Market Intelligence in Hypercompetitive Mobile Platform Ecosystems: A Pricing Strategy

Hoang D. Nguyen, Kajanan Sangaralingam, and Danny Chiang Choon Poo

Department of Information Systems, National University of Singapore, Singapore
{hoangnguyen,skajanan,dpoo}@comp.nus.edu.sg

Abstract. The recent years have seen a spurt of mobile developers in hypercompetitive mobile platform ecosystems. Yet, this is an unfair game where platform owners such as Apple, Google or Microsoft fence the information of their app store as top secrets. Our study, therefore, takes an important step in investigating the structure of rankings and sales revenue through 2,761 paid applications with weekly aggregated 32,109 observations to unveil a new indicator of market intelligence, *earning per download*. With the consideration of category effects, time effects and endogenity issues, our empirical results show that top-ranked paid apps can earn up to $7.80 per download. Our findings generate a number of insights for app developers to take actions in designing highly-ranked apps as well as manipulating prices, promotions or in-app purchases in order to unlock the full potential of their app sales.

Keywords: mobile apps, big data, ranking, pricing, power law, earning per download.

1 Introduction

With the increasing ubiquity of mobile devices, the world has been witnessing the boom of a new era of mobile applications ("apps"). On average, there are over 15,000 new apps launched weekly; and over 1.5 million apps are currently available on various mobile app stores such as Apple AppStore, Google Android Market, and Microsoft Phone Store [1–3]. In 2011, Apple Inc. announced their payment of $2.5 billion to app developers; and Gartner [4] forecasted a tenfold growth of mobile revenue between 2010 and 2014. These tremendous figures present the spectacular market of mobile apps with multifarious opportunities; however, it is a hypercompetitive and unfair mobile platform ecosystem where market information of million apps such as sales revenue and app demands remain the top secrets by platform owners [5].

The stiff competition requires app developers to adopt an appropriate pricing strategy in order to penetrate the app stores. There are four typical revenue models for mobile apps: paid, in-app purchases, in-app subscriptions or advertisement-based [6]. Each of them has unique advantages in promoting app sales; however, developers always face the trade-off between the demands and their pricing. Moreover, identifying market niches such as categories for publishing is daunting since it is extremely difficult for an app to get noticed in shoals of million apps. For instance, in Apple

F.F.-H. Nah (Ed.): HCIB/HCII 2014, LNCS 8527, pp. 350–359, 2014.

AppStore, there are only 240 mobile apps which win the laurels and become prominent to smartphone owners in short times [7]. Therefore, our study aims to reveal decisive factors which led to the success of a mobile app in these challenging markets.

In recent years, developing insights of market intelligence in mobile app stores has been drawing a number of research venues. Previous studies [5, 8] addressed competitive strategies in these markets by examining the takeoff and continued survival of apps. Motivated by their results, we investigated not only the various effects of dynamic attributes related to app positioning, developer actions and user engagement on app rankings, but also proposed a structural model of sales revenue and app rankings which accounted for numerous endogenity and heterogeneity issues. In this study, we conceptualized a new indicator of marketing intelligence, Earning Per Download (EPD). This brings us one step closer to the reality in mobile analytics and unlocks new directions for existing studies on estimating app downloads or sales revenue [9, 10]. Furthermore, our study provides useful visions for app developers to decide on their pricing strategy. We found that a top-ranked app yielded a gross up to $7.80 per download in the Apple AppStore during 2011.

The structure of the paper is as follows. In the immediately following section, we explain the data collection and our empirical models to investigate sales revenue and app rankings. After presenting our empirical results, we discuss our findings and conclude with a recommendation for future work.

2 Empirical Context, Conceptualization, and Data Collection

2.1 Background on Apple AppStore

In this paper, we primarily studied the Apple AppStore which is the foremost marketplace for mobile apps on iOS operating systems. With about half the market share of worldwide mobile app markets [11], Apple AppStore offers a convenient channel for developers to reach out to mobile users easily. It is a highly potential market; however, Apple does not publicly disclose the market information such the number of downloads, sales revenue or even their concealed formula for ranking apps.

