visits by 2.6. Health complexity and most demographic characteristics of patients do not have statistically significant effect on the change in office visit volume whereas patients' web preference affects the office visit volume significantly. Patients who prefer using the online service to phone calls when resolving simple issues related to their medical records or medications have, on average, 1 less office visit than their counterparts.

Table 1 – Demographic Comparisons

	eVisit User	Portal User	p-value
Patients	290	4,504	N/A
Pre-Office visit	4.83	4.54	0.265
Post-Office visit	4.12	3.65	0.013*
Average age	45.9	51.3	0.000***
Female (%)	74.8	58.7	0.000***
Married (%)	71.4	74.4	0.254
Single (%)	19.0	16.1	0.196
Fulltime employed (%)	64.1	58.8	0.074
Retired (%)	6.6	15.4	0.000***
Unemployed (%)	17.2	16.1	0.651
White (%)	96.2	93.9	0.107
Black (%)	0.7	0.7	0.967
Insurance (%)	95.5	95.1	0.745
Coverage (%)	22.8	16.4	0.005**
Increase in PP access	25.8	11.0	0.000***
Comorbidity condition	0.44	0.50	0.124
Web Preference	0.040	0.026	0.213
Zip code (%)	17.6	29.4	0.000***

Table 2 – Office visit frequency of eVisit users

	eVisit = 1	eVisit = 2	eVisit = 3	eVisit > 3
Patients	213	48	20	9
Pre-Office visit	4.58	5.42	4.80	7.67
Post-Office visit	3.96	4.67	3.70	5.89
Difference	-0.62	-0.75	-1.10	-1.78

Table 3 – Office visit frequency of non-users

	Change in Patient Portal Access				
	≤ 0	1–10	11–20	21–30	> 30
Patients	723	2,372	684	312	413
Pre-Office visit	4.86	3.93	4.67	5.36	6.67
Post-Office visit	2.50	3.19	4.03	4.86	6.75
Difference	-2.36	-0.74	-0.65	-0.50	0.08

Statistical significance: *p<0.05, **p<0.01, ***p<0.001

Table 4 – Estimates of influence of portal and eVisit usage

	OLS estimate	
Change in portal access	0.0250***	(0.005)
eVisit	-0.241*	(0.091)
Health complexity	-0.088	(0.127)
Web preference	-1.163*	(0.511)
Insurance	0.497	(0.229)
Cover	-0.021	(0.087)
Female	-0.078	(0.155)
Age	-0.013*	(0.005)
Patient demographics: fulltime, unemployed, retired, married, single, white, black, zipcode		
Observation	4,794	
Adjusted R-square	0.0263	

Table 5– DID estimates of effect of eVisit use

Explanatory variables	Coefficients	
Timeline	-0.147***	(0.018)
Patient type	0.026	(0.021)
Timeline× Patient type	-0.104**	(0.033)
Portal access	0.0006	(0.0004)
Patient characteristics		
Observation	9,588	

Notes: standard errors are in parentheses; Statistical significance: *p<0.05, **p<0.01, ***p<0.001

The DID estimates in Table 5 reinforce support for hypothesis 2. The eVisit service use (Product of Timeline and Patient type) is a significant factor affecting the volume of face-to-face encounters, whereas the intensity of patient portal usage is not significantly associated with the frequency of traditional encounters. eVisit users have, on average, 0.1 higher drop (post – pre) in the log frequency of face-to-face visits compared to non-users. Thus, DID analysis produces similar results as the previous model.

7. Discussion and Conclusion

This preliminary case study seems to indicate that both patient portal access and eVisits are significantly associated with the change in office visit volume, with higher magnitude in the case of eVisit usage. Increased frequency of portal use is related to higher number of office visits, and the higher the eVisit use, the lower the office visits. Since the data does not include the purpose of portal access, it is difficult to infer the cause and direction of influence between portal access and face-to-face encounters; each may affect the volume of the other, resulting in potential endogeneity effects. This is currently being investigated using simultaneous equation models. When physician heterogeneity is controlled, an additional eVisit is associated with decreased office visit, which indicates that there may be a substitution effect with eVisit. Additional DID analysis supports this conclusion.

An important limitation that needs to be emphasized is the number of practices in the study. Only one practice deployed the service with a 7-month pilot period ahead of the roll-out, even though the service was expanded to many more practices later. While this practice had three separate locations, all providing access to evisits, our study results are difficult to generalize with the current data. Ongoing research is examining these assumptions and questions with data from additional practices and locations.

Acknowledgments

We are grateful to the health system and the physicians, program director and data management staff associated with their e-health services for providing the data for this study and sharing insights on the operational process.

8. References

1. Margolinus D, Bodenheimer B. Transforming Primary Care: From Past Practice to The Practice of The Future. Health Affairs 2010; 29 (5): 779-784.
2. Jung C, Padman R (2014), "Disruptive Digital Innovations in Health Care Delivery: The Case for Patient Portals and Online Clinical Consultations", Invited Book Chapter, forthcoming in The Handbook of Service Innovation (Editors: R. Agarwal, W. Selen, EG. Roos, R. Green).
3. Eysenbach G. What is e-health? J Med Internet Res 2001; 3 (2): e20.
4. Wilson EV, Lankton NK. Modeling Patients' Acceptance of Provider-delivered E-health. JAMIA 2004; 11 (4): 241-248.
5. Weingart SN, Rind D, Tofias Z, and Sands DZ. Who uses the patient internet portal? The PatientSite experience. JAMIA 2006; 13 (1):91–95.
6. Liederman EM, Lee JC, Baquero VH, and Seites PG. Patient-Physician Web Messaging: The Impact on Message Volume and Satisfaction. JGIM 2004; 20 (1): 52-67.
7. Adamson SC, Bachman JW. Pilot Study of Providing Online Care in Primary Care Setting. Mayo Clinic Proceedings 2010; 85 (8): 704-710.
8. Carrell D, Ralston J. Variation in Adoption Rates of a Patient Web Portal with a Shared Medical Record by Age, Gender and Morbidity Level. AMIA Annu Symp Proc 2006; 871-876.
9. Jung C, Padman R, Paone S, and Shivchik G. Who are Portal Users vs. Early Adoptrs? A Preliminary Analysis. AMIA Annu Symp Proc 2011; 1070-1079.
10. Padman R, Shevchik G, Paone S, Dolezal C, Cervenak J. eVisit: A Pilot Study of a New Kind of Healthcare Delivery. Stud Health Technol Inform 2010; 160: 262-6.
11. Kumar A, Telang R. Does the Web Reduce Customer Service Cost? Empirical Evidence from a Call Center. Information Systems Research 2012; 23 (3): 721–737.
12. Palen TE, Ross C, Powers JD, Xu S. Association of Online Patient Access to Clinicians and Medical Records with Use of Clinical Services. JAMA 2012; 308 (19): 2012-2019.
13. Hoffmann JP. Generalized Liner Models: An Applied Approach. Boston: Pearson Education; 2004.

Users' Roles and Involvement in Online Health Communities —a Social Support Perspective

Xi Wang and Kang Zhao
The University of Iowa
xi-wang-1@uiowa.edu, kang-zhao@uiowa.edu

Abstract

Online health communities (OHCs) have become a major source of social support for people with health problems. Members of OHCs interact online with others who face similar problems and are involved in different types of social supports, including information support, emotional support and companionship. Using a case study of an OHC among breast cancer survivors, we study users' roles in different types of social supports and their involvement in the community. Machine learning techniques are first used to identify the types of social support embedded in each post. Then we generate each user's contribution profile by aggregating the user's activities in different types of social support and reveal the role the user plays in the OHC. Moreover, by comparing online activities for users with different roles, we illustrate that users' roles in seeking and providing various social supports are related to their levels of involvement in an OHC.

Keywords: Social Support, Online Health Community, User Profiling, Machine Learning, Text Mining.

1. Introduction

Nowadays more and more people use Internet to satisfy their health-related needs. According to a study by the Pew Research Centre, 80% of adult Internet users in the U.S. use the Internet for health-related purposes. Among them, 34% read health-related experiences or comments from others (Fox 2011). Compared with traditional health-related websites that only allow users to retrieve information, online health communities (OHCs) increased members' ability to interact with peers facing similar health problems and as a result better meet their immediate needs for social support. It is estimated that 5% of all Internet users participated in an OHC (Chou et al. 2009).

While people use OHCs for a wide range of needs, obtaining psychosocial support is one of the key benefits of involvement in OHCs (Kim et al. 2012; Rodgers and Chen 2005). Research has found that such support can help patients adjust to the stress of living with and fighting against their disease (Dunkel-Schetter 1984; Qiu et al. 2011; Zhao et al. 2014) and is a consistent indicator of survival (McClellan et al. 1993). An OHC also serves as an outlet for users' emotional needs and improve their offline life (Maloney-Krichmar and Preece 2005).

Literatures on social support suggest that OHCs mainly feature three types of social support: informational support, emotional support, and companionship (a.k.a., network support) (Bambina 2007; Keating 2013). *Informational support* is the transmission of information, suggestion or guidance to the community users (Krause 1986). The content of such a post in an OHC is usually related to advice, referrals, education and personal experience with the disease or health problem. Example topics include side effects of a drug, ways to deal with a symptom, experience with a physician, or medical insurance problems. *Emotional support*, as its name suggests, contains the expression of understanding, encouragement, empathy affection, affirming, validation, sympathy, caring and concern, etc. Such support can help one reduces the levels of stress or anxiety. Companionship or network support consists of chatting, humor, teasing, as well as discussions of offline activities and daily life that are not necessarily related to one's health problems. Examples include sharing jokes, birthday wishes, holiday plans, or online scrabble games. Companionship helps to strengthen group members' social network and sense of communities.

Then is a user's involvement in different types of social support related to her/his involvement in the OHC? A previous study showed that receiving more emotional support is positively correlated with users' longer life span, while receiving informational support is negatively correlated (Wang et al. 2012). However, the study did not consider companionship or differentiate users' behaviors in seeking and receiving support. In this research, we addressed three research questions regarding social support and user involvement in OHCs: (1) Can we use machine learning techniques to detect the seeking and provision of three types of social support embedded in interactions among users; (2) Are there any patterns of users' involvement in different types of social support activities? Or in other words, do users play different roles with regard to social support? and (3) Are different patterns in social support involvement related to users' involvement in an OHC? Our research considered all three major types of social support in OHCs and tried to reveal users' roles in seeking and providing support. The outcome of this research has implication for building and sustaining an active OHC through better thread/post recommendations and community management.

2. Detection of social support from texts

2.1 Dataset and the taxonomy of social supports

In this research, we used Breastcancer.org as a case study. It is a very popular peer-to-peer OHC among breast cancer survivors and their caregivers. With more than 140,000 registered users, the website provides various ways for its members to communicate, including discussion forum, private messaging, friend subscription, listserv, etc. We designed a web crawler to collect data from its online forum, which has 73 discussion boards. Our dataset consists of all the public posts and user profile information from October 2002 to August 2013. There are more than 2.8 million posts, including 107,549 initial posts. These posts were contributed by 49,552 users.

As we mentioned earlier, informational support, emotional support, and companionship are three major types of social supports in OHCs. Thus for each post, we need to determine whether it was seeking informational support (SIS), providing informational support (PIS), seeking emotional support (SES), providing emotional support (PES), or simply about companionship (COM). Note that we did not differentiate seeking and provision of companionship, because the nature of companionship is about participation and sharing. By getting involved in activities or discussions about companionship, one is seeking and providing support at the same time. It is also possible that a post could belong to more than one of the categories above. Table 1 lists example posts for each category and a post that belongs to two categories.

2.2 Data annotation and feature engineering

As it is almost impossible to label all 2.8 million posts manually, we used classification algorithms to decide what kind(s) of social support each post is about. To train the classification algorithm, we leveraged human annotated data. We randomly selected 1,333 (54 initial posts and 1,279 comments) out of our dataset. After basic training on the aforementioned five categories of social supports (SIS, PIS, SES, PES, COM), five human annotators were asked to read each post and decide whether the post is related to one or more categories of social supports.

Table 1 Example posts for types of social support

Social Support Category	Examples
Companionship (COM)	*Kelly Have a wonderful time in Florida, enjoy the sun and fun. Heather I'm loving her new CD. Didn't recognize any of the songs at first, but there are a few now that I find myself singing the rest of the day.* *This game has the poster making a new 2 word phrase starting with the second word of the last post Example: Post : Hand out Next poster: Out cast Next poster: Cast Iron Next poster: Iron Age Now let's begin the game~ Age Old*
Seeking Informational Support (SIS)	*Where do you buy digestive enzymes and what are they called?*
Seeking Emotional Support (SES)	*I feel like everyone else's lives are going forward, they have plans, hopes, aspirations because they feel. I am one of those not yet out of the woods. I was also someone who could never get cancer. I was a good person, exercised, ate well. Good people don't get sick. I have taken the step of antidepressants, they mitigate the damage, but do not block the pain or sadness I feel.*
Providing Informational Support (PIS)	*I had surgery Aug05 for bc recurrance. B4 surgery I had 33 IMRT rads, prior to that had 4A/C & 4 Taxol. I had bc in 2000 & had 37 rads in same general area. Now, my surgery won't heal. Wound doc says there is adema or something on my sternum (shown on recent MRI). My wound has been draining since it broke open in Sept.*
Providing Emotional Support (PES)	*Hope you feel better soon, we are here! Prayers Hugs come from Massachusetts APPLE♥.*
Providing Informational Support (PIS) & Providing Emotional Support (PES)	*I am also the daughter of a 35 yrs BC survivor. Mom is just now going through some more Cancer - alas - they found it in her lung, but it is totally unlikely to be a follow-up of her old BC. I am 45, and was 43 at DX time, my mom was diagnosed at 38... and I am a BRCA2 carrier. Tina, one day at a time. Maybe you'll get good news - it is so hard to wait!!! It is also important to remember that - whatever it is, it is highly treatable, and that YOU WILL SURVIVE too!!! and life goes on after. It will take some time, but it goes on... see my picture? even the hair is back!!! Hugs to all. I am happy you all found your way here, it is a great site for exchanging information, learning and finding support.*

Table 2 The number of post in each category of social support in the annotated dataset.

Social Support Category	Number
Companionship (COM)	435
Seeking Informational Support (SIS)	96
Seeking Emotional Support (SES)	22
Providing Informational Support (PIS)	411
Providing Emotional Support (PES)	249

To control the quality of human annotations, we also added to the pool 10 posts that have been annotated by domain experts. For each post, we only accepted results from annotators whose performance on the 10 quality-control posts is among top 3. Results from the other two annotators were discarded. Then a majority vote was used to determine whether a post is related to a category of social support. Table 2 shows the results of the annotation process.

Table 3 Summary of features for the classifier.

