Canalization or Increased Diffusion? An Empirical Analysis on the Impact of the Recommendation System in the Mobile App Market

Charles Zhechao Liu
Univ. of Texas at San Antonio
charles.liu@utsa.edu

Mohsen M. Jozani
Univ. of Texas at San Antonio
mohsen.jozani@utsa.edu

Kim-Kwang Raymond Choo
Univ. of Texas at San Antonio
raymond.choo@utsa.edu

Abstract

Online retailers have increasingly adopted product recommendation systems as an effective tool to improve product visibility and promote sales. This study examines the impact of the recommendation system in the popular Google Play mobile app store. By analyzing a 60-day panel dataset with 235,638 observations from 9,735 apps, we investigate how the characteristics of the recommended apps relative to those of the focal apps affect the adoption of mobile apps in this volatile market. Our results show that the relative strength of the recommended apps over the focal app plays a key role in influencing the outcome of recommendations. Moreover, the heterogeneity of the recommendations as represented by the diversity of the popularity of the recommended apps is positively associated with a more even distribution of revenue in the market. These findings provide insights for mobile app market operators to enhance the design of their recommendation systems.

1. Introduction

The market for mobile devices has undergone a rapid growth in the last few years with an equally remarkable increase in the development and use of mobile apps [1]. With statistics showing that mobile users install up to 40 apps and on average spend 2 hours daily using these apps [2], it is not surprising to see that the revenue for mobile apps has reached $89 billion in 2016 and is projected to exceed $189 billion by 2020 [3].

As the leading mobile platform, the Android market has reportedly taken up more than 86% of the market share and attracted thousands of software developers to develop mobile apps on this platform [4]. The large number and diversity of apps available in this market and the constant evolution of the Android platform have made it challenging for most users to choose which apps to install and for the platform administrator

to design effective mechanisms to promote their apps [5]. In the mobile market, apps are primarily distributed through centralized marketplaces such as Android's Google Play Store and Apple's App Store. A mobile app marketplace is an online platform where users can download and update applications to increase the utility and productivity of their mobile devices [6]. As a convenient venue for developers to publish and advertise their apps, a centralized marketplace has significantly contributed to the diffusion and adoption of mobile apps. However, the recent explosive growth of the number of mobile apps has substantially changed the competitive landscape in these marketplaces. By March 2017, the number of apps on Google Play has surpassed 2.8 million and is growing rapidly daily [3].

Previous studies suggest that unlike most e-commerce websites, Google Play is a superstar market favoring mostly popular apps [7]. Due to potential network effects and herd behavior [8], consumers generally go after apps from well-known developers or apps that have successfully established their reputation in the market, leading to a self-reinforcing loop in which the strong gets stronger and the weak gets weaker [1]. Small and new app developers often find it challenging to compete against these dominating players. Even if they develop disruptively innovative apps, their ideas will soon get implemented by the large app tycoons before they can gain the momentum they deserve.

In light of the characteristics of the mobile app market, consumers have increasingly relied on other information to identify the apps that meet their needs. Among the various factors that help a mobile app stand out in the fierce competition, product visibility has been found to play a critical role in improving the exposure and acceptance of a mobile app [5], [9]. In most electronic markets, product reviews and recommendation systems have been implemented to facilitate the product search process and enhance the visibility of the product [10]. In the mobile app market, similar approaches have been adopted. Product reviews provide an objective assessment of the quality and

HICSS

features of the app, while the recommendation system facilitates and expedites decision making when a user is overwhelmed with multitude of choices [10]–[12]. The recommendation system can draw users' attention toward under-exposed or niche apps that they would have hardly noticed on their own [13], which is analogous to the shelf placement strategy adopted in brick and mortar retailing [14] and the concepts of 4P and 3C widely examined in the marketing literature [15], [16]. In doing so, the market creates a vast network of apps that are inter-connected via recommendations, essentially forming clusters of apps based on their similar characteristics.

Despite that recommendation mechanisms have long been implemented in the mobile app market, little is known about the performance and outcome of such mechanisms. Given the vastly abundant choices and much lower marginal cost of adoption (relative to that of a physical goods) in the mobile market, it remains unknown whether consumers are following the recommendations and, if so, whether the use of such systems are making the search process more efficient for consumers and profit-enhancing for the platform operators.

Since very little empirical research has been conducted on the impact of the recommendation systems in the mobile market, in this research we focus on empirically analyzing the content-based recommendation system adopted by the largest app store in the mobile market - the Google Play store, and examining whether the recommendation system alters the market structure and shapes the competitive dynamics in this leading mobile app market.

Specifically, our study seeks to answer the following questions:
1. Does the recommendation system increase the overall adoption of mobile apps in the market?
2. How do the differences in various app characteristics (i.e., review rating, number of downloads, popularity) between the focal app and the recommended apps influence the outcome of the recommendations - i.e. do they boost or overshadow the adoption of either the focal apps or the recommended apps, or both?
3. How does the heterogeneity of the app recommendations change the distribution of the app downloads and the equality (or inequality) of the mobile app market?