According to Venturedata [12], Apple's ranking mechanism were shaped based on a number of criteria. This study focuses on two important criteria: the amount of downloads and the grossing revenue. First, in each category, Apple published a top list of mobile apps where there exists high correlation between the ranking and the number of downloads of an app, conventionally named as "App ranking based on downloads". Second, the list of "App ranking based on grossing" was built based on the total sales revenue of apps.

The data of our study were collected using a crawler from MobileWalla [1]. The dataset has been utilized and audited independently in previous research [2, 13], thus its reliability and accuracy are very high.

There are two dominant types of data available: (i) Time-Invariant data, such as name, descriptions, and features of apps and developer, etc., and (ii) Time-Variant data, such as user ratings, ranks, and reviews that change continuously. We organized our variables according to the conceptualization of key factors impacting rankings and sales revenue: 1) app positioning-related variables, 2) developer actions-related variables, 3) user engagement-related variables, 4) app features-related variables, 5) other control variables, and 6) ranking data.

Table 1. Data Collected and Derived from iTunes App Store

Variable	Description	Type
App Positioning		
Category popularity	Total number of apps released in a given category till a current week.	Time-Variant
Competition	Number of similar apps from different developers till the current week during the study period. The similarity is calculated using TF-IDF distance.	Time-Variant
Developer Actions		
Price	Market price in USD.	Time-Variant
Frequency of updates	Number of versions released for an app till the current week.	Time-Variant
Price reduction	Flag indicates whether price was reduced, for example due to the marketing promotions.	Time-Variant
User Engagement		
Review score	Average user review score for current version.	Time-Variant
App Features		
Size	Application footprint in Mega Bytes (MB).	Time-Variant
Design for iPad	Is the app designed only for iPad? Yes = 1.	Time-Invariant
3G/4G connectivity	Flag indicates whether the app supports 3G/4G Connectivity. It is extracted from the platform compatibility list.	Time-Invariant
Controls		
Age	Age of an app (weeks since launch).	Time-Variant
Developer's experience	Number of apps developed by the app developer.	Time-Variant
Dependent Variables		
App rank based on downloads	Average download-based rank of an app in its popular category during the current week. The value ranged from 1 (highest) to 240 (lowest).	Time-Variant
App rank based on grossing revenue	Average gross revenue-based rank of the app in its popular category during a current week. The value ranged from 1 (highest) to 240 (lowest).	Time-Variant

2.2 Data Collection

During the study period of 9 months from 1st May 2011 to 31st Jan 2012, we tracked 2,761 mobile apps with 32,109 weekly aggregated observations. App ranks are maintained in various top charts: i) top download for free apps (with and without in-app purchases), ii) top download for paid apps (with and without in-app purchases), iii) top grossing for both free apps and paid apps (with and without in-app purchases). Our analysis focuses on ii) and iii).

The summary statistics of the mobile app data we collected and derived are presented in Table 2 and the correlation matrix is shown in Table 3.

Table 2. Descriptive Statistics (N = 32109)

Variable	Mean	Std. Dev.	Min	Max
Price	4.5133	7.2750	0.14	299.99
Review score	1.4671	1.9218	0	5
Popular category	755.5918	69.0068	329	1006
Age	10.4075	7.9928	0	34.14
Size	47.5099	179.2242	0.0049	1863.6797
Competition	5.0401	4.1749	0	10
Developer experience	6.0228	8.3509	1	95
Frequent updates	2.3515	1.8323	1	17
Price Reduction	0.3797	0.4853	0	1
Design for iPad	0.2910	0.4542	0	1
3G/4G connectivity	0.0164	0.1271	0	1

