Impact of free app promotion on future sales: A case study on Amazon Appstore

Harshal A. Chaudhari
Dept. of Computer Science, Boston University
harshal@cs.bu.edu

John W. Byers
Dept. of Computer Science, Boston University
byers@cs.bu.edu

ABSTRACT

Amazon's *Free App of the Day* program, aimed at improving app visibility using daily free promotions, is a compelling experiment in the 'economics of free'. In this study, we investigate the longer-term consequences of free app promotions on the performance of apps on Amazon Appstore. In particular, we quantify the causal impact of such promotions on apps' future download volumes, star ratings, and sales rank using a multi-level model. On average, apps see a surge in download volumes during such promotions, albeit accompanied by a short-term negative effect on its star ratings. On average, sales rank briefly improves but falls to pre-promotion levels within a few months. Our findings suggest that lower ranked apps are the biggest beneficiaries of these promotions, as they witness the most significant sales impact. In addition, we show the presence of a cross-market spillover effect of such promotions on the performance of the same apps on Google Playstore. Our results underscore a nuanced set of trade-offs for an app developer: do the benefits of running a promotion and boosting sales rank warrant the lost revenue and risk of lower user ratings in the long run?

1 INTRODUCTION

The introduction of mobile devices such as tablets and smartphones has changed the way in which we work, socialize, and communicate. One important component driving this revolution is the introduction of mobile apps, where an entire ecosystem, the mobile app economy, has been created and has grown to an unprecedented scale in just the past decade.

Today, there are four leading app stores: Google Playstore and Apple App Store, each with over two million apps; and more recently, the Windows Store and the Amazon Appstore, an app store for the Android operating system, each with over 600,000 apps . Such large volumes generate an intensely competitive environment for app developers, who are often competing for the attention of the same pool of customers. Thus, app developers, in collaborations with app stores or third party companies often advertise their products using both classical and more innovative marketing strategies. These strategies include price-discounted promotions; offering free lite versions of their app; and offering freemium models . However, the implications of such promotions are not clear, as each of these options run the risk of losing revenue, to customers who would have paid full price, or would have purchased the premium model, had a discounted version not been on offer. Indeed, it is easy to find blog posts or news discussing the negative effects of such promotions[1].

In our work, we examine one such promotion in detail: Amazon Appstore's *Free App of the Day*, both from the perspective of the Amazon Appstore and the app developers who participated. In this promotion program, on a daily basis, Amazon prominently displayed one new paid app from the app store for *free* download in a spot of high visibility on the store website[2]. Clearly, on the day of promotion, a participating app developer suffers short-term losses, as their app is given away for free, presumably including some customers who would have subsequently purchased the app, had the promotion not been in place. But a key selling point of this program that Amazon touts regards *long-term* improvement in sales for apps participating in this promotion. A primary mechanism that could drive future sales is that the promotion causes a significant increase in the short-term popularity of the app, which translates into improved sales rank, which in turn translates into improved placement in Amazon Appstore search results, and better future sales. The extent to which such an effect is operative would be observable within the Appstore itself. A secondary mechanism that could drive future sales is increased awareness and word-of-mouth: the increase in brand and app awareness from a promotion could have a broader secondary effect as new consumers are reached. This secondary effect, if operative, would be observable both within the Amazon Appstore, but also in other app stores. A rational app developer, whose goal is long-term revenue maximization, thus has to weigh the short-term downside against the longer-term benefits: for example, assessing whether the incremental revenue from the customers purchasing the app after the promotion and *as a consequence* of the promotion, will offset the revenue lost on the day of promotion, thereby resulting in net profits.

Turning to the perspective of the app store, Amazon Appstore's objectives behind the *Free App of the Day* program are complex, as the Appstore is a two-sided marketplace in a competitive market. From a market structure standpoint, Amazon is the number two player in the Android appstore market, in direct competition with the Google Playstore, the primary marketplace. But gaining market share against the Google Playstore necessitates becoming more attractive to both sides of the market: in this case, app developers and app purchasers. In some sense, attracting app developers is the easier side of the equation, as it is a relatively low-cost proposition

CS697 Submission, Boston

DOI: 10.1145/nnnnnnn.nnnnnnn

[1] See:https://www.developereconomics.com/freemium-apps-killing-game-developers or https://gigaom.com/2011/08/02/54805-reasons-not-to-be-amazons-free-app-of-the-day/
[2] In late 2015, Amazon shut down the *Free App of the Day* promotion and replaced it with *Amazon Underground* .

for app developers to multi-home, and market their app in multiple app stores simultaneously. With an increasing customer base and a relatively uncrowded marketplace, Amazon can exploit 'network effects' to attract high quality Android developers to not only publish their apps on Amazon Appstore, but also to use other Amazon cloud services in various functionalities of their app, thereby creating new revenue streams for Amazon. Here too, the Amazon Appstore's incentives are aligned with those of the participating app developers. However, attracting new customers away from Google Playstore is likely a more powerful incentive for Amazon to run *Free App of the Day*, as it directly increases market share and also opens up potential revenue streams for Amazon, in the form of app purchases, but also in-app purchases, advertising, and subscriptions. But building share on this side of the market may work against app developers, as doing so prioritizes short-term wins via maximizing free downloads.

Ultimately, the complex set of non-aligned objectives in a two-sided market like Amazon Appstore leaves us with several interesting questions: what are the long-term consequences of participating in deep discount promotions in the Amazon Appstore? Is Amazon's promise of increased post-promotion sales a mere marketing gimmick to convince app developers to participate in the program, or does it hold in practice? What role do various app characteristics play in determining the success of such a promotion? And last, does Amazon's promotion strategy have any cross-market effect, spilling over to other Android appstores like the Google Playstore? In this paper, we provide preliminary answers to these questions through the lens of a year-long dataset that we collected from the Amazon Appstore and the Google Playstore platforms.

