THE ECONOMIC ROLE OF RATING BEHAVIOR IN THIRD-PARTY APPLICATION MARKET

Completed Research Paper

Lin Hao
Foster School of Business
University of Washington
Seattle, WA 98195-3226, USA
linhao@uw.edu

Xiaofei Li
Business School
Sichuan University
Chengdu, Sichuan 610064,China
lixf531@gmail.com

Yong Tan
Foster School of Business
University of Washington
Seattle, WA 98195-3226, USA
ytan@uw.edu

Jiuping Xu
Business School
Sichuan University
Chengdu, Sichuan 610064,China
xujiuping@scu.edu.cn

Abstract

This paper explores the fundamental influence of consumer rating behavior on an emerging third-party software application market, mobile app market. In app market, consumers' ex ante belief on app utility essentially is determined by the app rating while at the same time the app rating itself is derived from the ex post utility obtained by purchased customers. We develop an analytical model which explicitly characterizes this bidirectional rating-utility conversion based on a newly introduced concept "reservation rating". After building this conversion process into the utility function, we examine the market equilibrium and show how change in consumer rating attitude, such as being severer in offering ratings or being less critical in accepting ratings, would affect the developers' optimal choices of app price and app quality level as well as the platform owner's optimal revenue sharing policy.

Keywords: Online Word of Mouth, Mobile commerce, E-business, Pricing IS

Introduction

The traditional third-party application market is now heading towards its next phase. It has been witnessed that a novel third-party application market, mobile app market, is currently experiencing its explosive growth despite the recent economic downturn. Apple App Store, the world's leading mobile app market launched in July 2008, reached total number of 10 billion downloads in Jan. 2011 while it was just 3 billion in Jan. 2010 (*Wall Street Journal* Jan. 2011). Kaufman Brothers L.P estimated that over $1 billion revenue was generated from more than 350,000 applications inside the Apple App Store in 2010 and Citibank expected this revenue figure to be $2 billion in 2011. Following Apple's move, Google, Microsoft, Research in Motion (RIM) and Amazon opened their own mobile app stores. Gartner Inc. forecasted the total revenue of entire mobile app market will hit $15.1 billion in 2011, almost tripled its $5.2 billion in 2010. It will further increase to $35 billion in 2014 according to International Data Corp. (IDC)'s projection. Meanwhile, the tremendous success in this novel market is initiating a new trend which may change the whole climate of future third-party application market. Enlightened by the success of Apple App Store, Apple extended the same business model to desktop and laptop applications. On Jan 6th 2011, Mac App Store was opened with more than 1,000 computer apps out of the gate. In just 20 days, Pixelmator, a small software company, achieved 1 million dollar sale on Jan 25th from its one single image processing application sold at unit price $29.99. Traditional giant software companies such as Autodesk also joined the Mac App Store. The market is so enticing that even Microsoft, Apple's major competitor, is considering bringing its Microsoft Office's Mac version onto Mac App Store. "It's something we are looking at", said Amanda Lefebvre, Microsoft's senior marketing manager (*PCWorld* Jan. 2011).

Concerning this newly emerging economy, the adverse selection caused by information asymmetry is a serious issue which could potentially dismantle the market (Akerlof 1970) due to the following reasons. First, software product belongs to experience goods of which consumers can hardly observe true quality *ex ante* (Shapiro 1985). Second, even if the "objective" true quality is observed, it could mean different utility for different consumer depending on the consumer's subjective valuation for quality. The discounted utility is not necessarily known to the consumer herself/himself *ex ante* (Chen and Xie 2008). Third, one nature of the app market, the mixture of individual app developers and organizational app developers, generates greater belief dispersion on quality than the situation in which vendors are relatively more homogenous. Thus, without signals to distinguish the quality of the apps, consumers are almost clueless about their own *ex post* utility so that hardly figure out the corresponding willingness to pay. Adverse selection starts to occur and will potentially turn down the market.

To counter adverse selection, the rating, as derived from consumers' *ex post* experience, is a good source for consumers to form the correct *ex ante* beliefs on *ex post* net utility and therefore determine their willingness to pay (Li and Hitt 2010, Sun 2010). Though it is not the only source, it is much more influential than other factors in the app market compared to being in other markets. This is firstly because in the app market a plenty amount of ratings are aggregated and available at one single site easily to be found, such as iTunes for Apple. So it would not suffer the problem that ratings are widely distributed in multiple sites such that evidence from one single site doesn't have strong power to convey the quality message to consumers. Secondly, consumers are allowed to rate at that site only after they purchase the app, which to a large extent avoids shilling behavior so that keeps high credibility for ratings. Third, most importantly, other signaling devices such as advertising and branding are usually relatively weak due to individual developers' budget constraint or lack of marketing capability. Therefore, rating is the main criterion consumers rely on in the app market. Besides the consumers, developers and platform owners also pay significant attentions on ratings. Evidence from real business practices shows that the rating is one of the hottest topics and the biggest concerns among developers. Tens of thousands of posts on iPhonedevsdk.com, one of the most popular online iPhone developer communities, are related to how to improve ratings and how ratings affect the app sales revenue. Furthermore, since the developer agrees upon a revenue sharing contract with the platform owner, the platform owner's revenue is also connected with ratings, which naturally makes it a concern for the platform owner. All the above facts suggest that in the app market ratings play a very significant role, if not a central role, in determining both the success of an app and the prosperity of the entire market.

Given its great importance, our study focuses on the impact of rating on the app market in terms of the consumer's choice of purchase, the developer's choice of app price and quality level, and the platform owner's choice of revenue sharing policy. We model the ratings by explicitly characterizing the consumer rating behavior and synthesizing it into the model at fundamental level, i.e., into the consumer's utility function. Then we derive the developer's optimal choice of app price and quality level based on the consumer's rating-dependent utility function. Ultimately the platform owner's decision on revenue sharing policy will also be associated with the consumer rating behavior due to the revenue sharing contract with the developer.