Table 3. Correlation Matrix

Variable		1	2	3	4	5	6	7	8	9	10	11
ln(Price)	1	1.00										
Review score	2	-0.05	1.00									
Popular category	3	-0.01	-0.16	1.00								
Age	4	0.03	0.12	0.06	1.00							
Size	5	0.26	0.06	-0.03	0.01	1.00						
Competition	6	-0.11	-0.04	0.09	-0.02	-0.07	1.00					
Developer experience	7	-0.06	-0.07	0.03	0.01	0.07	0.00	1.00				
Frequent updates	8	-0.03	0.22	0.05	0.46	-0.04	0.06	-0.05	1.00			
Price reduction	9	0.10	0.20	-0.07	0.06	0.13	-0.06	-0.03	0.14	1.00		
Design for iPad	10	0.23	-0.05	-0.02	-0.04	0.03	-0.04	0.00	-0.03	0.06	1.00	
3G/4G connectivity	11	-0.06	0.06	-0.02	0.01	-0.03	-0.03	-0.04	-0.01	0.04	-0.03	1.00

As reported in Table 3, we can observe that the correlations between attributes are in the acceptable range. The highest correlation is 0.458 between Age and Frequent Updates. As we controlled for time effects in our subsequent model, there is no serious issue for modelling in our dataset.

2.3 Empirical Modeling

The download-based ranks can be influenced by the process of App Store Optimization (ASO) through app features, app positioning, developer actions and keywords; thus, we derive the below model, given app i and week t. The ranks and prices have been log-transformed rather than actual values to model the non-linear relationships amongst them and other attributes.

$$\begin{aligned}\ln(Download\ Rank_{it}) &= \beta_3 \times \ln(Price_{it}) + \beta_4 \times (Price\ Reduction_{it}) \\ &+ \beta_5 \times (Review\ Score_{it}) + \beta_6 \times (Frequency\ Updates_{it}) \\ &+ \beta_7 \times (Category\ Popularity_{it}) + \beta_8 \times (Competition_{it}) \\ &+ \beta_9 \times (Size_{it}) + \beta_{10} \times (Age_{it}) \\ &+ \beta_{11} \times (Developer's\ Experience_{it}) \\ &+ \beta_{12} \times (Design\ for\ iPad_i) \\ &+ \beta_{13} \times (3G/4G\ Connectivity_i) \\ &+ (Category\ Dummies_{it}) + \propto_i + cons\end{aligned} \tag{1}$$

On the other hand, *Sales Revenue* can be computed based on Power Laws and log-log distribution [14] using the following equation:

$$\ln(Sales\ Revenue_{it}) = b_{01} + b_{02} \times \ln(Grossing\ Rank_{it}) \tag{2}$$

We similarly posit an equation which depicts the relationship between downloads and download-based ranks:

$$\ln(Download_{it}) = b_{11} + b_{12} \times \ln(Download\ Rank_{it}) \tag{3}$$

In this study, we conceptualize a key variable, *Earning per Download (EPD)*, which plays a crucial role in measuring the effectiveness of pricing strategies for app developers. *EPD* is calculated as follows:

i. In paid model:

$$EPD_{it} = Price_{it} \tag{4}$$

ii. In paid model with in-app purchases:

$$EPD_{it} = Price_{it} + InApp_{it}\text{ , where } InApp_{it} = (InApp\ Sales)_{it}/Download_{it} \tag{5}$$

iii. In free model with ad-supports

$$EPD_{it} = Ads_{it}\text{ , where } Ads_{it} = (Ads\ Revenue)_{it}/Download_{it} \tag{6}$$

iv. In free model with in-app purchases

$$EPD_{it} = InApp_{it}\text{ , where } InApp_{it} = (InApp\ Sales)_{it}/Download_{it} \tag{7}$$

We argue that *EPD* is computable as a function of prices and in-app purchases. The following equation was developed based on the definition of *EPD*:

$$Sales\ Revenue_{it} = Download_{it} \times EPD_{it} \tag{8}$$

Based on equations (2), (3), and (8), we posit that the grossing-based ranks are the result of pricing strategies which are reflected in the configuration of *EPD*:

$$\ln(Grossing\ Rank_{it}) = \beta_0 + \beta_1 \times \ln(Download\ Rank_{it}) + \beta_2 \times \ln(EPD_{it}) \quad (9)$$

$$\text{where: } \beta_0 = \frac{(b_{11} - b_{01})}{b_{02}},\ \ \beta_1 = \frac{b_{12}}{b_{02}},\ \beta_2 = \frac{1}{b_{02}}$$