Group	Features
Basic Features	Whether the post is an initial post in a thread
	Whether the post is a self reply
	Length of the post
Lexical Features	Whether the post contains URLs (Y or N)
	Whether the post contains emoticon(s)
	Number of numeric numbers
	Number of Pronouns (e.g., they, we, I)
	Whether the post contains the negation word(s) (e.g., not, never, no)
	Whether the post contains name(s) of city, state, country (U.S.A, Canada, etc.)
	Whether the post contains phrases related to possibility (you must, you might, she had better, etc.)
	Whether the post contains names of drugs related to breast cancer (From http://www.cancer.gov/cancertopics/druginfo/breastcancer)
	Whether the post contains breast cancer terminology (From http://www.breastcancer.org/dictionary)
	Whether the post contains verb related to advice (Need, require, recommend, etc.)
	Whether the post contains emotional words (Love, sorry, hope, worry, etc.)
	Whether the post contains words related to seeking behaviors (Anybody, question, wonder, etc.)
	Whether the post contains words related to daily life topics (Vacation, joke, run, walk, etc.)
Sentiment Features	Frequency of words with positive and negative sentiment
	Objectivity and subjectivity scores
Topic Features	Topic distributions derived from LDA

Users in OHCs may have different writing styles or linguistic preference to express themselves. To capture these characteristics, we examined each post and extracted various types of features for the classifier: basic features, lexical features, sentiment features, and topic features. Table 3 summarizes these features. Many of the features were picked specifically for classification in this context. For example, we included whether a post is an initial post as a feature because many users seek supports by starting a thread. Inside each post, the existences of URLs and emoticons are often related to informational and emotional supports respectively. Similar to the approach used by (Wang 2012), we also checked the usage of phrases in the format of <you/he/she + MODAL verb > to express possibilities, such as "you should", "she could". We considered "he" and "she" in addition to "you", because some posts were created by family members of cancer survivors. To identify the difference between "seeking" and "providing" supports, we included words related to seeking behaviour, such as "question", "wonder" and "anybody". We also hoped that words related to daily life topics and geographical locations can effectively detect companionship. Meanwhile, we used OpinionFinder (Wilson 2005) to find the overall sentiment, as well as subjectivity and objectivity of each post. Besides these hand-picked or dictionary-based lexicons, we also wanted to capture whether the usage of other words and phrases can contribute to the classification. Using unigrams and bigrams is too fine-grained and leads to a feature set with very high dimension. Thus we adopted an approach similar to (Wang 2012) and applied topic-modelling technique Latent Dirichlet Allocation (LDA, with k=20) (Blei 2003) to the content of all posts and generated 20 topics. For each post, LDA gave a topic probability distribution, indicating the probability of this post corresponding to each topic. Such distribution for each post was then included in the feature set.

2.3 Evaluation of the classification

Because there are five categories of social supports and a post may be related to more than one category, we built a classifier for each category. For the classification of each category of social support, we applied various classification algorithms on annotated posts and picked the best performing algorithm (using 10-fold cross-validation). Because posts seeking emotional support accounted for only a small proportion among annotated posts (22 out of 1,333), we oversampled posts seeking emotional support when building the SES classifier. Table 4 compares the performance of different algorithms for the five categories of social support. AdaBoost was chosen to classify COM, PES[1], PIS and SIS, while logistic regression was the best choice for SES. Overall, our classifiers achieved decent performance with accuracy rate over 0.8 in all five classification tasks.

Table 4 Performance of classification algorithms for five categories of social supports.

Social support	Results	Naïve Bayes	Logistic Regression	SVM (Poly Kernel)	Random Forest	Decision Tree	AdaBoost
COM	Accuracy	0.696	0.787	0.783	0.771	0.767	0.804
	AUC	0.839	0.817	0.768	0.848	0.75	0.852
PES	Accuracy	0.713	0.830	0.840	0.830	0.81	0.817
	AUC	0.823	0.787	0.681	0.825	0.687	0.817
PIS	Accuracy	0.753	0.813	0.823	0.767	0.779	0.801
	AUC	0.824	0.83	0.783	0.837	0.717	0.859
SES	Accuracy	0.893	0.901	0.970	0.967	0.963	0.963
	AUC	0.749	0.867	0.656	0.851	0.671	0.668
SIS	Accuracy	0.851	0.880	0.943	0.931	0.937	0.914
	AUC	0.893	0.803	0.745	0.86	0.766	0.869

Table 5 Total numbers of posts in each category of social supports.

Social support category	Number of posts
Companionship (COM)	932,538
Seeking Informational Support (SIS)	284,027
Seeking Emotional Support (SES)	227,188
Providing Informational Support (PIS)	1,034,682
Providing Emotional Support (PES)	497,096

After applying the best-performing five classifiers on the remaining of the 2.8 million posts, each post received 5 labels, each of which indicates whether the post belong to one of the five social support categories. The total numbers of posts in each category are listed in Table 5.

Intuitively, there are more posts to provide support than to seek support. This is what most would expect from a popular OHC with a large and active user base. About 37% of the posts provided informational support, making it the largest group among the five. In other words, providing information support is the most popular activity in the OHC. Companionship posts constitute the second largest group, which suggests that members of the OHC did form a strong sense of community and discussed many issues other than cancer. In addition, 197,956 posts were predicted to

[1] Although the results of accuracy and ROC area of random forest are slightly better than AdaBoost for the PES classifier, the random forest classifier has much worse recall and f-measure. Thus we decided to choose AdaBoost.

provide informational and emotional support at the same time, representing the largest group with more than one category of social support.

3. User profiling and roles

After estimating the nature of social support in each post, we can then build a profile for each user by aggregating her/his posts by their social support categories. We represented each user's social support involvement with a 1×5 a vector. Each element in the vector is the percentage of the user's posts in a social support category. For example, user Mary has published 10 posts, with 3 companionship posts, 4 posts providing emotional support, 2 posts providing informational support, 1 post seeking emotional support, and no post seeking informational support. Then she will have a vector of <0.3, 0.4, 0.2, 0.1, 0>.

With social support distribution vectors of 47,581 users, we applied the classic K-means clustering algorithm to divide users into k groups, so that the users with similar social support distributions would belong to the same cluster. To find the best grouping of users, we tested various K values (from 2 to 20) and clustering results with Davies-Bouldin Index (DBI) (Davies and Bouldin 1979). DBI is defined as Equation 1, where $D_{intra}(C_i)$is the average distance from all the other users in cluster C_i to the centroid of C_i, and $D_{inter}(C_i, C_j)$ is the distance between centroids of $C_i$and C_j. Euclidean distance was used for this study. Generally speaking, DBI prefers smaller groups, for the value of intra-cluster distance is lower in the smaller group, and penalizes short inter-cluster distances. Therefore, the solution with the lowest DBI provides relative balance of small clusters and long distances between every two clusters.

$$DBI = \frac{1}{k}\sum_{i=1}^{k} max_{j:i\neq j}\left\{\frac{D_{intra}(C_i)+D_{intra}(C_j)}{D_{inter}(C_i,C_j)}\right\} \qquad \text{(Equation 1)}$$

We summarized the DBIs for different K values in Table 6. K=7 yielded the lowest DBI value and hence the best clustering results. Centroids for each of the 7 clusters are shown in Table 7.

Table 6 The DBIs for the K-means clustering with various K values

K	DBI	K	DBI
2	1.485806117	12	0.932705779
3	1.183743056	13	0.914857805
4	1.147831469	14	1.148624229
5	1.002816698	15	0.94766141
6	0.962159462	16	0.915504995
7	0.89111499	17	0.895295641
8	0.977535018	18	0.907029696
9	0.960697173	19	0.935044276
10	0.940555275	20	1.001204328
11	0.904557568		

From Table 7, we can see that, intentionally or not, OHC users do have different patterns in social support involvement and thus play different roles in the community. Some users' posts focused on one major category of social support. For example, users in cluster 0 published an average of 96.55% of their social support posts are to provide informational support. They obviously act as *information providers* in the community. Similarly, cluster 1 is for *community builder*s with 64.92% of the supports in companionship, and cluster 4 consists of *emotional sup-*

port providers. The two smallest clusters are for seekers: cluster 3 for *information seekers* and cluster for *emotional support seekers*. Meanwhile, users in cluster 2, the largest cluster of the seven, are *all-around contributors* with relatively balanced profiles in each social support category. Cluster 5 represents a group of *information enthusiasts*, who focus mainly on informational support, both seeking and providing.

Next we investigated how users in different groups are involved in the OHC. The level of involvement was measured by two metrics: productivity (i.e., a user's total number of posts) and life span (i.e., the number of days between a user's first and last post). Fig 1(a) compares the distributions of productivity for users in the 7 clusters. The curves suggest that community builders in cluster 1, albeit a small group of users, and all-around contributors in cluster 2 are the most productive members. By contrast, those who mainly seek support (informational or emotional) in clusters 3 and 6 published fewer posts than others. Fig 1(b) points to similar conclusions: those in clusters 1 and 2 stayed with the community for the longest time, while support seekers in clusters 3 and 6 have relatively short life span. Overall, those who are more actively involved in companionship tend to get involved in the community, while those who only seek support are more likely to "churn". Also, emotional support providers in cluster 4 are more involved that information providers in cluster 0.

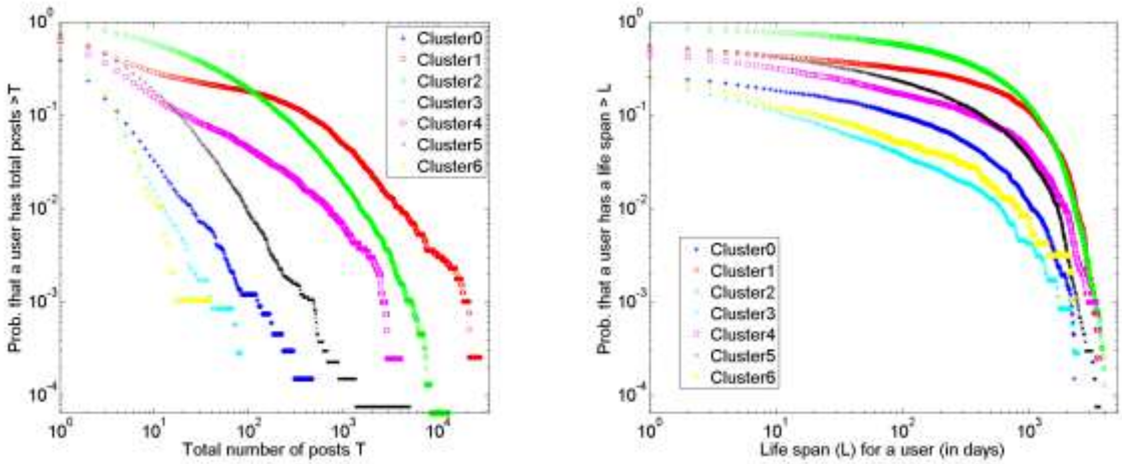
Table 7 Centroids of user clusters

Social Support	All users	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
COM	0.1126	0.0042	0.6492	0.1271	0.0154	0.0504	0.0408	0.0404
PES	0.1178	0.0074	0.0833	0.1511	0.0053	0.612	0.0315	0.0351
PIS	0.4422	0.9655	0.1277	0.4762	0.0152	0.2394	0.4369	0.0325
SES	0.0743	0.0067	0.0349	0.1245	0.0107	0.0481	0.0494	0.5868
SIS	0.2531	0.0162	0.1049	0.1211	0.9534	0.0501	0.4414	0.3052
# of users	47581	6647	3923	15336	3502	3994	13225	954
% of users		14%	8%	32%	7%	8%	28%	2%

To better clarify the differences among the clusters, we use Two-sample Kolmogorov-Smirnov test (K-S test) to compute the statistical gaps between every two clusters in both productivity and life span. Two-sample K-S test, which is used to compare whether two one-dimensional probability distributions are different, is defined as equation 2.

$$D_{n,n'} = \sup\left|F_{1,n}(x) - F_{2,n'}(x)\right| \qquad \text{(Equation 2)}$$

Where $F_{1,n}(x)$ and $F_{2,n'}(x)$ are the empirical distribution of a metric for two groups. The closer the result is to 0, the more likely the two samples are drawn from the same distribution. In Table 8, the upper triangular matrix (shaded area) shows the K-S statistics for productivity between every pair of user clusters in, and the lower triangular matrix shows the K-S statistics for life span. For both metrics, the difference between clusters 2 and 3 is the greatest, which is consistent with what we observed in the Fig 1. In addition, p-values for all K-S tests are less than to 0.001, suggesting statistically significant differences among all clusters' distributions of both involvement metrics.



(a) Numbers of posts *(b) Life span*

Figure 1.Complementary cumulative distributions of involvement metrics for the users in different clusters

Table 8 K-S statistics on involvement metrics for each pair of user clusters. The shaded area is for the comparison on number of the posts. The unshaded area is for life span. All values are significant at p=0.001 value.

	Cluster 0	C 1	C 2	C 3	C 4	C 5	C 6
Cluster 0	-	0.329	0.660	0.056	0.235	0.358	0.062
C 1	0.278	-	0.348	0.321	0.148	0.178	0.301
C 2	0.602	0.348	-	0.673	0.465	0.409	0.653
C 3	0.080	0.320	0.665	-	0.219	0.317	0.071
C 4	0.213	0.148	0.457	0.230	-	0.124	0.191
C 5	0.330	0.112	0.363	0.334	0.117	-	0.286
C 6	0.062	0.305	0.650	0.041	0.200	0.306	-

4. Conclusions and future work

This research analyzed users' patterns of involvement in different types of social support and how such patterns are related to their involvement. Using an OHC for breast cancer as a case study, we built classification models to detect the nature of social support in each post. After aggregating each user' posts, we grouped users based on their social support involvement patterns and discovered seven different user roles in the OHC. By comparing involvement among users in different groups, we found that those with high level of involvement are actively involved in companionship. In other words, the more users talk about issues that are not directly related to health, the more likely they will stay and contribute more. Also, similar to the study by (Wang et al. 2012), our results also illustrated that those who provide a lot of emotional support tend to be more involved than those who only provide information support.

The outcome of our study can shed lights on the design and management of an OHC. For example, a thread/post recommender that leverages users' roles in the community can help information providers quickly find threads that are seeking information. To keep an OHC active and sustainable, community managers may want to encourage or initiate companionships activities, such as birthday greetings, holiday plan discussions, gardening tips, online scrabble games, etc.

Admittedly, this research is still at an early stage. We only examined the relationship between users' own contributions and their involvement in OHC. We are planning to explore how the exposure to different types of social support correlates with users' involvement. We are also interested in whether a user's role changes over time. In addition, building a predictive model of involvement would also be an interesting direction for future research.

References

Bambina, A. 2007. *Online social support: the interplay of social networks and computer-mediated communication*, Youngstown, N.Y.: Cambria Press.

Chou, W. Y., Hunt, Y. M., Beckjord, E. B., Moser, R. P., and Hesse, B. W. 2009. "Social media use in the United States: implications for health communication," *J Med Internet Res* (11:4), p. e48.

Dunkel-Schetter, C. 1984. "Social Support and Cancer: Findings Based on Patient Interviews and Their Implications," *Journal of Social Issues* (40), pp. 77–98.

Fox, S. 2011. "The Social Life of Health Information, 2011," Pew Research Center's Internet & American Life Project.

Keating, D. M. 2013. "Spirituality and support: a descriptive analysis of online social support for depression," *Journal of religion and health* (52:3), pp. 1014–1028.