To answer these questions, we briefly review the related literature and present our theory and research hypotheses, followed by a description of our data and research methodology. Then we summarize the results from the empirical analysis and conclude the paper with directions for future research.

2. Theoretical background

In recent years, the growing literature on mobile app has examined various technical, social and business aspects of this emerging ecosystem to understand factors that lead to the prosperity of the market. Valuable insights can be drawn from examining market characteristics such as app acquisition and usage (i.e. [17], popularity trends (e.g. [1]), determinants of success (e.g. [9]), and revenue strategies (e.g. [5]). Among these themes, revenue models and success factors of mobile apps have been key topics of interest in the mobile app literature due to their impact on profits [6]. Since the creation of the market, platform operators and a lot of the mobile app developers have embraced a freemium revenue model that differs from that of the traditional software products [9]. One of the most important characteristics of the freemium revenue model is that it offers users an opportunity to try their apps before paying for it. Since a lot of apps can be personally evaluated, product review, the traditional indicator of product quality in the electronic markets, has been superseded by product visibility as the most critical determinant of success in the mobile app market [5]. With the extremely large number of apps available in the market and the very limited amount of time a user can spend on evaluating an app, catching the attention of the user has become a critical prerequisite for adoption [9].

Platform operators such as Google and Apple have also realized the importance of product visibility and implemented different mechanisms to promote it. For example, both markets offer various ranking lists in their app stores to allow users to browse a variety of apps based on app category or their recent popularity. In addition, a set of similar apps are recommended to users as they are browsing an app that shares some similar features. Theoretically, these recommendations should improve the overall exposure of the apps in the market and provide more competing opportunities for less well-known apps. However, Google Play has been found to produce a "winner-take-all" market outcome [7], where 10% of all apps reportedly received 70% to 90% of all downloads [1], a result that contradicts the prediction derived from other long tail markets that also exhibit extensive product variety and low search cost [9]. This phenomenon leads us to investigate how the app recommendations work in the mobile market and if they are playing the same role that has been documented in other markets.

A recommendation system is a computer-mediated system configured to form a large web of inter-connected products to help consumers find a product that meets their requirements and/or encourage them to purchase additional products to achieve cross-selling

[12], [18]. When the recommendation system is implemented, each product in this network is connected to a group of products that share similar attributes and the network position of each product, as determined by the number of its incoming and outgoing links, affects the overall product demand [19]. These recommendation systems have recently become popular in online environments, such as social media, e-commerce and mobile app markets [20]. They have been shown to increase product sales, product diversity, customer satisfaction, and customer loyalty, as well as providing a better understanding of customer needs [21], [22].

The use of online recommendation systems has been examined extensively in the literature. Several studies have focused on the recommendation system on Amazon.com since it is one of the most successful implementations of the recommendation systems [18], [20]. These studies examine the co-purchase recommendation system on Amazon Bookstore using either SNA (Social Network Analysis) or PageRank as the measure of importance for each product page [10], [14], [19]. In general, they found that the position of a product on the recommendation network affects its overall demand. Extending these studies, Lin et al. [18] examine the role of both incoming and outgoing links on product demand on the co-purchase network of Tmall, a popular Chinese ecommerce website. The authors found that the diversity and stability of the outgoing links decrease the demand for the focal product (the anchor product based on which the recommendations are made).

Building on these work, in this paper we analyze a similar recommendation system in the Google Play mobile app market. In particular, instead of using the standard network measures such as income or outgoing links adopted in the extant literature, we will characterize the relative position of an app in a recommended network through directly measuring the differences between the focal app and the associated recommended app in terms of their key characteristics such as quality and popularity, which is an approach that has not been used in other studies.

Moreover, in addition to influencing the demand for an app, the use of the recommendation system can potentially change the market structure, which in turn, transform the competitive dynamics. Prior studies have applied the theory of Long Tail to determine the impact of product recommendations on market concentration [7], [14]. Their findings indicate that the use of recommendation system positively affects the long tail distribution and consequently the market structure.

Given these findings, we attempt to go beyond the long tail distribution and examine whether the use of the recommendation system improves the

competitiveness of the relatively newer and less well-known apps in the market. To this end, we will use the Gini Coefficient [23] to examine the download distribution within the mobile app market. Gini Coefficient, originally introduced as a measure of inequality in income distribution, has often been used to study demand and revenue distribution. As illustrated in Figure 1, we will generate the Lorenz Curve for each app category and calculate the corresponding Gini Coefficient using the number of downloads computed from the app's sales rank. Then, we will examine the relationship between the diversity of the app recommendations and the Gini Coefficient within each app category to determine if a more diverse recommendation portfolio reduces the inequality in the mobile app market.
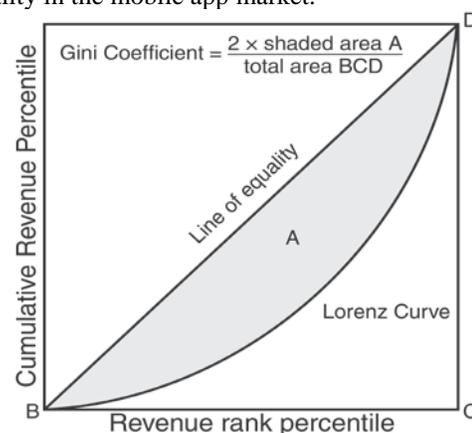
Figure 1. Lorenz curve and Gini coefficient in mobile app market