Knowing the prices of paid-only apps, top-download ranks for paid apps, and top-grossing ranks for all apps, EPD for paid apps with in-app purchases can be inferred using the following formula:

$$EPD_{it} = e^{\left(\frac{1}{\beta_2}\ln(Grossing\ Rank_{it}) - \frac{\beta_1}{\beta_2}\ln(Download\ Rank_{it}) - \frac{\beta_0}{\beta_2}\right)} \quad (10)$$

where: $EPD_{it} = Price_{it} + InApp_{it}$ for paid apps with in-app purchases

3 Data Analysis and Results

We analyze our structural model using two-stage regressions on the panel data in order to account for both endogenous variables: download-based ranks and grossing-based ranks.

In the first stage, we estimate the app ranking based on downloads using the Hausman-Taylor model as the hybrid model of both Fixed Effects (FE) and Random Effects (RE). The estimator allows us to capture unobserved individual heterogeneity and estimate the effects of both time-variant and time-invariant attributes. We also checked for panel-level autocorrelation, heteroskedasticity, Hausman's test and the effect of outliers to ensure robust and unbiased results. Table 4 summarizes the results of the app ranking model estimation.

The following are our findings from the first stage regression:

- **Developer actions:** As observed in Table 4, the effects of Price and Price Reduction are highly significant. When the Price is high, the app rank or the amount of download tends to be worst; however, Price Reduction such as having a short-term promotion can be a strategic factor to improve the app rank. Besides, releasing frequent updates of the app would also lead to a superior ranking.
- **User engagement:** The effect of good ratings on the app demand where highly-rated apps would probably draw more attentions from mobile users; thus they achieve better rankings in our model.
- **App features:** Mobile apps which are solely designed for iPad tend to be ranked better in terms of downloads. This suggests app developers to take advantages of the larger screen of the tablet rather than blowing up smartphone screens. Furthermore, as mobile devices are getting improved considerably on storage size and connectivity; these features are no longer the concerns for app users.
- **App positioning:** The effect of category popularity is significant where a category with a larger number of apps shows stiffer competition. For example, news, reference or sports categories are highly potential for publishing new apps; while, games, photo & video, or utilities are hypercompetitive to achieve better rankings. Table 5 reports the effects of categories on download-based app rankings.

Table 4. Estimation of App Ranking based on downloads

ln(Download Rank)	*Coefficient*		*z*	*P>z*
ln(Price)	β_3	0.3340190	26.12	0.000
Price Reduction	β_4	-2.5275620	-5.36	0.000
Review Score	β_5	-0.0234028	-10.15	0.000
Frequency Updates	β_6	-0.0048836	-1.48	0.139
Category Popularity	β_7	0.0001563	3.44	0.001
Competition	β_8	0.0074840	1.56	0.119
Size	β_9	0.0000002	1.02	0.309
Age	β_{10}	0.0213924	38.01	0.000
Developer's Experience	β_{11}	0.0119192	8.62	0.000
Design for iPad	β_{12}	-0.1288068	-2.25	0.025
3G/4G Connectivity	β_{13}	0.3072291	1.47	0.141
Category Dummies		(shown in table 6)		
_cons		5.2987120	22.49	0.000
Number of observations: 32,109, number of apps: 2,761, R^2 = .7123				

Table 5. Effects of categories on App Download Rankings

Category	*Coefficient*	*P>z*	*Category*	*Coefficient*	*P>z*
Books	-0.5595051	0.008	News	-1.011974	0.000
Business	-0.3238371	0.137	Photo & Video	-0.1083963	0.637
Education	-0.1993899	0.381	Productivity	-0.2158032	0.318
Entertainment	-0.4565721	0.024	Reference	-1.026778	0.000
Finance	-0.5922686	0.003	Social Networking	-0.9275086	0.000
Games	-0.1427003	0.584	Sports	-0.9738315	0.000
Lifestyle	-0.93245	0.000	Travel	-0.7910265	0.000
Medical	-0.9302471	0.000	Utilities	-0.3871226	0.060
Music	-0.7792271	0.000	Weather	-0.8063683	0.000
Navigation	-0.8063852	0.000			

In the second stage, we estimate the equation (9) where download-based ranks are computed as the residuals of the first stage regression and *EPD* is equivalent to Price for paid-only apps as in the equation (4). We performed various models such as Pooled OLS, Random Effect, or Fixed Effect; and Hausman Test to justify the effectiveness and unbiasness of our ultimate Fixed Effect estimator. Moreover, we corrected the variance–covariance by applying the accurate mean squared error.

