IMPACT OF ONLINE REVIEWS ON MOBILE APP SALES: OPEN VERSUS CLOSED PLATFORM COMPARISON

Chihoon Song, Department of Marketing, Sogang Business School, Sogang University, Seoul, Korea, suezor@sogang.ac.kr

Kyungdo Park, Department of Marketing, Sogang Business School, Sogang University, Seoul, Korea, kyungdo@sogang.ac.kr

Byung Cho Kim, Department of Logistics, Service and Operations Management, Korea University Business School, Korea University, Seoul, Korea, bkim@korea.ac.kr

Abstract

A rapid growth of the mobile app market has been identified as a key driver of the successful proliferation of the mobile devices such as smartphones and tablet devices. In this paper, we identify the success factors that contribute to the sales of mobile apps, grounded on the word-of-mouth literature. While open innovation is generally accepted as an effective solution for traditional goods development, the optimal platform choice for app development is not yet justified. Given this controversy, we examine the optimal platform choice for app development from a consumer’s perspective. Our findings indicate that closed platform outperforms open platform in the mobile app market. Interestingly, the interaction effect turns to be significant, meaning that the impact of online rating on app sales is greater under open environment than under closed environment. This implies that consumers depend on ratings more heavily when search cost is higher in the open environment under which apps are published without going through screening.

Keywords: Mobile App, Online Rating, Open Innovation, Closed Innovation, Platform
1 INTRODUCTION

The rapid development of the mobile networks and smartphones has changed our daily lives. One of the factors that contributed to the proliferation of the smartphones is the corresponding app market. According to a recent survey conducted by a Chicago-based consulting firm named “Gravity Tank”, 35 percent of the smartphone users think that smartphone apps changed their lifestyle, 54 percent say, their perception of the mobile phones was changed due to apps, and 67 percent answer that they purchased smartphones to use certain apps. (Kim and Kim 2010). The positive two-way interaction of the smartphone markets and the app markets has expedited the diffusion of smartphone in the global telecommunication market.

While there exist several popular app stores around the world, two app stores are market leaders, including Apple’s iTunes Store and Google’s Play Store (Müller et al. 2011). Interestingly, these two firms are very different in every aspect of the business, from the operating systems on which the apps run, to the management policies of the app stores. Apple plays a closed platform strategy, with which Apple has a high level of control on every single part of the process, including the development of the operating systems, selection of the apps, and the sales of the selected apps through its distribution channel, iTunes store. Apple is well known for its closed and secretive process in general. For example, Apple developed most parts of the iPhone’s hardware and software in-house, which led to the filing of more than 200 patents including multi-touch screen, scrolling, and zooming (Remneland-Wikhamn et al. 2011). Apple’s high level of control on the app selection and sales can be described in the following process. When an app developer wants to sell her app through iTunes Store, she should register and pay $99 one-year-based registration fee. Apple then applies its pre-screening process before posting the app on its iTunes Store. Potential security problems such as malicious codes, illegal contents or functional defects are checked and the marketability of the app is evaluated. After going through the process, Apple approves and posts the app on the iTunes Store. If the developer refuses to follow Apple’s policy, she is restricted for further updates of the posted app and sales of the new app. Due to its high level of control, Apple’s strategy is called a closed platform strategy.

On the other hand, Google plays a very different strategy. When Google released the Android operating system in November 2007, they said, “Android platform will be made available under one of the most progressive, developer friendly open-source licenses, which give mobile operators and device manufacturers significant freedom and flexibility to design products” (Open Handset Alliance 2007). This open nature of the development process influenced the way Google manages its app store. While iPhone app developers must use Mac OS and the software development kit provided by Apple, Android-based app developers can use any software development kit. Plus, Google skips the prescreening and verification process, implying that anyone is allowed to post any app on Google’s Play Store without much restriction. With these reasons, Google’s strategy is called an open platform strategy.