Our analyses show that participation in the Amazon *Free App of the Day* program is positively associated with increased sales volumes on the Amazon Appstore. Higher sales also lead to increased customer reviews. However, they run the risk of attracting customers who review the apps more critically than those who paid full price, in the spirit of the Groupon effect [10]. We show the presence of a differential impact of promotions on different apps, based on their perceived quality, with low-ranked apps being the biggest beneficiaries of such promotions. We also provide evidence suggesting that extensive marketing campaigns by Amazon do leads to large word-of-mouth and social media engagements for the promoted apps, thereby creating observable spillover effects in other appstores. Our findings extend the understanding of the use of discounted promotions in app marketplaces. Our results will provide useful insights to app developers on how to derive maximum effectiveness of their appstore marketing campaigns.

2 RELATED WORK

Our work connects to several recent streams of research in the marketing community. One line measures and models various key aspects of the app ecosystem. Notably, Ghose and Han [19] develop a structural model to estimate consumer demand for mobile apps based by quantifying their preferences for different app characteristics. Liu *et al.* [23] study the impact of freemium strategies on sales volumes and app revenues on Google Playstore, while Cheng and Tang [13] study similar strategies in software markets.

While the abundance of choices, coupled with low transaction and switching costs apps makes it difficult to tease apart the effects of new marketing strategies like Amazon's *Free App of the Day*, the existing literature shows that customers in the app economy make adoption decisions based on two key factors: app visibility and app quality [1]. One effective way to improving visibility at low cost is by being featured in lists like *highest earning apps, top new apps, editors' choices*, etc. All the appstores, including Amazon Appstore, populate many such lists on basis of sales rank, thus making it a key metric. Guided by the earliest work of Brynjolfsson *et al.* [8] in establishing the relationship between online book sales and sales rank on Amazon.com, researchers have estimated the parameters of the relationship between downloads and sales rank on various appstores using publicly available data [18]. These relationships have been used by Chevaliear and Goolsbee [14] to analyze price elasticity and by Ghose and Sundararajan [20] to study product cannibalization. These studies offer a sound theoretical foundation for hypotheses we investigate in our research.

Another line of related work highlights the economic significance of ratings, rankings, and reviews for both online and traditional marketplaces. Luca [24] showed that a one-star increase in a restaurant's rating on Yelp results in a 5-9% increase in revenue. Researchers studying Groupon [10, 15] have shown that while daily deals websites produce a surge of new customers for retail businesses, on average, they negatively impact the reputation of those businesses, as measured through Yelp ratings. Askalidis [2] studies the impact on sales of large scale promotion on the Apple App Store and Google Play.

In addition to product visibility, product quality is an important factor during adoption decision by consumers. On the Amazon Appstore, the visibility and quality of an app are determined by their sales rank, number of reviews, and displayed user ratings. The relatively short life-cycles of apps make it difficult for app developers to build up their brands. Hence, customers usually rely on app characteristics and their ex ante awareness developed via online word-of-mouth, user ratings and reviews, while making purchase decisions. Zhu and Zhang [27] study the impact of online reviews on the sales of gaming apps, and Chang *et al.* [11] study the impact of heterogeneity in customer preferences while making purchase decisions. We employ a similar methodology to these works to ascertain the presence of a similar heterogeneity in the impact of promotion based on the consumer biases in perceived app quality.

Lastly, our study also relates to studies of spillover effects. For example, Erdem and Sun [17] empirically study the cross-category spillover effects of advertising in umbrella brands However, we know of no similar study that empirically observes cross-market spillover effects.

3 DESIGN OF EMPIRICAL STUDY

The *Amazon Appstore for Android* is a third-party appstore for the Android operating system, operated by Amazon.com. It was launched in March, 2011 and is now available in nearly 200 countries. At the time of the launch it had about 3,300 apps; the number has increased significantly since then to nearly 334,000 apps at the time of this study. Similar to Amazon.com, the appstore apps are

Table 1: Summary Statistics of Amazon Appstore.

	Treatment		Control		Overall	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Price (USD)$^{(L)}$	2.31	1.45	2.58	3.08	2.57	3.05
File Size (megabytes)$^{(L)}$	44.28	66.10	35.72	78.70	35.93	78.42
Description Length (characters)$^{(L)}$	6.79	0.78	6.81	0.79	6.81	0.79
Number of Screenshots	7.23	2.64	6.62	2.87	6.64	2.87
App Age (months)	26.61	17.06	29.12	17.07	29.06	17.08
User Review Count$^{(L)}$	5.44	1.41	2.86	1.49	2.93	1.55
User Rating	3.97	0.50	3.66	0.93	3.67	0.92
Observations	1619		62545		64164	

Note: The sample period is from February, 2015, to December, 2015.
$^{(L)}$ denotes Logarithm of the variable.

sold via two channels – website interface and a smartphone app. Amazon.com offers the same selection of apps over both its channels. Because we are unable to distinguish the app downloads over the website channel from the ones over the smartphone app, we are limited to identifying the effects of only the app characteristics that are common to both the channels.

3.1 Free App of the Day Promotion

One of the most high-profile features of the *Amazon Appstore for Android* is the *Free App of the Day*, or FAD. The primary benefit for the apps participating in the FAD promotion is a spot of very high visibility, on both the channels. Along with it, Amazon uses its marketing machinery to promote the participating apps by making Facebook posts or tweets on their official Twitter account. As these posts get picked up by various bloggers and other such platforms, the promotion is only further amplified. The benefits of the promotion continue long after the the app's time in FAD spotlight is over at the end of the day. It finds a spot in the *'Most Recent Free Apps of the Day'* shoveler on both channels.

In addition to the increased direct visibility, the app continues to get post-FAD exposure throughout the appstore due to Amazon's recommendation system. Because of the increase in app downloads typically associated with FAD, the promoted apps show up on the product details pages of other apps under the *'Customers Who Bought This Item Also Bought'* feature. An increase in app downloads also translates into a higher *'Amazon Bestsellers'* list, further improving post-FAD exposure.

