Electronic Commerce Research and Applications 17 (2016) 173–190

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/ecra

# Revenue models, in-app purchase, and the app performance: Evidence from Apple's App Store and Google Play

Paolo Roma *, Daniele Ragaglia

*DICGIM – Università degli Studi di Palermo, Viale delle Scienze, Palermo 90128, Italy*

ARTICLE INFO

*Article history:*
Received 19 May 2015
Received in revised form 15 March 2016
Accepted 16 April 2016
Available online 18 April 2016

*Keywords:*
Mobile app market
Online distribution
Revenue model
E-commerce
Empirical analysis

ABSTRACT

In this paper, we empirically examine how the revenue model (paid, free, or freemium) adopted for a given app affects the app revenue performance as measured by the app daily revenue rank. We also study the impact of in-app purchase on this measure of performance. Moreover, we study how such relationships are contingent upon the distribution platform where the app is marketed as well as the type of category to which the app belongs. We test our hypotheses relying on a large sample of top grossing apps from the two major app stores, namely Apple's App Store and Google Play. Our findings reveal that in the Apple's App Store, paid and freemium models are equivalent and both are more effective than the free model in terms of app revenue performance. On the other hand, in Google Play no significant differences between paid and free revenue models emerge, whereas the freemium model is shown to be less effective even than the free model. Moreover, while in-app purchase is shown to positively influence the app revenue performance in Apple's App Store, this effect is reversed in Google Play. Finally, the type of category is also shown to influence the effects of the revenue model and in-app purchase (the latter to a lesser extent) on the app performance.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Software developed for mobile phones has been around for well over a decade. Indeed, in the early 2000s, before the term 'application store' was popularized, the distribution of mobile content was dominated by the Mobile Portal business model. Mobile Portals were mostly managed by Mobile Network Operators (MNOs), which constructed a highly centralized model (Kuo and Yu 2006). However, the mobile content ecosystem was at first unclear and did not attract enough developers and users to really soar, possibly due to the fact that the technology was not mature. This relatively stable context was dramatically shaken in 2008 by Apple, which, by launching the app store, introduced a new distribution paradigm in mobile commerce: the application store. An application store is essentially an online distribution platform from which users can download software applications for mobile devices to increase the utility associated to their usage.

Mobile applications (apps, hereafter) are typically developed by third parties, which can be either software houses or individuals. App stores operate as online two-sided markets, generating mutual advantages to all involved actors (Hagiu 2007). By means of developers, the platform owner, e.g., Apple, can exploit indirect network externalities that increase the value of its own devices and/or operating systems (OS). In fact, the higher the number of apps running on a device, the higher the potential functionalities of such device. On the other hand, developers are interested in distributing their apps via app stores, because this allows them to reach a multitude of consumers worldwide that they might not be able to reach on their own. Finally, consumers derive higher utility from the presence of a higher number of developers in the app stores as they have larger product variety available for purchase. The revenue-sharing rule adopted by Apple (and also by all major stores) implies that, for each transaction, 70% of the revenue goes to the developer, while the platform owner retains 30% of it.[1]

Nowadays a multitude of applications have become everyday "tools" in the life of people who want to be connected 24/7. As a matter of fact, according to Venturebeat, the app market has shown an astonishing growth stepping from less than $10 billion

* Corresponding author.
*E-mail addresses:* paolo.roma@unipa.it (P. Roma), daniele.ragaglia@unipa.it (D. Ragaglia).

[1] Note that the revenue sharing rule relates to how the revenue generated from each transaction involving apps is allocated between developers (70%) and app store owner (30%). This is different from the revenue models that developers can adopt to monetize from their apps. The present study focuses on the latter.

http://dx.doi.org/10.1016/j.elerap.2016.04.007
1567-4223/© 2016 Elsevier B.V. All rights reserved.

annual revenue in 2011 up to estimated $70 billion by 2017 (Takahashi 2014). Other sources provide even rosier predictions suggesting that the global mobile market will reach $150 billion by the end of 2015 (Ghose and Han 2014). To take advantage of the great potential of this market, in the last few years, numerous mobile device makers, such as RIM, Samsung, etc, have followed Apple's move. Such a rapid proliferation of app stores has involved not only traditional players of the smartphone industry, but also important new entrants such as Google, which launched its Android mobile Operating Systems (OS) and made the Android Market (later rebranded as Google Play) available to app users in 2008. However, as a result of great market consolidation, nowadays the scene is dominated by only two players: Apple's App Store and Google Play, which have become by far the most popular stores, retaining together almost 90% of the market (Gartner 2013). In particular, although Apple's App Store still generates higher revenue, Google Play delivers 60% more app downloads than Apple (Wallenstein 2015). These exemplificative figures as well as the online nature of the transactions in the app market demonstrate the relevance of this market in the context of Electronic Commerce (EC).

In this highly dynamic and competitive environment app developers are required to make several important decisions. For instance, they have to choose the type and the number of apps to market, which mobile operating systems to develop for and, hence, which app store to target, and finally which revenue model to choose for each app. These decisions are naturally pivotal for developers as they determine the success or the failure of apps (and their developers) in the market. However, in contrast to the huge popularity, the academic world has started devoting attention to the dynamics of success in this market only recently. For instance, a few initial empirical studies have focused on app demand estimation (Carare 2012; Garg and Telang 2013; Ghose and Han 2014) or purchase intention of paid apps (Hsu and Lin 2015). In addition, focusing on a case study, Halaburda et al. (2011) have examined the rationale behind the choice of the platform to develop for. However, no empirical studies have instead investigated the impact of the revenue model chosen for an app on its performance in the market. It is extremely important to shed light on this issue for two main reasons. From a practical viewpoint, the choice of the revenue model is perceived as one of the most difficult and crucially important decisions by app developers in order to succeed in the market, as demonstrated by the numerous industry articles focusing on this issue (Sourcebits 2014; Pappas 2013; Wilcox 2013; Munir 2014). As a matter of fact, a multitude of app developers struggle with the choice of the appropriate revenue model for their apps as the mobile app market has become very competitive and many developers suffer from low sales (Sourcebits 2014). From a literature viewpoint, our study is important because, although the performance implications of revenue models in the context of information goods and/or EC have been extensively discussed in the literature, still no univocal indications can be drawn. In fact, in this respect, contrasting views have emerged with some early studies supporting the economics of free with advertising (Anderson 2009), and other studies documenting the recent trends toward the use of models that require a payment to all or some users (i.e., paid or freemium) for online content provision (Pauwels and Weiss 2008). Therefore, further research is needed to better understand the effectiveness of different revenue models, especially with regard to a business setting such as the app market, where the choice of the revenue model and the relative economic consequences have never been investigated in previous literature in spite of being so pivotal for the fate of an app and its developer. The empirical investigation developed in this study contributes to fill this gap. Specifically, it aims at shedding light on:

1. The impact of existing app revenue models (namely, free, paid, freemium) on the app revenue performance, as measured by the app daily revenue rank, and whether (and how) this impact depends on the store (Apple's App Store vs. Google Play) where the app is marketed and the nature of the app itself (i.e., the category to which the app belongs).
2. The revenue performance implications of adopting the practice of in-app purchase (i.e., purchase of additional app features directly inside the app at incremental prices), and whether (and how) these revenue implications depend on the store where the app is marketed and the nature of the app itself.

With regard to the first point, in line with previous studies (e.g., Pauwels and Weiss, 2008) and motivated by the large popularity of free apps, we characterize our empirical study as a comparison of payment-based (paid and freemium) versus free models. That is, we formulate our hypotheses mainly as a comparison between payment-based (paid and freemium) versus free models to be consistent with prior literature. However, our analysis also takes into account the comparison between paid and freemium models (which is indeed included in one of our hypotheses). By testing our hypotheses through different statistical methods, we are able to compare the revenue performance of all the considered revenue models (free, paid, and freemium) and thus answer to our research question on revenue models for apps. With regard to the second point, we endeavor to understand whether the adoption of the in-app purchase practice in addition to the chosen revenue model can yield a revenue performance advantage to developers, and thus formulate our hypothesis accordingly. By providing empirical evidence of the effectiveness of in-app purchase, we also contribute to the literature on versioning (Shapiro and Varian 1999; Bhargava and Choudhary, 2001, 2008), as in-app purchase is indeed a peculiar form of versioning. Finally, with regard to both first and second points, it is important to understand whether and how the effects of revenue models and in-app purchase decisions depend on distribution platform and app nature. Thus, we formulate hypotheses on whether (and how) the store (Apple's App Store vs. Google Play) and the app category influence the relationship of revenue models and in-app purchase with revenue performance.[2] As explained later in detail, different platforms may be accessed by consumers with different characteristics (e.g., willingness to pay). As a result, developers' ability to monetize on their "creatures" depends not only on the choice of the revenue model (and in-app purchase) but also on whether and how the given revenue model (and the practice of in-app purchase) fits with the environment (Apple vs. Google ecosystems) where the app is marketed and the nature of the app itself.

To our purposes, we utilize an *ad-hoc* panel dataset of randomly selected apps obtained by collecting data from the two major app stores, Apple's App Store and Google Play, for a period of twenty weeks. To increase reliability and robustness of our findings, we test our hypotheses by using two different statistical approaches, namely regression analysis and Mann–Whitney non-parametric tests performed on opportunely matched subsamples. By way of anticipation, our findings reveal that there is no absolute dominance of a specific revenue model. Indeed, the effect of the revenue models is strongly contingent upon the distribution platform where the given app is marketed. Specifically, we find that in the Apple's App Store, paid and freemium models are shown to be more effective than the free model in terms app revenue

[2] Given the large number of app categories commercialized in major app stores, we do not formulate a priori hypotheses on the relationship between revenue models and app revenue performance for each specific category. Rather, posit that the category certainly plays a role in shaping such relationship and postpone the discussion related to specific categories to the result and practical implications sections.

performance. In this store, paid and freemium models yield statistically similar revenue performance. On the other hand, in Google Play no significant differences between paid and free revenue models emerge, whereas the freemium model is shown to be less effective even than the free model. With regard to in-app purchase, the effect of this practice is also shown to depend on the distribution platform. Specifically, while in-app purchase is shown to positively influence the app revenue performance in Apple's App Store, this effect is reversed in Google Play. We connect all these interesting results to the cross-store differences in terms of average consumer willingness to pay (Halaburda et al. 2011; Ghose and Han 2014). Finally, the type of category is also shown to influence the effects of the revenue model and in-app purchase (the latter to a lesser extent) on the app revenue performance. In this case, we explain that the suitability of a revenue model depends on the app category because different types of apps naturally differ in terms of complexity, needs and segments to be satisfied. We utilize our findings to provide useful indications to support app developers in their decisions about which revenue model to adopt for their apps and whether to enable in-app purchase or not.

The remainder of this paper is organized as follows. In Section 2 we introduce and discuss our hypotheses by connecting them to the existing literature and anecdotal evidence from the industry world. In Section 3, the empirical analysis is presented. Specifically, we describe the dataset, the explanatory variables and the statistical methods used in this paper. In Section 4 we present and discuss the empirical results. Finally, in Section 5 we provide the implications of our findings for theory and practice and conclude discussing the limitations of this study and some lines for future research.

## 2. Theory and hypotheses

The aim of this paper is to study the effects of revenue models and in-app purchase on app revenue performance, as measured by the app daily revenue rank, and how such effects are influenced by the distribution platform where the app is marketed and the type of category to which the app belongs. Accordingly, we ground on the vast literature on the choice of online revenue models to compare paid and freemium revenue models, which typically rely on payments from all or some users, with the free revenue model, which instead crucially relies on the presence of alternative revenue streams, mainly advertising. By using arguments from the versioning literature, we also formulate a hypothesis on the effect of the presence of in-app purchase. Finally, we formulate a number of hypotheses that suggest how these relationships are contingent upon the distribution platform as well as the app category. In this case, we rely on both theoretical arguments and evidence from the industry world.

Before introducing our hypotheses, it is important to explain our choice of considering daily revenue ranks as the measure of app revenue performance. This is because revenue ranks are the only publicly available information on revenues generated by apps. In fact, app stores provide rankings (so called top grossing rankings) based on the revenue generated by each app on a daily basis (Carare 2012; Garg and Telang 2013; Ghose and Han 2014), but do not reveal (for competitive reasons) any information on the actual amount of revenue. In addition, to the best of our knowledge, no commercial database provides this type of data. Hence, using top revenue ranks appears the only viable approach to measure the app revenue performance. This choice is further supported by the copious body of literature that has empirically established a relationship between ranks and actual sales in online commerce (Brynjolfsson et al. 2003; Chevalier and Goolsbee 2003) and specifically in the context of the app market (Garg and Telang 2013). As discussed later in greater detail, we utilize such relationship in formulating our regression model, whereas we simply use the ranks for the Mann–Whitney tests on matched subsamples.

### 2.1. The relationships between revenue models and revenue performance: free vs. paid vs. freemium

The study of different revenue models and their effect on firms' revenue performance has been extensively carried out in the context of information goods and/or EC by a vast and multidisciplinary literature (Amit and Zott 2001; Gallaugher et al. 2001; Afuah and Tucci 2003; Chai et al., 2007; Amit and Zott 2008; Enders et al. 2008; Pauwels and Weiss 2008; Clemons 2009; Teece 2010; Wirtz et al. 2010; Zott et al. 2011; Halbheer et al. 2014; Wagner et al. 2014). For instance, Pauwels and Weiss (2008) have empirically investigated the revenue consequences of moving from a pure free strategy to a paid/freemium strategy for online content providers, demonstrating that the gain or loss in revenue after the change in the revenue model depends on a number of factors such as whether the price is high relatively to the willingness to pay of users, whether all contents are affected by the change in the strategy, and whether the momentum of new free subscriptions has been completed. Similarly, Enders et al. (2008) have explored how different revenue models perform in the context of online social networks, concluding that the choice of free versus paid types of revenue models hinges upon the number of users, their willingness to pay as well as their trust.

