Spring 6-10-2017

THE EFFECT OF TEXTUAL PRODUCER-GENERATED DESCRIPTIONS ON DEMAND OF MOBILE APPLICATIONS

Michael Scholz
University of Innsbruck, Austria, Michael.Scholz@uibk.ac.at

Lauri Frank
University of Jyväskylä, Finland, lauri.frank@jyu.fi

THE EFFECT OF TEXTUAL PRODUCER-GENERATED DESCRIPTIONS ON DEMAND OF MOBILE APPLICATIONS

Research paper

Scholz, Michael, University of Innsbruck, Austria, Michael.Scholz@uibk.ac.at

Frank, Lauri, University of Jyväskylä, Finland, lauri.frank@jyu.fi

Abstract

We analyze the impact of different app description characteristics on app demand on the basis of panel data for six months and 1081 distinct apps. We use several text mining techniques in order to operationalize the descriptions' textual characteristics. The extracted variables are then used in an econometric investigation to examine their impact on apps' downloads. Our results provide evidence that app descriptions have an effect on demand. Apps with upfront price should be described in a neutral tone. Apps without an upfront price but with in-app purchase option should be offered with rather short descriptions that are written in a formal and subjective style.

Keywords: Mobile Apps, App Descriptions, Demand Estimation, Text Mining.

1 Introduction

Smart phones showed a high growth rate in the last years (Srivastava and Misra 2016) which can be attributed to the mass of available smart phone apps and the resulting utility of these apps. Based on a report by comScore (2016), around 197.4 million U.S. citizens owned a smart phone in December 2015. This corresponds to a penetration rate of 79.3%. The high penetration of smart phones has attracted many app developers. According to Statista (2017), around 2.2 million apps were available for iOS smart phones and approximately 2.8 million apps were available for Android smart phones in March 2017. Recent research, however, shows that only a few of these apps generate really high revenues that allow to pay a team of professional app developers (Garg and Telang 2013).

Researchers and especially app developers are thus interested in identifying the determinants of app demand. Recent research has analyzed the effect of several variables such as price, customer reviews and an app's file size on an app's demand (Ghose and Han 2014; Lee and Raghu 2014). Most of these variables are exogenous and hard to control for an app developer. Price, for example, is set with respect to the prices of competing developers, the file size is a result of the development process and thus not an endogenous variable, and customer reviews are written based on the customers' experiences. Mobile app developers mainly control their app's pricing model, features and how these features are described. The description of an app is hence an endogenous variable and can be modified by app developers in order to improve their apps' demand. Evidence on whether and how an app's textual description influences its demand is missing so far[1]. There is, however, evidence that textual producer-generated descriptions influence the price and the demand of products sold online (Dimoka et al. 2012).

[1] Note that there exist several websites that offer guidelines on how to write an app description. However, there is no scientific evidence for these guidelines and no scientific publication investigating the impact of app descriptions on demand.

Apps can be offered either with an upfront price or for free. In order to generate revenues with free apps, developers can embed ads in their apps or offer the opportunity to buy additional features. Such in-app purchases (IAP) are a variant of the freemium strategy (Liu et al. 2014); consumers can use the app for free, but need to pay a premium price when they are interested in some special features. Recent research has shown that the pricing strategy has an impact on an app's demand and the effect of other variables, such as price discounts, is moderated by the pricing strategy (Ghose and Han 2014).

We investigate the effect of textual app descriptions on app demand depending on the pricing strategy in order to answer the following research question:

RQ: What is the effect of textual app descriptions on app demand and does this effect vary between apps with upfront price and apps with IAP option?

More specifically, we use several linguistic measures (e.g., readability, polarity and formality) to characterize app descriptions and to identify those characteristics of app descriptions that affect an app's demand. We compare the impact of the linguistic description characteristics between apps with an upfront price and apps with IAP option. We use a panel data set that includes the top 200 ranked apps in each of the two pricing options. Several data about the apps were collected from appshopper.com for six month. Our results show that the polarity of a textual description significantly influences an app's demand if this app is offered with an upfront price. Description length, readability, formality and objectivity have no effect on the demand of apps with upfront price. Polarity does not affect the demand of apps with IAP option. Description length, readability, formality and objectivity, however, show a marginal effect on the demand of apps with IAP option. We also show that the number of substitutes for an app has a significant positive effect on the demand of apps with upfront price but a significant negative effect on the demand of apps with IAP option.