3.2 App Selection for Promotion

An interesting feature of the *Free App of the Day* is that the promoted apps are selected by Amazon from proposals submitted by developers themselves. Some of the factors taken into account while evaluating proposals are the appeal of the app to wide audience, size of the app, number of downloads, plans for marketing outside the appstore, etc. [3] Robustness checks which address the resulting bias from self-selection appear in the full version [12].

[3]https://developer.amazon.com/blogs/post/Tx2CE37E42FQM8M/Submitting-Your-App-for-FAD-Consideration.html

3.3 Data Description

In this section, we provide an overview of the major datasets used in our analysis: Amazon and Google appstore data, and FAD promotion history.

3.3.1 Amazon Appstore Data. We collected app profile data from the web interface of Amazon Appstore, while relied upon a third-party Amazon price tracker website `Keepa.com` for collecting daily price and sales rank of every app, from February, 2015 till December, 2015. It includes 23,882 distinct apps from the paid apps sections of the appstore. For every app, we further collected the entire history of publicly displayed user reviews, including submission date, review text, and star-rating, constituting a total of 800,000 user reviews. Thus, our dataset includes *daily* panel data on app sales rank, price, app characteristics and user review data. We capture an exhaustive list of app-related information provided to a user while browsing through the appstore. The observed app characteristics in our sample include app file size, release date, version, textual description, in-app purchase option, number of screenshots, number of permissions, maturity level, category, developer, minimum Android version supported by the app and number of apps provided by the same app developer.

3.3.2 FAD Promotion History. In order to obtain the FAD promotion history, we relied on Amazon Appstore's official Twitter account. Amazon used this account to daily inform its followers about which app was being promoted. There were 794 promoted app over a period of almost two year, from July 2013 to August 2015. To minimize confounding effects of multiple promotions on our analysis, we remove the apps promoted in our observation period, which had already been promoted in past. Thus, for the remaining 179 distinct apps, that participated in FAD promotion exactly once, we record their date of promotion.

Combining Amazon Appstore data with FAD promotion history, we present relevant summary statistics in Table 1 with 'Treatment' apps corresponding to the ones participating in FAD promotion. More detailed summary statistics are available in [12].

3.3.3 Google Playstore Data. We used techniques from image classification and text similarity to help find likely cross-listed FAD-promoted apps on the Google Playstore. Out of the original 794 FAD promoted apps, we found 720 very high-confidence matches. Due to the sheer volume of apps on the Google Playstore, Google does not maintain a sales rank across the entire appstore, only choosing to do so at a subcategory level. From publicly available data on `AppAnnie.com` and `AppBrain.com`, we collected sales rank history of 566 of the FAD promoted apps, across 52 different subcategories. In addition, we also collected a total of 480,000 publicly available user reviews and meta-data for these apps.

3.4 Hypotheses

Existing literature in economics and marketing science predicts that the consumers use external information to supplement their ex ante awareness of products while making purchasing decisions [16, 21]. Amazon's FAD promotion, lowers the cost incurred by the consumers when searching for external information regarding promoted apps, and is thereby expected to affect sales patterns [9]. In this section, we will formulate our hypotheses on how the Amazon's FAD promotion can lead to changes in apps' sales rank and user ratings patterns.

3.4.1 Impact within Amazon Appstore: Amazon's FAD promotion is a unique kind of recommendation tool, that not only provides the product for free, but also decreases the search costs drastically by providing 'directed' links, that take consumers directly to the product pages on the appstore. We hypothesize that such promotions may lead to sales trends with exceptionally high weights for the promoted apps. At the same time, in the spirit of the 'Groupon effect' studied by Byers *et al.* [10], we hypothesize that FAD promotion runs the risk of attracting consumers who review the promoted apps more negatively than those who purchase the same apps at full price. Hence, impacts of FAD promotion on longer-term sales and ratings is the central research question studied in this paper.

Now, while describing the various variables from our data, we provide brief theoretical explanation of how they play an important role in the analysis of our hypothesis. The file size of the apps tends to increase with increasing sophistication and utility, leader to larger download times. Users can incur higher data costs for downloading such apps. This is not only one of the factors influencing Amazon's choice of FAD promoted apps, but may also impact the number of downloads during promotions, even if the users do not have to pay for the app itself. We use the app release date to track the age of an app. As the app gets older, its sales rank tends to increase while average user rating decreases. Furthermore, Amazon prefers to promote relatively newer apps. To maximize the usable variation in different time-variant variables in our dataset, we aggregate them at a monthly frequency, with respect to app age in months. App developers periodically release new versions of their apps to introduce new features or in response to user feedback. Thus, the number of versions of an app is likely to be an indicator of its quality and functional maturity, both of which affect app demand and user ratings[4]. Detailed analysis explaining rationale behind including each variable is provided in the full version of the paper [12].

[4]User reviews/ratings on the newer versions of the app may cause an app developer to change certain app characteristics, thereby endogenously determining some of the observed app characteristics. As Amazon does not provide a history of app version updates for every app, we cannot control for these particular unobserved characteristics, although they may be correlated with the observed characteristics.

3.4.2 Cross-Market Spillover: In the app economy, consumers are more likely to be aware of popular apps across different appstores, than niche apps specific to their primary appstores i.e., Apple iTunes or Google Playstore. This is because, social media and technology bloggers serve as external sources of information, and play a key role in setting trends for popular apps, as evident in the example of Pokemon Go, a virtual reality based smartphone game [3]. We believe that FAD promotion can improve ex ante awareness for promoted apps among the users of Google Playstore via advertising and word-of-mouth referrals. In fact, Amazon's marketing strategy ensures that consumers can actively perform specific searches on Google Playstore using exact app names. These types of specific searches, or "directed searches" [25], take consumers directly to the app profile page, helping them quickly locate it. As a result, we hypothesize the presence of a cross-market spillover effect of FAD promotion on the Google Playstore, on the sales rank of promoted apps. On the other hand, while installing a FAD promoted app, the consumers on the Google Playstore make full priced purchases, and are unlikely to be 'experimenting' like their Amazon counterparts. Hence, we do not expect these consumers to leave overly critical user reviews for their purchases. Consequently, we do not expect presence of a strong cross-market spillover effect on user ratings of the promoted apps.