We account that the consumer rating behavior consists of two processes. One is to offer ratings, also referred as *ex post* ratings, based on *ex post* user experience which can be quantified by the consumer's received net utility after purchase (Kuksov and Xie 2010). The other is to accept *ex post* ratings as translating them into *ex ante* perceived net utility. Combining these two processes we eventually construct a bidirectional rating-utility system in which we call the former as utility-to-rating process and the latter as rating-to-utility process. Instead of not taking into account the rating behavior as previous literature (Li and Hitt 2010; Sun 2010), we characterize the consumer rating behavior in both process. To model the subjective rating behavior in the utility-to-rating process, we assume a positive linear relationship between *ex post* rating and received net utility, which we call as rating function. For the rating-to-utility process, we introduce a new concept, reservation rating. Reservation rating is defined as the *ex post* rating which signifies zero received net utility in consumers' *ex ante* perception. Based on reservation rating, a linear function between *ex post* rating and *ex ante* perceived net utility is established with potentially different slope and intercept than rating function.

Our paper contributes to the existing literature in the following two aspects. First, this paper proposes an analytical framework which fundamentally synthesizes the consumer rating behavior with economic utility theory by constituting a bidirectional rating-utility system. While we take app market as our research context, the bidirectional rating-utility system methodology can be applied in any market where consumer rating plays an important role in consumer's purchase decision. The concept of reservation rating invented for rating-to-utility process is an attempt to systematically model how consumers interpret ratings into net utility. Moreover, in our paper, we provide a justification from the economic utility theory angle to show the principle that over-leniency is detrimental to quality, which is empirically revealed by management behavioral literatures.

Second, based on above rating-utility system, our paper studies the impact of consumer subjective rating behavior on the consumer's choice, the developer's choice, and the platform owner's revenue sharing policy for the newly emerging app market. We provide results which are coincided with the evidences in real business practice and generate important managerial implications for the platform owners and developers in the app business, including how the change in consumer rating behavior would change their profit and how they should optimally react to the various consumer rating behaviors.

The rest of the paper is organized as follows. In the next section we present a review of the related literatures. This is followed by the model and findings. Then, we conclude the paper by summarizing the results and providing discussions about the directions for future research.

Literature Review

The literature on effects of online word-of-mouth has been proliferating rapidly during the last ten years. Most of them are empirical work in the context of movie and book industry, focusing on the effects of online word-of-mouth (WOM) in predicting or influencing sales revenue (Chevalier and Mayzlin 2006; Dellarocas 2004; Liu 2006; Duan et al. 2008). Dellarocas (2004) demonstrated that the online rating is a useful proxy for WOM in movie industry and it serves as one of the predictors for the movie's total revenue. Dellarocas and Narayan (2006) identified three metrics of online word-of-mouth: valence, variance and volume, in which valence is usually denoted by the average numeric average rating, variance is usually denoted by its statistical variance or entropy (Godes and Mayzlin 2004), and volume is counted as the number of ratings. Liu (2006) showed that the online WOM has significant explanatory power for box office revenue while the volume is stronger than valence. Duan et al. (2008) revealed that movies' online WOM valence significantly affects the volume which in turn determines the box office performance. For book market, Chen et al. (2004) found that the consumer rating is only a predictor for the book sales. However, Chevalier and Mayzlin (2006) showed that improvement of the online WOM valence would increase the book sales based on the data from Amazon.com and BarnesandNoble.com.

Forman (2008) also demonstrated that raters offer the rating more positively to the review with identity information and disclosure of identity information would increase the sales.

Researcher also investigated the connection between consumer ratings and sales in the context of other markets such as beer, DVD and video games (Clemsons 2006, Hu et al. 2008, Zhu et al. 2010). Little attention has been paid to the software market. One reason might be the online software selling has not been widely established until the app market appears. Zhou and Duan (2010) stated that evidence from CNET download.com shows increase in product variety strengthens the impact of positive consumer reviews but weakening the impact of negative ones. It also showed that positive professional reviews leads to more software download.

Given the above connection between the online WOM and the product sales, it is the natural next step for researchers to conceive firms' optimal strategy to adapt such connection. A growing body of literature has been devoted onto this topic and most of them are on analytical side. Chen and Xie (2005) showed that firm should choose advertising instead of price to adapt consumer review when sufficient consumers value the product's horizontal features. Dellarocas (2006) demonstrated how firms' shilling behavior, i.e. post anonymous messages that exalt their products on the purpose of changing the consumers' perception, will influence the firms' profits and consumers' surplus. Chen and Xie (2008) revealed that when and how the sellers should adjust their marketing communication strategy by changing product attribute information to adapt the consumer reviews. Kuksov and Xie (2010) studied the firm's optimal pricing and whether the firm should give an unexpected frill to early customers to boost their product experience. Li and Hitt (2010) both analytically and empirically showed that uni-dimensional ratings are more correlated with the value of the product rather than quality of the product. Firms need to account for price effects and could better serve the consumers by setting up the review system which explicitly separates the perceived value and perceived quality. Jiang and Chen (2007) considered both consumer reviews and consumer ratings and found that firms may have incentive to under-change in the early period.

Most of above firms' adaption strategies are implicitly based on the underlying assumption that ratings can signal the underlying true quality (or true value) to novice consumers at least in an expectation sense. However, it seems to be not always true, at least not always in a precise manner. The topic concerning how precise the ratings reflect the underlying true quality, an issue involved with consumer rating behavior, becomes increasingly popular in rating-relate research area. Hu et al. (2006) empirically showed that in presence of under-reporting, i.e., only extremely satisfied or extremely unsatisfied consumers would rate, consumers may not extract the true quality from the ratings if only valence (mean value) is known. Sun (2010) analytically demonstrated that consumers in subsequent period can figure out the exact product's true quality from the distribution of the ratings given by consumers in previous period. Li and Hitt (2008) showed that later consumers' perception on the quality may be biased because of the biased ratings left by early buyers whose preferences on quality are quite different from later ones'. Hu (2009) summarized two self-selection bias, purchasing bias which is partially mentioned in Li and Hit (2008), and under-reporting bias mentioned in Hu (2006), as two reasons which potentially lead to biased perception on true quality under certain circumstances. Moe and Trusov (2010), Moe and Schweidel (2011) also empirically showed that consumer rating behavior is significantly affected by previously posted rating. Lee (2009) demonstrated that social imitation and learning affect can influence user rating generation.