Kim, E., Han, J. Y., Moon, T. J., Shaw, B., Shah, D. V., McTavish, F. M., and Gustafson, D. H. 2012. "The process and effect of supportive message expression and reception in online breast cancer support groups," *Psycho-Oncology* (21), pp. 531–540.

Krause, N. 1986. "Social support, stress, and well-being among older adults," *Journal of gerontology* (41:4), pp. 512–519.

Maloney-Krichmar, D., and Preece, J. 2005. "A multilevel analysis of sociability, usability, and community dynamics in an online health community," *ACM Trans. Comput.-Hum. Interact.* (12), pp. 201–232.

McClellan, W., Stanwyck, D., and Anson, C. 1993. "Social support and subsequent mortality among patients with end-stage renal disease," *Journal of the American Society of Nephrology* (4), pp. 1028–1034.

Qiu, B., Zhao, K., Mitra, P., Wu, D., Caragea, C., Yen, J., Greer, G. E., and Portier, K. 2011. "Get Online Support, Feel Better--Sentiment Analysis and Dynamics in an Online Cancer Survivor Community," in *Proceedings of the Third IEEE Third International Conference on Social Computing (SocialCom'11)*, Boston, MA, pp. 274–281.

Rodgers, S., and Chen, Q. 2005. "Internet Community Group Participation: Psychosocial Benefits for Women with Breast Cancer," *Journal of Computer-Mediated Communication* (10:4), pp. 00–00.

Wang, Y.-C., Kraut, R., and Levine, J. M. 2012. "To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups," in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, CSCW '12, New York, NY, USA, pp. 833–842.

Zhao, K., Yen, J., Greer, G., Qiu, B., Mitra, P., and Portier, K. 2014. "Finding influential users of online health communities: a new metric based on sentiment influence," *Journal of the American Medical Informatics Association: JAMIA* (online first).

Perils of Uncertainty? The Impact of Contextual Ambiguity on Search Advertising Keyword Performance

Vibhanshu Abhishek	Jing Gong	Beibei Li[1]
Carnegie Mellon University	Carnegie Mellon University	Carnegie Mellon University
vibs@andrew.cmu.edu	jingg@andrew.cmu.edu	beibeili@andrew.cmu.edu

Abstract

In this paper, we explore how the contextual ambiguity of a search can affect a keyword's performance in search advertising. We propose an automatic approach of categorizing keywords and examining keyword ambiguity based on topic models from machine learning and computational linguistics. We quantify the effect of contextual ambiguity on keyword click-through performance using a hierarchical Bayesian model, and validate our study using a novel dataset from a major search engine containing information on click activities for 12,790 keywords across multiple product categories from over 4.6 million impressions. We find that consumer click behaviors vary significantly across keywords. Moreover, keyword category and contextual ambiguity significantly affect such variation. Specifically, higher contextual ambiguity can lead to higher click-through rate (CTR) on top-positioned ads, but the CTR decays faster with position. Our study has the potential to help advertisers design keyword portfolios, and help search engines improve the quality of sponsored ads.

1. Overview

Sponsored search advertising has become an important marketing channel for businesses today. When a consumer issues a query on a search engine using a keyword, the search engine identifies a list of advertisers who are bidding on the keyword. It subsequently presents an appropriate list of ads based on factors such as bids placed by the advertisers and their historical performance. The ability to present consumers ads tailored to their search context (as indicated by the keywords) considerably increases the likelihood that they will click on these ads.

However, even though a search keyword provides an indication of a consumer's search context, consumers from varied contexts might use the same keyword for searching. For example, a consumer who searches for the keyword "Mars" may be interested in astronomy and the planet Mars, or may be interested in buying chocolates and candies from the confectionery company Mars, or may be looking for a local chain of grocery stores in metropolitan Baltimore, Maryland. Therefore, the search engine faces ambiguity in predicting the consumer's search context. In comparison, some keywords are specific and do not have a variety of meaning, for example, "antivirus." Consumers who search for "antivirus" and advertisers who bid on "antivirus" are likely to refer to the same product. Because keywords can have either a narrow or broad context, they might have varying appeal to consumers, and their performance might depend on the ambiguity in their context.

In this paper, we want to understand the interplay between a keyword's context and consumers' search behavior. More specifically, we wish to ascertain how the breadth of a keyword's context might affect consumer behavior and keyword performance. On the one hand, prior literature in search theory suggests that as the uncertainty in the quality of search results increases, users are

[1] Author names are in alphabetic order.

more likely to search (e.g.,Weitzman 1979, Ghose et al 2013), because higher uncertainty leads to higher variance and more diversity in the alternatives, and users believe they will be more likely to find an alternative with high value during their search. On the other hand, consumer psychology theories suggest that as the alternatives become less relevant, users are more likely to abandon their search (e.g., Jacoby et al. 1974; Dhar and Simonson 2003), because users tend to get overwhelmed and discouraged by the complexity of information, and therefore lose their interest in the search results. In reality, keyword contextual ambiguity can result in both higher diversity in ad quality and higher probability of ad irrelevancy. Therefore, how keyword contextual ambiguity would affect consumer click behavior is unclear. To explore this question, we use a rich dataset from a major search engine to perform a cross-category analysis and examine which of these two opposing effects dominates in the context of search advertising.

In this study, we propose an automatic way of categorizing keywords and examining keyword contextual ambiguity based on topic models from machine learning and computational linguistics, and quantify the effect of contextual ambiguity on keyword click-through performance using a hierarchical Bayesian Model that allows for topic-specific effect and nonlinear position effect. We validate our study using a novel dataset from a major search engine that contains information on consumer click activities for 12,790 distinct keywords across multiple product categories from over 4.6 million impressions from August 10, 2007 to September 16, 2007.

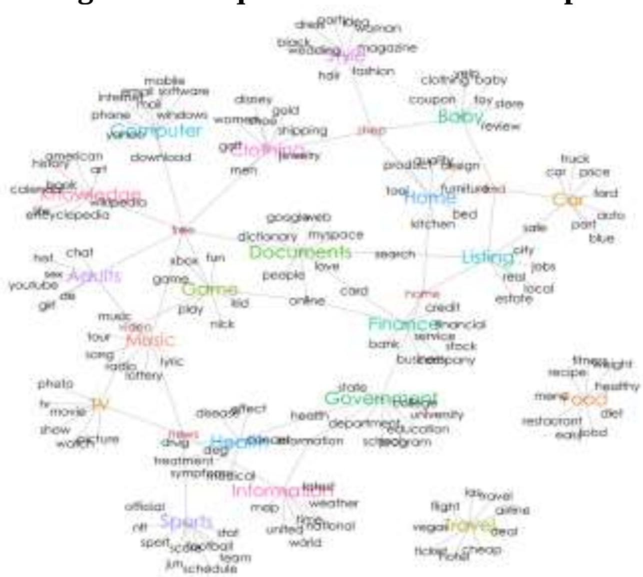
We find that consumer click behaviors vary significantly across keywords, and keyword category and the contextual ambiguity of the keywords significantly affect such variation. Specifically, higher contextual ambiguity can lead to a higher click-through rate (CTR) on top-positioned ads, but the CTR tends to decay faster with position. Therefore, the overall effect of contextual ambiguity on CTR varies across positions. Moreover, we also find significant interplay between keyword category and screen position. In particular, the distribution of CTR among different screen positions varies significantly across keyword categories. For example, the CTR of keywords in certain categories, such as “adult,” “home,” or “style,” is more evenly distributed across different positions compared to other categories, such as “baby products,” “finance,” and “travel.” These results suggest that position effects appear to be more significant for certain product categories than others.

This paper makes the following contributions. First, we demonstrate how machine learning tools such as topic models can be used to extract semantic characteristics of keywords based on large-scale text analytics. Second, we expand the search advertising literature by examining how keyword attributes affect keyword performance in multi-category search advertising. Most previous studies obtain data from a single advertiser and thus measure keyword performance from the perspective of one specific firm (e.g., Ghose and Yang 2009; Agarwal et al. 2011). The availability of detailed consumer search data for keywords across multiple advertisers and categories allows us to measure click performance at position level and identify keywords that have the potential to generate clicks – even at lower ad positions. Third, we find that consumer click behaviors vary significantly across keywords, and such variation can be explained by keyword category and the contextual ambiguity of keywords. Our empirical analysis offers new insights into consumer search behavior in the context of multi-category sponsored search, which can help advertisers select and evaluate keywords more effectively. Interestingly, our analysis suggests that consumers follow a two-step process while clicking on search ads. The effect of different keyword characteristics, such as entropy, differs in the propensity to start clicking and traversing the list of search ads. As a result, we reconcile the two opposing theories on consumer search and paves the path for richer consumer search models that can incorporate our findings.

2. Modeling Contextual Ambiguity

The major challenge in examining the impact of a keyword's contextual ambiguity on consumer click behavior is how to quantify such ambiguity. In this study, we model the contextual ambiguity of each keyword based on latent Dirichlet allocation model (LDA; Blei et al. 2003), a probabilistic topic model. Topic models are unsupervised algorithms from machine learning and natural language processing that aim to extract hidden topics from unstructured text data. To estimate the LDA model, we first construct a corpus of documents that store the top 50 Google organic search results of keywords. Based on the documents, we estimate the LDA model with a Gibbs sampler and obtain the topic probabilities. The most frequent words identified for the 20-topic model are presented in Figure 1, where topics are color coded. For convenience, we assign a label to each topic (e.g., "Sport," "Music," "Food") based on its high-frequency words. For example, documents related to "style" often contain words such as "dress," "party," "woman," and "fashion," and so on.

Figure 1. Frequent Words in Each Topic



We propose using topic entropy to measure keyword ambiguity. A keyword with higher entropy tends to relate to a broader range of topics (more ambiguous). Formally, let $\hat{\theta}_{kt}$ denote the posterior probability that keyword k belongs to topic t. We measure topic entropy as follows:

$$TOPIC\_ENTROPY_k = -\sum_{t=1}^{T} \hat{\theta}_{kt} \log(\hat{\theta}_{kt}), \tag{1}$$

where T is the total number of topics.

3. A Hierarchical Bayesian Model of Keyword Performance

To capture the impact of keyword characteristics on CTR and how CTR decreases with positions, we propose a hierarchical Bayesian model that allows for topic-specific effect and flexible nonlinear specifications. Following Abhishek and Hosanagar (2013) and Feng et al. (2007), we model the CTR for an ad at position p for keyword k that belongs to topic t as:

$$CTR_{kpt} = P(click_{kp} = 1 \mid topic_k = t) = \alpha_{kt} \gamma_{kt}^{p-1}, \tag{2}$$

where α_{kt} captures the baseline CTR, and γ_{kt} captures the change of CTR with positions. We assume that $\alpha_{kt} = \frac{\exp(\tilde{\alpha}_{kt})}{1+\exp(\tilde{\alpha}_{kt})}$, and $\gamma_{kt} = \frac{\exp(\tilde{\gamma}_{kt})}{1+\exp(\tilde{\gamma}_{kt})}$.

The total number of clicks is assumed to follow a Binomial distribution. :

$$P(CLICK_{kp} \mid IMP_{kp}, topic_k = t) \tag{3}$$
$$= \binom{IMP_{kp}}{CLICK_{kp}} CTR_{kpt}^{CLICK_{kp}} (1 - CTR_{kpt})^{IMP_{kp} - CLICK_{kp}},$$

where *IMP* denotes the total number of times consumers search for a particular keyword, and *CLICKS* measures the total number of clicks a particular keyword receives. To incorporate keyword heterogeneity, we assume that $\tilde{\alpha}_{kt}$ and $\tilde{\gamma}_{kt}$ follow a normal distribution:

$$\begin{pmatrix} \tilde{\alpha}_{kt} \\ \tilde{\gamma}_{kt} \end{pmatrix} \sim MVN(\mu_{kt}, \Phi), \tag{4}$$

where $\mu_{kt} = (\mu_{kt}^{(\alpha)}, \mu_{kt}^{(\gamma)})'$. To capture the impact of keyword characteristics, we further assume that the mean effects $\mu_{kt}^{(\alpha)}$ and $\mu_{kt}^{(\gamma)}$ are as follows,

$$\mu_{kt}^{(\alpha)} = \beta_{0t}^{(\alpha)} + X_k' \beta^{(\alpha)},$$
$$\mu_{kt}^{(\gamma)} = \beta_{0t}^{(\gamma)} + X_k' \beta^{(\gamma)},$$

where $\beta_{0t}^{(\alpha)}$ and $\beta_{0t}^{(\gamma)}$ denote the topic specific intercept terms and X_k is a vector of characteristics for keyword *k* including *TOPIC_ENTROPY, NUM_WORDS, BRAND, LOCATION, LOG_TRANS, AVG_AD_QUALITY, AVG_NUM_AD, and LOG_IMP*. *NUM_WORDS* denotes the number of words in the keyword. *BRAND* and *LOCATION* denote the presence of brand and location information in the keyword string. We use *LOG_TRANS* measures a keyword's transactional intent. An important factor that determines keyword performance is the quality of ads, which is captured in *AVG_AD_QUALITY*. *AVG_NUM_ADS* measures the average number of competing advertisers during an impression.

We assume that the intercept terms are drawn from a multivariate normal distribution as follows,

$$\begin{pmatrix} \beta_{0t}^{(\alpha)} \\ \beta_{0t}^{(\gamma)} \end{pmatrix} \sim MVN(\mu_0, \Omega_0), \tag{5}$$

We have multivariate normal priors on $\beta^{(\alpha)}$, $\beta^{(\gamma)}$, and inverse-Wishart priors on Φ and Ω_0.

To capture the topic-level effect, we incorporate the topic distribution associated with each keyword estimated from LDA. We have the log-likelihood function as follows:

$$LL = \sum_k \sum_p \log\left(\sum_t \hat{\theta}_{kt} P(CLICKS_{kp} \mid IMP_{kp}, topic_k = t)\right).$$

4. Major Findings

We estimated the model using 70% of the keywords chosen at random. We used Markov Chain Monte Carlo (MCMC) method for estimation and ran two MCMC chains, each with 70,000 iterations. The results are presented in Table 1.

First, the coefficient of topic entropy on α is positive and statistically significant, indicating that keywords with higher topic entropy are associated with higher overall CTR. This finding is consistent with the effect of "choice uncertainty" on pre-purchase search activities as observed in behavioral studies (Urbany et al. 1989). A consumer who searches with a higher entropy keyword may face higher uncertainty about alternatives of a consideration set, and may be more likely to explore a few ads to formulate a better understanding of alternatives. In this regard, our finding also concurs with literature on consumer search (Weitzman 1979) which suggests that as the uncertainty in the search results increases, users are more likely to search, because users believe they will be more likely to find an alternative with an extreme value in the utility distribution during their search. Another explanation for the positive effect of ambiguity on α can be provided by the organic results shown for the search. The search for ambiguous keywords leads to a varied set of organic results by construction, which might force users to click on ads to fulfill their search intent.