The below table reports partial elasticity of download ranks and *EPD* on the grossing ranks.

Table 6. Estimation of App Ranking based on grossing revenue

ln(Grossing Rank)		Coefficient	t	P>z
ln(Est. Download Rank)	β_1	0.9478962	5.16	0.000
ln(EPD)	β_2	-0.4327684	-2.91	0.004
_cons	β_0	0.8889208	1.18	0.237
Number of observations: 32,109, number of apps: 2,761, R^2 = .7341				

There is the significant effect of earning per download in which a one percent increase in *EPD* results in 0.43% better in the grossing-based rank. On the other hand, an increment in the download-based rank is associated with an increment of 0.95 in the grossing-based rank; thus, the lower number of download leads to the decline in sales revenue, however, at a diminishing rate. These effects demonstrate the trade-off between the demand and the *EPD* for an app; nevertheless, there is an appropriate value of earning per download where the app unseals the full potential of its position.

Based on the equation (10), we depict the estimated *EPD* as follows:

$$EPD = 7.7992964x\frac{Grossing\ Rank^{-2.3107048}}{Download\ Rank^{-2.1903083}}$$

Table 7 lists several estimated values of *EPD*.

Table 7. Estimation of Earning Per Download

No.	*Ranking based on downloads*	*Ranking based on sales revenue*	*Estimated Earning Per Download*
1.	1	1	7.80
2.	10	10	5.91
3.	20	25	3.25
4.	30	40	2.66
5.	50	70	2.24
6.	70	100	2.05
7.	100	150	1.76
8.	100	180	1.15

The above numbers provide a guideline for developers to infer an appropriate revenue model and price settings given the known demand. According to the results, a paid app can earn up to $7.80 per download when it ranked no. 1 in both download-based and grossing-based lists.

For illustration, the estimated app no. 4 in Table 7 (ranked 30^{th} in the download-based list and 40^{th} in the grossing-based list) should be profited up to $2.66 per download in which the price should be set at $1.99 and $0.67 should be earned from in-app purchases.

Similarly, in order to crack into top 10 in the top grossing list, an app which is ranked at 10^{th} in the top paid list should earn at least $5.91 per download. Thus, the app developer should consider to fix the app price at $5.99 or to set a lower price, along with introducing in-app purchases to draw more impressions.

Figure 1 shows the average prices and earning per download of apps at ranks between 1 and 50 during the study period.

In comparison to app prices, earnings per download for high-demanded apps are higher which rootle out the effects of in-app purchases in the light of converting new users into sales revenue.