Another important aspect of rating behavior which influences appraisal accuracy is the systematic rater error. While having not been noticed in the context of online WOM, it has been studied in the management literature for decades (Borman 1977; Kane 1994, Kane et al. 1995; Saal et al. 1980; Yun et al. 2005). Managers will make different inference on one employee's productivity when assessing the same sheet of his performance rating. Having the random rating error removed, this subjectivity has been demonstrated to be systematic (Kane 1994) if no control effort is made upon its systematic sources. Obviously this is the case for app market since consumers on the app market are under no control for their characters like maturity, conscientiousness and degree of sophistication in using apps, which could serve as such systematic resources. Kane (1994) also summarized the rating error into multiple categories and point out that leniency and severity are two important ones which could be potentially systematic. Kane et al. (1995) demonstrated that rating leniency is the most troublesome rating error and found it to be a relatively stable response tendency. Spence and Keeping (2010) suggested that when managers give performance rating to their employees, more experienced managers are associated with lower ratings, which is less degree of leniency. Berger et al. (2010) empirically showed that under the situation that employees' bonus payments are associated with ratings, a forced distribution requirement on ratings actually lead higher productivity than that under unforced distribution which is more skewed towards the higher rating values than forced one. This is especially intriguing to our research since by analogy consumers in our context are more or less alike managers who rate developers' apps and consequently developers' revenue are associated with those ratings.

The Model

Suppose the platform owner, the app developer and the app consumer are the three players on the app market. A three-stage dynamic game theoretic model is constructed here to study the equilibrium of the market. At the first stage, the platform owner determines and publicizes the developer's revenue-sharing percentage. By observing this percentage at the beginning of the second stage, the developer chooses either not to adopt and then exit the market, or to adopt and then determine the optimal quality level and optimal app price. If the developer chooses not to adopt the game ends. Otherwise, at the third stage, consumers decide whether to purchase or not. If purchased, they rate the app based on the received net utility. The third stage is repeated such that after a sufficient time period the rating becomes steady. The goals of the platform owner and the developer are to maximize their own benefit in this steady-state respectively. A developer would choose to participate when the profit is greater or equal to zero. Consumers would make the purchases when the expected net utility is greater or equal to zero.

Before deriving the propositions regarding the equilibrium, we elaborate several essential preliminaries which serve as the foundation of our model analysis.

Rating Function

We assume the linear relationship between the received net utility and the *ex post* rating. Since the rating is bounded between 0 and 1, the mathematical formula is given by:

$$rating = \min\left\{1, \max\left\{0, ku + r_0\right\}\right\} \tag{1}$$

u is the received net utility. r_0 is the *ex post* zero net utility rating which is an important concept in measuring the degree of severity. Low r_0 indicates that consumers are severe with respect to offering ratings. k is the rating-utility conversion rate. It represents the consumer's sensitivity of the *ex post* rating on difference in received net utility. The rating function models the subjectivity of rating behavior. For example, in the utility-to-rating process under a typical 10-point rating system, it is rather the consumer's subjective opinion to rate $r_0 = 4.0$ or $r_0 = 6.0$ when received net utility is exactly zero.

The Received Net Utility

In the similar spirit of Chen and Xie 2008, we partition the consumers into two groups. One consists of all the high valuation consumers who consider the app match their taste so that they appreciate the quality. The other consists of all the low valuation consumers who consider the app mismatch their taste so that they don't appreciate the quality. Then the received net utility of these two groups, which is associated with the app's "objective" true quality level q, its price p and the consumer's valuation for quality, are given by:

$$u_{matched} = q - p, u_{unmatched} = -p\ . \tag{2}$$

We denote b as the fraction of consumers who belong to high valuation group and then the fraction for the low valuation group is $1-b$. It is worthwhile noticing that b is an indicator of the developer's marketing performance. Large b represents high marketing performance since most of consumers are "matched". We suppose $b \in (0,1)$.

The Reservation Rating

Followed the definition of "reservation rating" in the introduction section, we would further explain how it fits into the economic utility theory. Reservation rating is not only intuitively a bar which the app needs to pass in order to be considered for purchase, but also fundamentally affects the consumers' willingness to pay. It works as a "ruler origin" to measure the *ex ante* perceived net utility from an app. For instance, with regard to a 5-star rating system, when a consumer with reservation rating of 3.5 stars observes an app with current rating of 4.5 stars, she/he would expect some positive net utility from the app. And a maximum rating of 5 stars would elevate this expectation. In other words, combining the reservation rating and the current *ex post* rating, the consumer will figure out their *ex ante* perceived net utility.

In the base model, the reservation rating r_R is assumed to be homogenous among all consumers. Suppose r is the *ex post* rating given from high valuation consumers. The *ex ante* perceived net utility of a consumer conditional on she/he belongs to high valuation group is given by:

$$u_e = \frac{r - r_R}{k} \tag{3}$$

In order to distinguish the degree of criticism between utility-to-rating and rating-to-utility processes, we call that consumers with low r_0 are "severe" and consumers with high r_R are "critical". Since people are generally more critical in accepting ratings than offering ratings, we assume that $r_R > r_0$.

Next, we derive the consumer's expected net utility $E(U)$. We suppose that both the true quality level and the type of valuation are unknown to the consumer before she/he experiences the app. However, through examining the distribution of app's *ex post* rating she/he would discover the probability of being in the high valuation group is b and the probability of being in the low valuation group is $1-b$. Hence, the expected net utility is given by:

$$E(U) = (1-b)\cdot(-p) + b \cdot u_e \tag{4}$$

when $E(U) \geq 0$, the consumer would make the purchase.