The remainder of the paper is organized as follows. We review related work in the next section and explain how this paper complements existing research. Section 3 presents our research model which we investigate with data from Apple's app market. Collection and preparation of our data are described in Section 4. Section 5 presents the empirical evaluation of our model. We discuss the results with implications for research and practice in Section 6.

2 Related Work

Various studies from marketing, information systems, and economics are of interest to our study. Specifically, two streams appear relevant to investigate the impact of textual producer-generated descriptions on app demand: first, studies that investigate the impact of textual product-related descriptions or opinions on the economic success of products; and second, studies that analyze the impact of several determinants of an app's demand.

2.1 Effect of Product-Related Descriptions on Demand

Studies investigating the economic impact of textual product-related descriptions help to identify text features that might determine an app's demand. Consumers judge the quality of an app by evaluating product information (Netzer and Srinivasan 2011). The length of product-related texts has been used as proxy for the amount of information that is transported with the texts (Ghose and Ipeirotis 2011; Mudambi and Schuff). Longer texts are assumed to be more informative. Mudambi and Schuff (2010), for example, found that the text length of customer reviews is positively correlated to reviews' perceived helpfulness.

The polarity or sentiment in which a product-related text is written has often been found to have an economic effect (Ludwig et al. 2013; Sonnier et al. 2011; Chintagunta et al. 2010). The polarity of a user-generated text indicates the satisfaction with the product whereas the polarity of a marketer-generated text shows how strong the provider is going to advertise its product (Scholz et al. 2013). Sonnier et al. (2011), for instance, cite evidence that neutral and positive comments on products have a positive effect on sales in short as well as in long term. Negative comments have, as expected a negative effect on sales. Scholz et al. (2013) show that positive and neutral user comments on Facebook fanpages of firms positively affect the number of Facebook users, but only neutral comments do have an effect on the conversion rate of those Facebook users that went to the online store of the firm.

Ghose and Ipeirotis (2011) as well as Scholz and Dorner (2013) demonstrated that the readability of online customer reviews influences the reviews' perceived helpfulness. Readability here expresses the reading level a recipient of a text must have to understand the text. The lower the reading level the easier a text can be understood also by younger consumers.

Other text characteristics such as objectivity, entropy, the number of spelling errors or the amount of adjectives have been also used in order to investigate economic effects of product-related texts such as customer reviews (Scholz and Dorner 2013; Ghose and Ipeirotis 2011). In summary, recent research has developed and applied several measures for textual characteristics that help analyzing the impact of textual producer-generated descriptions, customer reviews and social media comments on demand.

The findings of these studies indicate that textual product-related descriptions are used by consumers in order to diagnose a product's quality prior to purchase. Most of these findings, however, rely on the investigation of consumer- rather than producer-generated textual descriptions.

2.2 Determinants of App Demand

A few studies already investigated the impact of several determinants on the demand of smart phone apps. Carare (2012) analyzed the impact of app rankings on demand using a reduced-form model. Garg and Telang (2013) showed how app demand can be inferred from publicly available app rankings from Apple App Store and Google Play. Ghose and Han (2014) investigated the effect of several mobile app characteristics and demonstrated that demand increases with IAP option but decreases with in-app advertisement. The authors further found that a price discount is more effective in Google Play compared to Apple App Store. File size is negatively correlated to app demand whereas the number of screenshots for an app as well as the description length both positively influence app demand (Ghose and Han 2014). Lee and Raghu (2014) analyzed the effect of different factors on the survival of an mobile app in the top-grossing 300 chart. They provide evidence that offering apps across multiple categories is one of the most important success factors. Apps without an upfront price have a higher probability to survive in the top 300 charts than apps with an upfront price (Lee and Raghu 2014). Review volume as well as review valence have been also found to positively influence survival in the top 300 (Lee and Raghu 2014) and app demand (Ghose and Han 2014). Salo et al. (2013a) model the adoption of mobile apps by using the diffusion of innovations (DOI) concepts and conclude relative advantage, ease of use and observability as being significant for the use intention. Salo et al. (2013b) investigate camera-based applications and show that the use of mobile services may have both utilitarian and hedonic reasons. Liu et al. (2014) provide evidence that the free version of an app improves demand of the paid version if the free version gets a lot of positive customer ratings. Khalid et al. (2015) showed that app consumers most often complain about functional errors but stop using an app mostly due to privacy and ethics concerns and hidden costs. Salo and Frank (2015) conclude that users' complaints of their negative experiences of mobile use depend on the situational context: Indoor negative experiences are more often communicated than experiences outdoors or in vehicles.