Overall, the message provided by the extant literature is that numerous factors can strongly influence the comparison of different revenue models in terms of revenue performance, and thus there is no univocal preference of one revenue model over the other ones. However, no prior study has addressed such comparison in the context of mobile apps. By virtue of its peculiar features, which include, among others, the possibility to sell apps of completely different nature (e.g., games versus utility or navigation apps) in the same e-marketplace as well as the possibility to customize the app, the app store distribution model allows developers to adopt a variety of online revenue models and strategies for app monetization. However, this also requires careful decisions to developers, as choosing the right revenue model for the given app is crucial for the economic success of the app itself.

We can identify at least three main models utilized by developers in the app stores. First, some developers adopt a free app strategy, releasing the app for free, and mostly aim at monetizing by means of advertising and, to less extent, through non-personally identifiable data selling. We refer to it as a *free* revenue model. Second, other developers market only paid apps and mostly aim at monetizing by means of app sales. We refer to this option as a *paid* revenue model. Finally, some other developers combine "free" and "paid" models, by offering both free and paid app versions. Under this model, a low valuable version, which may include ads and has limited or time-expiring functionalities/features, is offered for free. As explained later, in this case the low valuable version can work as a trial offered to users to experiment with the product before committing to upgrade upon payment to the high valuable version, where the above restrictions are significantly reduced, if not totally eliminated. As a result, some users will purchase and use the high valuable version, whereas those who decide not to upgrade will keep using the low valuable version (if not time-expiring) and be subject to (more) intense advertising. We refer to this option as a *freemium* revenue model (Niculescu and Wu 2011; Wagner et al. 2014).

In the following subsections, we compare these three revenue models commonly adopted for marketing apps. Specifically, we formulate our hypotheses comparing paid and freemium models with the free model. This is because, in line with previous studies

**Table 1**
Summary of the hypotheses.

| Hypothesis | Hypothesis statement |
|---|---|
| H1a | Paid apps are associated with better daily revenue ranks than free apps. |
| H1b | Freemium apps are associated with better daily revenue ranks than free apps. This improvement in the daily revenue rank is even higher than that obtained using the paid revenue model, so that freemium apps are associated with better daily revenue ranks than paid apps as well. |
| H2a | It is more likely that paid apps are associated with better daily revenue ranks than free apps in the Apple's App Store. |
| H2b | It is more likely that freemium apps are associated with better daily revenue ranks than free apps in the Apple's App Store. |
| H3 | The effect of the different types of revenue models on app daily revenue rank depends on the app category. |
| H4 | Apps enabling in-app purchase are associated with better daily revenue ranks. |
| H5 | The option of in-app purchase is more likely to be effective in terms of daily revenue rank in the Apple's App Store than in Google Play. |
| H6 | The effect of in-app purchase on app daily revenue rank depends on the app category |

on the comparison of revenue models in online markets (e.g., Pauwels and Weiss 2008), we aim to provide developers with insights on whether they should opt for revenue models (paid or freemium) that require payments from all or some of the users for monetization, or utilize the free model thus crucially relying on other revenue streams, e.g., advertising, in order to monetize from apps. Moreover, contrasting evidences have emerged from the industry world. As a matter of fact, while the great popularity of the free model in the app market has supported since the beginning the belief that such revenue model should be preferred in this context[3], recent industry figures suggest that payment-based models (paid and freemium) allow developers to better monetize on apps (Lunden 2013). Therefore, we try to shed light on the correctness of these opposite evidences by means of rigorous econometric analysis. At any rate, our statistical analysis considers also the comparison between paid and freemium models (which is indeed included in one of our hypotheses). For the sake of clarity, we also summarize the hypotheses in Table 1.

### 2.1.1. Paid versus free revenue models

As already explained, under a paid app revenue model revenue mostly accrues from the price each user pays to download the app. On the other hand, the rationale behind the adoption of a 'free app' revenue model is that developers can heavily rely on other revenue streams, i.e., on revenue generated from transactions with other parties. Indeed, free apps work as two-sided markets (Rochet and Tirole 2003; Parker and Van Alstyne 2005). In our context, apps might attract not only app users, but also third parties who are interested in app users, when they are available in large numbers, for business purposes. Some intuitive examples of third parties are usually firms that are interested in advertising their products inside the app, i.e., advertisers, but also those in search of consumer information for their marketing purposes, i.e., market info seekers. In this case, app developers could release the app for free to users to create a large base of consumers (and thus large base of data) that can be very appealing to such third parties. As a result, app users would be subsidized, whereas third parties end up paying handsomely to advertise inside the app or to obtain non-personally identifiable information, thus allowing developers to profit from these revenue streams.

[3] According to Gartner (2013), 90% of downloads in Apple's App Store and Google Play are related to free apps.

Numerous sources have argued that digital products should be accessed for free by final users due to the fact that they can be produced and commercialized at zero marginal cost and can generate revenue from other streams, e.g., advertising (Anderson 2009). This helps explain the large adoption of this revenue model also for apps (Gartner 2013), which rightly belong to the category of information goods. However, others have argued that many firms selling digital products on the Internet have changed their revenue model by moving from advertising-only revenue model to payment-based models at least for part of their product offering (Pauwels and Weiss 2008). According to theory, the reason is threefold. First, the free revenue model requires a huge amount of users (critical mass) to attract a sufficient number of interested third parties, e.g., primarily advertisers but also info seekers, able to ensure developers' profitability (Canzer 2006; Laudon and Traver 2007; Enders et al. 2008). In the context of apps, this implies an extremely high number of app downloads to allow developers to profit from advertising or also info selling. Second, the increased competition from search engine advertising and other similar forms has made advertising less and less profitable to providers of online digital content (Pauwels and Weiss 2008). Third, numerous studies have highlighted the decline in the effectiveness of advertising on the Internet, which has reduced the appeal of this type of revenue model (Clemons 2009; Zott et al. 2011). Hence, it appears intuitive that releasing only a free app may turn out to be exceedingly risky. On the other hand, the paid revenue model allows developers to profit directly from app sales without the need of reaching a huge amount of downloads to make advertising and similar revenue streams really effective. Indeed, in spite of the fact that the free revenue model is commonly utilized for app commercialization, industry statistics strongly report that paid apps generate higher revenue than free apps and are predicted to play a more prominent role also in the future (Gartner 2013; Lunden 2013). Based on these theoretical arguments and industry evidence, we posit that paid apps should still guarantee a better revenue performance and thus be associated with better daily revenue ranks than apps for which developers have chosen a free model and thus have decided to rely mostly on advertising. Accordingly, we formulate the following hypothesis:

**Hypothesis 1a (H1a).** *Paid apps are associated with better daily revenue ranks than free apps.*

### 2.1.2. Freemium vs. free revenue models

The freemium revenue model combines aspects of the above two models as both free and paid versions are provided for the given app. In this case, firms might take advantage of the presence of heterogeneous customers in the app market and profit from price discrimination as well as advertising. Indeed, the free version could be appealing to lowly valuable customers, who can usually tolerate the "nagging" amount of advertising and the presence of limited or time-expiring features, whereas highly valuable customers usually upgrade to the paid version to avoid the hassle of advertising and have access to more features. Besides favoring market segmentation, the presence of a free version (in addition to the paid version) serves also as a mechanism to favor product trialability. In fact, in this case, the role of the free version is to let customers test the basic functionalities of the product and resolve the uncertainty about the real value to them, prior to committing to purchase and upgrade to the paid version (Rogers 1983; Moore and Benbasat 1991; Gallaugher and Wang 2002).

As we have already discussed, numerous studies focusing on online markets have documented the increasing tendency of information goods providers to move from pure free model to payment-based models (Pauwels and Weiss 2008; Clemons 2009;

Zott et al. 2011), given the difficulty in establishing a sufficient user base to make the advertising revenue model sustainable. However, some authors underscore the fact that online consumers have been used to receive many contents for free and thus may expect to still have access for free (Wagner et al. 2014). Thus, moving directly from a completely free revenue model to full paid revenue model may result in a revenue loss under certain conditions (Pauwels and Weiss 2008). In the context of apps, this is also likely to be the case as consumers have been used to have access to free apps since the launch of app stores (Gordon 2013). On the other hand, by favoring trialability, price discrimination as well as enabling alternative revenue streams such as advertising, the freemium revenue model can better balance between the benefits of both paid and free apps. Therefore, *ceteris paribus*, the freemium revenue model should not only result in a better app revenue performance as compared with the free revenue model, but also the increase in the performance relatively to the free revenue model should be even higher than that obtained via the paid revenue model, thus making it the most preferable revenue model. Accordingly, we formulate that freemium apps should enjoy better daily revenue ranks than both free and paid apps:

**Hypothesis 1b (H1b).** *Freemium apps are associated with better daily revenue ranks than free apps. This improvement in the daily revenue rank is even higher than that obtained using the paid revenue model, so that freemium apps are associated with better daily revenue ranks than paid apps as well.*

### 2.2. Interaction effect between revenue models and store

The above hypotheses examine the general role of revenue models in the app revenue performance as measured by the daily revenue rank. However, it is conceivable that the distribution platform where the given app is marketed could play a role in shaping such hypothesized effects. Indeed, the app market is characterized by the unique relationship between the mobile OS/device market and the app market. Once consumers select their favorite OS/mobile device (e.g., Apple or Android), they can only rely on the sponsored platform to source their apps, unless they change their mobile device (which happens rarely, according to Ghose and Han 2014). This implies that the characteristics of the customer base in a given app store strictly reflect those of the customer base of the associated product, i.e., the mobile device and the relative OS. As different mobile device makers and/or OS providers are likely to target different segments based on their product quality and marketing capabilities, it follows that different app stores will be accessed by different types of consumers. In particular, with regard to the two major mobile device ecosystems, it is widely recognized that Apple targets exclusively the high-end of the market, whereas the sales of Android devices are mostly determined by low-end segments of the market (Halaburda et al. 2011; Ghose and Han 2014; Edwards 2014; Nerney 2014). This implies that consumers accessing the Apple's App Store are on average more valuable than those in Google Play consumers. As a matter of fact, recent industry evidence shows that, although Google Play generates 70% more app downloads as compared with Apple's App Store, the total revenue is still 70% lower (Sims 2015). Previous research on online digital content has suggested payment-based revenue models (e.g., paid or freemium models) are more likely to be suitable in presence of consumers displaying sufficiently large willingness to pay for the product (Enders et al. 2008; Pauwels and Weiss 2008). On the other hand, the free revenue model, which relies on additional revenue streams such as those generated from advertising, can perform better in presence of low user willingness to pay but very large user base (Enders et al. 2008; Pauwels and Weiss 2008). Therefore, we should expect that the hypothesized preference of payments-based models (paid and freemium) over the free revenue model should be particularly emphasized in Apple's App Store, whereas it should be at least mitigated in Google Play. In turn, this implies that an increase in the revenue performance due to the adoption of a paid or a freemium model as compared to the free model is more likely to be observed in the Apple's App Store, by virtue of the higher willingness to pay of users accessing this store. Hence, our hypotheses state:

**Hypothesis 2a (H2a).** *It is more likely that paid apps are associated with better daily revenue ranks than free apps in the Apple's App Store.*

**Hypothesis 2b (H2b).** *It is more likely that freemium apps are associated with better daily revenue ranks than free apps in the Apple's App Store.*

### 2.3. Interaction effect between revenue models and app category

We also investigate the role of the app category in shaping the hypothesized relationship between the revenue model adopted for an app and its revenue performance, as measured by the daily revenue rank. Indeed, there exists a large body of literature demonstrating that product sales on the Internet depend on the type of product and its characteristics (Phau and Poon 2000; Grewal et al. 2004). Particularly, products differing in terms of differentiation or in the degree of sophistication/complexity are likely to result in different level of purchases on the Internet (Grewal et al. 2004; Chang et al. 2005). Similarly, prior literature has documented that different types of products are likely to attract different types of consumers in terms of purchase behavior and preferences (Reibstein 2002; Levin et al. 2003; Korgaonkar et al. 2006).

Based on these arguments, it is straightforward that, by attracting different consumer segments as well as by satisfying different types of needs, different product categories are associated with different consumer willingness to pay. Therefore, it can be easily argued that not all revenue models are suitable for all product categories. In particular, the free revenue model should be likely to perform better for product categories for which users are not very much willing or have not been traditionally used to spend money, but from which developers can still derive adequate revenues due to their feature of being attractive to large number of users and, as a consequence, to highly valuable third parties such as advertisers. Examples include news and magazine categories, for which online providers have traditionally adopted the free revenue model (Pauwels and Weiss 2008). With regard to apps, these arguments suggest that the hypothesized increase in the revenue performance guaranteed by these payment-based models as compared with the free model should be at least mitigated (if not eliminated or even reversed) in presence of news and magazine apps. On the other hand, the paid or the freemium revenue models should be more suitable for product categories that are likely to be complex and satisfy more sophisticated needs, thus delivering considerable value to customers. For this kind of products, consumers are naturally more willing to spend money (Wang et al. 2005). When there is no uncertainty about the value of the given type of product to them, they will afford to pay immediately and therefore, the full paid revenue model is likely to be the most effective. When, instead, such uncertainty exists and is relevant, they are more likely to prefer first to experiment with the basic features of the product before committing to purchase the full version of the product (Gallaugher and Wang 2002). In this case, the freemium revenue model is likely to be the most effective to monetize (Cheng et al. 2015). Utility applications are indicative examples

of categories for which users are willing to pay (sometimes after experimenting with the product), because of their added value to consumers (Wang et al. 2005). Based on these considerations, it appears clear that the app category will shape the efficacy of revenue models in terms of app revenue performance. However, given the large number of app categories commercialized in major app stores (in our sample, for instance we count 12 categories), we do not formulate a priori hypotheses on the relationship between revenue models and app revenue performance for each specific category. Rather, we maintain a more general view and posit that the category certainly plays a role in shaping such relationship. From a theoretical perspective, it is indeed more interesting to understand whether the category is influential in the relationship between revenue models and app revenue performance, rather than predicting the specific impact of revenue models for each detailed category. Therefore, we postpone the discussion related to specific categories to the practical implications section and formulate our next hypothesis as follows:

**Hypothesis 3 (H3).** *The effect of the different types of revenue models on app daily revenue rank depends on the app category.*

### 2.4. The relationship between in-app purchase and revenue performance

More and more developers provide app users who have downloaded their apps at a given price (or free) with the opportunity to purchase additional features, e.g., additional levels, or credits in case of a game, directly inside the app. This practice is known as in-app purchase. However, it is important to note that in-app purchase does not necessarily coincide with the freemium revenue model, which implies the release of free version in addition to the paid version. Rather, it is an additional strategy that can be pursued by developers to better segment and price discriminate among their customers to the scope of increasing profits. In fact, in-app purchase provides users with the opportunity to upgrade to better features and/or download additional functionalities with additional price charge, irrespective of whether a free version has been released for the app in addition to the paid version. In other words, it can be well the case that only the paid version has been released for the app (for instance at a price equal to $0.99) and in-app purchases are offered to add additional features at incremental prices. Essentially, in-app purchase can exist also in absence of a free version, which instead is necessary part of the freemium revenue model.