The effectiveness of the open innovation strategies has been well supported in the literature. Chesborough (2003) argued that the open innovation model has many advantages compared to the closed innovation models through case evidences, such as Xerox. The advantages of firms using open innovation incorporate new ideas into their business and rescue the projects that initially seem to lack promise but turn out to be surprisingly valuable. That explains the transition of the many firms from the closed to the open innovation model. Almirall and Casadesus-Masanell (2010) identified the conditions under which open innovation outperforms closed innovation. They find that open innovation is more effective than closed innovation regardless of the complexity when partnerships are flexible. When partnerships are fixed, open innovation works better than closed innovation for low and medium-low levels of complexity. Boudreau (2010) studied the relationship between two open strategies and their different influence on the rate of innovation, using data on 21 handheld computing systems (1990-2004). Two open strategies are granting outsiders access to the platform (i.e., opening the complement) and giving up some control over the platform itself (i.e., opening the platform). He
showed that using first opening strategy produced up to 4.6 times faster than the system’s preceding closed policy in the rate of introduction of new devices and second opening strategy increased the innovation rate by roughly 20 percent.

Despite the long-lasting support for the superiority of open innovation for traditional goods, it has been argued that an opposite direction, i.e., closed innovation being superior to open innovation in the context of app development. According to a recent survey on the app developers’ satisfaction (Open-First 2010), app developers’ satisfaction level is higher in the closed environment than in the open one and the consumers are more loyal to the app store with a closed platform than to the one with an open platform.

In this paper, we investigate whether the superiority of the open model over the closed model is valid in the app market. We empirically examine the two leading app stores, identify the factors that affect app sales, and investigate the influences of the key drivers on app sales with a focus on the platform strategy, i.e., open versus closed. The contributions we aim to make to the literature are twofold. Firstly, while most of the existing studies compare open innovation with closed innovation from a technology developer’s perspective, we look at the consumer side. That is, we measure the effectiveness of the open and closed innovation as app sales which well reflect how consumers evaluate the different ways of innovation. Secondly, we investigate the role of online rating in the app sales process. While numerous studies examined the effectiveness of the online rating on the sales of various products, it has not been studied in the context of mobile app yet. In addition, we study the differential effects of the online rating under open and closed models, which we believe has not been explored. Practically, our findings give implications to app store managers, developers and consumers. Especially, app store managers may use our results as a guideline to design an effective ecosystem for app developers.

The rest of the paper is organized as follows: In Section 2, we set up hypotheses and provide theoretical background for each of those. Section 3 introduces data and methodology we use and Section 4 presents the results. Section 5 concludes the paper.

2 THEORETICAL FOUNDATION AND HYPOTHESES DEVELOPMENT

2.1 Consumer Rating

Consumers tend to search the information about product quality before making a purchase decision. The degree of uncertainty may be higher online since consumers do not have a chance to touch and feel the products as they do at a brick-and-mortar store. According to Uncertainty Reduction Theory (Berger and Calabrese 1975), consumers engage in uncertainty reduction efforts to mitigate and eliminate the risk associated with the uncertainty and to maximize the outcome value when they lack knowledge of a product or the outcomes of consuming that product (Hu et al. 2008). Thus, consumers who perceive uncertainty will have incentive to actively search for the information about the product to reduce uncertainty (Dowling and Stealin 1994).

One of such searching activities is reading other consumers’ reviews. Liu (2006) argues that the reviews made by consumers are perceived as a more credible and more trustworthy source than other sources of information. Consumer ratings are perceived to represent the level of satisfaction. Especially for the experience goods whose quality is hard to evaluate before experiencing it, consumers make efforts to find a proxy for quality. Chevalier and Mayzlin (2006) show that consumer ratings are positively related to book sales, which is confirmed in the video game industry by Zhu and Zhang (2010). Certainly, mobile apps are experience goods for which rating information can be helpful. Thus, we presume that there exists a positive relationship between rating and app sales, which is captured in the following hypothesis.
H1: Consumer ratings and monthly app downloads are positively related.