Recent research has shown evidence for the influence of several factors on app demand. To our knowledge, no previous study has investigated the impact of textual producer-generated descriptions (except the length of the descriptions) on app demand. With respect to the findings of the first stream of studies, textual producer-generated descriptions as another information source might also be used prior to purchase in order to better diagnose the quality of apps. This paper thus contributes to recent research by using textual features, such as polarity and readability, to analyze how app descriptions affect demand.

3 Research Model

Apps having an upfront price force consumers to thoroughly deliberate on whether they should invest money in the app or not. Quality cues such as app descriptions hence might be more important for apps with upfront price. Apps without an upfront price but IAP option can be seen as rather search than experience goods because consumers can diagnose such apps' quality before investing money. App descriptions for these apps might not be so important and consumers might expect descriptions that are shorter and easier to understand for apps without an upfront price but an IAP option. Since longer texts are likely to transport more or deeper product information (Mudambi and Schuff 2010), we hypothesize that the demand of apps with upfront price increases with the description length.

H1a: The longer an app's description, the higher are its demand if the app has an upfront price.

Descriptions of apps without an upfront price but IAP option do not need to transport much information about the app's quality. These apps' descriptions should only inform a prospective consumer about the apps' features. Apps with IAP option are for free, so that consumers might prefer diagnosing the quality of these apps by directly testing and using the apps rather than reading lengthy product descriptions. Online consumers, for example, value a free trial even more than a money back guarantee to diagnose the quality of a low or medium risk product (Tan 1999). A textual producer-generated description only has the purpose to allow consumers to identify those apps that fit to their need (e.g., scanning PDF files, reading twitter messages). We thus hypothesize:

H1b: The shorter an app's description, the higher are its demand if the app has no upfront price but an IAP option.

Since app descriptions are written by the app providers (the developer in most cases), they are unlikely to be negative. Strongly positive descriptions might be interpreted as rhapsodic and not very confident (Lee and Ma 2012). We hence hypothesize that neutral app descriptions lead to a higher app demand than positive app descriptions independent of the app's pricing strategy.

H2: Apps with a neutral description attract larger demand than apps with a positive description.

Several mobile apps are dedicated to children and teenagers. The GSM Association and the Mobile Society Research Institute within NTT DOCOMO Inc. reported in 2013 that 30.8% of all children that participated in their survey download or use mobile apps (GSMA and NTT DOCOMO 2013). A lower readability level thus makes it easier for this group of consumers to understand the content of a mobile phone app. This might ultimately result in a higher number of consumers downloading an app. We thus hypothesize:

H3: The lower the readability level of an app's description, the higher is the app's demand.

Texts that should be unambiguously understandable have been found to be written in a rather formal style with many nouns, adjectives, prepositions and articles (Heylighen and Dewaele 2002). Texts including many verbs, adverbs, pronouns and interjections are typically only understandable in the presence of contextual information (Heylighen and Dewaele 2002). App descriptions that are written in a formal style might hence be rather easy to understand since only low contextual information (e.g., technical information) are required. Hence, we hypothesize:

H4: The higher the formality of an app's description, the higher is the app's demand.

A more subjective style has been found to increase the helpfulness of product reviews (Scholz and Dorner 2013). A review was here defined to be more subjective if it contains a higher amount of personal and possessive pronouns. A higher amount of such pronouns therefore increases the perceived subjectivity of an app's description. A text that is written in a subjective style better transports the spirit and opinion of the producer of the text. Prospective consumers might find an app description in rather a subjective style more confident. We thus hypothesize that a higher subjectivity of an app's description leads to a higher demand.

H5: The more subjectively an app's description is written, the higher is the app's demand.

In the next section, we will describe the data collected in order to evaluate the proposed hypotheses.