From a theoretical perspective, in-app purchase can be assimilated to versioning and upgrading. Numerous theoretical studies have studied the economic implications of the use of a versioning strategy for information goods, suggesting that versioning can be optimal in presence of consumer heterogeneity (Shapiro and Varian 1999; Bhargava and Choudhary, 2001, 2008). For instance, Bhargava and Choudhary (2008) demonstrate the optimality of versioning when the lower quality version is designed in a way that, relative to their valuations for the high-end version, highly valuable consumers have a lower relative valuation for the lower quality than do lowly valuable consumers (Bhargava and Choudhary, 2008). More recently, Wei and Nault (2013) study a setting where consumers can purchase low or high quality versions of the product and, in case they opt for the low quality version, they still have the opportunity to upgrade to the high quality version at a certain price. They show that versioning is optimal under this version-to-upgrade setting, with some consumers that will upgrade to the high quality version after purchasing the low quality version. The same authors show that versioning is optimal also when consumers have both group and individual tastes for quality (Wei and Nault 2014). Finally, versioning has been found optimal in presence of piracy (Wu and Chen 2008).

Overall, the above theoretical studies predict that under certain conditions pursuing a versioning strategy is suitable for information goods as it helps profit from price discrimination. Basic industry figures seems to suggest that this is likely to be the case in the context of apps as in-app purchase has been successful in the Apple's App Store, and it has reached large popularity in Google Play as well (Gartner 2013). Thus, both from theoretical and industry evidence viewpoints, it seems reasonable to expect that, *ceteris paribus*, apps enabling in-app purchase are more likely to be associated with higher revenue as a result of a finer-grained segmentation in the market, and thus be better ranked than apps offering no in-app purchase. Accordingly, we formulate the following:

**Hypothesis 4 (H4).** *Apps enabling in-app purchase are associated with better daily revenue ranks.*

### 2.5. Interaction effect between in-app purchase and store

The above hypothesis suggests that we should expect better revenue performance, as measured by the app daily revenue rank, for apps enabling in-app purchase. Here, similarly to the case of revenue models, we argue that the extent of the effect of in-app purchase on app revenue performance should be contingent upon the distribution platform where the app is commercialized. As a matter of fact, the same theoretical arguments discussed above for the revenue models apply also to the case of in-app purchase. In fact, in-app purchase as a practice of price discrimination requires heterogeneous consumers with some consumers willing to spend significantly more than others to upgrade to the more valuable app functionalities/features. Thus, in-app purchase is more likely to be effective when the consumer segments displaying large willingness to pay are of considerable size (Enders et al. 2008). As we have highlighted above, consumers accessing the Apple's App Store are on average more valuable than those in Google Play (Ghose and Han 2014; Sims 2015). Hence it is conceivable to expect that in-app purchase is more likely to be effective in terms of revenue performance in the Apple's App Store than in Google Play, given the presence of more valuable users in the former store. Hence, we formulate our fifth hypothesis as follows:

**Hypothesis 5 (H5).** *The option of in-app purchase is more likely to be effective in terms of daily revenue rank in the Apple's App Store than in Google Play.*

### 2.6. Interaction effect between in-app purchase and app category

In addition to the role of the distribution platform, we argue that the category to which an app belongs also influences the relationship between in-app purchase and the app revenue performance as measured by the app daily revenue rank. Once again, the same arguments discussed for revenue models apply also to the case of in-app purchase. In fact, as discussed earlier, to guarantee enhanced profitability, in-app purchase requires the presence of considerably large consumer heterogeneity in terms of willingness to pay, with some consumers willing to spend more for upgraded features. Intuitively, products of very different categories have different nature, which may result in user segments with significantly different purchase behavior, willingness to pay, needs to satisfy (Levin et al. 2003; Grewal et al. 2004; Reibstein, 2002; Wang et al. 2005). Therefore, consumer heterogeneity in terms of willingness to pay is likely to differ across categories. Consequently, price discrimination can be more effectively practiced

for categories where heterogeneity is more emphasized, while being less suitable for categories attracting more homogenous consumers. Similarly to the case of revenue models, given the large number of app categories commercialized in major app stores, we do not formulate a priori hypotheses on the relationship between in-app purchase and app revenue performance for each specific category. Rather, we formulate our next hypothesis by positing that the category influences the relationship between in-app purchase and app revenue performance, leaving room for the discussion on specific categories in the practical implications:

**Hypothesis 6 (H6).** *The effect of in-app purchase on app daily revenue rank depends on the app category.*

## 3. The empirical analysis

### 3.1. Data and Variables

In order to test our hypotheses, we collected data of apps for smartphones by weekly exploring the Italian version of the two major app stores, namely Apple's App Store and Google Play. Specifically, every Friday, we recorded data from a *random* sample of apps retrieved from the top 200 grossing apps ranking publicly available in these two stores.[4] In case an app was no longer listed among the top 200 apps in the relative store in subsequent observations, the actual ranking was retrieved from appannie.com, a well-known website providing the entire history of official app rankings and statistics. The choice of considering top app rankings is in line with previous literature, given that the majority of apps available in the app stores are definitely not displayed to consumers (Carare 2012). It is also important to clarify that, according to the available sources, top grossing app rankings are based on the total daily (not cumulated) revenue of an app including revenue from sales, in-app purchase and advertising (Carare 2012; Garg and Telang 2013; Ghose and Han 2014). The fact that the rankings are based on daily, rather than aggregate, app revenue allowed us to concentrate on a specific day and thus reduce collection effort that was mostly carried out manually, while still maintaining a sufficiently large observation period. As a matter of fact, the observations utilized in the present paper are related to all Fridays in the period going from October 19th, 2012 to March 8th, 2013 (20 weeks in total). There are three reasons behind this choice. First, we wanted to reduce the extremely significant data collection effort, which was very intense as the collection process could only be performed manually for most of the variables of interests and the controls. Second, after preliminary analysis we noticed that observing data once a week allowed us to appreciate changes in the rank of apps. Third, the choice of Friday as a representative day of the week is due to the fact in a preliminary analysis it largely reflected the behavior of the other days. Therefore, there would not be much difference if other days of the week were chosen. At any rate, as we discuss in the concluding section, future research could look at different days or incorporate more days in a week by diminishing the number of weeks of observation.

In our preliminary analysis, we *randomly* selected 50 apps from each of the two top 200 grossing rankings, so that we had initially 100 apps to be observed. However, when the observation period actually began, 9 apps from App Store and 1 app from Google Play were no longer among the top 200 grossing apps of the respective stores. Therefore, we *randomly* added 10 (9 from the App Store and 1 from Google Play) apps sharing the same characteristics of those apps. Recording data from the two stores for all the 20 weeks would yield a balanced panel dataset of 2200 observations related to 59 apps from app store and 51 apps from Google Play. However, during our period of observation, four apps were eliminated by the store where they were available for download, thus leading to a slightly unbalanced panel dataset counting 2177 observations.

Finally, it is important to underline that our sample does not include e-service apps, i.e., those apps operating as a new channel to facilitate consumers' purchase of a certain product/service that has been normally transacted via seller's website or physical store. Examples include apps marketed by airlines. These types of apps are naturally free, as the developer does not market the app to make money out of it. Rather, these apps serve to facilitate service acquisition and increase customer loyalty. These types of apps were never in the top 200 grossing rankings during our period of observation. Also, it is important to highlight that our sample includes only the paid version of freemium apps, and never the relative free version, as the latter did not appear in the top 200 grossing ranking during our observation period. In fact, all free apps in our sample do not have a corresponding paid version, whereas some paid apps in our sample have a corresponding a free version. Therefore, we study the effect of the freemium model on the daily revenue rank of the paid version, not of the free version, which, based on our data, was not as successful as the paid version.

We defined a set of variables of interest and some controls to test the formulated hypotheses and recorded the relative data. They are reported in Table 2, which provides a brief description and the modalities of all the variables. Table 3 shows the descriptive statistics. In the interest of length, we only discuss descriptive statistics related to our variables of interest. First, with regard to our dependent variable, the average revenue rank of apps in our sample is 240, with the highest rank being the first position and the lowest rank being the 1000th position, with no relevant changes across the two stores, i.e., Apple's App Store and Google Play.[5] With regard to our main independent variables, not surprisingly, releasing free apps is the most popular top grossing revenue model (53.88% of the occurrences), followed by the paid (27.38%) and freemium (18.74%) revenue models, respectively, as shown in column A of Table 3. In addition, developers of more than two third of the apps utilize the practice of in-app purchase. Moreover, from columns B and C of Table 3, only minor differences seem to emerge with the regard to the distribution of revenue models as well as in-app purchase across the two stores, except for the freemium revenue model, which accounts for 28.71% of the observations in Google Play, while accounting for 10.22% in Apple's App Store. With regard to the app category popularity, column A of Table 3 shows, as expected, that Games are undoubtedly the most popular app category with about 57.74% of our observations, whereas all the remaining categories display percentages less than 10%. Interestingly, columns B and C of Table 3 also show high similarity of the two stores (Apple's App Store and Google Play) in terms of distribution of apps across categories. Such high similarity, together with our choice of using a randomly selected sample from the two stores, increases our confidence that the impact of any potential bias due to endogenous decisions of developers is very limited, at least in our sample. In fact, one may argue that certain categories are more popular in one store rather than the other one as a result of developers' decisions to develop and commercialize certain apps in certain stores. On a theoretical basis, this may have impact on the reliability of our findings if certain categories turn out to be systematically more successful than others. However, given that the distribution of the categories

[4] Considering a random sample naturally helps reduce potential selection bias in the choice of the top grossing apps. Moreover, we later provide evidence that other potential bias concerns are very limited (if not absent at all) in our sample.

[5] Note that our sample did consist of apps initially ranked among the first 200 apps by revenue (either when selecting the initial sample or at the moment of first observation). However, during the period of observation some apps could naturally lose positions and be ranked in lower positions. This explains why the average rank is actually 240 and the minimum rank is 1000.

**Table 2**
Variables description.

| Variables | Description |
|---|---|
| *Dependent variable* | |
| Rank | Positive integer variable indicating the rank of the given app in the given week in the specific store |
| *Variables of interest* | |
| Free | Binary variable equal to 1 if the app is free in the given week and no paid version is available in the given week of observation; 0 otherwise. This variable is chosen as a baseline in the regression analysis and thus it is removed due to perfect collinearity with *Freemium* and *Paid* |
| Freemium | Binary variable equal to 1 if both paid and free versions (the latter either time-expiring or with limited functionalities, and possibly displaying advertisements) are available in the given week of observation; 0 otherwise |
| Paid | Binary variable equal to 1 if the app is paid and no free version is available in the given week of observation; 0 otherwise |
| In-app | Binary variable equal to 1 if the app enables in-app purchase of optional features in the given week; 0 otherwise |
| *Control variables* | |
| App category | 12 binary variables (Customizations; Education; Entertainment; Games; Healthcare & Fitness; Money & Finance; Music; News & Magazines; Photo & Video; Social Network; Travel & Navigation; Utility), each equal to 1 if the app belongs to the respective category; 0 otherwise. The dummy *Games* is chosen as a baseline in the regression analysis and thus it is removed due to perfect collinearity with the other category dummies |
| Company fame | Binary variable equal to 1 the app developer is a developer with fully established reputation worldwide; 0 otherwise. Based on revenue information and worldwide recognition, we identify 9 top developers in our sample, i.e., Apple, Disney, Electronic Arts, Gameloft, Popcap, Rockstar Games, Sega, Ubisoft, and Zynga. We also include Garmin, Marvel Entertainment, Norton by Symantec and TomTom due to their huge popularity |
| Developer type | Binary variable equal to 1 if the app is developed by a firm; 0 if developed by individual(s) |
| High app rating | Binary variable equal to 1 if, in the given week and store, the app rating was greater or equal to 4 stars. The high threshold is due to the fact that the great majority of apps in our sample had no low rating |
| New entries | Positive integer variable indicating the number of new entries in the top 200 ranking in the given week in the specific store |
| Number developer Apps | Positive integer variable indicating the number of apps marketed by the developer of the given app in the given week in the given store |
| Size | Continuous variable measuring the size (in Mbytes) of the app in the given week |
| Store | Binary variable equal to 1 if the app is observed in the App Store; 0 if observed in Google Play |
| Time since market launch | Positive integer variable measuring the time (in months) since the market launch of the given app |

across the two stores is similar, we can be confident about the limited (if any) impact of potential bias concerns. Columns D-N of Table 3 show that in-app purchase is utilized under all three revenue models (i.e., paid, free, freemium) adopted by developers. Specifically, almost all free apps in our sample enable in-app purchase (out of 53.88% observations of free apps, 53.28% are related to free apps with in-app purchase), whereas this practice is implemented in approximately half of paid apps (out of 27.38% observations of paid apps, 13.63% are related to paid apps with in-app purchase) and only 9.8% of freemium apps (1.84% out of 18.74%). In this case, there are some differences between the two stores, especially with regard to the percentages of paid and freemium apps giving users the in-app purchase opportunity. Finally, Table 3 provides detailed information about the distribution of free, paid and freemium revenue models as well as in-app purchase across different categories. Without any loss of generality, in the regression analysis, the variables *Games* and *Free* are considered as baselines for category and revenue model variables and thus are removed from the regression analysis given their perfect collinearity with respective category and revenue model variables (Roma et al., 2013). More importantly, note that our preliminary analysis suggests that there are no serious concerns of multicollinearity that can affect the results. Finally, note also that the introduction of the interaction terms as well as the dummy *Store* allows us to correctly test our hypotheses by mitigating potential issues implied by the fact that our dependent variable, i.e., the daily revenue rank, is a relative measure of revenue performance and there might exist differences of the two platforms in terms of total revenue generated by each store.