2.2 Number of Ratings

Number of consumers who participated in rating inherently gives information about popularity of the product (Chen and Xie 2008; Yang et al. 2012). Moreover, ratings with more raters are believed to be more credible. Also, studies on herding suggest that consumers tend to follow the crowd when they do not trust the private information (Banerjee 1992). Given that people do not like uncertainty, consumers may consider apps with more ratings as less uncertain products. Previous studies show the positive relationship between the volume of word of mouth (WOM) and product sales (Godes and Mayzlin 2004; Liu 2006; Um 2011). Especially, Um (2011) shows that the amount of reviews on apps posted in Apple’s App store have positive impact on the daily consumer demand. We test this hypothesis in apps of any type, i.e., iPhone apps and Adroid apps.

H2: Number of ratings and monthly app downloads are positively related.

2.3 Developer Experience

When consumers are making app purchase decision, they may face product uncertainty as well as developer uncertainty. Consumers purchase products from anonymous sellers online without face-to-face interaction. Because there are a lot of developers from novice to expert who want to sell apps, consumers cannot easily estimate the developer’s ability to produce a good-quality app, which generates uncertainty. In order to reduce of uncertainty and to choose the appropriate app, consumers attempt to use a variety of information sources to make a good decision. The type of information provided from Google Play store and Apple app store is slightly different, but mostly similar. For example, both app stores provide description, other apps developed by the developer, date of update, size of app and user reviews such as rating, the number of ratings. One of the sources that consumers can use as a good indicator of low uncertainty about app quality is the developer’s experience, which can be measures with the number of published apps developed by the developer of interest. Existing studies show that reduced uncertainty about the seller decreases uncertainty about the product (Dimoka and Pavlou 2008). Thus, we presume that consumers prefer apps developed by experience developers with many publications.

H3: Number of published apps by a same developer and monthly app downloads are positively related.

2.4 Number of Days

Numerous apps are released daily. Consumers tend to prefer new apps to old apps since new apps tend to come with new features. Jung and colleagues (2012) argue that the life cycle of apps has been shortened. Also Zhu and Zhang (2010) argue that the average life cycle of all games is approximately 33 months, but on average, more than 50% of game sales occur within the first four months after a game’s release. Thus, because monthly app download is likely to decrease overtime according to preceding discussion, we hypothesize the following:

H4: Number of days in the market and the monthly downloads are positively related.

2.5 Platform Type

Müller et al. (2011) examine the transaction cost for the consumers at different transaction phases, including information phase, agreement phase, execution phase and after sales phase. We suggest that the extent of the consumer transaction cost differs depending on the type of platform. On the Android platform where the platform provider’s involvement is relatively weak and thus we call it a “open platform”, consumers suffer from higher level of uncertainty, thus they need to incur higher search cost to hedge the risk, which can be a significant portion of the transaction cost. On the other hand, all
apps submitted to Apple app store need to go through a touch prescreening. This shows high level of Apple’s involvement in the app market, so we call it a “closed platform.” This helps consumers reduce their time and efforts in searching for good apps, since they may believe that minimal quality is guaranteed. Also, because developer can write apps only through the Software Development Kit (SDK) provided by Apple, the risk regarding suitability of device and app can be reduced. These mechanisms can lower the transaction cost. According to Hu et al. (2008), higher transaction cost will result in lowered sales in online marketplace. Therefore, we hypothesize the following:

H5: Monthly app downloads are higher in the closed platform than in the open platform.