Second, topic entropy has a negative and significant impact on the decay parameter γ, suggesting that keywords that have higher topic entropy witness a larger decrease in CTR with position. This indicates that on average while consumers are more likely to click ads associated with more ambiguous keywords on the top positions, they are more likely to click fewer ads positioned lower on the screen once they start the search. Reduced consumer search activity associated with ambiguous ads can be attributed to search costs theory (Stigler 1961) and information overload theory (Iyengar and Lepper 2000, Dhar and Simonson 2003). The well-established search costs literature has demonstrated that higher search costs can lead to lower search intensity (Stigler 1961; Weitzman 1979). Meanwhile, previous research has shown both theoretical and empirical evidence that information overload can discourage consumers and lead to consumer search termination due to loss of interests in the search results (Iyengar and Lepper 2000, Ghose et al. 2013). In the context of sponsored search, due to the high heterogeneity in the quality of the ads generated by the ambiguous keywords, consumers might find it cognitively costly to process these ads to pick the relevant ones (hence high search costs). In addition, a highly diverse set of ads might introduce an increasing amount of new information to consumers during the search process. Such information may lead to an overload for consumers due to their cognitive limitation and non-negligible search costs, and may discourage consumers from clicking on the ads as consumers proceed down the list. This finding is also highly consistent with the effects of knowledge uncertainty found by Urbany et al. (1989). In particular, more ambiguous keywords may occur in search contexts where consumers' knowledge uncertainty is high. The lack of knowledge can in turn increase the barrier for consumers to evaluate choices, leading to much higher search costs for consumers during the search process.

Table 1. Estimation Results

	Posterior Mean	Posterior SD	Posterior Mean	Posterior SD
	α_{kt} (Baseline Effect)		γ_{kt} (Position Interaction Effect)	
TOPIC_ENTROPY	0.064***	(0.021)	-0.145***	(0.023)
NUM_WORDS	0.019	(0.014)	0.055***	(0.016)
BRAND	0.064**	(0.025)	-0.161***	(0.030)
LOCATION	-0.041*	(0.025)	-0.193***	(0.030)
LOG_TRANS	-0.012	(0.010)	0.047***	(0.011)
LOG_IMP	-0.019**	(0.009)	-0.059***	(0.011)
AVG_NUM_AD	0.372***	(0.009)	0.256***	(0.010)
AVG_AD_QUALITY	28.722***	(0.412)	-7.170***	(0.540)

In addition, other keyword characteristics also significantly predict the baseline CTR. Specifically, brand-related keywords are associated with higher overall click propensity, while location-specific and popular keywords are associated with lower overall click propensity. Meanwhile, the impact of position tends to be smaller for longer and transaction-related keywords. However, position tends to have a stronger effect for brand-related, location-related, and popular keywords.

Figure 2 illustrates how CTR changes with position by topic. Our results suggest the position effect on CTR is heterogeneous across different topics. For example, baby-related keywords tend to attract higher CTR than listing-related keywords at the top position, but the CTR decreases quickly at lower positions. This pattern we observe for baby-related keywords might be driven by the fact that consumers are extremely interested in baby-related ads, which leads to a high CTR at the first position, but their search intent is satisfied after the first few clicks. On the other hand, searches related to “adult” “home,” and “style” start out with fewer clicks on ads in the top position but continue to receive clicks on ads in lower positions. In fact, we observe that the search depth for “adult” is the largest among all the categories considered in our analysis.

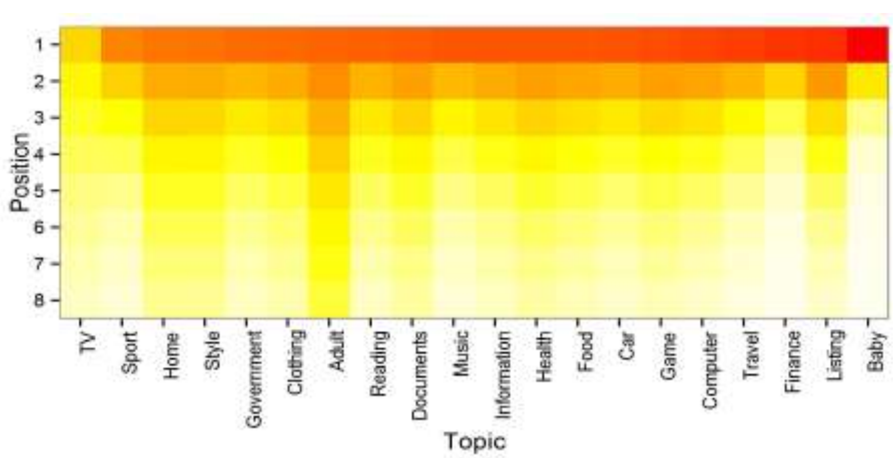
4. Conclusion and Discussion

In this paper, we study the effect of semantic characteristics on keyword’s CTR and try to provide insights into consumer behavior in the context of search advertising. First, the large-scale, cross-category analysis enables us to generate new insights in the search advertising literature. Second, we are able to exploit machine learning techniques to extract semantic characteristics of keywords given the large volume of keywords available to us. Third, we introduce a new keyword characteristic, “topic entropy,” that measures the ambiguity in the semantic meaning of a keyword. Our results suggest topic entropy and keyword categories are significant predictors of keyword performance and potentially affect consumer click behavior.

Our research has several managerial implications for advertisers and search engines. First, advertisers who are interested in driving traffic through clicks should be cautious about adding ambiguous keywords to their portfolios. Furthermore, if they need to choose some ambiguous keywords, they might be better off bidding aggressively and trying to attain higher positions. The search engines should also be cognizant of the topic ambiguity while designing their sponsored search strategies. For example, search engines may incorporate keyword characteristics while evaluating ad quality. Second, given that consumers behave differently across categories, advertisers can design their advertising strategies by allocating resources appropriately across products. Third, advertisers can use insights about how consumers respond to different keyword characteristics in designing better campaigns.

Figure 2. CTR by Topic for Sample Keywords

(Darker red colors represent higher CTR values and light yellow colors represent lower CTR values.)



References

Available at http://www.andrew.cmu.edu/user/beibeili/References-CSWIM2014.pdf.

IT, Ambidexterity, and Firm Performance: An Exploration

Liang Li
Univ. of International Business and Economics
liliang2013@uibe.edu.cn

Wei Zhang
Univ. of Massachusetts Boston
Wei.zhang@umb.edu

Abstract

In this paper, we explicate the concept of IT-based ambidexterity. We argue that through IT routinization and IT-initiated strategic alignment, IT can lead to organizational ambidexterity and subsequently superior firm performance. We identify a few technological antecedents of IT-based ambidexterity, and include a short discussion on possible research method to verify the theoretical framework and a brief projection on potential contributions of the proposed study.

Keywords: Information Technologies, Ambidexterity, Firm Performance

1. Introduction

Businesses that compete in today's fast changing, cutthroat market often have to excel in seemingly opposite fronts: They need to exploit what they know already to be competitive today and explore what they need to know to survive in the future (March 1991); they must align all activities within and across business units toward same organizational goals and be ready to adapt to constant changes in their environments (Gibson and Birkinshaw 2004); or they ought to fine-tune existing products and business models through incremental innovations as well as introducing revolutionarily new products and services through radical innovations. An organization's capability in successfully managing the trade-off between the two seemingly conflicting demand is referred to as organizational ambidexterity (Tushman and O'Reilly 2006).

As research has shown that ambidextrous organizations are better positioned to enjoy sustainable competitive advantage and superior performance over long term (Gibson and Birkinshaw 2004), Organizational ambidexterity has become an active research area in recent years (Raisch et al. 2009). Researchers have identified three paths to ambidexterity (Raisch and Birkinshaw 2008): structural solutions, contextual solutions, and leadership-based solutions. However, to the extent that information technologies (IT) has become an important strategic tool for organizations (Aral and Weill 2007), little research has addressed the impact of IT on organizational ambidexterity.

In this paper, we argue that in addition to previously identified paths to ambidexterity, IT can be another important solution that leads to organizational ambidexterity. We begin with elaborating IT-based ambidexterity and explicating its relationship with firm performance. We then identify a few technical antecedents of IT-based ambidexterity and argue that IT-based ambidexterity mediates the relationship between the identified antecedents and firm performance. A brief discussion on possible research method and potential contributions the proposed study can make conclude the paper.

2. IT-Based Ambidexterity

2.1 Organizational Ambidexterity

Organizational ambidexterity refers to an organization's capability in successfully managing the seemingly conflicting needs and executing the seemingly contradictory strategies (Raisch and Birkinshaw 2008; Raisch et al. 2009). Literature has described such conflicting needs and contradictions through multiple theoretical lens (Raisch and Birkinshaw 2008). While the most studied pair is exploitation and exploration (March 1991) – whereas exploitation refers to reusing what's already known for short-term profits and exploration refers to learning what's new for long-term survival, other examples of such contradictions that entail organizational ambidexterity include incremental and radical innovation (Atuahene-Gima 2005), alignment and adaptation (Gibson and Birkinshaw 2004; Im and Rai 2013), and efficiency and flexibility (Li et al. 2013).

Being able to simultaneously "exploit existing competencies and exploring new opportunities (Raisch et al. 2009; p.685)" allows ambidextrous organizations to enjoy long-term success. Consequently, many research efforts have been on how organizations can cultivate this pivotal capability. Literature identified three solutions to organizational ambidexterity (Raisch and Birkinshaw 2008). Among them, structural solutions were identified the earliest. They rely on spatial separation or parallel structures. Through spatial separation, organizations create separate organizational units and equip them with different strategy, policy, and culture for either exploitation or exploration (Duncan 1976). With parallel structures, organizations foster a secondary, often informal structure in addition to a formal primary structure in a same organizational unit. With the primary structure focusing on the main tasks, the secondary structure – e.g., task groups, informal networks, communities of practice – deals with non-routine tasks, to promote innovations, and to share and create new knowledge (McDonough and Leifer 1983).

Through contextual solutions, organizations foster the right organizational context to empower each organizational member to make autonomous decisions and hence afford the overall organizations the ambidexterity to handle diverse, competing strategic requirements (Gibson and Birkinshaw 2004). Leadership-based solutions focus on top management (Lubatkin et al. 2006), believing that for organizations to be ambidextrous the top management must excel in managing ambidexterity: Top management must acknowledge and embrace paradoxes, contradictions, and conflicts inherently in their organizations before they can sense the existence of contradictions and make balanced strategic decisions.

2.2 IT-Based Ambidexterity

While research has shown IT can be used to promote contextual ambidexterity (Im and Rai 2013), the direct link between IT and organizational ambidexterity has not been investigated despite a plethora of practical and theoretical hints to the close relationship between them. Practically, IT was fundamental to ambidextrous movement *mass customization* – which emphasizes both efficiency and flexibility – that became popular in 1990s. Theoretically, even if IT are designed to meet certain requirements, researchers have long argued against viewing IT as artifacts that are used in the prescribed ways to meet predefined goals and viewing users as robotic operators who could only accept IT as it is and only act on instructions (Bogers et al. 2010). The inherent flexibilities in IT, IT usage, and IT users make IT solution to ambidexterity intriguing.

On one side, IT has permeated into many, if not all, aspects of businesses. Most business functions have now been computerized, and business processes digitized, thus making IT an integral part to business routines. Such routines embody tacit knowledge accumulated by organizations (Crossan et al. 1999) and are considered a key organizational capability (Winter 2003). Once established, routines are regularly and repeatedly enacted by organizational members (Feldman and Pentland 2003), and can ultimately influence firm performance (Hoffer Gittell 2002).

As IT becomes integrated into business routines and using IT becomes integrated into performing business functions, users would try to make sense of IT and figure out how to take advantage of the functionalities afforded by the IT (Goh et al. 2011). With the integration emerging, new routines would capture the best practices with IT and embody what users know about how to best use the IT to fulfill business functions. When new routines stabilize and are extended to more users, IT could then be used to provide guidance and lead the users to perform a set of prescribed operations to fulfill business functions (Li and Mao 2012), thus becoming an ideal tool for companies to standardize operations and disseminate the best practice embedded in the IT. In this sense, as IT and business functions merge, the tight coupling between IT and business makes IT an instrumental tool for exploitation to profit from what's already known.

On the other side, researchers in strategic management have long argued that to remain competitive today, businesses must attain dynamic capability, "the ability to integrate, build, and reconfigure internal and external competencies to address rapidly changing environments (Teece et al. 1997, p517)." In other words, dynamic capability allows companies to quickly adapt to fast changing business environment. Research has linked IT to dynamic capability, as IT help companies to excel at sensing the changes in business environment, acquiring and integrating knowledge, and coordinating activities to cope with changes (El Sawy and Pavlou 2008).

Perhaps more importantly, IT not only help companies to react to changes, but also allow companies to proactively drive changes. In the seminal work on strategic alignment between business and IT, Henderson and Venkatraman (1993) argued that while it is common for business strategy to dictate IT strategy and process, as the strategic importance of IT rises, IT could potentially underlie the design of business strategy and process. Visionary managements would need to understand how IT can give birth to new products and services, and proactively adjust business strategies and recalibrate business processes to take full advantage of the IT capability to stay ahead of competition.

Thus IT-based ambidexterity refers to the balance between routinization of IT and IT-initiated strategic alignment. While the former allows firms to maximize short-term efficiencies by multiplying mature business routines and spreading best practices among users, the latter enables them to explore emerging technologies for the long-term interests in today's fast changing market environments. Thus we propose,

P1. Higher level of IT-based ambidexterity is associated with higher level of performance.

2.3 Technological Antecedents of IT-based Ambidexterity

While many factors can give rise to ambidexterity (Raisch and Birkinshaw 2008), previous research have not been discussed technological factors. In this study we explore how technological characteristics of IT can affect IT-based ambidexterity. IS research long acknowledges that IT characteristics still play important roles in shaping technology acceptance, diffusion (Rogers 2003) and usage (Leonardi 2011). At least, technologies with certain characteristics can affect how easy (or how difficult) it is to fulfill certain tasks (Leonardi 2011).

We consider three IT characteristics that may be particularly relevant to ambidexterity: scalability, trialability, and configurability. In the context of our study, we refer scalability to the degree to which IT can easily and satisfactorily accommodate the fluctuation in IT service demands (Laitinen et al. 2000). Once a routine is established and stabilized, IT with high scalability can help multiply the routine to as many users as needed, thus enhancing the level of IT routinization (e.g. Simparel 2013).

Trialability is the "degree to which an innovation may be experimented with on a limited basis" (Rogers 2003, p258). Through trials, companies can assess IT's strengths and weakness in the environment where the IT is to be used, predicting and verifying the outcome (Watanabe and Kondo 2003). Companies can also try changes in a limited scope (for example, to a limited number of users or for a limited number of customers) before they are fully deployed. Such test drives allow companies to evaluate changes to existing IT without incurring permanent changes to systems, facilitating companies to experiment with the systems so that they can find the best possible fit between IT and business processes.

We refer configurability to the ease with which IT can be customized, modified, and assembled to adapt to the ever changing business environment. IT with high configurability such as cloud-based technologies allow organizations flexibility in molding IT to satisfy their business requirements (Iyer and Henderson 2010), which afford companies opportunities to not only integrate IT with their business process (Elbanna 2006) but also use IT to drive innovations and change the competition landscape.