### 3.2. Methods

We utilize two different empirical approaches to test our hypotheses and show that our major findings are essentially consistent across these approaches, thus increasing our confidence about the main implications of this study. Specifically, our first approach is to test the hypotheses via a standard random effect regression model for panel data. The second approach consists of testing our hypotheses by means of the Mann–Whitney non-parametric test, which is performed after applying the Propensity Score Matching (PSM) method (Rosenbaum and Rubin, 1983; Rosenbaum and Rubin, 1984). The PSM technique is important due to the univariate nature of the Mann–Whitney test. Indeed, it serves to generate subsamples of apps that are similar in all the other relevant characteristics not included in the Mann–Whitney test and thus can help mitigate potential bias concerns related to the use of such test. In the following subsections we provide the description of our two approaches, while in Section 4 we present and discuss the findings showing that the two approaches yield consistent results.

#### 3.2.1. Regression analysis

As our dataset is a slightly unbalanced panel and the number of statistic units (apps) is quite large compared to the observation period (number of weeks), three basic regression models, namely pooled OLS, fixed effects and random effects models, are usually suggested in the econometrics literature (Wooldridge 2002). However, the fixed effects model should not be appropriate a priori in our setting because some variables of interest, e.g., store and categories, are time invariant. They would be eliminated due to perfect collinearity if fixed effects model were adopted. Therefore, we preliminarily compare pooled OLS and random effects models to analyze the effects of all explanatory and control variables. In all the cases, the Breusch-Pagan Lagrange Multiplier test strongly indicates the presence of random effects (as reported below Table 4). Therefore, to opportunely test our hypotheses, the following standard random effects regression model is considered in this paper (for the sake of presentation we report the model including the interactions of the revenue models with the dummy *Store*; the model including the interactions of the revenue models with the category dummies can be derived similarly):

$$Rank_{ijt}^{-a} = \alpha + (\beta_1 + \beta_2 \cdot Store_j) \cdot Paid_{ijt} + (\gamma_1 + \gamma_2 \cdot Store_j) \cdot Freemium_{ijt} + (\delta_1 + \delta_2 \cdot Store_j) \cdot In-app_{ijt} + \Phi \cdot Controls_{ijt} + m_{ij} + \varepsilon_{ijt} \quad (1)$$

In (1), subscript $i$ indicates the app, $j$ the store, and $t$ the week of observation. The exponent $a$ is the shape parameter of the power law function linking app sales (in value) to app daily revenue rank. The power law relationship between sales (in value) and ranks has been empirically demonstrated by the copious body of literature in online commerce (e.g., Brynjolfsson et al. 2003; Chevalier and Goolsbee 2003) and also specifically in the context of the app

**Table 3**
Descriptive Statistics.

| | *A* | *B* | *C* | *D* | *E* | *F* | *G* | *H* | *I* | *L* | *M* | *N* | *O* | *P* | *Q* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | *Free* | | | *Paid* | | | *Freemium* | | | *In-app* | | |
| | Total | Apple | Google | Total | Apple | Google | Total | Apple | Google | Total | Apple | Google | Total | Apple | Google |
| Free | 53.88% | 57.32% | 49.85% | | | | | | | | | | | | |
| Paid | 27.38% | 32.45% | 21.44% | | | | | | | | | | | | |
| Freemium | 18.74% | 10.22% | 28.71% | | | | | | | | | | | | |
| In-app | 69.59% | 71.55% | 67.30% | 53.28% | 56.22% | 49.85% | 14.47% | 13.63% | 15.45% | 1.84% | 1.70% | 1.99% | | | |
| Games | 57.74% | 59.11% | 56.13% | 45.02% | 47.70% | 41.87% | 8.68% | 9.71% | 7.48% | 4.04% | 1.70% | 6.78% | 53.05% | 56.22% | 49.35% |
| Social network | 4.59% | 5.11% | 3.99% | 3.67% | 3.41% | 3.99% | 0.00% | 0.00% | 0.00% | 0.92% | 1.70% | 0.00% | 4.59% | 5.11% | 3.99% |
| Money & finance | 0.92% | 0.00% | 1.99% | 0.00% | – | 0.00% | 0.00% | – | 0.00% | 0.92% | – | 1.99% | 0.00% | – | 0.00% |
| Photo & video | 4.59% | 6.81% | 1.99% | 0.60% | 1.11% | 0.00% | 3.99% | 5.71% | 1.99% | 0.00% | 0.00% | 0.00% | 0.92% | 1.70% | 0.00% |
| Entertainment | 0.92% | 1.70% | 0.00% | 0.92% | 1.70% | – | 0.00% | 0.00% | – | 0.00% | 0.00% | – | 0.92% | 1.70% | – |
| Education | 3.67% | 3.41% | 3.99% | 0.00% | 0.00% | 0.00% | 1.84% | 3.41% | 0.00% | 1.84% | 0.00% | 3.99% | 0.00% | 0.00% | 0.00% |
| Healthcare & fitness | 0.92% | 0.00% | 1.99% | 0.00% | – | 0.00% | 0.00% | – | 0.00% | 0.92% | – | 1.99% | 0.00% | – | 0.00% |
| Music | 2.76% | 5.11% | 0.00% | 0.00% | 0.00% | – | 0.00% | 0.00% | – | 2.76% | 5.11% | – | 0.00% | 0.00% | – |
| News & magazines | 3.67% | 5.11% | 1.99% | 2.76% | 3.41% | 1.99% | 0.92% | 1.70% | 0.00% | 0.00% | 0.00% | 0.00% | 2.76% | 3.41% | 1.99% |
| Utility | 7.35% | 6.81% | 7.98% | 0.00% | 0.00% | 0.00% | 3.67% | 6.81% | 0.00% | 3.67% | 0.00% | 7.98% | 0.00% | 0.00% | 0.00% |
| Customizations | 3.67% | 1.70% | 5.98% | 0.00% | 0.00% | 0.00% | 1.84% | 1.70% | 1.99% | 1.84% | 0.00% | 3.99% | 0.92% | 1.70% | 0.00% |
| Travel & navigations | 9.19% | 5.11% | 13.95% | 0.92% | 0.00% | 1.99% | 6.43% | 3.41% | 9.97% | 1.84% | 1.70% | 1.99% | 6.43% | 1.70% | 11.96% |

| Binary variables | *Descr. Stat.* | | | *Binary Variables* | *Descr. Stat.* | | |
|---|---|---|---|---|---|---|---|
| | Total | Apple | Google | | Total | Apple | Google |
| Store | 53.93% | 100.00% | 0.00% | Developer Type | 97.24% | 98.30% | 96.01% |
| Company fame | 21.77% | 22.06% | 21.44% | High App Rating | 78.41% | 75.21% | 82.15% |

| Continuous/discrete variables | *Descriptive Statistics* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Total* | | | | *Apple* | | | | *Google* | | | |
| | Mean | Std.D. | Min | Max | Mean | Std.D. | Min | Max | Mean | Std.D. | Min | Max |
| New entries | 44.71 | 15.88 | 0.00 | 90.00 | 48.55 | 12.44 | 0.00 | 60.00 | 40.21 | 18.13 | 0.00 | 90.00 |
| Number developer apps | 22.15 | 37.63 | 1.00 | 168.00 | 28.16 | 47.57 | 1.00 | 168.00 | 15.12 | 18.27 | 1.00 | 69.00 |
| Size | 141.55 | 354.71 | 0.02 | 2355.20 | 187.12 | 373.67 | 1.00 | 2007.04 | 88.21 | 323.27 | 0.02 | 2355.20 |
| Time since market launch (months) | 13.17 | 9.11 | 0.00 | 45.00 | 16.13 | 11.01 | 0.00 | 45.00 | 9.70 | 4.01 | 1.00 | 25.00 |
| Rank | 239.91 | 282.18 | 1.00 | 1000.00 | 244.45 | 299.47 | 1.00 | 1000.00 | 234.59 | 260.25 | 2.00 | 1000.00 |

Note that the statistics are computed using all 2177 observations for columns designated as "Total", whereas are computed using only the observations from Apple's App Store (1174) for the columns designated as "Apple" and only those from Google Play (1003) for the columns designated as "Google". Also note that the sign – indicates that it does not make sense to compute the statistics. For instance, the percentage of free apps in the category "Money & Finance" in Apple's App Store cannot be computed because there no apps in such category in Apple's App Store.

market (Garg and Telang 2013). According to these relevant studies, the use of the power law function linking sales and ranks is a powerful way to capture the revenue performance in absence of direct knowledge on revenue (which is indeed our case). Therefore, we take advantage of the strong evidence provided by previous literature on the power law relationship between sales and ranks and adopt this transformation as a dependent variable, instead of using the mere rank. Specifically, in the paper we present the results derived from our regression analysis for different values for the parameter *a*, which are consistent with the recent work of Garg and Telang (2013) conducted specifically in the app market. At any rate, it is noteworthy that the use of other values for this parameter leads to robust results. Moreover, note that when taking the alternative approach (i.e., the Mann–Whitney tests on matched subsamples), we simply use the mere ranks (instead of the power law transformation). This allows us to verify robustness of the findings obtained using the regression model.

With regard to the other elements in (1), $Controls_{ijt}$ stands for the vector of control variables reported in Table 2 ($\Phi$ is the vector of the relative coefficients), which we utilize under the different models we perform. On the other hand, $m_{ij}$ and $e_{ijt}$ are the app-store random effect and error terms, respectively.

### 3.2.2. Mann–Whitney tests

Our second empirical approach is to perform the Mann–Whitney test for each of our variables of interest, after opportunely controlling for the remaining relevant factors that may influence the app revenue performance, as measured by the app daily revenue rank. The Mann–Whitney test (sometimes also referred to as Mann–Whitney–Wilcoxon test) is the non-parametric alternative to the independent sample *t*-test. It is used to test whether two samples come from the same population against the alternative hypothesis that the observations in one of the two samples tend to have higher values than the observations in the other sample. Specifically, this test compares two subgroups (e.g., paid apps vs. free apps, or apps with in-app purchase vs. apps without in-app purchase) in terms of the dependent variable (the revenue rank in our case) by ranking the units of the two subsamples and computing for all possible combinations the number of times that higher values (i.e., better revenue ranks) are observed in the first subsample and the number of times that higher values (i.e., better revenue ranks) are observed in the second subsample. The minimum of these two numbers is then compared with the threshold values of statistical tables of the Mann–Whitney test or with its normal approximation ($z$) as implemented by major statistical software packages. If the $p$-value corresponding to the $z$ value is higher than 0.05 (or 0.1 if we want to be less restrictive), the hypothesis that the two subsamples come from the same distribution cannot be rejected. Vice versa, if the corresponding $p$-value is lower than 0.05 (or 0.1), we can conclude that one of the two subsamples has larger values (i.e., better rank in our case) than the other. The test also provides the probability that one subsample is associated with larger values (better revenue ranks in our case) than the other subsample.

**Table 4**
Results under Random Effects regression models ($a = 0.6$).

| | | Model 1 | | Model 2 | | Model 3 | |
|---|---|---|---|---|---|---|---|
| | Variable | Coefficient | Std. errors | Coefficient | Std. errors | Coefficient | Std. errors |
| Controls | Developer type | −0.0070 | (0.0211) | 0.0093 | (0.0158) | −0.0213 | (0.0216) |
| | Number developer apps | −0.0014*** | (0.0003) | −0.0014*** | (0.0003) | −0.0014*** | (0.0003) |
| | Company fame | 0.0177** | (0.0065) | 0.0211** | (0.0069) | 0.0226** | (0.0075) |
| | Size | 0.0000 | (0.0000) | 0.0000 | (0.0000) | 0.0000 | (0.0000) |
| | New entries | −0.0001** | (0.0000) | −0.0001** | (0.0000) | −0.0001** | (0.0000) |
| | Time since Mkt launch | −0.0029*** | (0.0008) | −0.0030*** | (0.0008) | −0.0030*** | (0.0008) |
| | High app rating | 0.0029 | (0.0062) | 0.0024 | (0.0063) | 0.0026 | (0.0063) |
| | Store | 0.0739** | (0.0236) | −0.0680** | (0.0225) | 0.0848** | (0.0258) |
| | Social network | −0.0528* | (0.0222) | −0.0531* | (0.0226) | −0.0631** | (0.0228) |
| | Money & finance | 0.0265+ | (0.0158) | 0.0155 | (0.0156) | 0.0495** | (0.0155) |
| | Photo & video | −0.0534* | (0.0264) | −0.0486* | (0.0215) | −0.1344** | (0.0430) |
| | Entertainment | −0.1076*** | (0.0307) | −0.1144*** | (0.0323) | −0.1206*** | (0.0326) |
| | Education | −0.0163 | (0.0166) | −0.0054 | (0.0171) | −0.0446+ | (0.0228) |
| | Healthcare & fitness | −0.0234 | (0.0169) | −0.0345* | (0.0165) | −0.0007 | (0.0161) |
| | Music | −0.0633+ | (0.0368) | −0.0162 | (0.0350) | −0.0535 | (0.0365) |
| | News & magazines | −0.0265 | (0.0317) | −0.0216 | (0.0277) | −0.0163 | (0.0331) |
| | Utility | −0.0280 | (0.0268) | −0.0135 | (0.0215) | −0.0935** | (0.0360) |
| | Customizations | 0.0006 | (0.0366) | −0.0232 | (0.0314) | 0.0733** | (0.0228) |
| | Travel & navigations | −0.0283 | (0.0193) | 0.0022 | (0.0170) | −0.0053 | (0.0375) |
| *H1a/b* | Freemium | 0.0468* | (0.0210) | −0.0479* | (0.0208) | −0.0380 | (0.0253) |
| | Paid | 0.0330*** | (0.0089) | −0.0098 | (0.0171) | −0.0203 | (0.0282) |
| *H2a/b* | Freemium X store | | | 0.0828** | (0.0322) | | |
| | Paid X store | | | 0.0460* | (0.0184) | | |
| *H3* | Freemium X Social | | | | | 0.0948** | (0.0295) |
| | Freemium X education | | | | | 0.0591* | (0.0270) |
| | Freemium X utility | | | | | 0.1255** | (0.0378) |
| | Freemium X customizations | | | | | −0.0469 | (0.0309) |
| | Freemium X travel & navigations | | | | | −0.0342 | (0.0399) |
| | Paid X photo & video | | | | | 0.0623* | (0.0280) |
| | Paid X news & magazines | | | | | −0.0916* | (0.0426) |
| | Paid X travel & navigations | | | | | 0.0215 | (0.0283) |
| *H4* | In-app | 0.0573** | (0.0216) | −0.0523** | (0.0193) | −0.0030 | (0.0259) |
| *H5* | In-app X store | | | 0.1524*** | (0.0277) | | |
| *H6* | In-app X photo & video | | | | | 0.0665 | (0.0418) |
| | In-app X customizations | | | | | −0.1194** | (0.0376) |
| | In-app X travel & navigations | | | | | −0.0050 | (0.0382) |
| | Constant | 0.0573** | (0.0216) | 0.1539*** | (0.0340) | 0.1407*** | (0.0391) |