In summary, although the impact of recommendation systems in e-commerce has been widely studied, to the best of our knowledge, this is the first study that investigates the role of recommendation systems in mobile app market. Furthermore, by proposing our unique angles on the direct competition between the focal app and the recommendations, and on the distribution of app adoption, our study will provide important theoretical and practical guidance on how the design of the recommendation system can lead to a more sustainable growth of the mobile app market.

3. Hypotheses and theory development

As the official market for the popular Android platform, Google Play offers a homepage for each app that displays a wide range of information such as app title, description, version, review rating, rating count, number of installs, and developer information. More importantly, regardless of the device used to access the app's homepage, Google Play shows a list of "Similar" apps that share some common features with the app

being viewed (the focal app). This essentially creates a network of recommendations that customers can use to compare apps and identify the app that best meets their search criteria.

Given that consumers will be able to see the recommended apps that otherwise would have been difficult to discover, the recommendation effectively increases the visibility of a mobile app and the probability of download. A good recommendation system in an app store can update the recommended list dynamically based on the constantly-changing market trends and shifting user preferences, creating a more vibrant market where even less well-known developers and less popular apps also have a chance to compete for user attention and actual downloads. Therefore, consistent with the findings of the prior studies that examine the effect of recommendation system on demand or revenue, we expect that there is a positive relationship between the frequency of presence on the recommendation list and the sales performance of the app:

H1: All else being equal, the frequency of the recommendations is positively associated with the sales performance of a mobile app.

In addition to the boosting effect on the recommended app's sales performance, the presence of the recommendations also leads to a direct competition between the focal app and the recommended apps. In electronic markets, product canalization [20] often occurs when multiple listings are simultaneously competing for the limited time and attention a shopper can devote in the product search process. A central question of interest is that, does the use of the recommendation system lead to a zero-sum game in which the sales of the recommended apps increases at the expense of the decreasing sales of the focal app, or does the use the recommendation system increase the overall demand for all similar products across the market due to the positive spillover effect?

The extant literature has largely addressed these questions at the market level, with specific focuses on how the network structure of the recommendation systems affects the outcome of the recommendations [14], [18], [19]. However, we adopt a different perspective in this study as we choose to focus on the individual app level rather than at the market level. We believe that mobile users on average can only spend very limited amount of time search for apps and the effect of recommendations occurs locally only among the apps that have been exposed to the user. Therefore, we expect the answers to these questions vary across apps and depend critically on the criteria of the comparison that are of the most importance to mobile app consumers. We argue that the relative outcome of the recommendations to be moderated by the nature of the competition, as driven by the comparison between the characteristics of the apps. Specifically, when the focal app has a relatively lower perceived quality (as reflected in their review rating), or are less well-received by consumers (as reflected in their number of downloads) relative to the apps being recommended, there is a greater chance that a consumer will choose the recommended app over the focal app, resulting in a cannibalization of the focal app. Conversely, if the focal app outperforms the recommended app in terms of either perceived quality or user adoption, the recommended apps will have little impact on the adoption of the focal app. Therefore, we hypothesize that:

H2a: All else being equal, the effect of recommendation is stronger when the recommended apps have higher average quality than the focal app.

H2b: All else being equal, the effect of recommendation is stronger when the recommended apps have lower average price than the focal app.

H2c: All else being equal, the effect of recommendation is stronger when the recommended apps enjoy higher average number of downloads compared to the focal app.

Other than influencing the adoption of the apps, the presence of the recommendation system can present a wider range of options to consumers, leading to a greater diversity in consumer's choice set. Studies had suggested that app stores exhibit a spillover effect that causes users to download multiple apps at the same time [1], even if some of them are not the ones that the user was originally searching for. Heterogeneous recommendations can lead to a greater extent of diversity both within and across category. This can encourage developers to diversify their app portfolio, and users can be exposed to niche app categories that are under-promoted. Therefore, based on similar research of product heterogeneity in electronic markets [18], we expect that, when the characteristics of the recommended apps are more diverse, which happens when apps from different price levels and different popularity levels are being recommended, it will result in a more evenly distributed impact of the recommendation system. Such heterogeneity will also help the minority apps in the market and potentially reverse the self-reinforcing loop that makes the strong grow stronger. Hence, we predict that:

H3a: Heterogeneity in the prices of the recommended apps is negatively associated with the inequality in the adoption of the mobile apps.

H3b: Heterogeneity in the popularity of the recommended apps is negatively associated with the inequality in the adoption of the mobile apps.