We note that scalability, trialability, and configurability present conflicting challenges to IT: It is simply expensive for companies to acquire IT that are simultaneously scalable, trialable, and configurable. For example, cloud-computing is consider to have high scalability and trialability (Iyer and Henderson 2010), but with limited configurability. Due to limited resource, companies likely have to make trade-offs between these desirable but competing characteristics when building their IT infrastructure (Simparel 2013). Thus we propose,

> *P2. The balance between IT characteristics – e.g. scalability, trialability, and configurability – is associated with IT-based ambidexterity.*

2.4 Mediation Effects

Finally, we argue that IT-based ambidexterity mediates the relationship between the identified technological antecedents and firm performance. After all, research on the business value of IT – exemplified by the "productivity paradox" – has long cast doubt on the direct link between IT and firm performances. Moreover, the current trend in the commoditization of IT makes IT itself increasing not a differentiating factor (Carr 2003). Hence the direct link between IT and firm performance is unlikely.

Still, companies using the same or similar IT can perform vastly differently. To help account for this apparent discrepancy, we need to take a closer look at how IT is actually used rather than just what IT is used. In the context of the reported study, we argue that it is the organizational ambidexterity cultivated on the basis of the technological characteristics of the IT that makes the differences in performance: The conflicts between different technological characteristics – scalability, trialability, and configurability in this study – can create tensions on how IT benefit the company if they are not used to form ambidexterity. Therefore, we propose,

P3. IT based ambidexterity mediates the relationship between IT characteristics – as captured by scalability, trialability, and configurability – and firm performance.

Proposed Research Method

The aforementioned research framework is merely preliminary. The constructs and propositions need to be further developed, refined, and operationalized to subject them to empirical validation. The resultant research model will be tested with both qualitative and quantitative data. Qualitative data collected through field observations and interviews will allow us to gain more first-hand insights into the complicated phenomenon under study and provide us with the opportunity to verify and refine our theoretical reasoning. Quantitative data collected through surveys, on the other hand, allow us to test the hypotheses in a positivist way. Triangulating the findings from both methods also brings more confidence in the validity of the findings.

Discussion

In this paper, we propose a new perspective on organizational ambidexterity, focusing on how IT could give rise to organizational ambidexterity and how IT-based ambidexterity can affect firm performance. Through this study, we aim to deepen our understanding in not only ambidexterity but also the strategic value of IT. We are also optimistic of the practical contributions of the proposed study, hoping that the results can shed new lights on what companies can do to enhance their organizational ambidexterity.

Acknowledgement

This project was supported by “the Fundamental Research Funds for the Central Universities” in UIBE (13QD07).

References

Aral, S., and Weill, P. 2007. "It Assets, Organizational Capabilities, and Firm Performance: How Resource Allocations and Organizational Differences Explain Performance Variation," *Organization Science* (18:5), pp. 763-780.

Atuahene-Gima, K. 2005. "Resolving the Capability—Rigidity Paradox in New Product Innovation," *Journal of Marketing* (69:4), pp. 61-83.

Bogers, M., Afuah, A., and Bastian, B. 2010. "Users as Innovators: A Review, Critique, and Future Research Directions," *Journal of Management* (36:4), pp. 857-875.

Carr, N.G. 2003. "It Doesn't Matter," *Harvard Business Review* (81:5), pp. 41-49.

Crossan, M.M., Lane, H.W., and White, R.E. 1999. "An Organizational Learning Framework: From Intuition to Institution," *Academy of Management Review* (24:3), pp. 522-537.

Duncan, R.B. 1976. "The Ambidextrous Organization: Designing Dual Structures for Innovation," in *The Management of Organization*, 1, R.H. Killman, L.R. Pondy and D. Sleven (eds.). New York: North-Holland, pp. 167-188.
El Sawy, O.A., and Pavlou, P.A. 2008. "It-Enabled Business Capabilities for Turbulent Environments," *MIS Quarterly Executive* (7:3).
Elbanna, A.R. 2006. "The Validity of the Improvisation Argument in the Implementation of Rigid Technology: The Case of ERP Systems," *Journal of Information Technology* (21:3), pp. 165-175.
Feldman, M.S., and Pentland, B.T. 2003. "Reconceptualizing Organizational Routines as a Source of Flexibility and Change," *Administrative Science Quarterly* (48:1), pp. 94-118.
Gibson, C.B., and Birkinshaw, J. 2004. "The Antecedents, Consequences, and Mediating Role of Organizational Ambidexterity," *Academy of Management Journal* (47:2), pp. 209-226.
Goh, J.M., Gao, G.G., and Agarwal, R. 2011. "Evolving Work Routines: Adaptive Routinization of Information Technology in Healthcare," *Information Systems Research* (22:3), pp. 565-585.
Henderson, J.C., and Venkatraman, N. 1993. "Strategic Alignment: Leveraging Information Technology for Transforming Organizations," *IBM systems journal* (32:1), pp. 4-16.
Hoffer Gittell, J. 2002. "Coordinating Mechanisms in Care Provider Groups: Relational Coordination as a Mediator and Input Uncertainty as a Moderator of Performance Effects," *Management Science* (48:11), pp. 1408-1426.
Im, G., and Rai, A. 2013. "It-Enabled Coordination for Ambidextrous Interorganizational Relationships," *Information Systems Research* (25:1), pp. 72 - 92
Iyer, B., and Henderson, J.C. 2010. "Preparing for the Future: Understanding the Seven Capabilities Cloud Computing," *MIS Quarterly Executive* (9:2), pp. 117-131.
Laitinen, M., Fayad, M.E., and Ward, R.P. 2000. "The Problem with Scalability," *Communications of the ACM* (43:9), pp. 105-107.
Leonardi, P.M. 2011. "When Flexible Routines Meet Flexible Technologies: Affordance, Constraint, and the Imbrication of Human and Material Agencies," *MIS Quarterly* (35:1), pp. 147-168.
Li, L., and Mao, J.-Y. 2012. "The Effect of Crm Use on Internal Sales Management Control: An Alternative Mechanism to Realize Crm Benefits," *Information & Management* (49:6), pp. 269-277.
Li, L., Zhang, W., and Mao, J.-Y. 2013. "The Business Value of Crm Technology: From the Perspective of Organizational Ambidexterity," *AMCIS 2013*, Chicago, IL, U.S.A.
Lubatkin, M.H., Simsek, Z., Yan, L., and Veiga, J.F. 2006. "Ambidexterity and Performance in Small- to Medium-Sized Firms: The Pivotal Role of Top Management Team Behavioral Integration," *Journal of Management* (32:5), pp. 646-672.
March, J.G. 1991. "Exploration and Exploitation in Organizational Learning," *Organization Science* (2:1), pp. 71-87.
McDonough, E.F., and Leifer, R. 1983. "Using Simultaneous Structures to Cope with Uncertainty," *Academy of Management Journal* (26:4), pp. 727-735.
Raisch, S., and Birkinshaw, J. 2008. "Organizational Ambidexterity: Antecedents, Outcomes, and Moderators," *Journal of Management* (34:3), pp. 375-409.
Raisch, S., Birkinshaw, J., Probst, G., and Tushman, M.L. 2009. "Organizational Ambidexterity: Balancing Exploitation and Exploration for Sustained Performance," *Organization Science* (20:4), pp. 685-695.
Rogers, E.M. 2003. *Diffusion of Innovations*, (5 ed.). New York: Free Press.

Simparel. 2013. "Using ERP Systems to Fuel Operational Scalability," *Apparel Magazine* (55:1), pp. 4-8.

Teece, D.J., Pisano, G., and Schuen, A. 1997. "Dynamic Capabilities and Strategic Management," *Strategic Management Journal* (18:7), pp. 509-533.

Tushman, M.L., and O'Reilly, C.A., III. 2006. "Ambidextrous Organizations: Managing Evolutionary and Revolutionary Change," *California Management Review* (38:4), pp. 8-30.

Watanabe, C., and Kondo, R. 2003. "Institutional Elasticity Towards It Waves for Japan's Survival—the Significant Role of an IT Testbed," *Technovation* (23:4), pp. 307-320.

Winter, S.G. 2003. "Understanding Dynamic Capabilities," *Strategic Management Journal* (24:10), pp. 991-995.

Does *mianzi* really matter?

Yonghong Sun
School of Management
Xi'an Jiaotong University
simonsun@mail.xjtu.edu.cn

Yuewen Liu
School of Management
Xi'an Jiaotong University
yuewenliu@mail.xjtu.edu.cn

Abstract

Mianzi is a Chinese word which refers to the image or reputation one person could have in the presence of others. Previous studies have illustrated the importance of mianzi in affecting individual performance or behavior. Different from previous studies mainly using laboratory experiments, we use a field study to examine the mianzi effect. In this study, we let 60 students of a course submit their final homework papers via a social network platform, so that their papers will be exposed to their classmates. In this way, we try to examine the impact of the presence of others on the students' performance in their homework. The results show that the students with relatively more friends in the class perform better than those with relatively fewer friends, the students receiving more attention from others perform better than those receiving less, and the students using real photos as their digital profile pictures perform better than those who do not. The results imply that mianzi really matters.

Keywords: *mianzi*, evaluation apprehension, social facilitation, audience, coactors

1. Introduction

Mianzi is a Chinese word. It refers to the image or reputation one person could have in front of others. In order to keep *mianzi*, people have to mind their behavior in the presence of others. The most obvious example is that players in sports games try to play well when they are being watched by a large audience.

Based on review of prior work, Zajonc (1965) proposed a hypothesis regarding how the presence of others affect individual performance. He hypothesized that the presence of others, as spectators or as coactors, enhances the emission of dominant responses, which in turn facilitate or impair individual performance. He argued that the presence of others may raise the individual general arousal or drive level. If the task is simple or well learned, the presence of others would facilitate the individual performance. On the other hand, if the task is complex and unfamiliar, the presence of others would impair the individual performance. By observing maze and runway performances of cockroaches under solitary and social conditions, Zajonc, Heingartner, and Herman (1969) provided evidence for their theory of social facilitation. Though most evidences on which Zajonc's thesis is based are from nonhuman animal's experiments, a large amount of subsequent research which use human as subject support his thesis as well(see (Bond & Titus, 1983; Geen & Gange, 1977) for a literature review.).

Cottrell (1972) alleged that the performance change in the presence of others is due to that individual feels apprehensive over implicit or explicit judgment of his performance by those observers. Cottrell, Wack, Sekerak, and Rittle (1968) compared university students performance in a pseudo recognition task under three conditions, performed alone, performed in the presence of 2 persons who were not spectators, and performance before an audience of 2 passive spectators. They found that students in the presence of spectators exhibited performance improvement, but others didn't, including those students who performed in the presence of 2 persons who were not spectators. So, individual performance will be affected only if he feels evaluation apprehension in the presence of others. White, Mitchell, and Bell Jr (1977) showed the supportive evidence that people with high evaluation apprehension produce more than people with low evaluation apprehension.

To sum up, in the presence of other persons, as spectators or as coactors, evaluation apprehension produces arousal which may enhance or impair individual performance. People have to act to avoid losing *mianzi* (or face) or fight for *mianzi*. In this study, we use a field study to examine the *mianzi* effect. We let 60 students of a Microeconomics course submit their final homework papers via a social network platform, so that their papers will be exposed to their classmates. In this way, we try to examine the impact of the presence of others on the students' individual performance.

The rest of the paper is organized as follows. In section 2, the shortcomings of previous studies are summarized based on analyzing related literature. The method we used in present study is introduced in section 3. Results are presented in section 4. We conclude this study in the last section.

2. Related Literature

First, prior literature examining the effect of the presence of others mainly measures individual performance in terms of speed and accuracy (see (Bond & Titus, 1983; Geen & Gange, 1977) for a literature review.). Actually, these studies can be divided into two stages. Before the work by Zajonc (1965), animals were mainly used as subjects, and social facilitation effect was measured in terms of speed or accuracy that these animals can make under different experimental conditions. After Zajonc (1965), more and more researchers shifted attention from nonhuman to human. Nevertheless, except for the study by Gallupe et al. (1992), speed or accuracy was still used as performance measure to examine the effect of the presence of others. Gallupe et al. (1992) found that larger groups via electronic brainstorming can generate more unique ideas and more high-quality ideas. Thus, how the presence of others affect individual other performance, such as creativity, need further investigation.

Second, prior studies examining the effect of the presence of others are in the context of face to face presence. However, computer mediated presence are becoming more and more universal nowadays. On one hand, some scholars argued that commuter mediated communication would lose a lot of social cues such that social effect is reduced to a large extent (see (Kiesler, Siegel, & McGuire, 1984) for example.). On the other hand, Nunamaker, Dennis, Valacich, Vogel, and George (1991) concluded that one benefit to use EMS (electronic meeting systems) is that "to pause and reflect on information and opinions of others during the meeting and serves as a permanent record of what occurred." So it is plausible that people pay more attention to their image-building in the computerized setting especially when they know each other offline since their behavior or performance will be recorded forever. Given these two competing effects, we need to further explore how computer mediated presence affect individual performance.

Third, recent research on this issue has been extended to other contexts, such as strategic game environments, but there still lack of field evidence. In previous experiments subject performance was only affected by his own behavior or choice. We found two recent studies where individual performance was affected not only by his own choice, but also by others' choice. Charness, Rigotti, and Rustichini (2003) compare the choices of players in a 2-person game experiment with or without an audience. They found that when audience is present, players are more likely to choose the strategy favorable to the audience. Andreoni and Bernheim (2009) provided an explanation for the norm of 50-50 division even in the dictator game. Different from previous explanation that people care about fairness, they argued that people like to be perceived as fair. That is to say, people care about their social image in the presence of audience such that they voluntarily cede exactly half to another individual in the dictator game. Undoubtedly these laboratory experiments expand our knowledge on environments where social facilitation may also play a role, but filed evidences are still very rare and desired.

The present study overcomes the shortcomings of previous studies mentioned above. First, we investigate how the presence of others affects creative activities. In this study we require students to finish an assignment for which they need to make full use of their creativity. Second, we examine how computer mediated presence affect individual performance. Third, we provide field evidence for the presence effect in a strategic environment where students interact with other and their performance depends on others' performance.

3. Method

3.1 Subject

Our subjects are year one graduate students at School of Management, Xi'an Jiaotong University, China. We teach them Microeconomics in their first semester after enrolment starting from September every year. These students come from different universities, studied different majors before, and have different background knowledge of Microeconomics.

3.2 Design and procedure

We need to teach two times one week. One is two hours, the other is three hours. Totally, we have to teach 48 hours for this course. Students are evaluated in two ways. First, they need to take a quiz once a week. Each time these quizzes start half an hour before the three hour class is finished. Students are required to answer questions related to the content that they just have learnt in the class. Totally there are ten quizzes, two points is the full mark for each quiz, and we give students a score for each quiz. The total scores of quizzes are used for final evaluation. Second, once we finish teaching the course, each student is required to submit an article within two weeks. In the article, students need to write a short story to analyze some real life phenomenon from Microeconomics perspective. We almost have no any restrictions on this homework except that the stories should be interesting and insightful.