Note: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Number of obs.: 2177. Number of apps: 110. Standard errors are robust to heteroskedasticity and serial correlation. Breusch-Pagan Lagrangian multiplier test for random effects: $\chi^2$ 1(2)[3] = 2422.92 (2685.16) [2462.1], $p = 0.0$ ($p = 0.0$) [$p = 0.0$] for models 1 (2), and [3], respectively. Note that all the interactions not reported in the table were deleted either because of perfect multicollinearity or because of the absence of the given revenue model and/or in-app purchase in the selected category.

The Mann–Whitney test would be ideally suitable for our analysis as it allows us to statistically compare couples of subsamples of apps utilizing different revenue models and compare the two subsamples of apps enabling in-app purchase and apps without in-app purchase. However, the main issue behind the use of Mann–Whitney is that such test considers only one variable at a time. That is, differently from a regression analysis, it tests the effect of a variable without taking into account other covariates that can also potentially affect the dependent variable. In setting where data are not derived from a controlled experiment, it is crucial to control for all covariates that are relevant in determining the dependent variable. This is because it may happen that results will be considerably biased, if excluded covariates display certain degree of correlation with the variables considered in the analysis. For instance, a simple comparison of the revenue performance of paid apps vs. free apps via the Mann–Whitney may not reveal the true effect of the revenue model, and thus it may yield totally biased results if the revenue model displays certain degree of correlation with other variables (e.g., app category or the number of apps marketed by the given developer, etc...). Indeed, the Mann–Whitney may end up attributing a certain effect to the considered variable (paid vs. free revenue models), while in reality such effect is due to factors that are not controlled for in the analysis. The regression analysis helps mitigate (if not solve) this problem because it allows including in the regression model a wide range of binary, discrete, and continuous variables, in addition to that of our interest. The Mann–Whitney is instead a univariate type of analysis. To render the Mann–Whitney really suitable to our purposes and avoid considerable bias concerns, we then need to apply the test to subsamples of our main sample that are generated by opportunely controlling for the other relevant factors not considered in the non-parametric test. This can be done by using the Propensity Score Matching (PSM) technique (Rosenbaum and Rubin, 1983, 1984). This class of methods is commonly utilized when researchers want to evaluate the impact of a treatment applied to statistical units in absence of controlled experiments. In this case, units receiving a treatment (for instance, apps for which the developer has adopted the paid revenue model) and those receiving the alternative treatment (for instance, apps for which the developer has adopted the free revenue model) may differ not only in their treatment status (paid vs. free) but also in other characteristics potentially correlated to both treatment status and the outcome of interest (app revenue performance in our case). To avoid the biases that this may generate, matching algorithms find a non-treated (or alternatively treated) unit that is similar to a treated unit across several dimensions (or characteristics)

by computing a propensity score that capture such overall similarity. Essentially, these matching algorithms associate two statistical units different in terms of treatment (the variable object of study), but being similar for the remaining relevant characteristics. For instance, if we aim at comparing paid vs. free apps (where paid revenue model is the treatment and the free revenue model is the non-treatment or the alternative treatment), the PSM technique couples paid apps with free apps based on the similarity they show in terms of the remaining relevant characteristics, such as the size of the app, the category, the type of developer, and all the other variables that we controlled for in the regression analysis. After applying the PSM technique, we are provided with a subsample consisting of couples of free and paid apps that are similar in terms of the relevant characteristics except for the revenue model (paid vs. free). At this point, the Mann–Whitney test can be performed on the subsample to understand which revenue model yields better revenue performance. Indeed, after applying the PSM technique we can be sufficiently confident the comparison would not yield biased results as the other relevant characteristics have been controlled for before running the non-parametric test and thus the apps in the subsamples considerably differ for the revenue model but not for the other relevant characteristics.

Therefore, our second approach consists of performing Mann–Whitney tests for all the variables of interest on subsamples generated applying the PSM technique in order to reduce the risk that our results will be biased. Specifically, we performed the Mann–Whitney test on one-to-one matched subsamples to compare paid vs. free apps, freemium vs. free apps, freemium vs. paid apps, apps with in-app purchase vs. apps without in-app purchase in both Apple's App Store and Google and for each category (when comparison can be made). With regard to the PSM technique, to ensure a more rigorous one-to-one matching, we applied PSM matching within strata. That is, we considered exact matching within app categories and store, and then within each of these strata, we considered all the other remaining characteristics to further match treated and non-treated apps based on the propensity score measure. For instance, we matched a game app only with the most similar game app (based on the propensity score). Also, we imposed a common support (overlap condition) with no-replacement by removing treatment observations for which the propensity score was higher than the maximum or less than the minimum of the score of non-treatment observations.

## 4. Results

We present the results of our regression models in Tables 4–6. Each table refer to a different value of parameter a. Specifically, a is set to be equal to 0.6, 0.8 and 0.4 in Tables 4–6, respectively. Each of these tables includes three different models. Model 1 does not include any interaction variable, Model 2 include the set of interactions of the revenue models and the in-app purchase with the distribution platform, i.e. the dummy variable *Store*. In Model 3 we include the set of interactions of the variables related to the revenue models and in-app purchase with the categories' variables.[6] To better understand the results in Tables 4–6, it is helpful to recall that in Model 2 the interaction terms of the revenue models and in-app purchase with the variable *Store* reflect the effect of such variables in Apple's App Store (*Store* = 1), whereas the direct terms of such variables reflect the effect of the same variables in Google Play (*Store* = 0). Regarding the interactions of the revenue models and in-app purchase with category variables, they should be interpreted as the effect of such variables in each respective category.

In Table 7, we instead present the results of the Mann–Whitney tests performed on matched subsamples. Specifically, we include the hypotheses to be tested and the relative variables, the *z*-statistics and the corresponding *p*-value, which indicates whether the null hypothesis that two considered subsamples come from the same distribution can be accepted or not. In addition, it is provided the probability that one subsample has larger values (i.e., better revenue ranks) than the other, thus providing, in case of rejection of the null hypothesis, indication on which subsample displays better revenue ranks.

To present and discuss the results we take into account both types of analyses as they are shown to be largely consistent with each other. The positive and significant coefficient of the *Paid* dummy in Model 1 of Tables 4–6 would seem to suggest that the paid revenue model is associated with higher top grossing ranks than the free revenue model. In the same vein, the positive coefficient of the *Freemium* dummy would seem to suggest that this revenue model has a positive impact on app revenue performance as compared with the pure free model.

However, from Model 1 it is not clear whether this result is valid irrespective of the store where apps are marketed or rather the effect depends (and thus may change depending) on the store. To better shed light on this result (related to H1a and H1b) as well as to test hypotheses H2a and H2b, we introduce interaction terms of the revenue models variables with the *Store* dummy (as done in Model 2). The regression results reported in Model 2 of Tables 4–6 show that these effects actually strongly depend on the distribution platform where the given app is commercialized. Specifically, the positive and significant coefficients of the interaction terms of both paid and freemium models with the dummy *Store* suggest that in the Apple's App Store both freemium and paid models are confirmed to improve the revenue performance as compared with the free model. This is further confirmed by the Mann–Whitney tests performed in Table 7. Indeed, in the Apple's App Store we can strongly reject the hypothesis that paid and freemium apps have similar revenue performance to that of free apps (*p*-values equal to 0.003 and 0.000). Rather, the fact that probability of observing better revenue ranks for free apps is quite lower than ½ in both cases (last column of Table 7) confirms that paid and freemium apps are associated with better ranks and thus superior revenue performance than free apps. No significant difference in terms app revenue ranks is instead shown between paid and freemium apps in the Apple's App Store. Indeed, the non-significance of the Mann–Whitney test performed comparing matched subsamples of freemium and paid apps implies that we cannot reject that the two revenue models lead to statistically similar revenue performance in the Apple's App Store.

As for the store Google Play, Model 2 in Tables 4–6 show that paid and free revenue models do not differ significantly in terms of app revenue performance. In fact, after introducing the interaction, the coefficient of the *Paid* dummy, which indeed represents the effect of the paid revenue model in Google Play as compared with the free revenue model, is not significant. In this case, the Mann–Whitney test in Table 7 provides a different result, showing that we can reject the hypothesis that paid and free apps display similar revenue performance (*p*-value equal to 2.18%). Indeed, it is shown that free apps are less likely to display better revenue ranks than paid apps. However, given that the two types of analyses provide contrasting results, no clear indication can be drawn with regard to the preference of paid vs. free models. More interestingly, the comparison of freemium and free revenue models in

[6] We do not present the model including both sets of interactions simultaneously because this would determine a regression model with a very large number of interactions (given the number of observations at our disposal), which, as it is well known, unavoidably increases multi-collinearity, thus resulting in reduced significance of coefficients. Therefore, in this case, it is reasonable to rely on separate models to study and provide the relative implications on how decisions on revenue models and in-app purchase are contingent upon the distribution platform as well as the app nature. At any rate, the results under the full model are shown to be reasonably robust, in spite the diminished levels of significance in some circumstances.

**Table 5**
Results under Random Effects regression models ($a = 0.4$).

| | | Model 1 | | Model 2 | | Model 3 | |
|---|---|---|---|---|---|---|---|
| | Variable | Coefficient | Std. errors | Coefficient | Std. errors | Coefficient | Std. errors |
| Controls | Developer type | −0.0097 | (0.0304) | 0.0171 | (0.0255) | −0.0278 | (0.0340) |
| | Number developer apps | −0.0019*** | (0.0004) | −0.0019*** | (0.0004) | −0.0019*** | (0.0004) |
| | Company fame | 0.0262** | (0.0099) | 0.0324** | (0.0105) | 0.0347** | (0.0115) |
| | Size | 0.0000 | (0.0000) | 0.0000 | (0.0000) | 0.0001 | (0.0001) |
| | New entries | −0.0002*** | (0.0001) | −0.0002*** | (0.0001) | −0.0002*** | (0.0001) |
| | Time since Mkt launch | −0.0044*** | (0.0009) | −0.0045*** | (0.0009) | −0.0045*** | (0.0009) |
| | High app rating | 0.0041 | (0.0093) | 0.0030 | (0.0094) | 0.0033 | (0.0094) |
| | Store | 0.0944*** | (0.0267) | −0.1264*** | (0.0332) | 0.1090*** | (0.0289) |
| | Social network | −0.0614* | (0.0254) | −0.0619* | (0.0259) | −0.0785** | (0.0251) |
| | Money & finance | 0.0489* | (0.0245) | 0.0326 | (0.0243) | 0.0859*** | (0.0260) |
| | Photo & video | −0.0715+ | (0.0383) | −0.0626* | (0.0311) | −0.2046*** | (0.0579) |
| | Entertainment | −0.1314*** | (0.0323) | −0.1404*** | (0.0339) | −0.1504*** | (0.0337) |
| | Education | −0.0203 | (0.0264) | −0.0041 | (0.0265) | −0.0621* | (0.0318) |
| | Healthcare & fitness | −0.0310 | (0.0255) | −0.0474+ | (0.0250) | 0.0057 | (0.0265) |
| | Music | −0.0775 | (0.0564) | −0.0039 | (0.0538) | −0.0588 | (0.0571) |
| | News & magazines | −0.0189 | (0.0487) | −0.0121 | (0.0432) | −0.0027 | (0.0525) |
| | Utility | −0.0337 | (0.0406) | −0.0115 | (0.0330) | −0.1267* | (0.0526) |
| | Customizations | 0.0050 | (0.0539) | −0.0285 | (0.0465) | 0.1073*** | (0.0306) |
| | Travel & navigations | −0.0388 | (0.0284) | 0.0132 | (0.0444) | 0.0210 | (0.0516) |
| H1a/b | Freemium | 0.0839** | (0.0319) | −0.0652* | (0.0297) | −0.0547 | (0.0352) |
| | Paid | 0.0709*** | (0.0155) | −0.0086 | (0.0245) | −0.0192 | (0.0370) |
| H2a/b | Freemium X store | | | 0.1291** | (0.0448) | | |
| | Paid X store | | | 0.0857*** | (0.0262) | | |
| H3 | Freemium X social | | | | | 0.1601*** | (0.0414) |
| | Freemium X education | | | | | 0.0924* | (0.0392) |
| | Freemium X utility | | | | | 0.1828*** | (0.0558) |
| | Freemium X customizations | | | | | −0.0578 | (0.0464) |
| | Freemium X travel & navigations | | | | | −0.0653 | (0.0570) |
| | Paid X photo & video | | | | | 0.1068** | (0.0372) |
| | Paid X news & magazines | | | | | −0.1399* | (0.0658) |
| | Paid X travel & navigations | | | | | 0.0183 | (0.0358) |
| H4 | In-app | 0.0898** | (0.0314) | −0.0786** | (0.0281) | −0.0079 | (0.0358) |
| H5 | In-app X store | | | 0.2326*** | (0.0350) | | |
| H6 | In-app X photo & video | | | | | 0.1142+ | (0.0601) |
| | In-app X customizations | | | | | −0.1605*** | (0.0501) |
| | In-app X travel & navigations | | | | | −0.0173 | (0.0518) |
| | Constant | 0.1342*** | (0.0421) | 0.2746*** | (0.0444) | 0.2558*** | (0.0535) |

Note: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Number of obs.: 2177. Number of apps: 110. Standard errors are robust to heteroskedasticity and serial correlation. Note that all the interactions not reported in the table were deleted either because of perfect multicollinearity or because of the absence of the given revenue model and/or in-app purchase in the selected category.