2.6 Interaction Effect: Consumer Rating and Platform Type

Zhu and Zhang (2010) suggest one of contributing factors to the relationship between consumer rating and product sales is the difference in consumer’s reliance on online consumer reviews. We presume that both perceived uncertainty and platform strategy affect consumer’s reliance on online review. As stated above, Apple’ app store with closed platform strategy adopts mechanisms to reduce the consumer’s perceived uncertainty. Since the consumers considering purchase of the apps in the Apple’s app store perceive a low level of uncertainty, they do not have to rely too much on the online review. On the other hand, Google designed the Android market with a free market philosophy whereby the market regulates itself (Müller et al. 2011). In other words, Google Play store does not prescreen apps so that any app can be published freely. The market’s self-selection mechanism works in a way that apps with low popularity naturally disappear. However, from a consumer’s perspective, search cost is significant since it is a consumer’s job to screen and select risk-free apps. To reduce search cost, consumers depend more heavily on the rating information when they make their selection, and thus, we presume that the impact of ratings on downloads is greater under the open environment. On the other hand, under the closed environment, rating’s influence is weaker since search cost is significantly lower due to Apple’s prescreening process. Hypothesis 6 reflects the theory.

H6: The positive association between consumer rating and monthly downloads is weaker in the closed platform.

Figure 1 depicts our conceptual framework.
3 DATA AND METHODOLOGY

3.1 Data

Data on monthly app downloads and the number of developer-by-developer apps come from Xyologic (www.xyologic.net), a research company that collects, analyzes and shares app sales data. Xyologic releases over 220 reports, covering 4 platforms (Android, iPhone, iPad, and Windows phone) and 30 countries monthly. The reports contain detailed information about application such as country rank, previous rank, category, monetization type, monthly download numbers and so on. We gather data on consumer rating, number of ratings, and the release date from Apple and Google app stores.

We use app download data for Android and iPhone, from the period between July 8 and August 8, 2012. More specifically, we focus on the top 150 paid and top 150 free apps for Android and for iPhone in Korea. We collect reviews for each app on the list from the Apple app store and Google Play store. They provide explanation of each app with average user rating, number of ratings and the date of update. We collect such data regarding 600 apps from both app stores between September 11 and 12, 2012. In order to find the precise effects of consumer ratings we need to track the consumer rating daily. But since we can only get the monthly sales data, the best way to deal with this gap is to collect review data on August 8. However, the company did not release the sales reports on August 8. Furthermore, it is impossible to predict the 150 best-selling android or iPhone paid/free applications in advance, so a little time difference between downloads and ratings was unavoidable.

Raters use a scale ranging from 1 to 5 in both Apple and Google app stores, 1 being the worst and 5 being the best. But the unit of measure of consumer rating is different. 0.5 is the unit of measure for consumer ratings on Apple’s app store. On the other hand, 0.1 is the unit of measure on Google play store. If we do not convert the unit of consumer rating, it is difficult to uncover the precise effects of rating on sales. Therefore to derive accurate result, we match both units as 0.1.

We also include two other independent variables, number of published apps by the developer and number of days in the market. Data on the number of published application by developer comes from Xyologic and number of days in the market comes from the app stores. Because both app stores provide the date of update not release date, a little more efforts were made to measure the length of stay in the market. Because consumers leave their comments after downloading apps, we assume that the date of oldest comment is closer to the released date. So we search the date of oldest comment of each app and calculate the number of days in the market.

Our final data set contains of 600 apps that were available for the two platforms in July 2012. Of these, 58 apps were dropped since rating information is not available, and additional 28 were dropped due to price change.

3.2 Methodology

Our model aims to answer the two fundamental questions regarding the app market: (1) What are the factors that influence app downloads? (2) How different is the impact of consumer ratings on app downloads across different platforms? Data were analyzed using log-linear regression in the SPSS Program package. We use the research model as follow:

\[
\ln Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 D + \beta_6 X_1 D + \epsilon
\]

where Y is monthly downloads, as a dependent variable. \(X_1\) is a set of independent variables. Specifically, \(X_1\) is consumer rating, \(X_2\) is the number of ratings, \(X_3\) is the number of published apps by the same developer and \(X_4\) is the number of days in the market. To operationalize the platform strategy, we create a dummy variable indicating whether an application is sold at Apple’s app store or Google play store. The dummy variable, \(D\), takes the value of 1 when the app is sold at Apple’s app store and it is 0 if the app is sold at Google play store. Because our conceptual framework
suggests that the consumer rating variable is conditional on the type of strategy, we model the interaction effect. $\beta_i$ is the coefficient capturing the effects of those variables and $\epsilon$ indicate the error term.