The developer's profit is given by:

$$u_D = ps - hq^2 \tag{5}$$

where h is the developer's cost rate on quality and $h > 0$. The quadratic form of the overall cost hq^2 represents the diminishing return of investment on quality (Choudhary 2007). The condition for the developer to participate is $u_D \geq 0$. The developer seeks to maximize her/his profit by choosing optimal p and q given the demand side's constraint $E(U) \geq 0$.

The platform owner's revenue is given by:

$$u_P = p(1-s) \tag{6}$$

where s is bounded between 0 and 1. The platform owner's goal is to find the optimal s^* which maximizes u_P.

It is worthwhile mentioning that in our model the platform owner's cost is neglected. Regarding app hosting cost, we believe it is already covered by a fix annual fee paid by the developer. For example, Apple iOS platform charges $99 for annual membership to allow a developer upload her/his apps on the Apple App Store. We also speculate this membership fee may not be the major revenue source. The evidence is nowadays roughly 350,000 apps are available on the Apple App Store. Even under the most extreme case that one app correspond to one developer, the annual revenue from this fix membership fee is 35 million at most, which account for less than 4% of the total platform owner's app sales revenue. Meanwhile, regarding this membership fee, once the developer decides to participate, it is the sunk cost which doesn't affect their optimal choice of quality level and price.

Table 1 summarizes the notations of the base model. In the succeeding part of model extensions, new notations will join under the same stylized fashion and will be clarified at corresponding part of the paper.

Table 1 Model parameters and decision variables

Parameters	
k	Rating-utility conversion rate
r_0	*Ex post* zero net utility rating
r_R	Reservation rating
r	*Ex post* rating from high valuation group consumers
b	Fraction of consumers in the high valuation group
h	Developer's cost rate on quality
Decision variables	
q	Quality level of the app
p	Price of the app
s	Developer's revenue sharing percentage

Proposition 1. Given the revenue sharing percentage s, the developer's optimal choice of price p^* and quality level q^* are:

Region 1-1 (self-driven): where $b > b_1$ and $0 < h \le h_1 s$,

$$p_1^* = \frac{b(1-r_R)}{k(1-b)}, q_1^* = \frac{br_0 - r_R b - r_0 + 1}{k(1-b)}; \tag{7}$$

Region 1-2 (platform owner driven): where $b > b_1$ and $h_1 s < h \le h_2 s$,

$$p_2^* = \frac{1}{2}\frac{sb^2}{h} + \frac{b(-r_R + r_0)}{k}, q_2^* = \frac{1}{2}\frac{sb}{h}; \tag{8}$$

Region 1-3 (poor marketing): where $b \le b_1$ and $0 < h \le h_3 s$,

$$p_3^* = \frac{b(1-r_R)}{k(1-b)}, q_3^* = \frac{br_0 - r_R b - r_0 + 1}{k(1-b)}. \tag{9}$$

In all other regions of (b,h) the developer cannot make non-negative profit. The above thresholds are:

$$b_1 = 1 - \frac{1-r_R}{r_R - r_0}, \ h_1 = \frac{1}{2}\frac{bk(1-b)}{br_0 - r_R b - r_0 + 1}, \ h_2 = \frac{1}{4}\frac{bk}{r_R - r_0} \text{ and } h_3 = \frac{kb(1-b)(1-r_R)}{(br_0 - r_R b - r_0 + 1)^2}.$$

Proposition 1 shows that the developer's optimal choice of price and quality level are divided into three possible cases, depending upon her/his marketing performance b and cost rate h.

Generally speaking, higher quality level q will lead to higher *ex post* ratings from high valuation consumers. It would increase subsequent consumers' *ex ante* perceived net utility as well as the consumer's willingness to pay. Thus, the developer can gain from charging a higher app price p to exploit the additional willingness to pay. For the developer in Region 1-1, because of the low cost rate ($0 < h \le h_1 s$), the marginal gain from improving app quality is always greater than its marginal cost until r

reaches its maximum. Once the *ex post* rating hit the maximum allowed rating, the developer's further efforts on increasing consumer's received utility will no longer be reflected by the *ex post* ratings, which means the further improvement on quality becomes unobservable to subsequent consumers in the rating-to-utility process. Therefore, her/his optimal quality level q_1^* is the one at which high valuation consumers will offer the maximum rating (Define r^* as the r under optimal price and quality level and notice that $r^* = k(q_1^* - p_1^*) + r_0 = 1$) . In Region 1-2, when her/his marketing performance is also good but the cost rate is high ($h_2 s \geq h > h_1 s$), marginal cost from improving app quality increases faster than that in Region 1-1 so that it will surpass the marginal gain before the maximum $r = 1$ is realized. Thus the r in Region 1-2 will be between r_R and 1. In Region 1-3, the optimal price p^* and quality level q^* are of the same analytical form as Region 1-1. It indicates when the developer's marketing performance is poor and the cost rate is relatively low ($0 < h \leq h_3 s$), the developer's optimal choice is to choose the quality level at which $r = 1$. The intuition behind is that since only a few consumers who belong to high valuation type would appreciate the quality and rate positively after purchase, the developer need to provide sufficient satisfaction to them, making them rate as high as possible to counteract the low ratings from mismatched consumer group. Otherwise because the overall rating is too low the subsequent consumer's *ex ante* perceived net utility as well as their willingness to pay would be too low.

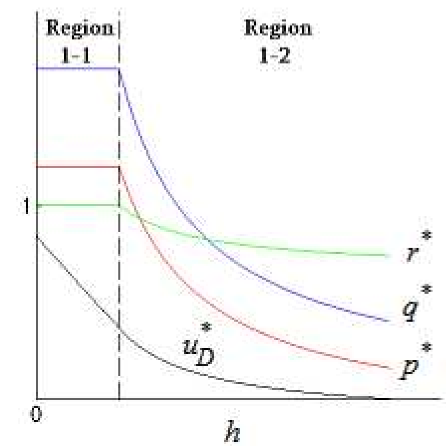

Figure 1 Developer's optimal choices given s

Figure 1 depicts the developer's p^*, q^*, r^*, u_D^* in Region 1-1 and Region 1-2. It shows all of them are non increasing function of the developer's cost rate h .