4. Data and research methodology

4.1. Sample characteristics

To test our hypotheses, we collected data from ten different app categories from the Google Play store on a daily basis for 60 consecutive days between November 2nd, 2016 and Dec 31st, 2016. Being the largest and the only official app store for the leading Android mobile platform, Google Play presents a unique opportunity to study the dynamics in the mobile app market. Using an automatic software agent, we collected detailed app level data such as app id, download rank, app categories, price, review rating, number of reviews, developer, the range of downloads, app type (paid vs. free), app features, app version, release date, in-app purchases, and the same set of data for a set of recommended apps. The final dataset consists of 235,638 observations in 60 days, covering 9,735 focal apps that appeared on our selected ranking lists and 42,977 apps that appeared on these focal apps' recommended lists.

Table 1: Descriptive statistics of key variables

Variable	Mean	Std.	Min.	Max.
Download Range	7042242.52	55676380.16	1	1000000000
PaidApp	0.42	0.49	0	1
Rank	231.09	149.72	1	504
TopRank Freqeuency	12.23	21.64	1	59
Rating	4.17	0.56	1	5
RatingCount	234042.48	1898984.204	1	57526695
Price	4.54	7.01	0	124.99
InApp Purchase	4.26	8.39	0.99	199.99
RecomApp Rating	4.18	0.51	1	5
RecomApp Price	5.12	10.36	0.01	299.99

*The statistics for download range is generated using the lower bound of the download range.

The dataset consists 4,480 (46.02%) free apps and 5,255 (53.98%) paid apps. A summary of the descriptive statistics of the data are shown in Table 1.

We chose to collect and analyze daily app data as the recommendations change daily and their effect takes place in real time. When an app is removed from the recommendation list, consumers may not see the app anymore and they cannot follow the link to examine it further. Hence, a daily longitudinal dataset allows us to effectively capture the impact of the recommendation system.

4.2. Estimating the number of downloads

In the literature, the sales performance variable (the dependent variable for H1, H2a, H2b, and H2c) has been measured by both the download rank and the actual number of downloads. The download rank data are readily available on Google Play and were collected by our software agent and included in the dataset. However, even though some prior studies have found that there exists an inverse relationship between sales rank and actual sales, which can be used to recover the sales data [24], others have argued that the validity of this approach depends on the assumption that the sales rank is computed from the actual sales figures [25]. Since this assumption is difficult to verify in the Google Play market where ranking algorithm remains largely unknown[1][25], [26], we decided to use the number of download as the measure for a mobile app's sales performance and only use the rank variable as an independent variable.

Given this measurement choice, the next issue we need to address is that Google Play does not publish the exact number of downloads and only provides statistics on the range of the number of downloads (i.e., 100-500, 500-1000, 1,000-5,000, …, etc.), which presents a great challenge on our data analysis. Upon monitoring the data published in the Google Play store, we found that due to high consumer demand for mobile apps, a substantial number of apps move from one download range to the other within a relatively short period of time (i.e., a couple weeks or even a few days in some popular app categories). This important observation leads us to conduct a separate analysis to recover the number of downloads for the apps in our sample. Specifically, we scanned through the entire sample and identified dates on which a particular app moved from one download range to the next level. For example, if an app moved from the "500 to 1000" download category to the "1000 to 5000" download category on Nov 15, 2016, then we can confidently infer that this particular app has received at least 1000 cumulative downloads by Nov 15. In other words, by monitoring changes in the download category on given

---

[1] Unlike other e-retailers, Google Play does not rank its apps solely based on download numbers. Factors such as rating, rating counts, and retention rate have all been found to influence the rank.

dates, we are able to recover the cumulative number of downloads for a subset of the apps in our dataset.[2]

Next, given that it has been widely established that there exists a positive correlation between the number of downloads and the number of reviews for products in electronic markets [7], [27], which we fully expect to be carried over to the mobile app market, we can uncover the relationship between the number of downloads and the number of reviews (labeled as *ratingCount* in Table 1). Once we obtain this parameter, we can generalize the relationship to the rest of the observations in our sample for which we only have the number of reviews.

Based on the above description, we constructed a separate dataset that consists of 1,358 apps (14% of the total number of apps we have in our sample) for which we identified the exact download number of downloads due to the change in download ranges. Then, we estimated the predicted relationship by regressing the actual number of downloads on the number of reviews, controlling for factors such as app type (free vs. paid), app popularity, app category, etc. Since the relationship between the number of downloads and the number of reviews may change as more users adopt the app, we divided the apps in our sample into three categories: low, medium, and high, based on their popularity. These three dummy variables are multiplied with the number of review variable to form three interaction terms. As a result of this model specification, the coefficients for these interacting variables represent the proposed relationship for apps in these specific categories.

We ran both OLS and a fixed-effect estimation on the proposed model. The OLS estimator served as a baseline case while the fixed-effect model was introduced to examine if there are any systematic differences due to factors not captured by our model. Table 2 presents the results of both regressions.