Hawkins (1980) defines an outlier as an observation that significantly deviates from other observations as to arouse suspicion that it was generated by a different mechanism. The presence of outliers can lead to inflated error rates and substantial distortions of parameter and statistic estimates (Osborne and Overbay 2004). Because Mahalanobis distance and Cook’s distance are both frequently used to find the outliers (Newton and Rudestam 1999), we use Cook’s distance method, and we remove 3 outliers.

3.3 Summary Statistics

Table 1 presents summary statistics of our sample. Android users mostly download free apps, whereas iPhone users’ download of paid apps is significant. The reason for the large number of free app download for Android is because of the size of the user pool and the openness of the platform.

<table>
<thead>
<tr>
<th>Android Apps</th>
<th>Number of Observations</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free</td>
<td>134</td>
<td>761753.73</td>
<td>120038.90</td>
<td>270000.00</td>
<td>12423000.00</td>
</tr>
<tr>
<td>Paid</td>
<td>144</td>
<td>1611.81</td>
<td>2170.51</td>
<td>400.00</td>
<td>14200.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>iPhone Apps</th>
<th>Number of Observations</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free</td>
<td>131</td>
<td>196541.98</td>
<td>300407.27</td>
<td>54000.00</td>
<td>2441000.00</td>
</tr>
<tr>
<td>Paid</td>
<td>110</td>
<td>15215.45</td>
<td>17194.73</td>
<td>4600.00</td>
<td>102000.00</td>
</tr>
</tbody>
</table>

Table 1. Summary Statistics of Apps

Table 2 shows summary statistics of reviews in our sample. The data suggest that consumer rating is overwhelmingly positive in both app stores. Researchers have observed similar patterns in other contexts, such as book (Chaevalier and Mayzlin 2006) and game (Zhu and Zhang 2010). A t-test indicates that consumer ratings of apps for iPhone are significantly higher than those for Android. This result indicates that iPhone users have higher loyalty and satisfaction. The number of ratings is higher in the Android app store than in the iPhone app store. It seems that Android users have a higher incentive to share their experience with others than iPhone users possibly due to the high uncertainty of Android app quality.
Table 2. Summary Statistics of Ratings

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer Rating</td>
<td>4.12</td>
<td>0.45</td>
<td>2.35</td>
<td>4.83</td>
</tr>
<tr>
<td>Number of Ratings</td>
<td>54533.63</td>
<td>165771.14</td>
<td>9.00</td>
<td>1540046.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer Rating</td>
<td>4.40</td>
<td>0.46</td>
<td>1.65</td>
<td>4.94</td>
</tr>
<tr>
<td>Number of Ratings</td>
<td>9473.03</td>
<td>17974.12</td>
<td>8.00</td>
<td>150546.00</td>
</tr>
</tbody>
</table>

4 ANALYSIS AND RESULTS

In this section, we show and discuss the estimation results of our proposed model. Table 3 presents the regression results. Note that unadjusted $R^2$ is 0.132 and adjusted $R^2$ is 0.122.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unstandardized Coefficient</th>
<th>Standardized Coefficient</th>
<th>T</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>4.326</td>
<td>1.373</td>
<td>3.149</td>
<td>.002***</td>
</tr>
<tr>
<td>Consumer Rating</td>
<td>1.300</td>
<td>.333</td>
<td>.240</td>
<td>3.899</td>
</tr>
<tr>
<td>Number of Ratings</td>
<td>5.288E-06</td>
<td>.000</td>
<td>.254</td>
<td>5.806</td>
</tr>
<tr>
<td>Developer Experience</td>
<td>-.000</td>
<td>.001</td>
<td>.015</td>
<td>.361</td>
</tr>
<tr>
<td>Number of Days</td>
<td>-.002</td>
<td>.002</td>
<td>-.059</td>
<td>-1.412</td>
</tr>
<tr>
<td>Platform Type</td>
<td>5.869</td>
<td>2.047</td>
<td>1.134</td>
<td>2.867</td>
</tr>
<tr>
<td>Consumer Rating by Platform Type</td>
<td>-1.194</td>
<td>.479</td>
<td>-1.027</td>
<td>-2.494</td>
</tr>
</tbody>
</table>

* p < .10, ** p < .05, *** p < .01.