4 ECONOMETRIC MODEL

In this section, we specify the 'within-between' formulation of the multilevel models [5] to estimate the *causal* impact of the FAD promotion on the sales and user ratings' patterns of the promoted apps. In this study, we create a longitudinal dataset by tracking sales rank and review history of many apps over several months. Hence, our study has a hierarchical structure – repeated measurements at level-1, nested individual apps at level-2, which are further nested into separate categories at level-2. An alternative is a Fixed Effects (FE) approach, but this limits the effects that can potentially be studied, as they introduce dummy variables corresponding to higher levels of measurement in the hierarchical structure. In the full version of the paper, we show that the primary effects we identify in the Multilevel case are similarly present (with nearly identical magnitudes) in FE models.

4.1 Model Specification

Because the apps participating in the FAD promotion (treatment apps) were promoted on different days in our observation period, we have a multiplicity of "experiments" to exploit. Our empirical approach relies on contrasting the change in sales rank and user ratings of the treatment apps in a given period with those that did not get promoted in the same period (control apps).

We adopt an *individual growth model* or level-1 submodel that incorporates the linear change in sales rank with respect to age of an app. Following the within-between model from Bell and Jones [5], we also introduce the app-level mean and the centering term for age, a time-varying covariate, to separate the 'within' and 'between' effects of the variable, necessary for causal interpretation.

FAD promotion, lasting for exactly a day, acts as an intervention for the treatment apps, and introduces an abrupt discontinuity in the trajectory of app's sales rank (or user rating) over time. To postulate a post-promotion change in trajectory, we include a time-varying predictor, After_{ij} in the level-1 submodel that specifies whether and, if so, when each app experiences the discontinuity. Before an app i is promoted, $\text{After}_{ij} = 0$, and if and when, it gets promoted, After_{ij} becomes 1. We stipulate that level-1 residuals are drawn from an underlying normal distribution, $\epsilon_{ij} \sim \mathcal{N}(0, \sigma_\epsilon^2)$.

$$\text{Sales Rank}_{ij}^{(L)} = \pi_{0i} + \pi_{1i}\text{Age}_{ij} + \pi_{2i}\text{After}_{ij} + \epsilon_{ij} \tag{1}$$

where Age_{ij} is a (series of) time-variant value for app i. Because After_{ij} distinguishes the pre- and post-promotion epochs for app i, the growth parameter π_{2i} captures the magnitude of the instantaneous impact of promotion by permitting a discontinuity in the *intercept* of the trajectory. To create a post-promotion trajectory, that differs in *slope*, we include another predictor Post_{ij} which clocks age of an app from the day of its promotion. Before an app i is promoted, Post_{ij} is 0. On the day the app is promoted, Post_{ij} remains at 0. However, after that, its values begin to increase in concert with the primary temporal predictor, Age_{ij}. It is worth noting that because timing of promotion is app-specific, the cadence of Post_{ij} is also app-specific. Finally, we model the curvilinear change in trajectory post-promotion by adding Post_{ij}^2. The level-1 submodel become,

$$\text{Sales Rank}_{ij}^{(L)} = \pi_{0i}+\pi_{1i}\text{Age}_{ij}+\pi_{2i}\text{After}_{ij}+\pi_{3i}\text{Post}_{ij}+\pi_{4i}\text{Post}_{ij}^2+\epsilon_{ij} \tag{2}$$

While the level-1 submodel describes how each app changes over observational period, the level-2 submodel we now define describes how those changes differ across apps [7, 26]. To do so, we introduce app-level means of the time-variant variables while modeling the *intercept* term. If we let X_i be a vector representing the time-invariant app-specific characteristics, then we can simply include them in the level-2 submodel without the risk of introducing collinearity:

$$\pi_{0i} = \gamma_{00}+\gamma_{01}\overline{\text{Age}}_i+\gamma_{02}\overline{\text{After}}_i+\gamma_{03}\overline{\text{Post}}_i+\gamma_{04}\overline{\text{Post}^2}_i+\alpha\text{X}_i+\zeta_{0i} \tag{3}$$

where $\overline{\text{Age}}_i$ and $\overline{\text{After}}_i$ are the app-level means; as such, the time-invariant component of Age_{ij} and After_{ij} respectively [5]. After combining both the levels of the multi-level model, and some algebraic simplification, we can express a *composite* model as follows,

$$\begin{aligned}\text{Sales Rank}_{ij}^{(L)} \quad = \quad & \gamma_{00} + \pi_{1i}(\text{Age}_{ij} - \overline{\text{Age}}_i) + \pi_{2i}(\text{After}_{ij} - \overline{\text{After}}_i) \\ & +\pi_{3i}(\text{Post}_{ij} - \overline{\text{Post}}_i) + \pi_{4i}(\text{Post}_{ij}^2 - \overline{\text{Post}^2}_i) \\ & +\pi_5\overline{\text{Age}}_i + \pi_6\overline{\text{After}}_i + \pi_7\overline{\text{Post}}_i + \pi_8\overline{\text{Post}^2}_i \\ & +\alpha\text{X}_i + (\epsilon_{ij} + \zeta_{0i}) \end{aligned} \tag{4}$$

where $\pi_5 = \gamma_{01}-\pi_{1i}, \pi_6 = \gamma_{02}-\pi_{2i}, \pi_7 = \gamma_{03}-\pi_{3i}$ and $\pi_8 = \gamma_{04}-\pi_{4i}$ respectively. Residuals at both levels are assumed to be Normally distributed: $\epsilon_{ij} \sim \mathcal{N}(0, \sigma_\epsilon^2)$ and $\zeta_{0i} \sim \mathcal{N}(0, \sigma_0^2)$. Heteroscedasticity at the level-1 is explicitly modeled by including additional level-2 submodel.

$$\pi_{1i} = \gamma_{10} + \zeta_{1i} \tag{5}$$

[5]This app-level centering is different from centering around the grand mean, which has a different purpose: to keep the value of the intercept of model within the range of the data and to aid convergence. Although, by definition, the app-level mean centering ensures grand mean centering as well.