Table 2 shows that both models have good explanatory power over the proposed relationship. In particular, the three interaction terms that involved the key independent variable, *ratingCount* (number of reviews) were all significant in both models, indicating that there exists a strong positive relationship between the number of downloads and the number of reviews, and this relationship differs slightly across different types of apps (low, medium and high popularity) as expected. To determine which model gives us the best

approximation, we applied the coefficients of the three interaction terms back to the subsample which has the actual number of downloads. It turned out that the OLS estimates yielded the closest approximation to the true data. Hence the coefficients obtained from the OLS regression were used to compute the number of downloads for the rest of our sample.

Table 2: Number of downloads regression results[*]

Dependent variable *Number of Downloads*	Model 1 (OLS)	Model 2 (Fixed Effect)
PaidApp	-838.06[*]	-9.11[*]
Price	1268.06[**]	148.15
Rating	34.52	15.56
*RatingCount*Low*	5.37[**]	2.57[*]
*RatingCount*Med*	8.26[**]	3.82[*]
*RatingCount*High*	18.38[**]	3.48[**]
Fit Statistics (n=1,781)	0.81 (Adjusted R^2)	24.02[*] (*F Stat*)

[*]$p<5\%$,[**]$p<1\%$.

It is likely that the relationship between the number of downloads and the number of reviews also depends on other factors not captured by our model. However, given the high adjusted R square (0.81) in the OLS regression. We are confident that the set of parameter estimates we adopted provide a very good approximation to the true data and can be safely generalized to the entire sample.

4.3. Estimation method

Having obtained the number of downloads for all of the apps in our dataset, we proceeded to test H1 through H2c. To improve the validity of our results, we took the difference of the number of downloads variable and used these difference as our dependent variable. This procedure ensures that we are examining the changes in the number of downloads as a result of changes in the independent variables of our interest, and not their effects on the cumulative number of downloads (which could be influenced by a number of variables that were at work before the data collection time frame).

Since our dataset contains both cross-sectional and longitudinal data, it lends itself to estimators specifically designed for panel datasets. Following the well-established approach in econometrics [5], we addressed these issues by adopting the Generalized Least Squares (GLS) estimator with corrections for both cross-panel heteroscedasticity and within-panel autocorrelation. These corrections were incorporated because a Breusch-Pagan/Cook-Weisberg test for

---

[2] We tried to minimize the potential bias by collecting the data at the same time (midnight) each day. By doing so, we allow for the same time interval for downloads to accumulate. Therefore, even though download numbers could change at any time, the number collected reflects the number of downloads received within a constant 24-hour time window. In addition, there is relatively fewer downloads at midnight, further reducing any possible data discrepancy.

heteroskedasticity yields a high χ^2 across panels, indicating the presence of systematic variation across different app categories. Moreover, the Wooldridge test for autocorrelation also shows that substantial first-order autocorrelation (AR1) is present in our dataset, which is not surprising given the daily nature of the sample.

To test H1, we regressed the differenced number of downloads on the frequency of recommendations as measured by the number of times this app appears on other apps' recommendation lists, controlling for random effects such as price, review rating, and fixed effects such as app category and developer.

To test H2a through H2c, a comparison between each focal app and all of its recommended apps was conducted with regard to the difference in app rating, price, and the number of downloads. Then, the same dependent variable used in H1 will be regressed on the differences of review rating, price and app download number, controlling for a similar set of factors as used in testing H1. Note that for H2c, we are testing the influence of the total download numbers on a user's comparison between the focal app and the recommended app, hence the computed number of downloads was used instead of the difference in downloads, as the total number of download is a better measure for the popularity of a mobile app.

Finally, to test H3a and H3b, we focused on examining the degree of inequality in the adoption of paid mobile apps within different app categories. We decided to focus on paid apps only because the traditional measure of market inequality is based on income or revenue, whereas free apps do not generate upfront revenue and we were not able to observe the other revenue streams such as in-app purchases or advertising income through public sources. Hence, we limit our analysis to paid mobile apps only for which we can compute the revenue using their price multiplied by the number of downloads.

We set our unit of analysis at the category level instead of the market level for several reasons. First, Google Play is the largest mobile app market in the world and it is technically impossible to collect a sufficiently large sample to capture the true extent of revenue distribution in the market. Second, apps differ significantly in terms of their adoption across different categories (e.g., games enjoy far more downloads than a references app). Hence it is not meaningful to compare app adoption across different categories. Finally, we selected categories for which Google publishes the top categorical ranking list. This sampling approach allows us to include a representative body of the apps in the category and apply a common measure (i.e., categorical rank) to compare the adoption of similar apps.

To measure the dependent variable for H3a and H3b, the degree of heterogeneity in the prices and quality of the recommended apps, we computed the standard deviation of these variables for each focal app. Then we took the average of these standard deviations for all focal apps within the same category, which allows us to assess the category-level heterogeneity among all the recommendations.