Table 3. The Effect of Reviews on Monthly Downloads

The coefficients for the consumer rating and the number of ratings are both positive and statistically significant at the 1% level. Also the coefficient for interaction term is negative and statistically significant at the 5% level. Hypothesis 1 is supported based on our estimation results. Consumer rating and app downloads are positively related in our data set, which is consistent with the literature (e.g., Chevalier and Mayzlin 2006; Zhu and Zhang 2010). Hypothesis 2 is also supported by our estimation results. That is, our data show that the relationship between number of ratings and app downloads is positive and significant. Our research provides empirical validation for the argument
made by Kim et al. (2011) that word of mouth is the most important factor that drives app downloads. While previous studies such as Um (2011) focus on iPhone apps, we extend it to Android apps, adding generalizability to the findings.

Hypothesis 3 is not supported. One possible explanation is lack of detailed information from the current system. App store presents up to 4 other apps published by the same developer. If the consumer wants to get more accurate information about the developer’s level of expertise, they need to put the name of the developer in the search box which incurs extra time and effort. Thus, assuming that not all consumers exert efforts for thorough search, the information about the developer experience may not be properly used.

Hypothesis 4 is not supported. The coefficient for length of stay in the market shows the expected sign but turned out to be not statistically significant. Both app stores present the date of update instead of the release date. In the absence of the release date, we created a proxy value based on the date that the oldest comment was posted. Although this is a reasonably sound approach, there is a possibility that the gap between the release date and our proxy might lead to the result.

Hypothesis 5 is supported. Our data shows that monthly app downloads are significantly higher with the closed platform than with the open platform. This result is consistent with practitioners’ findings that consumers have a high level of satisfaction and loyalty in the applications on Apple’s App store (Changewave Research 2011).

Hypothesis 6, which is our main hypothesis, is also supported. This result indicates that the interaction effect between consumer rating and platform type is significant. Specifically, the positive relationship between consumer rating and app downloads is weaker with the closed platform than with the open platform.

5 CONCLUSION

This paper identifies key drivers of the mobile app sales with a focus on the openness/closedness of the environment for app development. We empirically examine the differential effects of the online rating on app downloads on the open and closed platforms. Our results indicate that closed innovation generates higher level of app sales than open innovation, unlike other markets where open innovation turns more effective than closed innovation. The effects of other independent variables are consistent with the existing literature in that online ratings and number of ratings are positively associated with app sales. This study contributes to the literature in the following ways. The findings indicate that closed innovation may be optimal for app development while open innovation is viewed as the solution in general as proven in the literature. This result may give a guideline to app store managers who design an optimal ecosystem for app development. Also, we show the effectiveness of the rating system in app sales, indicating that online ratings can serve as a good proxy of app quality which is often unobservable.

Our study deserves discussions about weaknesses. Firstly, there are still other factors that affect app sales, which are not captured by our model. For example, although we use online ratings with presumption that they are proxies of unobserved app quality, our model did not use actual or perceived quality levels of apps. Secondly, we collect our data based on one-month window. There is about a month lag between collection of app downloads data and collection of rating data. It would be better to add time dimension to our dataset. Also, app downloads data is not from the app stores. Although it would be more trustful if the data had been collected from the app stores, it might not be feasible.
References


Osborne, J.W. and Overbay, A. The power of outliers (and Why researchers should always check for them). Practical Assessment, Research and Evaluation, 9.


Xyo (2011). Quality of our app download estimates. Available at http://xyo.net/blog/quality-of-our-