The residuals part of the *composite* model now becomes $(\epsilon_{ij} + \zeta_{0i} + \zeta_{1i} \times \text{Age}_{ij} - \overline{\text{Age}}_i)$. This reveals two important properties about level-1 residuals: they can be both *autocorrelated* and *heteroscedastic* within-app. Like level-1 residuals, we make an assumption that level-2 residuals have an underlying bivariate normal distribution.

Now, π_{2i} is the 'within' effect and π_6 is the 'between' effect of the FAD promotion [4, 22]. We further extend it by modeling π_{2i} on level-2 to include app-specific time-invariant characteristics. This enables us to model and quantify the effect (if any) of app-specific characteristics on the impact of FAD promotion.

$$\pi_{2i} = \gamma_{20} + \beta\text{X}_i \tag{6}$$

We sequentially introduce and compare estimated fixed effects and variance components to identify which predictors explain most variation. Similarly, we drop variance and covariance terms when the null hypothesis cannot be rejected, for example, in Equation 6.

Because the sales rank and user ratings data is observed at a high frequency, serial correlation is a major concern. Following the recommendations of Bertrand *et al.* [6], throughout our analysis, we compute standard errors using the generalized Huber-White formula clustered at app level. This allows for arbitrary error correlations among the daily sales rank or user ratings observations.

4.2 Heterogeneous Impact of Promotion

Consumers downloading apps from Amazon Appstore do not make purchase decisions solely based on the price and the app characteristics which are fixed by the app developers; but they may use pre-existing biases or develop some regarding the quality of the apps based on the user reviews and sales ranks of the apps before the day of promotion. For example, we believe that two apps which offer the same core functionality may well experience very different impacts of FAD promotion, if their sales ranks and user ratings are different. Therefore, to check for such a heterogeneity of impact of promotion, we adopt a very conservative definition of "app quality" by segregating the promoted apps into three rank categories based on their average sales rank through our observation period. Apps with average sales rank in the rank 1 to 1984 form Rank Category 1, 1985 to 4573 form Rank Category 2 and the rest form Rank Category 3. To each of these rank categories, we add the control apps whose average sales ranks lie within the category boundaries. Finally, we introduce these rank categories as time-invariant variables in the model from Equation 4 in order to study the heterogeneity of FAD impact.

4.3 Cross-Market Spillover

As we do not have access to the data for control group of apps on the Google Playstore, we cannot specify a model that provides causal inference regarding cross-market spillover effect of FAD promotion. Hence, our model aims at identifying the correlational effects of FAD promotion and sales rank trends on the Google Playstore. Unlike the impact of FAD promotion on Amazon Appstore, we do not expect a longer duration impact on the Google Playstore. Hence, to maximize the usable variation, we code various variables at weekly frequency. Because each app is promoted on a different day on the Amazon Appstore, we create a categorical variable, Interval_{ij} that measures the offset in weeks from the day of promotion, for

Table 2: Impact of FAD promotion

	(1) Sales Rank$^{(L)}$		(2) Monthly Review Count$^{(L)}$		(3) User Rating	
Mean Effects:						
After	-0.282^{***}	(-3.74)	2.931^{***}	(33.22)	-0.160^{***}	(-5.71)
Age	0.027^{***}	(39.64)	-0.048^{***}	(-66.87)	-0.008^{***}	(-14.36)
Post	0.226^{***}	(7.83)	-1.088^{***}	(-35.96)	-0.010^{***}	(-2.76)
Post^2	-0.020^{***}	(-4.72)	0.113^{***}	(24.51)		
Between Effects:						
Age*(between)*	0.008^{***}	(10.76)	-0.005^{***}	(-15.80)	-0.005^{***}	(-6.06)
After*(between)*	-20.001	(-1.01)	24.181^{**}	(2.26)	-1.559	(-1.45)
Post*(between)*	14.326	(0.90)	-16.930^{**}	(-2.00)	0.541	(1.50)
Post^2*(between)*	-1.849	(-0.86)	2.216^{**}	(1.96)		
Interaction Effects:						
After × Size$^{(L)}$	0.142^{***}	(2.94)	-0.100^{*}	(-1.73)	-0.023	(-0.64)
After × Number of Screenshots$^{(L)}$	-0.327^{**}	(-2.17)	0.037	(0.17)	-0.040	(-0.37)
After × Description Length$^{(L)}$	-0.196^{***}	(-2.69)	0.066	(0.73)	0.024	(0.52)
Variance Components:						
var(Age)	0.001^{***}	(0.00)	0.000^{***}	(0.00)	0.002^{***}	(0.00)
var(Constant)	1.836^{***}	(0.06)	0.427^{***}	(0.02)	1.399^{***}	(0.04)
corr(Age, Constant)	-0.035^{***}	(0.00)	-0.008^{***}	(0.00)	-0.030^{***}	(0.00)
var(Residual)	0.086^{***}	(0.00)	0.154^{***}	(0.00)	0.026^{***}	(0.00)
Observations	64164		64164		63735	
AIC	75444		82666		24077	
BIC	75753		82974		24367	
Pseudo Log Likelihood	-37688		-41299		-12006	

Note: Referent level for maturity rating is 'All Ages'. $^{(L)}$ denotes Logarithm of the variable. Cluster-robust t-statistics (at app level) are shown in parentheses. Truncated version of the table due to space constraints.
Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

app i. This allows us to model the spillover effect as follows,

$$\begin{aligned}\text{Sales Rank}_{ij}^{(L)} \quad = \quad & \beta_0 + \beta_1 \text{Interval}_{ij} + \beta_2 \text{AppId}_i + \beta_3 \text{AppCategory}_i \\ & + \beta_4 \text{Time}_j + \beta_5 \text{AppId}_i \times \text{Time}_j + \epsilon_{ij} \qquad (7)\end{aligned}$$

where we have included fixed effects for each app, each week, as well as an interaction between the two. Similar to the previous model, we compute standard errors using generalized Huber-White formula, clustered at app level.