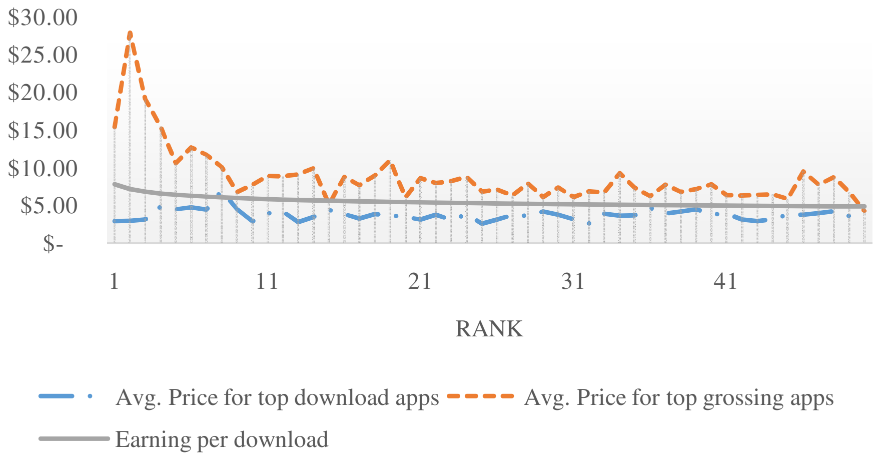

Fig. 1. Comparison of app prices and earning per download

4 Conclusion

Our study takes one step further in advancing mobile analytics on pricing strategies in the fast-growing environment of mobile apps. We proposed a structural model which is capable of generating reliable insights for market intelligence with publicly available data. Time and category effects, along with endogenity and heterogeneity issues are also considered in the model. Most importantly, we conceptualized a new indicator of market intelligence, Earning Per Download, which is useful for both app developers and researchers in the mobile industry.

There are several implications for app developers and publishers. First, we provided directive numbers for them to design effective pricing strategies and to unlock the full potential of their app positioning. Second, app developers should target their market niches based on our findings of category effects. Third, in AppStore, most of potential customers are forgiving and paying attentions on the latest review score; thus, releasing frequent updates is a perfect solution to gain customers' confidence and downloads. Last, designing apps for larger screen devices such as iPad would be strategic to capture shares in the app store.

This paper is not an end, but rather a beginning of forthcoming research. We note that app ranks are extremely volatile and being varied in the matter of hours; thus, crawling and matching of the data on a finer time scale rather than weekly basis are necessary. Furthermore, we are in the process of fetching additional in-app purchase data which would shed new lights on market intelligence of free apps with in-app purchases. By extending our model with them, app market best-kept secrets such as sales revenue or ranking mechanism would be publicly exposed.

References

1. Datta, A., Dutta, K., Kajanan, S., Pervin, N.: Mobilewalla: A mobile application search engine. In: Proc. 3rd Int. Conf. Mob. Comput. Appl. Serv. (2011)
2. Murphy, D.: Mobilewalla: The Market is Inches Away from One Million Mobile Apps, `http://www.pcmag.com/article2/0,2817,2397200,00.asp`
3. Shevchik, L.: Mobile App Industry to Reach Record Revenue in 2013 (2013), `http://blog.newrelic.com/2013/04/01/mobile-apps-industry-to-reach-record-revenue-in-2013/`
4. Inc., G.: Gartner Says Worldwide Mobile Application Store Revenue Forecast to Surpass $15 Billion in 2011, `http://www.gartner.com/newsroom/id/1529214`
5. Ramasubbu, N., Sangaralingam, K., Pervin, N., Dutta, K., Datta, A.: Surviving Hyper-Competitive, Unforgiving Platform Ecosystems: Examining Developer Strategies in iOS and Android Marketplaces. In: The 22nd Annual Workshop on Information Technologies and Systems, Orlando, Florida, USA (2012)
6. Kanada, P.: Which Can Be Best Mobile App Revenue Models?, `http://theappentrepreneur.com/mobile-app-revenue-models`
7. Apple: App Store Downloads on iTunes, `http://itunes.apple.com/us/genre/ios-games/id6014?mt=8`
8. Sangaralingam, K., Pervin, N., Ramasubbu, N.: Takeoff and Sustained Success of Apps in Hypercompetitive Mobile Platform Ecosystems: An Empirical Analysis. In: Proceedings of 2012 International Conference on Information Systems (ICIS)., Orlando, Florida, USA (2012)
9. Fang, F.: A method for mobile app download estimation (2013)
10. Garg, R., Telang, R.: Inferring App Demand from Publicly Available Data. MIS Q. (2012)
11. NPD Group Consumer Tracking Service, M.P.: OS Share of Smart phone sales (2011)
12. Venturedata: Apple Updates App Store Application Ranking Algorithm: Downloads a Greater Impact (2012)
13. Freierman, S.: One million mobile apps, and counting at a fast pace. New York Times (2011)
14. Goolsbee, A., Chevalier, J.: Measuring prices and price competition online: Amazon and Barnes and Noble. Quant. Mark. Econ. 1, 203–222 (2003)