Corollary 1.1 For the developer in Region 1-2, $\partial p^* / \partial s > 0$ and $\partial q^* / \partial s > 0$.

Corollary 1.1 shows that for the developer in Region 1-2, her/his optimal choice of quality level and price is driven by the revenue percentage she/he obtain. By increasing s , the platform owner can provide more incentive to the developer to improve her/his app quality.

For the developer in Region 1-1, optimal quality level is not affected by s . Therefore hereafter we give an intuitive name for Region 1-1, "self-driven region". Developers in this region are referred as self-driven developers. We give another name for Region 1-2, "platform owner driven" region. We refer developers in that region as "platform owner driven developers". For Region 1-3, we name it "poor marketing region" and the developers are referred to "poor marketing developers". Since poor marketing region is unlikely to be the main component of the market, we focus our attention on the self-driven region and platform owner driven region.

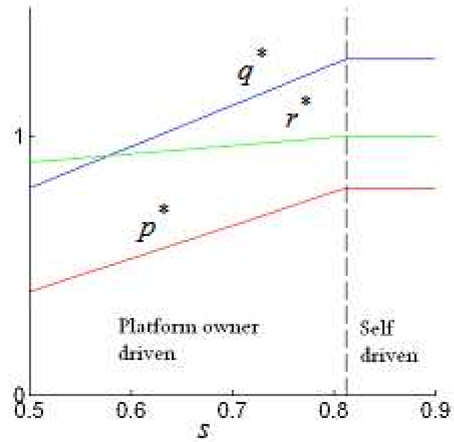

Figure 2 Developer's optimal choices as function of s

Notice that the boundaries of the regions ($h_1 s, h_2 s$) are proportional to revenue sharing percentage s which is set by the platform owner at the first stage of the game. This indicates that given a developer's cost rate h, whether she/he would choose to be in self-driven region or platform owner region, not only depends on her/his h and consumer rating behavior (k, r_0, r_R), but also depends the platform owner's choice of s. Figure 2 illustrate that as s increases, the developer with a given h could switch from platform owner driven to self-driven. More detailed, when the developer is in platform owner driven region, increasing in s would increase the q^* and p^* until the r hit the maximum. Then the developer enter the self-driven region and q^* and p^* are no longer function of s.

Corollary 1.2 In all regions, the developer's optimal price $\partial p^* / \partial r_R < 0$. In self-driven region, the optimal quality $\partial q^* / \partial r_R < 0$. In platform owner driven region, the optimal quality $\partial q^* / \partial r_R = 0$.

Corollary 1.2 shows how developer should optimally respond to consumer's more critical view in accepting ratings. When r_R increases, the original optimal price can no longer be justified by the rating. The developer in any of those regions need to decrease the app price. Platform owner driven developers would keep the same quality level while self-driven developers would choose a lower quality level. This is because in self-driven region, as long as this maximum $r = 1$ is realized, further quality improvement would not promote the consumers' perception on *ex ante* net utility. Since the price drops as stated above, the required quality level to achieve the maximum $r = 1$ decreases.

Corollary 1.3 In self-driven region, $\partial q^* / \partial r_0 < 0$ and $\partial p^* / \partial r_0 = 0$. In platform owner driven region, $\partial p^* / \partial r_0 > 0$ and $\partial q^* / \partial r_0 = 0$.

When the consumer becomes more severe in offering ratings, which means r_0 decreases, for the platform owner driven developer, the best way is to decrease the price. However, the self-driven developer's optimal response is to choose a higher quality level but keep the price unchanged. This is because when r_0 decreases, the r will no longer be maximum. Thus the self-driven developer regains the incentive to choose a higher quality since the consumer would be able to observe and appreciate such improvement again.

Corollary 1.3 is coincided in spirit with management literature that over-leniency is a significant issue. Berger et al 2010's discovered empirically in the corporate environment a forced distribution on performance ratings will lead to higher productivity. We show that when r_0 increases (more lenient), the "productivity" indicator q^* decrease. We make the analogy to the app market context and confirm the same principle through providing an explanation from the angle of economic utility theory.

Notice that the boundaries of the regions are affected by r_0, r_R . Figure 3 shows that when r_R increases, the platform owner driven region shrinks but self-driven region expands. For a developer in platform owner driven region, she/he may switch to self-driven when r_R increases. It also shows how such developer's choice of optimal quality level and optimal price change across different regions.

Figure 4 shows that when r_0 increases, both the total region expands and self-driven region expands. Similarly, the developer may switch from platform owner driven to self-driven region when r_0 increases.