5 EMPIRICAL RESULTS

Our empirical analysis focuses on four main questions:

- What is the impact of FAD promotion on the sales rank, number of reviews and user ratings of the promoted apps on Amazon Appstore?
- Do the time-invariant app characteristics of the promoted apps have an effect on the potential impact of FAD promotion?
- Is there a heterogeneity in the impact of FAD promotion based on 'quality' of the promoted app?
- Is there a cross-market spillover effect of FAD promotion on Google Playstore?

5.1 Impact of FAD promotion

In this section, we provide a formal analysis of fitting the model from Equation 4, with estimates reported in Table 2. The co-efficient of the After variable quantifies the immediate impact, experienced during the day of promotion, while co-efficients of Post and Post^2 variables help us better understand the longer term impact by describing the shape of the post-promotion trajectory of the dependent variable. We find that FAD promotion causes a 25% improvement immediately. However, post promotion, the sales rank starts falling at a significantly faster rate than it would have in the absence of any promotion. Comparing the co-efficients of Post and Post^2 variables, we observe that it takes around 3-4 months for the rate of fall of sales rank to stabilize to its pre-promotion rate.

To estimate whether the improvement in sales for a small period after the promotion is enough to offset the losses sustained due to free give-aways of the app during promotion, it is important to know the exact parameters of the Pareto distribution relationship between sales rank and actual sales volume. Estimating these parameters is beyond the scope of this study due to unavailability of actual sales data. Our analysis provides a framework to easily evaluate the net developer revenue, given these parameters. However, it should be noted that a net negative developer revenue does not always mean that a developer would suffer losses. Without conducting a counter-factual experiment, it is likely that we would overestimate the revenue lost at the time of promotion as it is impossible to know how many customers who downloaded the app for free would have otherwise purchased the app at its full price and contributed to the lost revenue.

Consistent with the sales rank trend, we observe that FAD promotion causes an abrupt 18-fold increase in the number of monthly reviews. Similar to sales rank, the number of monthly reviews keeps decreasing until after 4-5 months, at which point they stabilize to the pre-promotion values. However, consistent with our hypothesis, we find that the increased downloads in the month of promotion and subsequent abrupt increase in user reviews is achieved at a cost of a significant decrease in the average star rating. FAD promotion causes an abrupt decrease of 0.16 stars immediately after promotion, and increases the overall rate of decline of star ratings by up to 0.01 stars more every month.

We offer two potential explanations for the decline of star ratings: this could be because the users who download apps during FAD promotion are more likely to be experimenting with new apps. Such users may install an app simply because it is free, notwithstanding their actual needs, and review the app with low rating due to the app's perceived inability to impress them. An alternative explanation is offered via anecdotal evidence[6]. In case of apps that provide services via cloud infrastructure, the overwhelming increase in app usage during promotion may lead to poor quality of service due to inadequate resources, resulting into dissatisfied users who leave critical reviews with low star ratings.

Observing the co-efficients of the interaction terms, we find that some of the app-specific time-invariant characteristics affect the effectiveness of the FAD promotion. A 10% increase in app size results in 1.42% fall in the post-promotion immediate sales rank and a 1% fall in the number of monthly reviews in the month of promotion. One extra screenshot in the app profile page improves the sales rank immediately after FAD promotion by 4.5%. Similarly, a 10% increase the length of textual description also improve the effectiveness of FAD promotion by up to 2%. While the price of the app or presence of in-app purchase options within the app significantly affect the general trends on Amazon Appstore, they do not seem to affect the effectiveness of FAD promotion.

While our results provide insights into the impact of FAD promotion, they do not provide a conclusive answer whether an app developer should participate in a FAD promotion. Nevertheless, our analysis reveals that FAD promotion positively impacts sales ranks and the volume of reviews of the promoted apps comes with a cost of significant decline in star rating, underlining a nuanced set of trade-offs for the app developers.

5.2 Heterogeneous impact of FAD promotion

For brevity, we present only a visual summary of the heterogeneity in the impact of FAD promotion across rank categories in Figure 1, deferring analytical results to the full version [12]. In this figure, we plot the means of estimated dependent variables for each of the rank categories at 30-day intervals for five months prior through six months after the FAD promotion. The dashed lines in each figure represent robust 95% confidence intervals for each point estimate. The first row of Figure 1 shows that apps belonging to all rank

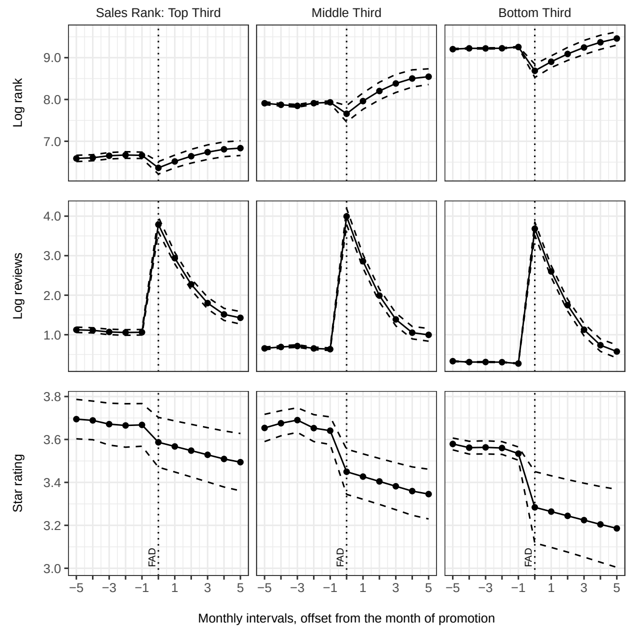

Figure 1: Heterogeneous impact of FAD promotion

categories experience an improvement in their sales rank during promotion. Moreover, the bottom third of promoted apps are the biggest beneficiaries of the promotion – they recoup substantial benefits for a longer duration as compared to the top third, which receive a minimal benefit for a shorter duration. The second row, depicting impact on the number of reviews further reinforces this result. Bottom third apps experience a larger increase in the number of monthly reviews as compared to the upper two thirds. However, they also (problematically) experience the most significant decline in star ratings, as evident in the third row of Figure 1.