Google Play leads to surprising results. Indeed, in Model 2 of Tables 4–6, the negative and significant coefficient of the *Freemium* dummy, which indeed represents the effect of the freemium model as compared with free model in Google Play, implies that the freemium model has actually a negative effect in Google Play. That is, this revenue model is actually less effective in terms of revenue performance than the free revenue model (and thus also in comparison with the paid revenue model). This result is fully confirmed for observations also by performing the Mann–Whitney test, which supports rejection of revenue performance equivalence and shows that the probability of observing better revenue ranks for free apps is approximately 0.9. The test also rejects the revenue equivalence between freemium and paid models in Google Play, showing that the probability of observing higher ranks for paid apps is quite high (0.775). We explain these results as the consequence of high product cannibalization in Google Play. Indeed, the presence of a free version can cannibalize the paid version in this platform given that Android users on average tend to limit their spending on apps (Halaburda et al. 2011; Ghose and Han 2014; Edwards 2014; Nerney 2014) and thus they are more likely to decide to utilize the free version without upgrading to the full paid version. Therefore, since the freemium revenue model relies on a considerable portion of users that will self-select to pay for the app perhaps after having experimented with the free version, its effectiveness will be greatly lower in a platform accessed by consumers with relatively low willingness to pay. In fact, in this case, the freemium revenue model will suffer from low app sales as the portion of paying users is not sufficiently developed, and from insufficient advertising revenues as the model is not fully centered on this revenue stream. Thus, the revenue performance of this model turns out to be inferior even to both paid and free models.

Overall, this first set of evidences suggests that H1a and H1b are not fully confirmed because the paid and freemium revenue models do not always lead to better revenue performance than free apps. Rather the preference for these models depends on the store where apps are marketed. Also, no clear dominance of freemium vs. paid revenue models is confirmed. On the other hand, the findings suggest that H2a and H2b are confirmed, as the superior revenue performance of paid and freemium revenue models is more likely to occur in the Apple's App Store as compared with the free revenue model. Indeed, in Google Play, freemium revenue model is shown to be even inferior to the free model, while not univocal indications can be drawn with regard to the comparison of paid and free models. In contrast, both freemium and paid revenue models are shown to be associated with better revenue ranks than free apps in the Apple's App Store.

**Table 6**
Results under Random Effects regression models ($a = 0.8$).

| | | Model 1 | | Model 2 | | Model 3 | |
|---|---|---|---|---|---|---|---|
| | Variable | Coefficient | Std. errors | Coefficient | Std. errors | Coefficient | Std. errors |
| Controls | Developer type | −0.0057 | (0.0147) | 0.0042 | (0.0098) | −0.0162 | (0.0142) |
| | Number developer apps | −0.0010*** | (0.0003) | −0.0010*** | (0.0003) | −0.0010*** | (0.0003) |
| | Company fame | 0.0107* | (0.0047) | 0.0128** | (0.0049) | 0.0139** | (0.0052) |
| | Size | 0.0000 | (0.0000) | 0.0000 | (0.0000) | 0.0000 | (0.0000) |
| | New entries | −0.0001** | (0.0000) | −0.0001** | (0.0000) | −0.0001** | (0.0000) |
| | Time since Mkt launch | −0.0020** | (0.0007) | −0.0020** | (0.0007) | −0.0020** | (0.0007) |
| | High app rating | 0.0019 | (0.0037) | 0.0016 | (0.0037) | 0.0018 | (0.0037) |
| | Store | 0.0563* | (0.0223) | −0.0318* | (0.0155) | 0.0641** | (0.0244) |
| | Social network | −0.0420* | (0.0192) | −0.0415* | (0.0192) | −0.0473* | (0.0197) |
| | Money & finance | 0.0137 | (0.0101) | 0.0056 | (0.0098) | 0.0273** | (0.0099) |
| | Photo & video | −0.0377* | (0.0189) | −0.0356* | (0.0156) | −0.0886** | (0.0336) |
| | Entertainment | −0.0826** | (0.0292) | −0.0877** | (0.0309) | −0.0911** | (0.0310) |
| | Education | −0.0113 | (0.0103) | −0.0047 | (0.0110) | −0.0296+ | (0.0157) |
| | Healthcare & fitness | −0.0151 | (0.0110) | −0.0233* | (0.0111) | −0.0018 | (0.0102) |
| | Music | −0.0478+ | (0.0254) | −0.0039 | (0.0538) | −0.0435+ | (0.0252) |
| | News & magazines | −0.0258 | (0.0207) | −0.0121 | (0.0432) | −0.0195 | (0.0206) |
| | Utility | −0.0212 | (0.0178) | −0.0115 | (0.0330) | −0.0659* | (0.0262) |
| | Customizations | −0.0014 | (0.0238) | −0.0285 | (0.0465) | 0.0477** | (0.0165) |
| | Travel & navigations | −0.0193 | (0.0222) | 0.0132 | (0.0444) | −0.0186 | (0.0279) |
| H1a/b | Freemium | 0.0244+ | (0.0138) | −0.0652* | (0.0297) | −0.0272 | (0.0181) |
| | Paid | 0.0115+ | (0.0063) | −0.0086 | (0.0245) | −0.0208 | (0.0215) |
| H2a/b | Freemium X store | | | 0.1291** | (0.0448) | | |
| | Paid X store | | | 0.0857*** | (0.0262) | | |
| H3 | Freemium X social | | | | | 0.0524* | (0.0206) |
| | Freemium X education | | | | | 0.0362+ | (0.0185) |
| | Freemium X utility | | | | | 0.0830** | (0.0276) |
| | Freemium X customizations | | | | | −0.0348+ | (0.0201) |
| | Freemium X travel & navigations | | | | | −0.0124 | (0.0277) |
| | Paid X photo & video | | | | | 0.0382+ | (0.0209) |
| | Paid X news & magazines | | | | | −0.0584* | (0.0286) |
| | Paid X travel & navigations | | | | | 0.0233 | (0.0227) |
| H4 | In-app | 0.0359* | (0.0154) | −0.0786** | (0.0281) | −0.0018 | (0.0177) |
| H5 | In-app X store | | | 0.2326*** | (0.0350) | | |
| H6 | In-app X photo & video | | | | | 0.0404 | (0.0301) |
| | In-app X customizations | | | | | −0.0815** | (0.0278) |
| | In-app X travel & navigations | | | | | 0.0028 | (0.0273) |
| | Constant | 0.0366+ | (0.0222) | 0.2746*** | (0.0444) | 0.0859** | (0.0301) |

Note: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Number of obs.: 2177. Number of apps: 110. Standard errors are robust to heteroskedasticity and serial correlation. Note that all the interactions not reported in the table were deleted either because of perfect multicollinearity or because of the absence of the given revenue model and/or in-app purchase in the selected category.

The rationale behind these results is that Google Play is accessed by consumers characterized by relatively low willingness to spend money to download apps (Halaburda et al. 2011; Ghose and Han 2014; Edwards 2014; Nerney 2014). Hence, in this platform, the hypothesized preference of payments-based models over the free revenue model turns out to be mitigated and even reversed (in certain cases) because a large number of users would only consider downloading free apps. In contrast, Apple's App Store is on average accessed by more valuable consumers, who are not much worried about spending their money to download the apps they need. Thus, the paid and freemium revenue models are more effective in generating revenue to developers in Apple's App Store than in Google Play.

Before discussing our findings on how the effect of revenue models varies across app categories (H3), it is important to point out that not all the interactions between the revenue models and the app category (and similarly those between the in-app purchase and the app category) appear in Model 3 of Tables 4–6 and not all the categories are reported in Table 7 where present the Mann–Whitney tests. This is because in many cases not all the revenue models are utilized for all categories, at least in our randomly collected sample. In addition, some interactions are omitted due to perfect multicollinearity when only one revenue model is utilized in a given category. For instance, from Table 3 the freemium model is the only revenue model utilized for apps falling into *Music*, *Healthcare & Fitness* and *Money & Finance* categories, whereas this revenue model is never utilized *Photo & Video*, *Entertainment*, and *News and Magazines* categories. Also, in certain categories only two revenue models are utilized. Finally, note that the free revenue model is considered as a baseline in our regression analysis, thus all relative interactions are obviously not included in Tables 4–6. Table 7 provides the only categories for which a comparison of revenue models is possible in our sample. Thus, we refer to this Table to present this particular result, also in light of the fact that the results from Model 3 (reported in Tables 4–6) are mostly identical (unless explicitly pointed out). Indeed, both empirical approaches we utilize show that the effect of the revenue model can vary across app categories, thus providing support to our hypothesis *H3* overall. Specifically, we find that *News & Magazines* is a category where the paid revenue model does not perform as well as the free model because the hypothesis of equal revenue performance is strongly rejected and the probability of observing better revenue ranks for free apps is quite high (0.925). As we have explained when introducing *H3*, the rationale behind this finding is that these types of apps work well as two-sided markets and thus rely on other revenue streams, mostly advertising. Users have been

**Table 7**
Mann–Whitney tests for matched subsamples.

| Tested hypotheses Ho: rank(X) = rank (Y) | X | Y | z | *p*-value | Prob (rank (Y) is better than rank (X)) |
|---|---|---|---|---|---|
| *Apple App Store* | | | | | |
| Paid vs. free | Paid | Free | −2.955 | 0.003** | 0.379** |
| Freemium vs. free | Freemium | Free | −4.085 | 0.000*** | 0.235*** |
| Freemium vs. paid | Freemium | Paid | 0.863 | 0.3882 | 0.562 |
| In-app | In-app = 1 | In-app = 0 | −3.630 | 0.000*** | 0.349*** |
| *Google Play* | | | | | |
| Paid vs. free | Paid | Free | −2.294 | 0.0218* | 0.405* |
| Freemium vs. free | Freemium | Free | 8.669 | 0.000*** | 0.905*** |
| Freemium vs. paid | Freemium | Paid | 6.541 | 0.000*** | 0.775*** |
| In-app | In-app = 1 | In-app = 0 | 2.328 | 0.020* | 0.607* |
| *Freemium vs. free* | | | | | |
| Social | Freemium | Free | −5.411 | 0.000*** | 0.000*** |
| Travel & Navigations | Freemium | Free | 5.411 | 0.000*** | 1.000*** |
| Games | Freemium | Free | 4.469 | 0.000*** | 0.743*** |
| *Paid vs. free* | | | | | |
| News & Magazines | Paid | Free | 4.601 | 0.000*** | 0.925*** |
| Travel & Navigations | Paid | Free | 4.194 | 0.000*** | 0.887*** |
| Photo & Video | Paid | Free | −4.231 | 0.000*** | 0.000*** |
| Games | Paid | Free | −0.498 | 0.618 | 0.485 |
| *Freemium vs. paid* | | | | | |
| Education | Paid | Freemium | 3.360 | 0.001*** | 0.810*** |
| Utility | Paid | Freemium | 2.059 | 0.040* | 0.634* |
| Customizations | Paid | Freemium | 0.077 | 0.939 | 0.505 |
| Travel & Navigations | Paid | Freemium | −4.263 | 0.000*** | 0.223*** |
| Games | Paid | Freemium | −1.761 | 0.078+ | 0.402+ |
| *In-app* | | | | | |
| Photo & Video | In-app = 1 | In-app = 0 | 4.222 | 0.000*** | 0.890*** |
| News & Magazines | In-app = 1 | In-app = 0 | 3.679 | 0.000*** | 0.840*** |
| Customizations | In-app = 1 | In-app = 0 | 5.411 | 0.000*** | 1.000*** |
| Games | In-app = 1 | In-app = 0 | −3.873 | 0.000*** | 0.312*** |

The size of matched subsamples utilized for the Mann–Whitney test varies from a few hundreds observations to almost one thousand observations.

traditionally used to have access to these types of contents for free and thus are likely to expect to still have access for free (Wagner et al. 2014). Therefore, in spite of the general tendency of tilting away from the free revenue model for online content (Pauwels and Weiss 2008), our result suggests that for news and magazine applications, relying on advertising is still the most suitable strategy. By looking at Table 7, a similar result is observed for the category *Travel & Navigations*. However, in this case, the result is not consistent with that derived from the regression analysis, thus no univocal indication can be provided. In contrast, for the category *Photo & Video* both empirical approaches show that the paid model leads to superior revenue performance than the free model. The rationale is that customers are usually willing to pay for photo and video applications due to the considerable added value of these types of apps (Deloitte 2016).[7]

The freemium model is shown to be particularly effective as compared with the free model in case of *Social Networks* applications. The freemium model is utilized for *Social Network* apps only Apple's App Store. Therefore, the positive effect of the freemium model in Apple's App Store as compared with the free revenue model is particularly amplified in this category. No univocal indication can be obtained for the category *Travel & Navigations* as the free model is shown to be preferable from Table 7, but the corresponding interaction in Tables 4–6 is never significant. With regard to the comparison of paid and freemium revenue models, the freemium model should be preferred to the paid model in case of *Education* and *Utility* applications. Indeed, both regression analysis and Mann–Whitney tests provide this indication. As discussed when introducing *H3*, all these product categories are particularly complex and satisfy more sophisticated needs (e.g., education) than others, thus usually providing considerable value to customers. However, the complex nature of the needs these apps should satisfy often creates relevant uncertainty about the real quality delivered to users. In this case, consistently with our finding, the freemium revenue model is preferable as users are more likely to prefer first to experiment with the basic features of the product before purchasing the full paid version (Cheng et al. 2015). No univocal indication can be instead obtained for the category *Travel & Navigations* as the freemium model is shown to be more preferable than the paid model in Table 7, but the corresponding interactions in Tables 4–6 are never significant. The comparison between freemium and paid models is also shown to be not significant for the category *Customizations* under both types of analysis. Finally, the Mann–Whitney tests performed in Table 7 allow us to compare the performance of revenue models in the category *Games*. This comparison was not possible in the regression analysis, as we have chosen the category *Games* as a baseline. From Table 7, we note that for apps falling into this category paid and free models are equivalent and both should be preferred to the freemium model. The explanation could be due to cannibalization reasons as games are by far the most popular category of apps and thus they are more likely to attract a large mass of consumers who are not willing to upgrade to the paid version if they have the free version at disposal. Taken together, all these results confirm that certain revenue models are more suitable for certain types of apps, while being less appropriate for others.