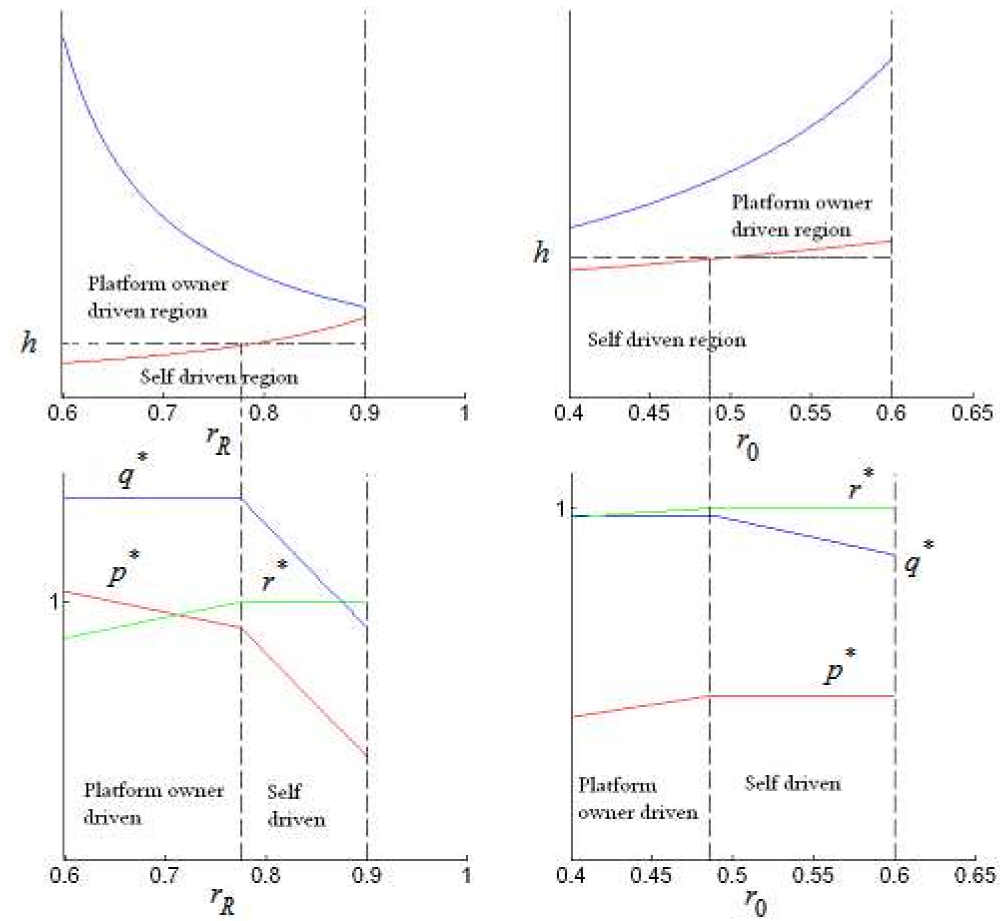

Figure 3 & 4 Region distribution and developer's optimal choices across regions

Proposition 2. The platform owner's optimal choice of developer's revenue sharing percentage s^* is

Region 2-1 (squeezing): when $b > b_1$ and $0 < h \le h_{21}$,

$$s_1^* = \frac{2h(1 + br_0 - br_R - r_0)}{bk(1-b)}; \tag{10}$$

Region 2-2-1 (encouragement): when $b > b_1$ and $h_{21} < h \le 2h_2/3$,

$$s_{21}^* = \frac{1}{2} + \frac{h(r_R - r_0)}{bk}; \tag{11}$$

Region 2-2-2 (retention): when $b > b_1$ and $2h_2/3 < h \le h_2$,

$$s_{22}^{*}=\frac{4h\left(r_R-r_0\right)}{bk}; \quad (12)$$

Region 2-3: when $b \le b_1$ and $0 < h \le h_3$,

$$s_3^{*}=\frac{h\left(1+br_0-br_R-r_0\right)^2}{b\left(1-r_R\right)k\left(1-b\right)}. \quad (13)$$

The above thresholds are:

$$h_{21}=\frac{1}{2}\frac{bk\left(1-b\right)}{2+br_0-br_R-r_0-r_R}.$$

Proposition 2 characterizes the sub game perfect Nash equilibrium (SPNE) of the game. It shows how the platform owner at first stage should offer different revenue sharing percentage s to different type of developers depending upon their cost rate h and marketing performance b. The SPNE take into account the developer's best response in the second stage. For example, when the platform owner believes its profit u_P is maximized if the developer chooses self-driven region at the second stage, the platform owner at the first stage chooses to offer s_1^* to incentivize the developer to do so.

Corollary 2.1 In all regions, the optimal revenue sharing percentage $\partial s^* / \partial h > 0$.

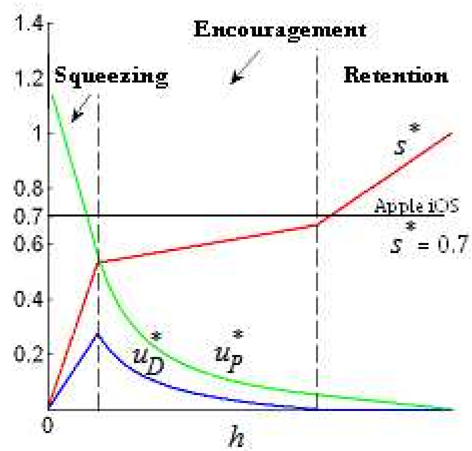

Figure 5 Platform owner's optimal choices of s

Corollary 2.1 suggests among all regions, the platform owner should increase the developer's revenue sharing percentage when their cost rate increases. Figure 5 depicts the optimal revenue sharing percentage s as a function of the cost rate h. Regarding Region 2-2-1 (encouragement) and Region 2-2-2 (retention), they essentially correspond to the platform owner driven region in Proposition 1. Figure 1 shows that when the platform owner driven developer's cost rate increases, the optimal quality level q^* decreases. Corollary 2.1 indicates that in this situation the platform owner should offer a higher revenue sharing percentage to drive the quality level back for platform owner's best interest. Although the platform owner's percentage share decreases because of giving more to the developer, it would eventually pay the platform owner back through the additional revenue gained from the consumer because of the higher quality app.

Regarding Region 2-1 (squeezing), it essentially corresponds to the self-driven region in Proposition 1. As we know the optimal quality q^* is not a function of revenue sharing percentage s for self-driven developers, the argument for the dynamics between s^* and h in Region 2-2-1 and Region 2-2-2 is no longer valid. The real situation is that if the platform owner knows that the developer has a lower cost rate

h in Region 2-1, the platform owner can squeeze the developer by giving a lower s because under such low s the developer still can get non-negative utility due to the low cost rate and it is still the developer's best interest to achieve the same q^* at which $r=1$ and charge the same p^*. Therefore the platform owner acquires more revenue from the decreased s .

Squeezing the developers may be dangerous since they may leave. One argument why low cost rate developer doesn't leave this platform even when receiving low revenue sharing percentage lies in our assumption that the platform owner is the monopoly or has the monopoly power over the market, such as predominance in consumer base. For example, Apple still grasps roughly 70 percent of the entire mobile app market and this number is 99.4 percentage two years ago. The developers are allured by the large number of potential customers. Another reason is that consumers using different platforms are usually separated, for example, chances are little that one consumer is both iPhone user and Andriod phone user. Since the cannibalization is a mild issue, the developer could earn more profit by releasing its app on another platform as long as she/he can earn positive benefit from that platform. For example, a famous game app, "Angry birds", has been made available for both iPhone and Andriod.