4 Data Collection and Preparation

4.1 Data Collection

We collected panel data from appshopper.com – one of the largest directories for iPhone, iPad and Mac apps that is independent of Apple – in order to answer our research question: What is the effect of textual app descriptions on app demand and does this effect vary between apps with upfront price and apps with IAP option? Appshopper.com provides daily updated lists of the top 200 Apple iPhone apps. We gathered data every second day from the list of the top apps with upfront price and the chart list of apps with IAP option between October 28, 2014 and April 25, 2015 (6 months). The data include 690 distinct apps with upfront price and 391 distinct apps with IAP option and the number of observations totals to 25,172. Our data set includes each app's sales rank, title, description, price, and the number of days the app has been listed as a top-200 app. Sales ranks are based on download data that are not publicly available.

We rely on log data in order to avoid several Type-1 and Type-2 errors one would expect when using survey data about the usage and effect of mobile app descriptions (de Reuver and Bouwman 2015).

4.2 Data Preparation

We used the estimation method proposed by Garg and Telang (2013) to estimate app demand based on app sales ranks. More specifically, the demand for an app i is calculated as

$$Demand = 52{,}958 \cdot Rank_i^{-0.944} \tag{1}$$

App descriptions were characterized by a number of variables including length, polarity, readability, formality, and objectivity. These variables were automatically extracted using text mining techniques (Kostoff and Geisler 1999). The length of an app's description was measured based on the word count and the sentence count.

We measured the polarity of app descriptions based on the sentiment dictionary by Hu and Liu (2004). Contexts C_w around each identified polarized word w are then formed based on the four words before w and the two words after w. Each word in a context was classified as neutral (w^{neu}), negator (w^{neg}), amplifier (w^{amp}) or de-amplifier (w^{dea}). If C_w contains a comma, only those words between the polarized word w and the comma are defined as context. The polarity of a description for app i is finally the sum of the polarity of its contexts divided by the square root of the word count n (Rinker 2013).

$$Polarity_i = \frac{\sum_{\forall C_w \in i} Polarity_{C_w}}{\sqrt{n}} \tag{2}$$

with

$$Polarity_{C_w} = \left(1 + \gamma\left(x^{amp} - x^{dea}\right)\right) Polarity_w (-1) \sum_{w \in C_w} w^{neg}$$
$$x^{amp} = \sum_{w \in C_w} x^{neg} w^{amp}$$
$$x^{dea} = \max\left(\sum_{w \in C_w} - x^{neg} w^{amp} + w^{dea}, -1\right)$$
$$x^{neg} = \left(\sum_{w \in C_w} w^{neg}\right) \text{modulo } 2$$

For estimating a description's readability, we chose ARI (automated readability index) which has been widely used to classify the readability of texts (Sawyer et al. 2008). ARI uses the number of characters, words, and sentences to form a readability score that approximates the grade that is needed to understand the underlying text.

$$ARI_i = 4.71 \frac{characters_i}{words_i} + 0.5 \frac{words_i}{sentences_i} - 21.43 \tag{3}$$

Grade level 1 (lowest level) corresponds to ages 6 to 8 whereas grade 12 (highest level) corresponds to the reading level of a 17 years old pupil.

Nouns, adjectives, prepositions, and articles typically occur in texts that require low context information to be correctly understood by a reader. These texts are classified as formal texts (Heylighen and Dewaele 2002). Informal texts (i.e., texts requiring large contextual information), in contrast, are characterized by many pronouns, adverbs, verbs, and interjections. Based on the estimation method by (Heylighen and Dewaele 2002), we approximate the formality of an app's description as follows.

$$Formality_i = 0.5 \left(\frac{formals_i - contextuals_i}{formals_i + contextuals_i + conjections_i} + 1 \right) \tag{4}$$

with

$$formals_i = nouns_i + adjectives_i + prepositions_i + articles_i$$
$$contextuals_i = pronouns_i + adverbs_i + verbs_i + interjections_i$$

Objectivity is computed as 1 minus the proportion of personal and possessive pronouns (Scholz and Dorner 2013).

$$Objectivity_i = 1 - \frac{personal\_pronouns_i + possessive\_pronouns_i}{words_i} \tag{5}$$

Table 1 shows that apps with an upfront price attract significantly less consumers ($p < 0.001$) than apps without an upfront price but an IAP option. At an average price of $US 2.59 developers gain on average $US 5366 per day with an app that is listed in the top 200. Apps with IAP option are only more profitable if each consumer of such an app spends at least 86 cents in in-app purchases on average. Garg and Telang (2013) cite evidence that the revenue gained from an app without an upfront price is only 16 cents per download based on in-app purchases. This leads to an estimated average revenue of $US 999 per download and day for apps with IAP option that are listed in the top 200.