4.4. Gini-coefficient

The key independent variable for H3a and H3b is the inequality of app adoption in a given category. Following Oestreicher-Singer and Sundararajan [14], we computed the within-category Gini coefficient based on the Lorenz curve [23], a well-established measure for income distribution in a given population. In our analysis, the horizontal axis of the Lorenz curve is the (ascending) ranking order (in percentile) of a mobile app's revenue (which is obtained from price multiplied by the number of downloads) within each app category, and the vertical axis is the cumulative app revenue as a fraction of the total revenue for all app up to the associated rank order.

More specifically, if we rank apps $N=$ {1, 2, 3, ..., n} into an ascending percentile ρ, where $0 < \rho \leq 1$, and compute the revenue as a percentile $y(\rho)$ of the cumulative revenue. Then the Lorenz curve can be plotted as a function of $L(\rho, y(\rho).)$. Accordingly, Gini Coefficient will be computed as twice the area between the Lorenz curve and the diagonal line (the line of equality) that starts from the origin and ends at the upper right corner with coordinates (1, 1). The higher the Gini-Coefficient, the higher the inequality in the market and the larger the shaded area. In contrast, when the market has no disparity, which happens when everyone receives the same revenue, the Lorenz curve will overlap with the diagonal line of equality, resulting in a zero value for the Gini-Coefficient and no shaded area in Figure 1.

Upon generating the values for both the dependent variable and independent variable, we ran a similar GLS regression on our data, with a specification for fixed-effects. This specification is adopted as all of the data points generated for this analysis are at the app category level and the category characteristics are likely to have a strong impact on the predicted relationship between the dependent variable and independent variable.

5. Results

We summarized the results of the two different regressions in Table 3. The independent variables used

are listed on the left column, and the two columns on the right correspond to the two different models used to test the three sets of hypotheses. Note that unit of analysis also differs across these two models. H1 and H2a through H2c were tested at the app level and H3a and H3b were tested at the app category level.

Table 3: Number of downloads regression results[*]

Hypotheses Tested:	H1, H2a – H2c	H3a and H3b
Dependent Variables:	Difference in Number of downloads	Gini-Coefficient
Constant	45.59[**]	.021[**]
RecomFrequency (H1)	.92[**]	N/A
Price	-15.08	.0317[**]
Rating	30.63[*]	.0044
Rank	-55.93[**]	.0005[**]
Paid (binary)	-8.26[**]	N/A
RecomRatingDiff (H2a)	13.14[**]	N/A
RecomPriceDiff (H2b)	2.52	N/A
RecomDownloadDiff (H2c)	0.0085[**]	N/A
Hetergeneity in Price (H3a)	N/A	0.18[*]
Hetergeneity in Downloads (H3b)	N/A	-0.0073[**]
Estimator	GLS (RE)	GLS (FE)
Fit Statistics	912.10[**] (χ^2)	44.02[**] (F Stat)

[*]$p<5\%$, [**]$p<1\%$. *The coefficients for control variables are omitted due to space limit.*

The middle column of Table 3 shows that there is a strong positive relationship between the frequency of recommendation and the daily incremental downloads, and that the effect of recommendation is significantly stronger when the recommended apps have a higher review rating, or when they have a higher number of downloads than the focal app, thus supporting H1, H2a and H2c.

Surprisingly, the price difference between the focal app and the recommended app does not seem have a significant influence on the number of downloads, suggesting that H2b is not supported.

A further examination of the coefficient obtained from the regression shows that, for an average app, each additional appearance on the recommendation list is associated with .92 more daily downloads. It is worth noting that this number only represents the medium effect for an app with average characteristics. This number is expected to be higher for apps that are on the top-ranking list or with a higher rating, and lower for apps that are not as popular or have a low

review rating. In addition, a one unit difference between the recommended app's review rating and that of the focal app is associated with 13.14 more downloads for the recommended app, whereas each 1000 download number difference is associated with 8.5 more downloads for the recommended app.

With regard to hypotheses H3a and H3b. The right column in Table 3 indicates that only the coefficient estimate for the heterogeneity in the download numbers of the recommended apps is significant and helps to reduce the revenue disparity in the mobile app market, thus supporting hypothesis H3b. The coefficient for the price heterogeneity variable has the negative sign and is not significant, thus rejecting H3a. To provide an overview of the results of our analysis, we summarized the outcome of the hypotheses testing in Table 4.

Table 4: Number of downloads regression results[*]

Hypothesized Relationship	Result
H1: Higher frequency of recommendation leads to more downloads	Supported
H2a: the effect of recommendation is stronger when the recommended apps have higher average quality than the focal app	Supported
H2b: the effect of recommendation is stronger when the recommended apps have lower average price than the focal app	Not Supported
H2c: the effect of recommendation is stronger when the recommended apps have higher average downloads than the focal app	Supported
H3a: Greater heterogeneity in the prices of the recommended apps leads to lower Gini-Coefficient (more equal distribution).	Not Supported (reverse sign)
H3b: Greater heterogeneity in the popularity of the recommended apps leads to lower Gini-Coefficient (more equal distribution).	Supported

6. Discussion

The results presented in the preceding section reveal interesting dynamics in the Google Play market where a mobile app recommendation system is implemented. Consistent with our theoretical conjecture, the presence of the app recommendations leads to more downloads of the recommended apps. Such a boosting effect is stronger when the recommended apps have higher review rating or download numbers than the focal app, suggesting that mobile app users do not simply follow the recommendations provided by Google Play. Instead, they carefully take advantage of the rich information available on the app's homepage before making their adoption decision. Moreover, among the various kinds of information present on the

recommended list, price is not a dominant factor when mobile users compare the recommended apps with the focal app, which is contrary to what we hypothesized.