We set up a group in the online community, Renren[1]. Students are required to post their articles in the group. Like Facebook in USA, Renren is one of the most popular online communities in China. Users can post and share articles, photos and video clips so on in Renren. In particular, Renren advertise to freshman every year so that almost all university students have a Renren account and use it to communicate from time to time. So, for those students in our class, almost all of them have been users of Renren, and they post and share different things with their friends via Renren. After a student post his/her article in the group, his/her classmates in the group, and other friends who are not in the group, can read and comment his/her article. Furthermore, some students in the group have established friend relationship in Renren, but others may not have. We have an evaluation panel which consists of three professors who will independently evaluate the quality of students' submissions. The full mark for each article is 10 points. We average scores given by three professors as the final score for an article.

3.3 The audience and/or coaction condition

We try to examine how the presence and/or coaction of friends affect individuals' performance in their homework. Renren provides such kind of service that individual student can know how many times his profile is visited by his friends. We term it *nvisits* which represents how much attention his friends paid to him. Please note that his friends may be his classmates in our class, and maybe not. On the other hand, in the Renren group we established, different students have different number of friends. We term it *nfriends* which can denote how many friends he cares in the Renren group are preparing the homework together with him. The reason why different students have different numbers of friends in the group is that one student can freely choose to or not to invite others to be his friends. So some students may have many classmates as their friends in the group, but others may have just few.

3.4 Control variables

Moreover, we need to control those factors which might confound the *mianzi* effect. First, Renren encourages users to use real photo as their profile pictures. After a user has changed her profile picture to real photo and the photo is then certified by Renren's staff, Renren will mark the user as a star user. If a student chooses to be a *star User,* that means she is more likely to care about her *mianzi* in the social network platform. So we need to control this potential self-selection bias. Second, we also need to control those factors that might influence students' performance in their homework. These factors include their background knowledge of Microeconomics, their attitudes and interests in this course, their expectations on themselves etc. All the information to a large extent can be captured by students' quiz score. Additionally, though students can get some cues from others' homework, but they cannot copy or imitate from others too much. They have to prepare their own stories. So we don't need to worry that students' performance change is due to that they copy or follow others' behavior. After control above factors, we can now isolate the *mianzi* effect. Next the statistical results are presented.

4. Results

The descriptive statistics of variables are shown in Table 1. We have 60 subjects in total. Among the subjects, 73% are *starUser*s, which means 73% of the subjects use real photos as their profile pictures. The subjects have around 397 friends (*nFriends*) on average in RenRen website, and

[1] http://xiaozu.renren.com/xiaozu/620021711

have been visited around 2868 times (*nVisits*) on average. In the group, they have around 27 friends (*gfriends*), which is about half of the group members. The network of the subjects is illustrated in Figure 1. The figure shows that, while some subjects connect to nearly all the other subjects in the graph, some other subjects connect to merely one other subject. The subjects get an average score of 6.95/10 for their *homework*, and 14.27/20 for their *quizzes*.

Table 1, Descriptive Statistics

	N	Mean	S.D.
starUser	60	0.73	0.45
nFriends	60	396.77	250.92
nVisits	60	2867.75	5356.47
gFriends	60	27.00	12.51
homework	60	6.95	1.68
quiz	60	14.27	2.93

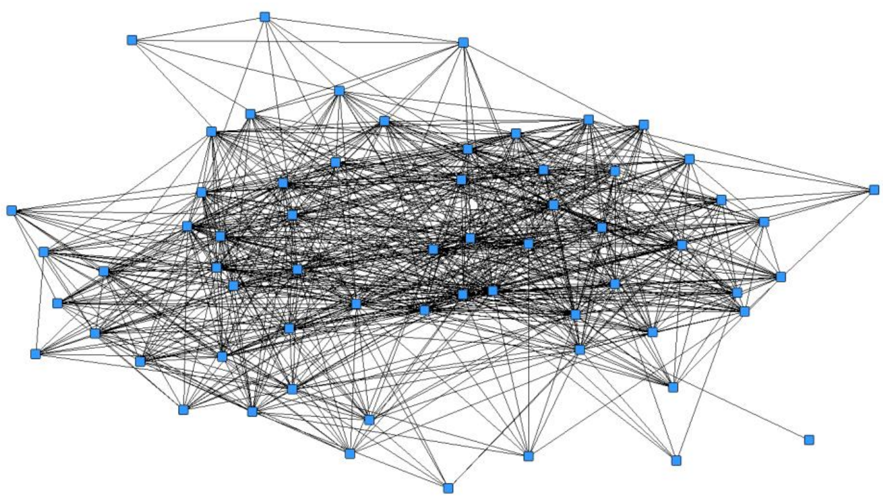


Figure 1, the Social Network

Table 2 illustrates the pearson correlations among the variables. The correlation coefficients show that, *nFriends* has a significantly positive correlation with *nVisits*. This seems to indicate that, the more friends a subject has, the more visits the subject gets. However, after we carefully examine the relationship between these two variables, we didn't find that there is a strong linear correlation between them, which is shown in the scatter plot in Appendix. Moreover, *nFriends* and *gfriends* are also significantly positively correlated. This may indicate that, an active subject who has more friends in Renren platform would have more friends in the group.

Table 2, Pearson Correlations

	homework	*starUser*	*nFriends*	*gfriends*	*quiz*
starUser	0.25**				
nFriends	-0.05	0.21*			
gFriends	0.18*	0.01	0.56***		
quiz	0.26**	-0.06	0.04	0.19*	
nVisits	0.08	0.17	0.77***	0.35***	-0.16

*Note: ***, **, and * indicate significance levels of p<0.01, p<0.05 and p<0.1 separately.*

The regression results are illustrated in Table 3, and model fitness indices are illustrated in Table 4. The regression coefficient of starUser is significantly positive (B=1.253, *p=0.006*), which show that the subjects who use their photos as their profile picture in RenRen.com may perform better in their *Homework*. This indicates that the subjects who use their real photos in RenRen.com value their *mianzi* more than those who do not use their real photos. The regression coefficient of *nFriends* is significantly negative (B=-0.005, *p=0.001*), while the regression coefficients of *gFriends* is significantly positive (B=0.044, *p=0.023*). The negative coefficient of *nFriends* seems to imply that those subjects with more friends feel like social activities, spend less time on studying, and therefore perform worse in their homework. The positive coefficient of *gFriends* show that, the more friends a subject has in the group, the more *mianzi* issue the subject may face, thus the better the subject perform. It is quite obvious that *quiz* positively affect the subjects' performance in homework. This indicate that those students who have better background knowledge in Microeconomics, more interest in this course, and higher expectation on themselves will performance better accordingly. The regression coefficient of nVisits is also positively significant (B=0.000, *p=0.008*), shows that the more visits a subject get, the better the subject perform. If a subject got more visits, that means he received more attention from others. In this case, he has to care more about his image and try to perform better.

Table 3, Regression Results

	B	S.E.	t	p	Sig.	Tolerance	VIF
C	3.536	1.047	3.379	0.001	***		
starUser	1.253	0.440	2.850	0.006	***	0.938	1.066
nFriends	-0.005	0.001	-3.440	0.001	***	0.297	3.371
gfriends	0.044	0.019	2.339	0.023	**	0.641	1.560
quiz	0.191	0.069	2.769	0.008	***	0.881	1.136
nVisits	0.000	0.000	2.777	0.008	***	0.371	2.697

*Note: *** and * indicate significance levels of p<0.01 and p<0.1 separately;*
The dependent variable is "homework".

Table 4, Model Fit

n	df	Mean Square	F	Sig.	R^2	Adj. R^2
60	5	10.363	4.864	0.001***	0.311	0.247

*Note: *** indicates a significance level of p<0.01.*

5. Concluding remarks

According to the social facilitation theory, the presence of others might improve or impair individual performance, which depends on whether individual are familiar with or well master the task they are performing. In this study, Microeconomics is indeed not an easy course, and the homework requires the students to make great efforts and fully use their creativity. However, the homework has been assigned to students in the first class, and they have a very long time to prepare. So it is plausible to expect that the presence of others in the online community should enhance their performance. The statistical results are consistent with our predictions.

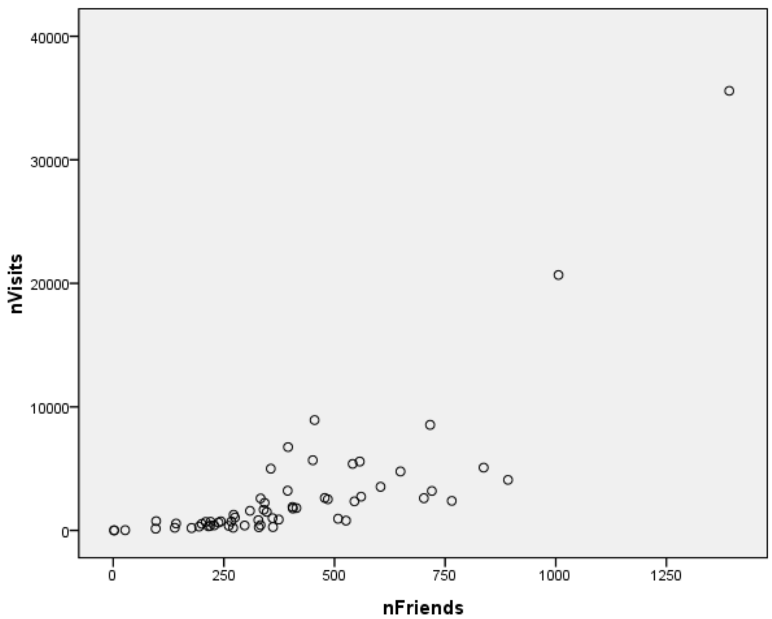
Furthermore, although it is somewhat difficult to quantify the *mianzi* effect, it doesn't mean this effect is trivial and can be neglected safely. Economists always think that incentive or punishment mechanism would monotonically influence people's behavior. However, the effectiveness of the mechanism could be enhanced or weaken due to the existence of *mianzi* effect. So *mianzi* effect should be taken into serious consideration in both theoretical research and practical decision making.

References

1. Andreoni, J., & Bernheim, B. D. 2009. Social Image and the Fifty-Fifty Norm: A Theoretical and Experimental Analysis of Audience Effects. ***Econometrica***, 77(5): 1607-1636.
2. Bond, C. F., & Titus, L. J. 1983. Social facilitation: A meta-analysis of 241 studies. ***Psychological Bulletin***, 94(2): 265-292.
3. Charness, G., Rigotti, L., & Rustichini, A. 2003. They are watching you: Social facilitation in institutions: Department of Economics, UC Santa Barbara, University of California at Santa Barbara, Economics Working Paper Series.
4. Cottrell, N. B., Wack, D. L., Sekerak, G. J., & Rittle, R. H. 1968. Social facilitation of dominant responses by the presence of an audicence and the mere presence of others. ***Journal of Personality and Social Psychology***, 9(3): 245-250.
5. Cottrell, N.B. (1972). Social Facilitation. In C. McClintock (ed.), *Experimental Social Psychology* (pp. 185–236). New York: Holt, Rinehart & Winston.
6. Gallupe, R. B., Dennis, A. R., Cooper, W. H., Valacich, J. S., Bastianutti, L. M., & Nunamaker, J. J. F. 1992. ELECTRONIC BRAINSTORMING AND GROUP SIZE. ***Academy of Management Journal***, 35(2): 350-369.
7. Geen, R. G., & Gange, J. J. 1977. Drive theory of social facilitation: Twelve years of theory and research. ***Psychological Bulletin***, 84(6): 1267-1288.
8. Kiesler, S., Siegel, J., & McGuire, T. W. 1984. Social psychological aspects of computer-mediated communication, 10 ed., Vol. 39: 1123-1134. US: American Psychological Association.

9. Nunamaker, J. F., Dennis, A. R., Valacich, J. S., Vogel, D., & George, J. F. 1991. Electronic meeting systems. ***Commun. ACM***, 34(7): 40-61.
10. White, S. E., Mitchell, T. R., & Bell Jr, C. H. 1977. Goal Setting, Evaluation Apprehension, and Social Cues as Determinants of Job Performance and Job Satisfaction in a Simulated Organization. ***Journal of Applied Psychology***, 62(6): 665-673.
11. Zajonc, R. B. 1965. Social Facilitation. ***Science***, 149(3681): 269-274.
12. Zajonc, R. B., Heingartner, A., & Herman, E. M. 1969. Social enhancement and impairment of performance in the cockroach. ***Journal of Personality and Social Psychology***, 13(2): 83-92.

Appendix



A Framework for Designing An Intelligent e-Health Management System

Jinwei Cao
University of Delaware
jcao@udel.edu

Pengzhu Zhang
Shanghai Jiao Tong University
pzzhang@sjtu.edu.cn

Abstract

Currently there is no formal, integrated framework to guide the design of an intelligent e-health system that communicates effectively with health care consumers and providers. In this paper we review key communication issues involved in the design of effective e-Health applications, apply human–computer interaction (HCI) principles, and propose a new framework to determine feature sets of prospective intelligent health management systems.

Keywords: e-Health, human-computer interaction, trust, adoption

1. Introduction

Electronic Health (e-Health) "is an emerging field in the intersection of medical informatics, public health and business, referring to health services and information delivered or enhanced through the Internet and related technologies. In a broader sense, the term characterizes not only a technical development, but also a state-of-mind, a way of thinking, an attitude, and a commitment for networked, global thinking, to improve health care locally, regionally, and worldwide by using information and communication technology.'' (Pawar et al. 2012) Many e-Health systems are developed recently, such as health information websites, online social support networks, interactive electronic health records, health decision support systems, and mobile health communication systems. These e-Health systems aim to increase consumer and provider access to relevant health information, improve the quality of care, and encourage the adoption of healthy behaviors (Kreps and Neuhauser 2010).

How to design an intelligent e-health system that communicates effectively with health care consumers and providers, however, is a question still does not have a standard answer so far. Many research review articles calls for e-Health tools that are easy to use, flexible, interactive, interoperable, and accessible for different audiences (Kreps and Neuhauser 2010). Although researchers have identified certain desired features of these tools (Kim and Chang 2007; Klasnja and Pratt 2012; Sillence et al. 2006), there is no formal, integrated framework to guide the design and development of an intelligent e-health system, which can provide the right information for different audiences at the right time, in the right place, and in the best ways to guide health care and health promotion.

To address the aforementioned gaps, we review key communication issues involved in the design of effective e-Health applications, apply human–computer interaction (HCI) principles, and propose a new framework to determine feature sets of prospective intelligent health management systems.

Specifically, we aim to answer the following research questions:

1. What are the design features that may enhance the trust and adoption of an intelligent e-health management system?
2. How to design the user interface of an intelligent e-health management system for effective tracking health information and presenting health management services?

2. Literature Review

2.1 Factors impacting e-Health system trust and adoption

Trust in health information systems has been studied extensively in the field and has been proved to be an important factor predicting the adoption of the system (Song and Zahedi 2007). A staged model of trust (Sillence et al. 2006) describes different stages of trust that develop over time, and gives a detailed list of design and content features that may affect the trust in different stages for web-based health systems. Particularly, users may use heuristic cues such as the website's visual appeal and layout/navigation to rapidly screen available websites, and then in the second stage use content features (e.g., whether the site has commercial background or not, whether the site has external links, etc.) to further evaluate the trustworthiness of a website. In the last stage, design and content features such as interactivity and personalized content can help with information integration and building long term trust.