Based on the above discussion, we give more interpretive names to the regions in Proposition 2. Region 2-1 is referred to as "squeezing region". Region 2-2-1 is named as "encouragement region" since the platform owner would increase s to encourage them to increase the quality of app when h becomes higher. Region 2-2-2 is referred to as "retention region" since those developers' profits are zero, which indicates they are on the verge to leave the platform. The platform owner should set a high s to retain them so that the platform owner can continually obtain revenue from apps they produce. The developer in squeezing region will be incentivized to choose self-driven region in second stage. The developer in encouragement and retention region will be incentivized to choose platform owner driven region in second stage. From Figure 5 we can see the developer whose cost rate is around the intersection of squeezing region and encouragement region can obtain the highest utility. Lower cost rate developer's utility will be squeezed. The extreme case is when the cost rate is close to zero the platform owner can almost take nearly all the revenue but leave very little share to the developer. The developer's participation can still be justified because developing such quality app cost almost nothing for her/him.

Corollary 2.2 In encouragement and retention region, $s^* > 1/2$

This coincides with our observation from the real business practice. For instance, Apple iOS platform offers $s=0.7$. For app market, given the fact that a large portion of the developers are individuals or development teams made up of several individuals, the platform owner may consider the majority of cost rates are on the higher side in encouragement or retention region, rather than in squeezing region. Therefore, it is the optimal strategy for the platform owner to offer a revenue sharing percentage greater than one half.

Corollary 2.3 The optimal revenue sharing percentage $\partial s^* / \partial r_R > 0$ in encouragement and retention regions, but $\partial s^* / \partial r_R < 0$ in squeezing region.

According to Proposition 1, when consumers become more critical in accepting ratings, the developer in encouragement or retention region would decrease the price. Such price drop decreases the total revenue. By raising an appropriate amount in s, the developer in encouragement region would choose a higher quality and then correspondingly increase the price, which compensates the decrease in platform owner's revenue share such that eventually benefits the platform owner.

The argument for squeezing region is a little counterintuitive. It states that in squeezing region when the consumer's reservation rating increases the platform owner prefers squeezing the developer more, rather than incentivizing the developer by offering more revenue share. The underlying logic is the following. Notice in Proposition 1, the boundary of self-driven region $h_1 s$ increases as r_R increase. For the developer in squeezing region, she/he would still choose self-driven region at second stage when r_R increases. The optimal quality, app price and the total app sale revenue are determined by the new r_R, not s. Given more s^* would not further promote the choice of quality level as the case in encouragement region but only subsidizing more revenue sharing to the developer. Instead, if the platform owner decrease the s, as long as decreased s still make the developer stay in self-driven region, i.e. maintain the choice of price and quality level, the platform owner would benefit from decreasing s. However, the platform owner cannot decreases significantly. Illustrated in Figure 2, given the developer's cost rate h, when s decreases significantly the developer originally in self-driven region may switch to platform owner driven (notice $h_1 s$ decreases).From the developer's perspective, it turns out to be that moving to platform owner region will incur greater loss than staying in self-driven region with less revenue percentage if there is only an appropriate reduction in s. Choosing between two "evils", the developer believes it is her/his best interest to stay in self-driven region. Hence the platform owner is able to squeeze.

Corollary 2.4 The optimal revenue sharing percentage $\partial s^* / \partial r_0 < 0$ in encouragement, retention and squeezing regions.

When r_0 increases, i.e., consumers generally offer higher ratings, the developer in encouragement or retention region would charge a higher price and keep the same choice of quality level. Higher price indicates higher revenue so that the platform owner can take out more shares of the additional revenue through decreased s. For the developer in squeezing region, the optimal price doesn't change but the optimal quality level decreases since consumers are more lenient in offering ratings. The argument why s^* decreases is similar to the case where r_R increases in squeezing region in Corollary 2.3.

Conclusion

In this paper, we parameterize consumer rating behavior into three parameters (k, r_0, r_R) and construct a bidirectional rating-utility system which incorporates these parameters into utility calculation. We apply such framework to the app market where ratings play a central role in determining the consumer's *ex ante* perceived net utility as well as their willingness to pay. We identify three types of developers, self-driven, platform owner driven and poor marketing developer, based decisions on optimal choice of quality level and price given such consumer's choice. We provide comprehensive analysis on how change in consumer rating behavior (r_0, r_R) would change developers' optimal choices of quality level and price. We also provide the dynamics of the platform owner's optimal choice of revenue sharing policy when (r_0, r_R) alter. We analytically identify the over-leniency issue from our rating-utility economic model, a famous issue which has been demonstrated in many behavioral management literatures, and then justify that it exhibits the characteristics shown empirically by those literatures.

As mentioned in introduction section, this paper provides a fundamental way to incorporate rating behavior into traditional utility-based analysis. In this sense, one possible direction of future research is to study the impact of rating behavior on other aspects of traditional industrial organization (IO) research such as competition and product differentiation. Many IO work rely on full information assumptions, such as the *ex ante* observable quality so as to utility. Signals such as ratings now can be embedded in utility theory therefore replicating those works on basis of rating behavior would be implementable. Research questions such as how to compete on ratings and how consumer rating behavior affects the market competition would be interesting. The impact of rating behavior on product differentiation strategy, such as how the developer can take advantage of ratings to implement the *ex ante* perception on product differentiation, would also be another intriguing topic to explore.

Another opportunity for extension is from behavioral angle. Extant behavioral research suggests that the heterogeneity of scaling in rating system, for example, 5-star scale rating system and 10-point scale rating system would induce discrepancy and bias in ratings behavior. An 8.0 in 10-point rating has potentially different meaning than 4-star in 5-star scale ratings system for the same person. Thus an interesting question is how the platform owner designs an appropriate rating scale to manipulate the consumer rating behavior on the purpose of approaching the optimal (k, r_0, r_R) demonstrated in this paper.