These findings confirm that the impact of the recommendation system is equally evident in the mobile app market as in other electronic markets. More importantly, offering a recommendation list may not be sufficient to maximize all the potential of the recommendation system. Based on the findings of our research, platform administrators should facilitate the comparison of the key characteristics such as review rating and the number of downloads between the focal app and the recommended app, as doing so will significantly leverage the power of the recommendation system.

Our results also show that the choice of the recommended apps can significantly influence the disparity of mobile app adoption in the market, especially within certain app categories. Specifically, recommending apps in different stages of their lifecycle helps to improve the visibility of these otherwise under-promoted apps. The support for H3b also indicates that these low visibility app may benefit more from being recommended by the platform than those superstar apps do, who most likely have already occupied a spot on one of the top-ranking list, or has been advertised heavily through other communication channels. If the objective of a platform administrator is to provide a marketplace with a more competitive environment and foster the sustainable growth of the market, they should try to promote these relatively young and less well-established apps and their developers to ensure that they will not be forced to exit the market due to the dominating apps attracting the majority of user attention in the market.

Furthermore, our results also indicate that the heterogeneity in the prices of the recommended app does not seem to have the expected influence on the revenue distribution in the market. This is consistent with the lack of support for H2b and shows that price is not the most important determinant in the mobile app market. Due to the relatively low prices of mobile apps, and more importantly, the wide availability of free apps, it is possible that mobile app users can easily find free apps that meet their needs, or find a needed paid app that offers the free version for them to trial. In either case, they don't have to sacrifice quality or use experience in return for a small savings in app price. This result, if proved to be true in similar mobile app markets, will have a profound impact on the developer's app promotion strategy, as they will have to focus on other product dimensions beyond price, which has traditionally been a key determinant of competition.

7. Conclusion and future research

Online retailers have increasingly relied on the use of recommendation systems to improve product visibility and promote sales. This study empirically examines the effect of recommendation system in the leading Google Play mobile app market, which is uniquely characterized with its multitude of app choices and the prevailing use of the recommendation system. We collect daily app data from the Google Play store and constructed a 60-day panel dataset that includes a wide range of data on app recommendations and their market performance. This rich dataset allows us to examine how the competition between the focal app and its recommendations affects their relative adoptions, and how the heterogeneity of the recommendations influences market inequality.

Our research will contribute to both practices and research in the increasingly popular mobile app domain, which has intrinsically different characteristics from other e-commerce markets in terms of reaction, flexibility, competition and product portfolio [9]. From an academic perspective, our study extends the recommendation systems literature and introduces the new angle of competition within recommendation system and inequality of adoption to the mobile app ecosystem. The new approach we developed to recover the actual number of downloads from a subset of the data observed in the market has important theoretical and empirical implications on research in electronic markets where actual sales data are difficult to obtain

From a practical standpoint, the result of our study will provide insights for mobile app market operators to enhance the design of their recommendation systems and provide tangible measures to evaluate the performance of these systems. Our findings can also benefit the developers to by providing them a better understanding of the competition and guidance on how to focus their efforts in areas that can best leverage their competitive advantages.

To the best of our knowledge, our study is among the first to empirically examine the impact of recommendation systems in the mobile app market. The results of our study open up a new arena for future research. It will be interesting to examine whether the use of the recommendation system leads to a zero-sum game in which the recommended apps gain market share at the cost of the focal apps, or whether it leads to a market expansion in which more downloads are achieved for both focal and recommended apps, or how superstar or highly rated recommended apps may influence app downloads, as all are questions of great importance to mobile app market stake holders.

References

[1] T. Petsas, A. Papadogiannakis, M. Polychronakis, E. P. Markatos, and T. Karagiannis, "Measurement, Modeling, and Analysis of the Mobile App Ecosystem," *ACM Trans Model Perform Eval Comput Syst*, vol. 2, no. 2, p. 7:1–7:33, Mar. 2017.

[2] "App Annie 2016 Retrospective — Mobile's Continued Momentum," *App Annie*, 2017. [Online]. Available: http://go.appannie.com/app-annie-2016-retrospective. [Accessed: 15-Jun-2017].

[3] "Number of available applications in the Google Play Store from December 2009 to March 2017," 2017. [Online]. Available: https://www.statista.com/statistics/266210/number-of-available-applications-in-the-google-play-store/. [Accessed: 15-Jun-2017].

[4] "Smartphone Vendor Market Share, 2016 Q3," *IDC*, 2017. [Online]. Available: http://www.idc.com/promo/smartphone-market-share/vendor. [Accessed: 15-Jun-2017].