Moving to the effect of the presence of in-app purchase, the positive and significant coefficient of this variable in Model 1 of Tables 4–6 would seem to suggest that the effect of enabling in-app purchase on revenue ranks is positive. However, to understand whether this effect holds irrespective of the distribution platform,

[7] Note that the freemium revenue model is never utilized for *News & Magazines* and *Photo & Video* categories, as it can be seen from Table 3.

we need to introduce the interaction with the *Store* dummy in the regression model or perform the Mann–Whitney test in each specific distribution platform. In this respect, Model 2 shows that this interaction term, which captures the effect of in-app purchase in Apple's App Store, is positive and significant. On the other hand, the dummy *In-app*, which, in this case, captures the effect of in-app purchase in Google Play, is negative and significant. The same result is observed by performing the Mann–Whitney test in Table 7. In this table, it is indeed shown that for both Apple's App Store and Google Play we can strongly reject the hypothesis that apps with in-app purchase display similar revenue performance to apps that do not offer this opportunity. However, the effect of in-app purchase is opposite in the two stores. In the Apple's App Store the probability to observe better revenue ranks for apps without in-app purchase is quite low (0.349), whereas it is quite higher than ½ (the exact value being 0.607) in Google Play. The fact that the effect of in-app purchase is opposite across the two stores and it is positive only in Apple's App Store implies that our hypothesis *H4* is not supported while hypothesis *H5* is confirmed. The rationale of this result should lie again on the relatively low willingness to pay of consumers accessing Google Play. In fact, theory (e.g., Bhargava and Choudhary 2008) suggests that the existence of two segments with distinct valuations for the high-end and low-end versions is necessary to achieve higher profitability under a versioning strategy, as compared with selling only one version. The highly valuable segment is on average not adequately developed in Google Play, while being relevant in the Apple's App Store (Ghose and Han 2014; Nerney 2014). Therefore, we observe that offering in-app purchase is an inferior strategy in the former store, while it is preferable in the latter store.

Regarding the role of in-app purchase among categories (*H6*), the results in Models 3 of Tables 4–6 and those in Table 7 show that the effect of such practice on app revenue performance changes across categories, although to a lesser extent as compared to the case of revenue models. This supports our hypothesis *H6*. In particular, the two types of analysis are consistent, especially with regard to the negative effect of in-app purchase for the category *Customizations*. This negative effect of in-app purchase might be explained by the fact that consumers usually perceive this type of apps, which are essentially wallpapers for smartphones, as lowly valuable so they are unlikely to spend money to upgrade to nicer features. In contrast, the Mann–Whitney test performed for the category *Games* shows that in-app purchase is beneficial in terms of revenue. As a matter of fact, the probability of observing high revenue ranks for apps with no in-app purchase is only 0.312 and the hypothesis of equivalence is strongly rejected. The rationale is possibly that game users perceive in-app purchase as a way to personalize the games by buying within the app new game characters, new credits and new challenges. Therefore, a considerable portion of them is willing to purchase additional features inside the app.

To facilitate the understanding of our findings, Table 8 summarizes our major results and the relative explanations we have provided in this section.

## 5. Implications and conclusions

This paper provides unique empirical evidence of the roles of the revenue models and in-app purchase in driving app revenue performance, as measured by app daily revenue rank. More importantly, it offers unique evidence of how the effects of these factors hinge upon the distribution platform where the app itself is marketed as well as the category to which the given app belongs. Therefore, our analysis has relevant implications for both existing literature and practitioners, which we discuss below.

### 5.1. Implications for the existing literature

Our findings add to the existing knowledge of the app market by revealing how the revenue performance is influenced by the choice of revenue models (and in-app purchase) and how this influence is strongly contingent on the distribution platform as well as the nature of the app. Previous studies on the app market have indeed mostly disregarded these aspects, which instead are perceived as pivotal by developers (Sourcebits 2014; Pappas 2013; Wilcox 2013; Munir 2014). Moreover, our study adds to the existing literature on the choice of revenue models in online markets (e.g., Pauwels and Weiss 2008), by shedding light on the revenue performance implications of revenue models in the specific context of the app market. In this environment several revenue models are commonly utilized, but performance implications have never been examined before. In this regard, our study offers novel "food for thought" to researchers in the field of e-commerce by documenting that the suitability of a revenue model depends on the distribution platform where the app is marketed as well as the nature of the app itself. As a matter of fact, our analysis reveals that both paid and freemium revenue models guarantee higher revenue performance than the free revenue model in the Apple's App Store. No significant differences in terms of revenue performance emerge between paid and freemium models in this store. The results for the Apple's App Store can be explained by the relatively high average willingness to pay of consumers accessing this store. Indeed, Apple consumers are on average not worried about spending money to buy apps or (in case of the freemium model) upgrade the full paid version after having tried the free version, if quality is delivered to them. The results for the Apple's App Store are consistent with the basic figures provided by industry (Lunden 2013) as well as with the observation that many online content providers have started tilting away from the free revenue model, which mainly relies on advertising, to favor more payment-based models (Pauwels and Weiss 2008; Clemons 2009; Zott et al. 2011).

Interestingly, the same results do not occur in Google Play. Rather, in this store, we find that the adoption of a freemium revenue model has a negative impact on app revenue performance as compared with the free revenue model. We have explained that this interesting result is caused by the large product cannibalization to which paid versions are exposed in presence of a free version, as a consequence of the relatively low willingness to pay of Google users on average. The cannibalization phenomenon is largely detrimental for the freemium model. Indeed, on the one hand, a few consumers will upgrade to full paid version so low revenue will be generated from it. On the other hand, not sufficient ads revenue will be generated from the free version because, differently from a pure free revenue model, the freemium revenue model is not fully centered on ads. As a result, in Google Play, the revenue performance of the freemium model will be inferior to those of paid and free models. No univocal indications from the two types of analysis emerge with regard to the comparison between paid and free models in Google Play. Perhaps, this is due to the fact that the absence of free version eliminates the risk of product cannibalization for the paid version and makes the paid model a reasonable alternative to the free model as developers of paid apps can still benefit from a small segment of highly valuable users that can afford buying a paid app. Our study also provides novel evidence on how the effects of revenue models can vary across app categories. That is, our findings show that these effects can be reinforced or weakened (or even reversed), depending on the specific category to which the given app belongs.

Finally, our findings on the role of in-app purchase add to the versioning literature (e.g., Bhargava and Choudhary, 2008) by providing unique empirical evidence on how a peculiar type of versioning strategy, which is indeed known as in-app purchase in

**Table 8**
Summary of the results and explanations.

| Hypotheses | Status | Explanation |
|---|---|---|
| H1a | Not confirmed | Hypothesis H1a is not confirmed as paid apps are not always associated with better daily revenue ranks than free apps. Indeed, in Google Play, no significant differences in revenue performance arise. The reason is due to the fact that users have relatively low willingness to pay. Hence, in spite of the general tendency of tilting away from the free revenue model in online markets, this model can still be as effective as the paid model in environments characterized by relatively low willingness to pay. |
| H1b | Not confirmed | Hypothesis H1b is not confirmed, as freemium apps are not always associated with better daily revenue ranks than free apps. Indeed, in Google Play, the opposite occurs, which makes the freemium revenue model inferior to both free and paid models. The rationale is that the presence of a free version in the freemium model can cannibalize the paid version in this platform given that Android users have on average limited willingness to pay and thus may decide to utilize the free version without upgrading to the full paid version. Also the free version will not be able to generate enough advertising revenue because, differently from the free model, the freemium model is not fully centered on generating revenue from ads. |
| H2a | Confirmed | Hypothesis H2a is confirmed because results show that it is more likely that paid apps are associated with better daily revenue ranks than free apps in the Apple's App Store. Indeed, paid apps are associated with better revenue rank than free apps in this store, whereas no significant difference emerges in Google Play. The reason is that Apple's App Store players are more willing to pay that Android users, thus they are less worried about spending their money if quality is delivered to them. |
| H2b | Confirmed | Hypothesis H2b is confirmed because results show that it is more likely that freemium apps are associated with better daily revenue ranks than free apps in the Apple's App Store. The rationale is the same as that provided for H2a. Moreover, no significant differences in terms of revenue performance between paid and freemium apps emerge in Apple's App Store. |
| H3 | Confirmed | H3 is confirmed as our results show that certain revenue models are more (or less) preferable in certain categories. This is because different categories satisfies different needs, attract different consumer segments, and are associated with different levels of uncertainty with regard to product value. |
| H4 | Not Confirmed | H4 is not confirmed as apps enabling in-app purchase are not always associated with better daily revenue ranks. Indeed, in Google Play the opposite occurs. That is, *ceteris paribus*, apps without in-app purchase leads to higher revenue performance. The rationale is that, according to theory, in-app purchase can be optimal in presence of two segments with distinct valuations for high-end version and low-end version. As the highly valuable segment is on average not adequately developed in Google Play, the in-app purchase strategy turns out to be inferior in this store. |
| H5 | Confirmed | H5 is confirmed because the option of in-app purchase is more likely to be effective in terms of daily revenue rank in the Apple's App Store than in Google Play. The rationale follows from above. Indeed, differently from Google Play, the highly valuable segment is largely developed in Apple's App Store due to the higher average consumers' willingness to pay in this store. |
| H6 | Confirmed | H3 is confirmed (although to a less extent as compared with H3) as our results show that in-app purchase can be more (or less) preferable in certain categories. The explanation follows that provided for H3. |

the app market, influences the revenue performance. Our findings suggest that in-app purchase is also a positive determinant of app revenue performance only in the Apple's App Store. As we have explained, this is because a segment with high willingness to pay is required to effectively practice price discrimination (Chen et al., 2014; Roma et al., 2014; Roma et al., 2015). While this segment is certainly present in Apple's App Store, it might not be well developed in Google Play, thus leading to a negative impact in terms of revenue performance in the latter store. Similarly to the revenue models above, the category is likely to influence the relationship between in-app purchase and app revenue performance, although to a lesser extent.

### 5.2. Managerial implications

Our findings have also a number of remarkable implications and guidelines for developers involved in revenue model and in-app purchase decisions. First, we inform developers that the preference for a revenue model strongly depends on the platform where the given app is marketed as well as on the category of the app itself. Therefore, developers should take into account these two factors when making a crucial decision for their business development such as the choice of the revenue model for their apps. According to our findings, developers in the Apple's App Store should consider following the general tendency illustrated by numerous recent studies on online content provision and take advantage of the demonstrated Apple users' high willingness to pay by adopting payment-based revenue models, rather than relying on advertising streams of revenue. In particular, developers of *Social Network* apps in Apple's App Store should be interested in the freemium revenue model as its impact is shown to be very strong for such category. On the other hand, our study warn developers from using the freemium model in Google Play because consumers in this store are on average less prone to spend money on apps. Therefore, the portion of them that will self-select to switch from the free version to the paid version is limited in this store. In addition, the revenues from advertising are insufficient to compensate the loss from sales given that the revenue model is not fully centered on this alternative stream. As a result, differently, from the Apple's App Store, developers in Google Play should avoid adopting this hybrid model and rather opt for either free or paid ones as they are shown to be equivalent in terms of app revenue performance. By using a paid model developers eliminate the risk of product cannibalization, as the free version is not released. Hence they can benefit from selling to a relatively small segment of users that are more willing to pay than others and do not have a free version alternative at disposal for download. Those commercializing free apps will instead benefit from the large number of users not willing to pay for apps but still able to bear the disadvantage of being subject to nagging advertising. The interesting result on the negative effect of freemium model in Google Play has another important implication for developers. Indeed, the extent of product cannibalization between free and paid versions arguably depends on the characteristics of the free version as compared with those of the full paid version. If the free version does not differ much from the paid version in terms of number of features and overall quality, the risk of cannibalization clearly increases. For numerous apps, it can be observed that, except for the presence of (more) advertising in the free version, the differences between the two versions are not as neat as they should be to avoid the detrimental effects of product cannibalization. Therefore, to render the freemium revenue model viable in an environment where the average consumer willingness to pay is relatively low, developers may rethink about their version design strategies and differentiate the two versions significantly.

With regard to the suitability of revenue models in specific categories, developers should take into consideration the freemium model for apps, such as those falling into *Education* or *Utility* categories, which are more complex or satisfy more sophisticated and valuable needs. For these types of value-added apps even Google users are more likely to commit to purchase the paid version after experimenting with the free one. Similarly, *Photo & Video* apps in both Google and Apple stores should be commercialized using

the paid model because consumers tend to assign high value to activities such as photo uploading and sharing (Deloitte 2016). For *Games* apps the freemium model may not be as effective as paid or free revenue models due to the fact game apps usually target mass app market and thus the majority of users are unwilling to upgrade to the paid version if they have a free version available. Finally, while in general payment-based revenue models should be preferred in Apple's App Store, developers of *News & Magazine* apps in both Apple and Google stores should not forgo the free revenue model as it may yield superior revenue performance than the paid revenue model (the freemium model is not utilized in this category, at least in our sample). In fact, these types of apps operate as two-sided markets, usually attracting large number of users and thus allowing developers to monetize from third parties such as advertisers.