From the perspective of a developer, it may be surprising that low ranked apps appear to be the biggest beneficiaries of promotion in terms of downloads. We speculate that the popular top apps have already captured market share and thus have less potential to attract new customers than relatively unknown low ranked apps. At the same time, low ranked apps are of poorer quality on average, and their exposure to a wider audience via promotion leads to many more critical reviews, and the subsequent steep decline in star ratings.

Analysis of the heterogeneous nature of the impact of FAD promotion allows us to look at the question of participation in FAD promotion in a more nuanced manner. It is evident that top app developers experience an increase in the number of downloads with close to no decline in star ratings, while the developers of the bottom ranked apps, notwithstanding the decline in star ratings, benefit more in short-term profits. For developers belonging to either of these categories, the improvement in sales through FAD promotion may outweigh the risk of long-term damage to the app reputation. However, the developers of the apps in middle category face a difficult tradeoff as to whether to prioritize short-term profits or long-term app reputation.

[6] https://blog.shiftyjelly.com/2011/08/02/amazon-app-store-rotten-to-the-core/

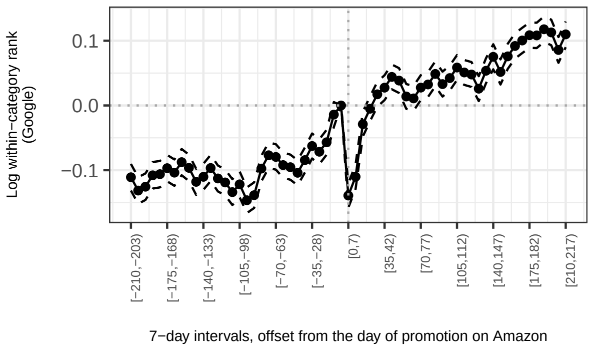

Figure 2: Cross-market spillover effect on sales rank

5.3 Cross-Market Spillover

We now consider the impacts of spillovers from FAD promotion on the Amazon Appstore to other appstores, namely Google. As described in Section 3.3.3, due to a large volume of apps, Google does not publicly display a uniform sales rank for every app across the entire appstore, choosing to do so only at the level of categories. Unfortunately, this limits our ability to quantify the magnitude of the cross-market spillover, as we cannot normalize the effect of FAD promotion across different subcategories without detailed information regarding app downloads for all apps. For brevity, we provide a visual summary of the cross-market spillover effect of FAD promotion on Google Playstore in Figure 2, obtained by fitting the model from Equation 7. We plot the estimates (β_1) for the categorical variable Interval which represents the offset in weeks from the day of promotion of an app on the Amazon Appstore along with a 95% confidence interval. We see evidence of an improvement in the categorical sales rank in the week of promotion, supporting the hypothesis of a cross-market spillover effect. The effects seems to last for a few weeks after the FAD promotion. One should not make strong inferences from this figure, however, for the reasons described above. Moreover, due to absence of control apps in the dataset, we cannot demonstrate a causal relationship between the FAD promotion and the observed effect. Nevertheless, we believe that the presence of such a striking trajectory of sales rank for different apps that are promoted on Amazon Appstore, exactly in the week of promotion, is likely a strong indicator of cross-market spillover and warrants further examination. This cross-market spillover effect is also supported anecdotally, e.g., the statistics provided by the developer *Tasharen Entertainment* in their blogpost detailing their experience during FAD promotion[7].

Although, we do not provide *causal* evidence supporting the cross-market spillover effect, our analysis and anecdotal evidence strongly supports the presence of such an effect. A plausible explanation of this spillover effect is that Amazon's aggressive marketing of the promoted apps is an attempt to attract new users to Amazon Appstore. However, after the end of FAD promotion, users who become aware of the promoted app perform 'directed searches' of the app names on their primary appstore i.e., Google Playstore to download the app, instead of downloading Amazon Appstore app and then purchasing the app over it. In future work, we propose to explore whether spillover effects on Google Playstore mirror those on Amazon (i.e., improved sales rank but lower ratings).

[7] http://www.tasharen.com/?p=4664

6 CONCLUSIONS

Appstores, like most traditional and online market platforms, are dominated by a few best-selling apps, with the long tail of other apps competing for visibility and attention of customers. However, in case of appstores, the absence of operational costs associated with inventory management, and the relative ease of running large scale online advertising campaigns has given rise to innovative marketing strategies. In this paper, we have examined a number of hypotheses to analyze the impacts of deep discounted promotions in the app economy. While there remain challenges to building a predictive model that could quantify the expected costs and benefits of such promotions, our modeling framework and empirical measurements provide a significant first step.

Our empirical results presented in Table 2 highlight that on average, all apps promoted in the Amazon *Free App of the Day* program experience a significant immediate improvement in the sales on account of improved visibility. However, the long-term effects of such a promotion strategy depend on the quality of the app. The improvement in the post-promotion sales volumes may not be sustained long enough to offset the lost revenue on the day of promotion, especially for the top apps. App developers should be cognizant that promotions lead to an abrupt increase in engagement of the users in form of reviews (both positive and negative), and on an average, cause negative impact on the reputation of app.