[5] C. Z. Liu, Y. A. Au, and H. S. Choi, "Effects of Freemium Strategy in the Mobile App Market: An Empirical Study of Google Play," *J. Manag. Inf. Syst.*, vol. 31, no. 3, pp. 326–354, Winter 2014.

[6] P. Roma and D. Ragaglia, "Revenue models, in-app purchase, and the app performance: Evidence from Apple's App Store and Google Play," *Electron. Commer. Res. Appl.*, vol. 17, pp. 173–190, May 2016.

[7] N. Zhong and F. Michahelles, "Google Play is Not a Long Tail Market: An Empirical Analysis of App Adoption on the Google Play App Market," in *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, New York, NY, USA, 2013, pp. 499–504.

[8] W. Duan, B. Gu, and A. B. Whinston, "Informational Cascades and Software Adoption on the Internet: An Empirical Investigation," *MIS Q.*, vol. 33, no. 1, pp. 23–48, 2009.

[9] G. Lee and T. S. Raghu, "Determinants of Mobile Apps' Success: Evidence from the App Store Market," *J. Manag. Inf. Syst.*, vol. 31, no. 2, pp. 133–170, Fall 2014.

[10] B. Leem and H. Chun, "An impact of online recommendation network on demand," *Expert Syst. Appl.*, vol. 41, no. 4, Part 2, pp. 1723–1729, Mar. 2014.

[11] V. Carchiolo, A. Longheu, M. Malgeri, and G. Mangioni, "Searching for experts in a context-aware recommendation network," *Comput. Hum. Behav.*, vol. 51, Part B, pp. 1086–1091, Oct. 2015.

[12] J. Goldenberg, G. Oestreicher-Singer, and S. Reichman, "The Quest for Content: How User-Generated Links Can Facilitate Online Exploration," *J. Mark. Res. JMR*, vol. 49, no. 4, pp. 452–468, Aug. 2012.

[13] S. Wang, M. Gong, H. Li, and J. Yang, "Multi-objective optimization for long tail recommendation," *Knowl.-Based Syst.*, vol. 104, pp. 145–155, Jul. 2016.

[14] G. Oestreicher-Singer and A. Sundararajan, "Recommendation Networks and the Long Tail of Electronic Commerce," *MIS Q.*, vol. 36, no. 1, pp. 65-A4, Mar. 2012.

[15] E. J. McCarthy, "Basic Marketing, Richard D," *Irwin Homewood IL*, vol. 181, 1964.

[16] R. Lauterborn, "New marketinglitany: 4Ps Passes, 4Cs Takeover," *Adverting Age*, 1990.

[17] R. M. Frey, R. Xu, and A. Ilic, "Mobile app adoption in different life stages: An empirical analysis," *Pervasive Mob. Comput.*, 2017.

[18] Z. Lin, K.-Y. Goh, and C.-S. Heng, "The Demand Effects of Product Recommendation Networks: An Empirical Analysis of Network Diversity and Stability," *MIS Q.*, vol. 41, no. 2, pp. 397-A10, Jun. 2017.

[19] G. Oestreicher-Singer and A. Sundararajan, "The Visible Hand? Demand Effects of Recommendation Networks in Electronic Markets," *Manag. Sci.*, vol. 58, no. 11, pp. 1963–1981, Jun. 2012.

[20] S. S. Li and E. Karahanna, "Online Recommendation Systems in a B2C E-Commerce Context: A Review and Future Directions," *J. Assoc. Inf. Syst. Atlanta*, vol. 16, no. 2, pp. 72–107, Feb. 2015.

[21] B. Pathak, R. Garfinkel, R. D. Gopal, R. Venkatesan, and F. Yin, "Empirical Analysis of the Impact of Recommender Systems on Sales," *J. Manag. Inf. Syst.*, vol. 27, no. 2, pp. 159–188, Oct. 2010.

[22] F. Ricci, L. Rokach, and B. Shapira, "Introduction to Recommender Systems Handbook," in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds. Springer US, 2011, pp. 1–35.

[23] C. Gini, "Measurement of Inequality of Incomes," *Econ. J.*, vol. 31, no. 121, pp. 124–126, 1921.

[24] A. Ghose, M. D. Smith, and R. Telang, "Internet Exchanges for Used Books: An Empirical Analysis of Product Cannibalization and Welfare Impact," *Inf. Syst. Res.*, vol. 17, no. 1, pp. 3–19, 2006.

[25] R. Garg and R. Telang, "Inferring App Demand from Publicly Available Data," *MIS Q.*, vol. 37, no. 4, pp. 1253–1264, Dec. 2013.

[26] F. Zhu and X. (Michael) Zhang, "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *J. Mark.*, vol. 74, no. 2, pp. 133–148, Mar. 2010.

[27] M. Harman, Y. Jia, and Y. Zhang, "App store mining and analysis: MSR for app stores," in *2012 9th IEEE Working Conference on Mining Software Repositories (MSR)*, 2012, pp. 108–111.