Based on Davis' Technology Acceptance Model (TAM), another study (Kim and Chang 2007) proposes a model to explain the impact of functional characteristics including "Information Search," "Usage Support," "Customization," and "Purchase & Security" on the adoption of health information systems. However the details about the functional characteristics are not available.

Building on DeLone and McLean's model of IS success, Chatterjee et al. (2009) examine a set of factors for success mobile work in healthcare. Similar to the work in (Sillence et al. 2006), they also group the factors based on the categories of "system quality", "content quality", and "service quality". However, out of their proposed features, they only found portability (system quality), task structure (content quality), spatial mobility (content quality), and system reliability and support (system quality) as the key success factors when incorporating mobile technology into the overall information systems strategy.

Several research studies also point out the importance of user characteristics and their impact to the adoption of e-health systems. For example, it was found that people with disabilities or serious chronic conditions, individuals trying to maintain healthy lifestyles, parents with small children, and the elderly or their caregivers are more likely to adopt personal health records (PHRs) (Archer et al. 2011). In addition, systems tailored to individual needs are found to have higher user acceptance and satisfaction (Archer et al. 2011; Klasnja and Pratt 2012).

Based on the above models we summarize a set of system design and content features for enhancing web-based and/or mobile health systems trust and adoption. We include a "user characteristics" dimension, including factors such as user's medical condition, and will use this dimension to adjust the system and content design.

2.2 User Interface Design for effective e-Health Solutions

Usability (user interface and support) is believed to be the key to the adoption and use of e-Health systems (Archer et al. 2011). Besides general design features listed in the section 2.1, since e-Health systems are dealing with patients with various kinds of disease, the user interface design also requires some special features to cope with disease-specific barrios, such as problems with flashing or moving objects, low contrast, inappropriate font size, poor navigational design, crowded or cluttered screens, and difficulty seeing red (Archer et al. 2011). For example, Lorenz and Oppermann (2009) listed design characteristics that are recommended for elderly users such as: font sizes between 36pt and 48 pt; horizontal and vertical grid alignment of all used elements; arrangement of the buttons at the bottom of the interface; one-level navigation instead of using menu structures; redundant user guidance by color-coding and blinking boxes; color-neutral displays for visual impaired users; and slow animation speed. In addition, design of e-Health systems need to keep privacy in mind – for example systems should include some non-identifying username to reduce the risk of inconsistent data and should not use audio without warning (Doherty et al. 2010).

For m-Health systems using mobile devices especially smart phones, multi-touch interfaces and context-aware hardware sensors make it possible to design highly interactive and customized applications (Liu et al. 2011). For example, when designing tracking tools using mobile devices, multiple input methods such as through text or voice could be applied, with the latter is more suited for elderly populations or patients with hand-eye coordination or eyesight problems. Studies also suggest that tracking results should be uploaded automatically, reminders could be automatically triggered, presentation of patients' data should be interactive and the improvement patterns could be depicted using 2D or 3D visualizations, and feedback from either clinicians or through algorithm should be quick and timely (El-Gayar et al. 2013; Liu et al. 2011).

Many studies agree that it is important to design the user interface tailoring to individual characteristics. In a study reviewing the interventions providing personally tailored SMS, researchers found tailoring variables including participant's name or nickname, nominated support person's name, age, gender, behavioral history, behavioral preferences, behavioral goals, behavioral barriers, previous SMS responses, and medical status (Fjeldsoe et al. 2009). A more detailed user ontology including demographical profile, life history, diagnosis, treatment, and treatment providers was used in designing successful mental health interventions (Doherty et al. 2010).

To design and develop a truly interactive, customized user interface for e-Health systems, end-users should be involved during system design and development (user-centered design) (El-Gayar et al. 2013). Gulliksen et al. (2003) identify 12 key principles for the development of user-centered systems. Similarly, ISO 13407 "Human-centered design processes for interactive systems" and ISO 20282 "Ease of operation of everyday products" provide guidance on human-centered design for everyday products and especially computer-based interactive systems (Lorenz and Oppermann 2009). Typically, a user-centered design includes cycles for specification, mock-ups, user feedback and refinements. The result of each single loop becomes an input to the subsequent loop (Lorenz and Oppermann 2009). A staged evaluation strategy can prove useful in this user-centered design process (Doherty et al. 2010).

We plan to apply user-centered design guidelines to design a personalized UI based on user characteristics. The initial design will be based on general and special design features aforementioned.

3. A Framework for Designing an Intelligent Health Management System

In this study we propose a framework (Fig. 1) for designing personalized, mobile health information management platform for enhanced adoption. How to assess the user characteristics (what information would be useful) and use them to customize the UI and communication channels would be the key.

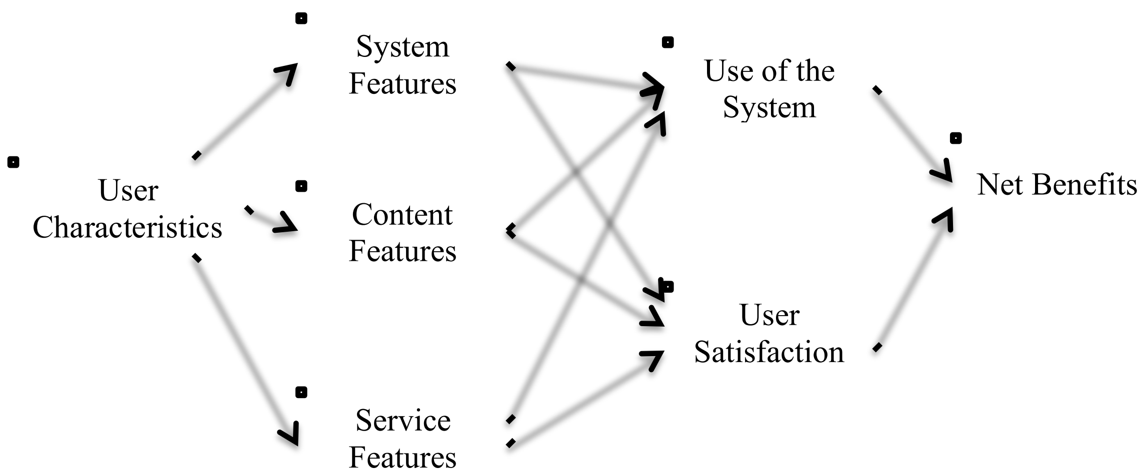


Fig. 1 Proposed Framework

In this new framework, we will add a new group of factors called User Characteristics, and the initial factors will include age, gender, and medical condition. We may add other factors based on literature review and focus group and/or survey results.

We believe the User Characteristics will affect the design of System Features, Content Features, and Service Features. Based on the literature review, we decide to include the following ones as the three sets of features.

- System Features: Visual appeal; Layout/navigation; Information search; Social identity cues; Advert; Brand; Portability; Interactivity.
- Content Features: Language style and tone; Site motivations; Content level; Task structure; Source knowledge; Cross referencing; Personalized content; Updated content; User generated content
- Service Features: System reliability; System support

The three sets of design features would then impact the use of the system and the user satisfaction to the system, and finally determine the benefits of the system. Similar to (Chatterjee et al. 2009)and (Liu et al. 2011), we could assess the benefits along multiple dimensions, such as:

improvement in clinical outcome, adoption of healthy behavior, and/or reduced disease management cost.

We will use questionnaires and focus groups to initiate the user-centered design process, develop a system prototype, and then use lab experiment to evaluate the different design features and the entire design framework. Qualitative methods such as interviews will be used to complement the experiment.

5. Conclusion

In this paper, we review key features that determine the trust and adoption of e-Health systems and key issues and principles involved in the design of effective e-Health applications. We then propose a new framework and a user-centered design process to determine feature sets of prospective intelligent health management systems. Such a formal, integrated framework is expected to guide the design of an intelligent e-health system that communicates effectively with health care consumers and providers.

References

Archer, N., U. Fevrier-Thomas, C. Lokker, K. A. McKibbon, and S. E. Straus. 2011, "Personal Health Records: A Scoping Review," *Journal of the American Medical Informatics Association : JAMIA* (18:4), Jul-Aug, pp. 515-522.

Chatterjee, S., S. Chakraborty, S. Sarker, S. Sarker, and F. Y. Lau. 2009, "Examining the Success Factors for Mobile Work in Healthcare: A Deductive Study," *Decision Support Systems* (46:3), pp. 620-633.

Doherty, G., D. Coyle, and M. Matthews. 2010, "Design and Evaluation Guidelines for Mental Health Technologies," *Interacting with Computers* (22:4), pp. 243-252.

El-Gayar, O., P. Timsina, N. Nawar, and W. Eid. 2013, "A Systematic Review of IT for Diabetes Self-Management: Are we there Yet?" *International Journal of Medical Informatics* (82:8), pp. 637-652.

Fjeldsoe, B. S., A. L. Marshall, and Y. D. Miller. 2009, "Behavior Change Interventions Delivered by Mobile Telephone Short-Message Service," *American Journal of Preventive Medicine* (36:2), pp. 165-173.

Gulliksen, J., B. Göransson, I. Boivie, S. Blomkvist, J. Persson, and Å Cajander. 2003, "Key Principles for User-Centred Systems Design," *Behaviour & Information Technology* (22:6), pp. 397-409.

Kim, D., and H. Chang. 2007, "Key Functional Characteristics in Designing and Operating Health Information Websites for User Satisfaction: An Application of the Extended Technology Acceptance Model," *International Journal of Medical Informatics* (76:11), pp. 790-800.

Klasnja, P., and W. Pratt. 2012, "Healthcare in the Pocket: Mapping the Space of Mobile-Phone Health Interventions," *Journal of Biomedical Informatics* (45:1), pp. 184-198.

Kreps, G. L., and L. Neuhauser. 2010, "New Directions in eHealth Communication: Opportunities and Challenges," *Patient Education and Counseling* (78:3), pp. 329-336.

Liu, C., Q. Zhu, K. A. Holroyd, and E. K. Seng. 2011, "Status and Trends of Mobile-Health Applications for iOS Devices: A Developer's Perspective," *Journal of Systems and Software* (84:11), pp. 2022-2033.

Lorenz, A., and R. Oppermann. 2009, "Mobile Health Monitoring for the Elderly: Designing for Diversity," *Pervasive and Mobile Computing* (5:5), pp. 478-495.

Pawar, P., V. Jones, B. F. Van Beijnum, and H. Hermens. 2012, "A Framework for the Comparison of Mobile Patient Monitoring Systems," *Journal of Biomedical Informatics* (45:3), pp. 544-556.

Sillence, E., P. Briggs, P. Harris, and L. Fishwick. 2006, "A Framework for Understanding Trust Factors in Web-Based Health Advice," *International Journal of Human-Computer Studies* (64:8), pp. 697-713.

Song, J., and F. Zahedi. 2007, "Trust in Health Infomediaries," *Decision Support Systems* (43:2), pp. 390-407.

Understanding organizational employees' information security policies compliance behavior: the critical roles of perceived justice and corporate ethical values

Yulong Ke	Wenli Li	Jin Yin
Dalian University of Technology	Dalian University of Technology	Dalian University of Technology
keyulong@mail.dlut.edu.cn	wlli@dlut.edu.cn	yinjin@mail.dlut.edu.cn

Abstract

This paper aims to understand how perceived justice, corporate ethical values influence organizational employees' commitment in turn shape employees' information policies compliance intention. Drawing on organizational commitment theory, we proposed a research model that integrated perceived justice, corporate ethical values and employees' organizational commitment. With 254 employee survey responses from 4 organizations, the results indicated that perceived justice, corporate ethical values were positive associated with organizational employees' commitment. Moreover, affective commitment and normative commitment had significant effect on employees' information security policies compliance intention. However, the impact of continuous commitment on compliance intention was not significant. Theoretical and management implication for organization information security management were discussed.

Keywords: Perceived Justice, Corporate Ethical Values, Organizational Commitment, Information security policies Compliance

1. Introduction

Due to facing the severe cyber threats, information security is consider as vital to the ongoing health and success of the organization and many organizations have been increasing investment to improve information security. The latest global information security survey reported by Ernst & Young indicate that 70% of organizations own information security policies at the highest organizational level(Kessel et al. 2013). Although information security policies tell the organization employees the right business process, what should do and what should be punished clearly and explicitly. Prior research reported that organization employees seldom comply with the information security policies and may choose to commit security policies violation for reasons of convenience in their day-to-day routine(Herath et al. 2009b; Von Solms et al. 2004). Thus, employees' information security policies compliance still remains a significant concern for organizations and much more research are needed to understand employees' information security policies compliance in order to encourage compliance in organizations(Herath et al. 2009a; Ifinedo 2013).

Scholars strive to understand the mechanism of information security policies compliance and violation behaviour by investigating information security on a numbers of theories, such as general deterrence theory(D'Arcy et al. 2009; Herath et al. 2009b; Hovav et al. 2012; Hu et al. 2011; Straub 1990), protective motivation theory(Ifinedo 2012; Johnston et al. 2010; Vance et al. 2012), theory of planned behaviour(Bulgurcu et al. 2010; Hu et al. 2012; Ifinedo 2012). Perceived justice which refers to employees' perceptions of fairness/unfairness in organizations is considered to be an important factor that influence employees' IT policies compliance(Xue et al. 2011). Crossler et al. (2013) also emphasize that how organizational justice influences compliance is an avenue to understand employees' information security policies compliance. However, the effect of perceived justice on employees' information security policies compliance has not been adequately studied. Moreover, to our knowledge, corporate ethical values considered to be a key component culture has not been investigated in the field of information security.

To bridge the research gap, draw on organizational commitment theory, this paper aims to investigate the effect of perceived justice and corporate ethical values on shaping the employees' information security policies compliance. Thus, for this paper, the research questions are:

RQ1: How does perceived justice and corporate ethical values influence employees' organizational commitment?

RQ2: How do they contribute to the formation of the employees' information security policies compliance intention?

The remainder of the paper is structured as follows: in the next section, prior relevant literature that related to organizational commitment theory are reviewed and present the theoretical development. Moreover, the research model and hypotheses are proposed. Following that, the research method, hypotheses testing and discussion are conducted. Finally, we draw a conclusion for the theoretical contribution and practical implication.

2. Theoretical Foundations

Commitment refers to “a psychological state that binds the individual to the organization” (Allen et al. 1990 P.14). Commitment can influence behaviour even in the situation without extrinsic motivation and it is difference from exchange-based forms of motivation or target-relevant attitudes(Meyer et al. 2001). Organizational commitment theory proposed by Allen et al. (1990), consisted of three component, labelled as affective commitment, continuous commitment, and normative commitment. Affective commitment refers to an emotional bind to, involvement or identification with the organization(Allen et al. 1990). Continuous commitment refers to costs benefit analysis associated with leaving the organization and normative commitment refers to moral obligation to the organization(Allen et al. 1990). Organizational commitment theory has been used by researchers to predict a variety of important employee outcomes, including job performance(Aryee et al. 2002), turnover intention(Poon 2012), website continuous usage(Li et al. 2006) and information security policies compliance(Herath et al. 2009b).