For appstores, long-term success depends on the satisfaction of both customers as well as app developers. There needs to be a complex trade-off between providing users with quality apps at low price, while at the same time mitigating potential losses to app developers' reputation and profits. Existing incentives to provide higher app visibility is most attractive to lower-ranked apps. Such practices may yield gains in market share for a short-term, but inhibit long-term customer retention.

Last, but not least, we find that increased app visibility on account of promotion increases brand awareness due to social media and word-of-mouth engagements. This effect not only drives future sales on the primary appstore i.e., Amazon Appstore, but also spills over across the markets onto other appstores like Google Playstore. This adds an additional complexity in measuring the *true* impact of promotions on the revenues of app developers. Indeed, a rational app developer should weigh the incremental revenue from across different appstores against short-term reputation damages, while assessing the merits of promotion.

Our findings contribute both to the academic literature and give guidance to practitioners in the mobile app marketplace. Our study contributes to the growing body of research that utilizes publicly available e-commerce data to empirically validate research questions. It extends the existing knowledge about promotion strategies in the app marketplace, while also validating existing theories about consumer behaviors during discounted promotions. Finally, for app developers, our study provides insight on influential factors and determinants of successful marketing strategies.

REFERENCES

[1] Chris Anderson. 2006. *The long tail: Why the future of business is selling less of more.* Hachette Books.

[2] Georgios Askalidis. 2015. The Impact of Large Scale Promotions on the Sales and Ratings of Mobile Apps: Evidence from Apple's App Store. *arXiv preprint arXiv:1506.06857* (2015).

[3] Stuart Barnes. 2016. Understanding Virtual Reality in Marketing: Nature, Implications and Potential. (2016).

[4] Brandon Bartels. 2008. Beyond" fixed versus random effects": a framework for improving substantive and statistical analysis of panel, time-series cross-sectional, and multilevel data. *The Society for Political Methodology* (2008), 1–43.

[5] Andrew Bell and Kelvyn Jones. 2015. Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods* 3, 01 (2015), 133–153.

[6] Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan. 2004. How much should we trust differences-in-differences estimates? *The Quarterly journal of economics* 119, 1 (2004), 249–275.

[7] Anthony S Bryk and Stephen W Raudenbush. 1987. Application of hierarchical linear models to assessing change. *Psychological Bulletin* 101, 1 (1987), 147.

[8] Erik Brynjolfsson, Yu Hu, and Michael D Smith. 2003. Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science* 49, 11 (2003), 1580–1596.

[9] Erik Brynjolfsson, Yu Jeffrey Hu, and Michael D Smith. 2006. From niches to riches: Anatomy of the long tail. (2006).

[10] John W Byers, Michael Mitzenmacher, and Georgios Zervas. 2012. The Groupon effect on Yelp ratings: a root cause analysis. In *Proceedings of the 13th ACM Conference on Electronic Commerce.* ACM, 248–265.

[11] Kwangpil Chang, Sivaramakrishnan Siddarth, and Charles B Weinberg. 1999. The impact of heterogeneity in purchase timing and price responsiveness on estimates of sticker shock effects. *Marketing Science* 18, 2 (1999), 178–192.

[12] Harshal A Chaudhari and John W Byers. 2017. Impact of Free App Promotion on Future Sales: A Case Study on Amazon Appstore. (2017). Full version of this submission. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3078067.

[13] Hsing Kenneth Cheng and Qian Candy Tang. 2010. Free trial or no free trial: Optimal software product design with network effects. *European Journal of Operational Research* 205, 2 (2010), 437–447.

[14] Judith Chevalier and Austan Goolsbee. 2003. Measuring prices and price competition online: Amazon. com and BarnesandNoble. com. *Quantitative marketing and Economics* 1, 2 (2003), 203–222.

[15] Benjamin Edelman, Sonia Jaffe, and Scott Duke Kominers. 2016. To Groupon or not to Groupon: The profitability of deep discounts. *Marketing Letters* 27, 1 (2016), 39–53.

[16] James F Engel, Roger D Blackwell, and Paul W Miniard. 1995. Consumer Behavior, 8th. *New York: Dryder* (1995).

[17] Tülin Erdem and Baohong Sun. 2002. An empirical investigation of the spillover effects of advertising and sales promotions in umbrella branding. *Journal of Marketing Research* 39, 4 (2002), 408–420.

[18] Rajiv Garg and Rahul Telang. 2012. Inferring app demand from publicly available data. *MIS Quarterly* (2012).

[19] Anindya Ghose and Sang Pil Han. 2014. Estimating demand for mobile applications in the new economy. *Management Science* 60, 6 (2014), 1470–1488.

[20] Anindya Ghose and Arun Sundararajan. 2006. Evaluating pricing strategy using e-commerce data: Evidence and estimation challenges. *Statist. Sci.* (2006), 131–142.

[21] Philip Kotler and Kevin Lane Keller. 2006. Marketing Management 12e. *New Jersey* (2006).

[22] Alastair H Leyland. 2010. No quick fix: understanding the difference between fixed and random effect models. *Journal of Epidemiology & Community Health* 64, 12 (2010), 1027–1028.

[23] Charles Z Liu, Yoris A Au, and Hoon Seok Choi. 2012. An empirical study of the freemium strategy for mobile apps: Evidence from the Google Play Market. In *Proceedings of 33rd International Conference on Information Systems.*

[24] Michael Luca. 2011. Reviews, reputation, and revenue: The case of Yelp. com. *Com (September 16, 2011). Harvard Business School NOM Unit Working Paper* 12-016 (2011).

[25] Wendy W Moe. 2003. Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. *Journal of consumer psychology* 13, 1-2 (2003), 29–39.

[26] David R Rogosa and John B Willett. 1985. Understanding correlates of change by modeling individual differences in growth. *Psychometrika* 50, 2 (1985), 203–228.

[27] Feng Zhu and Xiaoquan Zhang. 2010. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing* 74, 2 (2010), 133–148.