Our findings also inform developers that the cross-store differences in terms of consumer willingness to pay also determine a different impact of the in-app purchase strategy on app revenue performance in the two stores. Similarly to the case of revenue models, given that the willingness to pay of consumers is relatively large in Apple's App Store, developers in this store should consider the opportunity to segment users by charging them different prices based on the number and quality of features they choose for their apps via in-app purchase. A notable exception in the effect of in-app purchase in App Store is related to *Customizations* apps, for which users are reluctant to upgrade to enhanced features, given the low added value of this kind of applications. On the other hand, the in-app purchase strategy is shown to be always detrimental to developers in Google Play. Because of their lower willingness to pay on average, most of Google Play users will enjoy the lower price charged for the base version without upgrading to better features via in-app purchase, thus hurting developers' profitability. For such reason, developers may consider forgoing the adoption of the in-app purchase option in Google Play.

### 5.3. Limitations and future research directions

To the best of our knowledge, this is the first study examining the effect of revenue models and in-app purchase on the app revenue performance and how such relationships are influenced by the distribution platform and the app category. Therefore, there are some limitations, which open the room for future research directions.

First, our choice of considering daily revenue ranks as our measure of app revenue performance is due to the fact that revenue of apps are not publicly available and, to the best of our knowledge, no commercial database provides actual data. Hence, it appears reasonable to utilize revenue ranks, also by virtue of the relationship between sales and ranks, which has been largely demonstrated in online markets (Brynjolfsson et al. 2003; Chevalier and Goolsbee 2003) and specifically in the app market (Garg and Telang 2013). Nevertheless, future studies may consider actual revenue or profits as measures of performance when these data become available for the app market.

Second, our random selection has resulted in similar app category distribution across the two stores, which has increased our confidence that potential selection bias concerns due to market heterogeneity are unlikely to arise. However, future studies could consider applying other selection bias reducing techniques to further address this issue. Also future studies could focus on specific categories in order to gather more observations related to one specific category and thus further validate our findings, especially with regard to the impact of categories for which the two statistical approaches we utilized in the paper provide contrasting results.

Third, our study focuses on the Italian version of Apple Store and Google Play in a time span of twenty weeks. In order to increase the generality of our findings, future research could study the determinants of app revenue performance in other markets such as US, UK, etc. or look at a broader time span of observation. For instance, employing different samples collected in different years could allow further validation of our results.

Fourth, to reduce the data collection effort we have decided to consider one specific day of the week, i.e., Friday, instead of varying the day of observation every week. Future studies could validate our findings, for instance by observing apps more days in a week in a shorter time span of observation.

Finally, in addition to these avenues for improvement, there are a plethora of potential directions for future research on the app market as research in this field has just started. Particularly, the study of developers' entry strategies and the analysis of the platform choice for new app launch seem to be very promising directions for future research.

## Acknowledgments

The authors would like to thank the Editor-in-Chief Prof. R. J. Kauffman, the Senior Editor Prof. Maurizio Naldi, and the four anonymous reviewers for their valuable and insightful comments, which helped improve the paper significantly.

## References

Afuah, A., Tucci, C., 2003. Internet Business Models and Strategies. McGraw-Hill, Boston.
Amit, R., Zott, C., 2001. Value creation in e-business. Strateg. Manag. J. 22 (2), 493–520.
Amit, R., Zott, C., 2008. The fit between product market strategy and business model: implications for firm performance. Strateg. Manag. J. 29 (1), 1–26.
Anderson, C., 2009. Free: The Future of a Radical Price. Hyperion, New York.
Bhargava, H.K., Choudhary, V., 2001. Information goods and vertical differentiation. J. Manage. Inf. Syst. 18 (2), 89–106.
Bhargava, H.K., Choudhary, V., 2008. When versioning is optimal for information goods? Manage. Sci. 54 (5), 1029–1035.
Brynjolfsson, E., Hu, Y., Smith, M.D., 2003. Consumer surplus in the digital economy: estimating the value of increased product variety at online booksellers. Manage. Sci. 49 (11), 1580–1596.
Canzer, B., 2006. E-business: Strategic Thinking and Practice. Houghton Mifflin Company, Boston - New York.
Carare, O., 2012. The impact of bestseller rank on demand: Evidence from the app market. Int. Econ. Rev. 53 (3), 717–742.
Chai, K., Potdar, V., Chang, E., 2007. A survey of revenue models for current generation social software's systems. Lecture Notes in Computer Science, vol. 4704/2007. Springer, Berlin, pp. 724–38.
Chang, M.K., Cheung, W., Lai, V.S., 2005. Literature derived reference models for the adoption of online shopping. Inf. Manage. 42 (4), 543–559.
Chen, R., Gal-Or, E., Roma, P., 2014. Opaque distribution channels for competing service providers: Posted Price vs. Name-Your-Own-Price mechanisms. Oper. Res. 62 (4), 733–750.
Cheng, H.K., Li, S., Liu, Y., 2015. Optimal software free trial strategy: limited, time-locked or hybrid? Prod. Oper. Manage. 24 (3), 504–517.
Chevalier, J., Goolsbee, A., 2003. Measuring prices and price competition online: Amazon.com and BarnesandNoble.com. Quant. Market. Econ. 1 (2), 203–222.
Clemons, E.K., 2009. Business models for monetizing Internet applications and web sites: experience, theory, and predictions. J. Manage. Inf. Syst. 26 (2), 15–41.
Deloitte, 2016. Photo sharing: trillions and rising. Retrieved on February 11, 2016. http://www2.deloitte.com/global/en/pages/technology-media-and-telecommunications/articles/tmt-pred16-telecomm-photo-sharing-trillions-and-rising.html.
Edwards, J., 2014. These maps show that Android is for people with less money. Business Insider. Retrieved on September 24, 2015. http://www.businessinsider.com/android-is-for-poor-people-maps-2014-4?IR=T.
Enders, A., Hungenberg, H., Denker, H.-P., Mauch, S., 2008. The long tail of social networking: revenue models of social networking sites. Eur. Manage. J. 26 (3), 199–211.
Gallaugher, J.M., Wang, Y., 2002. Understanding network effects in software markets: evidence from web server pricing. MIS Q. 26 (4), 303–327.
Gallaugher, J.M., Auger, P., BarNir, A., 2001. Revenue streams and digital content providers: an empirical investigation. Inf. Manage. 38 (7), 473–485.
Garg, R., Telang, R., 2013. Inferring app demand from publicly available data. MIS Q. 37 (4), 1253–1264.
Gartner, 2013. Gartner says mobile app stores will see annual downloads reach 102 billion in 2013. Retrieved on September 24, 2015 from http://www.gartner.com/newsroom/id/2592315.
Ghose, A., Han, S.P., 2014. Estimating demand for mobile applications in the new economy. Manage. Sci. 60 (6), 1470–1488.

Gordon, M.E., 2013. The history of app pricing, and why most apps are free. Retrieved on September 24, 2015 from http://flurrymobile.tumblr.com/post/115189750715/the-history-of-app-pricing-and-why-most-apps-are.
Grewal, D., Iyer, G.R., Levy, M., 2004. Internet retailing: enablers, limiters and market consequences. J. Bus. Res. 57 (7), 703–713.
Hagiu, A., 2007. Merchant or two-sided platform? Rev. Network Econ. 6 (2), 115–133.
Halaburda, H., Gans, J., Burbank, N., 2011. Developing an App for That: Mobile Application Strategy (Banking)" HBS Case (711–415) and Teaching Note (711–535).
Halbheer, D., Stahl, F., Koenigsberg, O., Lehmann, D.R., 2014. Choosing a digital content strategy: how much should be free? Int. J. Res. Mark. 31 (2), 191–206.
Hsu, C.-L., Lin, J.C.-C., 2015. What drives purchase intention for paid mobile apps? – An expectation confirmation model with perceived value. Electron. Commer. Res. Appl. 14 (1), 46–57.
Korgaonkar, P., Silverblatt, R., Girard, T., 2006. Online retailing, product classifications, and consumer preferences. Internet Res. 16 (3), 267–288.
Kuo, Y., Yu, C., 2006. 3G Telecommunication operators' challenges and roles: a perspective of mobile commerce value chain. Technovation 26 (12), 1347–1356.
Laudon, K.C., Traver, C.G., 2007. E-Commerce: Business, Technology, Society. Pearson Prentice Hall, New Jersey.
Levin, A.M., Levin, I.P., Heath, C.E., 2003. Product category dependent consumer preferences for online and offline shopping features and their influence on multi-channel retail alliances. J. Electron. Commerce Res. 4 (3), 85–93.
Lunden, I., 2013. Gartner 102B app store downloads globally in 2013, $26B in sales, 17% from in-app purchases. Retrieved on September 24, 2015. From http://techcrunch.com/2013/09/19/gartner-102b-app- store-downloads-globally-in-2013-26b-in-sales-17-from-in-app-purchases/.
Moore, G.C., Benbasat, I., 1991. Development of an instrument to measure the perceptions of adopting an information technology innovation. Inf. Syst. Res. 2 (3), 192–222.
Munir, A., 2014. App monetization: 6 bankable business models that help mobile apps make money. Retrieved on September 24, 2015 from http://info.localytics.com/blog/app-monetization-6-bankable-business-models-that-help-mobile-apps-make-money.
Nerney, C. 2014. The Android one could gut the market for high-end Android phones. Retrieved on September 24, 2015. http://www.citeworld.com/article/2604668/mobile-byod/the-android-one-could-gut-the-market-for-high-end-android-phones.html.
Niculescu, M.F., Wu, D.J., 2011. When should software firms commercialize new products via freemium business models? Working Paper, Georgia Institute of Technology.
Pappas, A., 2013. Which apps make more money. Retrieved on September 24, 2015 from http://www.visionmobile.com/blog/2013/04/which-apps-make-more-money/.
Parker, G., Van Alstyne, M., 2005. Two-sided network effects: a theory of information product design. Manage. Sci. 51 (10), 1494–1504.
Pauwels, K., Weiss, A., 2008. Moving from free to fee: how online firms market to change their business model successfully. J. Market. 72 (3), 14–31.
Phau, I., Poon, S.M., 2000. Factors influencing the types of products and services purchased over the internet. Internet Res. 10 (2), 102–113.
Reibstein, D.J., 2002. What attracts customers to online stores, and what keeps them coming back? J. Acad. Mark. Sci. 30 (4), 465–473.
Rochet, J.C., Tirole, J., 2003. Platform competition in two sided markets. J. Eur. Econ. Assoc. 1 (4), 990–1029.
Rogers, E.M., 1983. Diffusion of Innovations. Free Press, New York.
Roma, P., Di Martino, G., Perrone, G., 2013. What to show on the wine labels: a hedonic analysis of price drivers of Sicilian wines. Appl. Econ. 45 (19), 2765–2778.
Roma, P., Zambuto, F., Perrone, G., 2014. Price dispersion, competition, and the role of online travel agents: Evidence from business routes in the Italian airline market. Transport. Res. Part E 69, 146–159.
Roma, P., Zambuto, F., Dominici, G., 2015. An empirical analysis of online price dispersion in the Italian airline industry. Int. J. Electron. Market. Retailing 6 (4), 339–358.
Rosenbaum, P., Rubin, D., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70 (1), 41–55.
Rosenbaum, P., Rubin, D., 1984. Reducing bias in observational studies using subclassification on the propensity score. J. Am. Stat. Assoc. 79 (387), 516–524.
Shapiro, C., Varian, H., 1999. Information Rules: A Strategic Guide to the Network Economy. Harvard Business Press, Cambridge.
Sims, G., 2015. Google Play Store vs. the Apple App Store: by numbers. Retrieved on September 24, 2015 from http://www.androidauthority.com/google-play-store-vs-the-apple-app-store-601836/.
Sourcebits, 2014. Choosing your app monetization strategy. Retrieved on September 24, 2015 from http://sourcebits.com/app-development-design-blog/choosing-app-monetization-strategy/.
Takahashi, D., 2014. Mobile apps could hit $70 billion in revenues by 2017. Retrieved on September 24, 2015 from http://venturebeat.com/2014/04/29/mobile-apps-could-hit-70b-in-revenues-by-2017-as-non-game-categories-take-off/.
Teece, D.J., 2010. Business models, business strategy and innovation. Long Range Plan. 43 (2–3), 172–194.
Wagner, T.M., Benlian, A., Hess, T., 2014. Converting freemium customers from free to premium – the role of the perceived premium fit in the case of music as a service. Electron. Markets 24 (4), 259–268.
Wallenstein, A., 2015. How Google Play is closing the app gap with Apple. Retrieved on September 24, 2015 from http://variety.com/2015/digital/features/google-play-apple-itunes-app-store-1201449326/.
Wang, C.L., Zhang, Y., Ye, L.R., Nguyen, D.D., 2005. Subscription to fee-based online services: what makes consumer pay for online content? J. Electron. Commerce Res. 6 (4), 304–311.
Wei, X., Nault, B.R., 2013. Experience information goods: "version-to-upgrade". Decis. Support Syst. 56, 494–501.
Wei, X., Nault, B.R., 2014. Monopoly versioning of information goods when consumers have group tastes. Prod. Oper. Manage. 23 (6), 1067–1081.
Wilcox, 2013. Are you using the right app revenue model? Retrieved on September 24, 2015 from http://www.developereconomics.com/apps-using-right-revenue-models/.
Wirtz, B.W., Schilke, O., Ullrich, S., 2010. Strategic development of business models: implications of the Web 2.0 for creating value on the Internet. Long Range Plan. 43 (2–3), 272–290.
Wooldridge, J.M., 2002. Econometric Analysis of Cross Section and Panel Data. The MIT Press, Cambridge.
Wu, S., Chen, P., 2008. Versioning and piracy control for digital information goods. Oper. Res. 56 (1), 157–172.
Zott, C., Amit, R., Massa, L., 2011. The business model: recent developments and future research. J. Manage. 37 (4), 1019–1042.