3. Research Model and Research Hypotheses

3.1 Employees’ organizational commitment and information security policies compliance

In this paper, organizational commitment refers to the tie that an employee attach with an organization(Herath et al. 2009b). Organization behaviour research has used organization commitment to predict a variety of important employee outcomes, such as job performance(Aryee et al. 2002), turnover intention(Poon 2012). In the field of information system, Li et al. (2006) use organizational commitment to predict online users’ website continuous usage. In the field of information security, Herath et al. (2009b) argued that employees’ organizational commitment influence employees’ effectiveness perception of ones’ action in turn influence employees’ information security policies compliance. The higher the organizational commitment, the more likely the employees will engagement in security behaviours and the less he/she will put the organization at a risk(Herath et al. 2009b). Thus, we propose the following hypothesis：

H(1/2/3): Affective commitment/ continuous commitment/ normative Commitment. is positively associated with information security policies compliance intention.

3.2 Perceived justice and employees' organizational commitment

Particularly, perceived justice which refers to employees' perceptions of fairness/unfairness in organizations is considered to be an important factor that influence employees' policies compliance(Xue et al. 2011). It is generally accepted that perceived justice consist of four dimension: distributive justice, procedural justice, interpersonal justice and informational justice(Ambrose 2002; Colquitt 2001). Distributive justice defines as the perceived fairness of outcome distributions, procedural justice focus on the decision-making processes, interpersonal justice emphasize on being treated respect and informational justice refers to the being provided with explanations for the decision(Xue et al. 2011). Xue et al. (2011) argued that have no effect on employee behavioural outcomes. Therefore, in this research, we conceptualize perceived justice as a formative second-order construct with three first order construct except interpersonal justice.

Prior research has reported that perceived justice significantly influence employee attitudes and behaviours(Colquitt 2001; Xue et al. 2011). Willison et al. (2013)argue that workplace disgruntlement may be an important motivation in relation to employee computer crime. Perceived justice is considered to be an important factor that influence employees' policies compliance(Xue et al. 2011).

Cohen-Charash et al. (2001) reported that procedural and distributive justice is positively associated with organizational commitment. Baker et al. (2006) also conduct a conclusion that procedural and distributive justice had positive impact on organizational commitment. When employees perceived that he was treated fairly, the more he will comply with the organization policies and rules. The more an employee perceived justice, the more valuable he will feel to a group and the more likely to identify to an organization. Thus, we propose the following hypothesis：

H(4a/4b/4c): Perceived justice is positively associated with affective commitment/ continuous commitment/ normative Commitment..

3.3 Corporate ethical values and employees' organizational commitment

Corporate ethical values refers to a subset of organizational culture, including various "formal" and "informal" systems of behavioural control(Treviño et al. 1998). Corporate ethical values can significantly impact an employee' behaviour(Baker et al. 2006). The formal systems include policies, reward system and the informal systems include beliefs and norms. Prior research has reported that corporate ethical values positive associated with job satisfaction, organizational commitment(Baker et al. 2006; Schwepker Jr 2001). Hu et al. (2012) also conclude that organization culture played the critical role on employees' information security policies compliance. Thus, we propose the following hypothesis：

H(5a/5b/5c): Corporate ethical values is positively associated with affective commitment/ continuous commitment/ normative Commitment.

4. Research Method

4.1 Measurement Items

The measurement items were adopt from prior existing literature. The items for distributive justice, informational justice and procedural justice were draw on Xue et al. (2011) and did a little modification. The scales for corporate ethical values were adapt from Valentine et al. (2011).The items for organizational commitment were based on Li et al. (2006) and Allen et al.

(1990). The scales for information security policies compliance were referenced from Bulgurcu et al. (2010).The items were using seven-point liker scale anchored by 1 (strongly disagree) and 7 (strongly agree).

4.2 Data Collection

In order to ensure the validity and reliability of the sample, the pre-test was performed. And then, we distributed a total of 300 paper-based questionnaires to the organization employees in the department of finance and accounting of four international Business Process Outsourcing company based in the northeast of China. A t-test was conduct and the result indicated that this was no difference between the four company samples. And 252 valid questionnaires was used to conduct the data analysis after removing some error and incomplete questionnaires.

5. Data Analysis

5.1 Measurement Model

SmartPLS (Ringle C M et al. 2005) was used to conduct the data analysis. There are two reason that partial least squares (PLS) was choose. One is that PLS do not require strict normal distribution of the sample(Chin 1998). The other one is that PLS can perform the formative second-order construct easily(Chin 1998). In our paper, perceived justice is a Reflective-Formative second-order construct which is consists of distributive justice, informational justice and procedural justice. The results of measurement model were shown in table 1.

Construct	Composite Reliability	Cronbachs Alpha	AVE	1	2	3	4	5	6	7	8
Corporate Ethical Values	0.904	0.787	0.824	**0.908**							
Distributive Justice	0.956	0.930	0.878	0.295	**0.937**						
Informational Justice	0.880	0.795	0.709	0.487	0.289	**0.842**					
Procedural Justice	0.925	0.879	0.805	0.393	0.333	0.609	**0.897**				
Affective Commitment	0.931	0.901	0.771	0.502	0.131	0.461	0.312	**0.878**			
Continuous Commitment	0.900	0.851	0.694	0.390	0.288	0.323	0.464	0.204	**0.833**		
Normative Commitment	0.884	0.800	0.719	0.532	0.190	0.460	0.441	0.517	0.424	**0.848**	
Compliance Intention	0.960	0.917	0.923	0.441	0.254	0.332	0.237	0.446	0.164	0.371	**0.961**

Table 1. Reliability，Validity and Correlation of the Latent Variable Scores

According to the guideline propose to evaluate the reliability and validity by (Fornell et al. (1981)), the composite reliability and Cronbachs Alpha are above 0.7 which indicate the reliability is adequate. The factor loading and the average variance extracted（AVE）all provided support for the adequate convergent validity. Moreover, the square root of average variance extracted are all above the inter-construct correlations which demonstrated a good discriminant validity

5.2 Structural Model

Following the measurement model testing, the structural model was also conducted. The variance explained by the research model was 28.7% for employees' information security policies compliance intention, 27.9% for affective commitment, 24.7% for continuous commitment and 34.2% for normative commitment.

The result indicated that the path from affective commitment (H1: $\beta=0.327, p<0.001$), normative commitment (H3: $\beta=0.212, p<0.01$) to information security policies compliance intention are significant. These indicate that the strong and positive emotional attachment and high moral obligation to an organization are an important role to employees' information security policies compliance. And affect commitment plays the predominating role. That reveals that in the field of information security, employees' identification plays the most important role in the security policies compliance. These findings are consistent with prior research(Allen et al. 1990; Li et al. 2006). In a study of Meyer et al. (2001) found that affective commitment plays the most important role association with behaviours. Yet the effect of continuous commitment (H2: $\beta=0.012, p>0.1$) on information security policies compliance intention is not significant. The finding also consistent with prior research(Allen et al. 1990) which found that continuance commitment was unrelated to organization employees outcomes. One reason for this may be when an employee chose to commit an computer crime or policies violation, the perceived benefit exceeded the cost which consist with the finds of Hu et al. (2011).

Moreover, perceived justice was found to have signification positive effects on affective commitment (H4a: $\beta=0.192, p<0.01$), continuous commitment (H4b: $\beta=0.359, p<0.001$) and normative commitment (H4c: $\beta=0.282, p<0.001$). These findings are consistent with prior research(Baker et al. 2006; Poon 2012). It indicates that the more an employee perceived justice, the more valuable he will feel to a group and the more likely to identify to an organization.

The effect of corporate ethical values on affective commitment (H5a: $\beta=0.404, p<0.001$), continuous commitment (H5b: $\beta=0.208, p<0.001$) and normative commitment (H5c: $\beta=0.389, p<0.001$) were all significant. The findings also consist with prior research(Baker et al. 2006; Valentine et al. 2011) which indicates that organizational culture plays an important role in the formation of employees' organizational commitment in turn shape employees' information security policies compliance.

6. DISCUSSION

6.1 Implications for Research and Practice

Results of our study indicate that organization commitment theory is a robust theory that can also be used to do research in the field of information system security in the organization context. Moreover, perceived justice and corporate ethical values are two important factors in shaping employees security behaviour. From a management perspective, creating justice climate and ethical culture have frequently been suggested as a means for deterring unethical behavior within the organization. This study finds that creating such justice climate and ethical culture may have additional benefits such as greater stronger organizational commitment and subsequently stronger information security policies compliance intentions. Therefore, much more actions should be taken to create justice climate and ethical culture.

7. CONCLUSIONS

Many organization have emphasized on the formal control, such as computer monitoring and punishment for the policies violation to ensure the information system security. However, the formal control are not sufficient and much more effect should be created on the justice climate and the ethical culture. We propose a research model integrated with organizational commitment, perceived justice and corporate ethical values. The results shows that perceived justice and corporate ethical values are all positive associated with organizational commitment in turn shape employees' information security policies compliance. It highlight that perceived justice and corporate ethics should take into account for the information security research. Meanwhile, it also highlight that creating justice climate and ethical culture play important role on employees' information security behaviour.

References

Allen, N. J., and Meyer, J. P. 1990. "The measurement and antecedents of affective, continuance and normative commitment to the organization," *Journal of occupational psychology* (63:1), pp 1-18.

Ambrose, M. L. 2002. "Contemporary justice research: A new look at familiar questions," *Organizational Behavior and Human Decision Processes* (89:1), pp 803-812.

Aryee, S., Budhwar, P. S., and Chen, Z. X. 2002. "Trust as a mediator of the relationship between organizational justice and work outcomes: Test of a social exchange model," *Journal of organizational Behavior* (23:3), pp 267-285.

Baker, T. L., Hunt, T. G., and Andrews, M. C. 2006. "Promoting ethical behavior and organizational citizenship behaviors: The influence of corporate ethical values," *Journal of Business Research* (59:7), pp 849-857.

Bulgurcu, B., Cavusoglu, H., and Benbasat, I. 2010. "Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs and Information Security Awareness," *Management Information Systems Quarterly* (34:3), pp 523-548.

Chin, W. W. 1998. "The partial least squares approach to structural equation modeling," *Modern methods for business research* (295:2), pp 295-336.

Cohen-Charash, Y., and Spector, P. E. 2001. "The role of justice in organizations: A meta-analysis," *Organizational behavior and human decision processes* (86:2), pp 278-321.

Colquitt, J. A. 2001. "On the dimensionality of organizational justice: a construct validation of a measure," *Journal of applied psychology* (86:3), p 386.

Crossler, R. E., Johnston, A. C., Lowry, P. B., Hu, Q., Warkentin, M., and Baskerville, R. 2013. "Future directions for behavioral information security research," *Computers & Security* (32:2), pp 90-101.

D'Arcy, J., Hovav, A., and Galletta, D. 2009. "User awareness of security countermeasures and its impact on information systems misuse: a deterrence approach," *Information Systems Research* (20:1), pp 79-98.

Fornell, C., and Larcker, D. F. 1981. "Evaluating structural equation models with unobservable variables and measurement error," *Journal of Marketing Research* (18:1), pp 39-55.

Herath, T., and Rao, H. R. 2009a. "Encouraging information security behaviors in organizations: Role of penalties, pressures and perceived effectiveness," *Decision Support Systems* (47:2), pp 154-165.

Herath, T., and Rao, H. R. 2009b. "Protection motivation and deterrence: a framework for security policy compliance in organisations," *European Journal of Information Systems* (18:2), pp 106-125.

Hovav, A., and D'Arcy, J. 2012. "Applying an extended model of deterrence across cultures: An investigation of information systems misuse in the U.S. and South Korea," *Information & Management* (49:2), pp 99-110.

Hu, Q., Dinev, T., Hart, P., and Cooke, D. 2012. "Managing Employee Compliance with Information Security Policies: The Critical Role of Top Management and Organizational Culture," *Decision Sciences*.

Hu, Q., Xu, Z., Dinev, T., and Ling, H. 2011. "Does deterrence work in reducing information security policy abuse by employees?," *Communications of the ACM* (54:6), pp 54-60.

Ifinedo, P. 2012. "Understanding information systems security policy compliance: An integration of the theory of planned behavior and the protection motivation theory," *Computers & Security* (31:1), pp 83-95.

Ifinedo, P. 2013. "Information systems security policy compliance: An empirical study of the effects of socialization, influence, and cognition," *Information & Management*:0).

Johnston, A. C., and Warkentin, M. 2010. "Fear Appeals and Information Security Behaviors: an Empirical Study," *Mis Quarterly* (34:3), pp 549-566.

Kessel, P. v., and Allan, K. 2013. "Under cyber attack: EY's Global Information Security Survey 2013," *Ernst & Young*).

Li, D., Browne, G. J., and Chau, P. Y. K. 2006. "An Empirical Investigation of Web Site Use Using a Commitment- Based Model," *Decision Sciences* (37:3), pp 427-444.

Meyer, J. P., and Herscovitch, L. 2001. "Commitment in the workplace: Toward a general model," *Human resource management review* (11:3), pp 299-326.

Poon, J. M. 2012. "Distributive Justice, Procedural Justice, Affective Commitment, and Turnover Intention: A Mediation–Moderation Framework1," *Journal of Applied Social Psychology* (42:6), pp 1505-1532.

Ringle C M, S, W., and A, W. 2005. "SmartPLS 2.0 (beta)," *Hamburg,Germany, http://smartpls.de*.

Schwepker Jr, C. H. 2001. "Ethical climate's relationship to job satisfaction, organizational commitment, and turnover intention in the salesforce," *Journal of Business Research* (54:1), pp 39-52.

Straub, D. W. 1990. "Effective IS security: An empirical study," *Information Systems Research* (1:3), pp 255-276.

Treviño, L. K., Butterfield, K. D., and McCabe, D. L. 1998. "The ethical context in organizations: Influences on employee attitudes and behaviors," *Business Ethics Quarterly*), pp 447-476.

Valentine, S., Godkin, L., Fleischman, G. M., and Kidwell, R. 2011. "Corporate ethical values, group creativity, job satisfaction and turnover intention: The impact of work context on work response," *Journal of Business Ethics* (98:3), pp 353-372.

Vance, A., Siponen, M., and Pahnila, S. 2012. "Motivating IS Security Compliance: Insights from Habit and Protection Motivation Theory," *Information & Management*.

Von Solms, R., and von Solms, B. 2004. "From policies to culture," *Computers & security* (23:4), pp 275-279.

Willison, R., and Warkentin, M. 2013. "Beyond deterrence: an expanded view of employee computer abuse," *MIS Quarterly* (37:1), pp 1-20.

Xue, Y., Liang, H., and Wu, L. 2011. "Punishment, Justice, and Compliance in Mandatory IT Settings," *Information Systems Research* (22:2), pp 400-414.

SPONSORS

CSWIM 2014 thanks the following sponsors for their support of the conference.

电子科技大学
经济与管理学院

电子科技大学 研究生院
University of Electronic Science and Technology of China

NSFC
National Natural Science
Foundation of China

CSWIM Website: www.cswimworkshop.org