Furthermore, we notice not only the characteristics of platform owner's optimal (k, r_0, r_R) vary between squeezing region and encouragement region but also does the optimal revenue sharing percentage. This discrepancy provides the platform owner the incentive to separate out these two regions into multi store fronts. Being aware that organizational developers like software companies usually have relatively low cost rate so that being more likely to site in squeezing region, the platform owner may want to set up an additional storefront, called “corporate app store”, in complementary to “individual app store” where individual developers are seated in. Then the platform owner may design different rating system and adopt different revenue sharing percentage to these two stores, for their own profit maximization. We believe as the app market expands rapidly, a greater degree of heterogeneity on app supply side would be expected so that such separation may be profitable.

Finally, this paper speculates the existence of reservation rating (r_R). This calls for empirical observation of such phenomenon by, for example, analyzing online WOM data. This paper can provide an economic structure to formulate potential structural econometric models. This research stream can lead to the answers to some interesting questions such as what factors cause the differences in individual reservation ratings.

References

Akerlof, G.A. 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics* (84:3), pp. 488-500.

Berger, J., Harbring, C., Sliwka, D. 2010. "Performance Appraisals and the Impact of Forced Distribution: An Experimental Investigation," *Working Paper*

Borman, W.C. 1977. "Consistency of rating accuracy and rating errors in the judgment of human performance," *Organizational Behavior and Human Performance* (20:2), pp. 238-252

Chen, Y., Xie, J. 2005. "Third-Party Product Review and Firm Marketing Strategy," *Marketing Sci.* (24:2), pp. 218–240

Chen, Y., Xie, J. 2008. "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix, " *Management Science* (54:3), pp. 477–491

Chevalier, J. A., Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*. Vol. XLIII, pp. 345–354

Choudhary, V. 2007. "Comparison of Software Quality Under Perpetual Licensing and Software as a Service," *Journal of Management Information Systems* (24:2), pp. 141–165.

Forman, C. 2008. "Examining the Relationship Between Reviews and Sales: The Role of Reviewer Identity Disclosure in Electronic Markets," *Information Systems Research* (19:3), pp. 291–313

Clemons, E. K., Gao, G., Hitt, L. M. 2006. "When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry, " *Journal of Management Information Systems* (23:2), pp. 149–171

Dellarocas, C., Farag, N.A., Zhang, X. 2004. "Using Online Reviews as a Proxy of Word-of-Mouth for Motion Picture Revenue Forecasting, " *Working Paper*

Dellarocas, C. 2006. "Strategic Manipulation of Internet Opinion Forums: Implications for Consumers and Firms, " *Management Science* (52:10), pp. 1577–1593

Zhou, W., Duan, W. 2010. "Online User Reviews and Professional Reviews: A Bayesian Approach to Model Mediation and Moderation Effects," *ICIS 2010 Proceedings*, pp. 256

Godes, D., Mayzlin, D. 2004. "Using Online Conversations to Study Word-of-Mouth Communication, " *Marketing Science* (23:4), pp. 545-560

Hu, N., Pavlou, P.A., Zhang, J. 2006. "Can online reviews reveal a product's true quality?: empirical findings and analytical modeling of Online word-of-mouth communication," *Proceedings of the 7th ACM conference on Electronic commerce*. ACM, New York, NY, USA, pp. 324-330

Hu, N., Liu, L., Zhang, J. 2008. "Do online reviews affect product sales? The role of reviewer characteristics and temporal effects," *Inf Technol Manage*, Vol 9, pp. 201–214

Hu, N., Zhang, J., Pavlou, P.A. 2009. "Overcoming the J-shaped Distribution of Product Reviews," *Communications of the ACM* (52:10), pp. 144-147

Jiang, B., Chen, P. 2007. "An Economic Analysis of Online Product Reviews and Ratings," *Working Paper*

Kane, J.S., 1994. "A Model of Volitional Rating Behavior, " *Human Resource Management Review* (4:3), pp. 283-310

Kane, J.S., Bernardin, H.J., Villanova, P., Peyrefitte, J. 1995. "Stability of Rater Leniency: Three Studies," *The Academy of Management Journal* (38:4), pp. 1036-1051

Kuksov, D., Xie, Y. 2010. "Pricing, Frills, and Customer Ratings," *Marketing Sci.* (29:5), pp. 925–943

Lee, Y.J., Tan, Y., Hosanager, K. 2009. "Do I Follow My Friends or the Crowds? Examining Informational Cascades in Online Movie Reviews," *Proceedings of WITS 2009*.

Li, X., Hitt, L. M.. 2010. "Price Effects in Online Product Reviews: An Analytical Model and Empirical Analysis," *MIS Quarterly* (34:4), pp.809-831.

Li, X. Hitt, L.M. 2008. "Self-Selection and Information Role of Online Product Reviews," *Information Systems Research* (19:4), pp. 456–474

Liu, Y. 2006. "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing* Vol. 70, pp. 74–89

Moe, W., Schweidel, D.A. 2011. "Online Product Opinions: Incidence, Evaluation and Evolution," *Working Paper*

Moe, W., Trusov, M. 2011. "Measuring the Value of Social Dynamics in Online Product Ratings Forums," *Journal of Marketing Research*. forthcoming.

Shapiro, C. 1985. "Optimal pricing of experience goods," *The Bell Journal of Economics* (14:2), pp 497-507

Sun, M. 2010. "How Does Variance of Product Ratings Matter? " *Working Paper*

Yun, G.J., Donahue, L.M., Dudley, N.M., McFarland, L.A. 2005. "Rater Personality, Rating Format, and Social Context: Implications for Performance Appraisal Ratings," *International Journal of Selection and Assessment* (13:2), pp 97-107

Zhu, F., Zhang, X. 2010. "Impact of Online Consumer Reviews on Sales:The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing*. Vol. 74, pp 133–148

Saal, F., Downey, R.G., Lahey, M.A. 1980. "Rating the ratings: Assessing the psychometric quality of rating data," *Psychological Bulletin (*88:2), pp 413-428