Variable	Apps with Upfront Price		Apps with IAP Option	
	Mean	SD	Mean	SD
Demand	2071.69	5069.71	6244.37	12729.40
Upfront Price (in $US)	2.59	2.04	0.00	0.00
Word Count	307.37	158.11	295.39	118.49
Sentence Count	23.95	16.02	23.67	10.93
Polarity	0.42	0.37	0.49	0.35
Readability	10.29	4.03	9.34	2.71
Formality (in %)	69.68	5.77	70.39	5.29
Objectivity (in %)	94.10	2.66	94.42	2.15
Observations	12,932		12,240	

Table 1. Summary of app characteristics

App descriptions differ only marginally between apps with upfront price and apps with IAP option. Descriptions of apps with IAP option are slightly easier to read than descriptions of apps with upfront price. The readability of descriptions for apps with IAP options corresponds to the reading level of a 15 years old U.S. teenager whereas the readability of descriptions for apps with upfront price corresponds to the reading level of a 16 years old teenager.

In the next section, we analyze the effect of the app descriptions on the demand of apps and compare this effect of the two different pricing strategies (i.e., upfront price and IAP option).

5 Analysis and Results

5.1 Analysis

In this subsection, we specify the linear mixed effects model used to analyze the effects of app descriptions on demand. We take the natural logarithm of the apps' demand as dependent variable because an absolute difference in the downloads of two apps becomes less important if the absolute number of downloads of both apps increases.

The demand of an app i might depend on the app's demand in the past due to a bandwagon effect (Carare 2012): the more an app has been downloaded, the more consumers assume a high quality of this app and also download it. In order to capture this bandwagon effect, we include the first-order lag of the demand and the natural logarithm of the number of days app i is listed in the top 200 as independent variable.

The price p_i of an app has been shown to reduce an app's utility for consumers and ultimately an app's demand (Ghose and Han 2014). An app's demand might also depend on the average price $\bar{p}_{i'}$ of those apps i' that are substitutes to i and the number of substitutes $\lambda_{i'}$. We identified substitutes based on app descriptions. More specifically, we generated a vector of bi-grams for each app's description and used these bi-gram vectors as input for a hierarchical clustering. Hierarchical clustering was implemented with average linkage function and the cosine measure was used to compute distances between the bi-gram vectors. We conducted several such hierarchical clusterings using multiscale bootstrap resampling in order to get robust app clusters (Shimodaira 2004; Scholz et al. 2016).

The characteristics of an app's description X_i include the description's length, readability, polarity, formality, and objectivity as explained in Section 4. The resulting regression model is given as follows.

$$Log(Demand_{i,t}) = \alpha + \beta_1 Log(Demand_{i,t-1}) + \beta_2 Log(Days_{i,t}) + \beta_3 p_{i,t} + \beta_4 \bar{p}_{i',t} + \beta_5 \lambda_{i',t} + \gamma X_i + \varepsilon_{i,t} \quad (6)$$

We estimated this model with a restricted maximum likelihood approach and random intercepts on the clusters of apps that are substitutes to each other. Variance inflation factors below 3 indicate absence of multicollinearity in our model. The results are presented and discussed in the next subsection.

5.2 Results

The results of our econometric model (see Table 2) show that an app's demand at t significantly and positively depends on the demand at the previous collection time $t-1$. The number of days an app i is listed as a top 200 app also positively affects the demand of i. Both effects can be interpreted as a bandwagon effect. Consumers download apps just because of a high download volume in the past.

Variable	Apps with Upfront Price		Apps with IAP Option	
	Estimate	Std. Error	Estimate	Std. Error
Intercept	0.738***	0.220	1.530***	0.191
Log(L_1(Demand))	0.815***	0.005	0.876***	0.005

Log(Days in Top 200)	0.028***	0.002	0.023***	0.002
Upfront Price	-0.011***	0.002	–	–
Number of Substitutes	0.002***	0.000	-0.001***	0.000
Average Price of Substitutes	0.017*	0.006	-0.001	0.002
Length	0.001	0.001	-0.001*	0.000
Polarity	-0.054***	0.012	0.001	0.010
Readability	0.001	0.001	-0.002	0.001
Formality	-0.002	0.001	0.003**	0.001
Objectivity	0.238	0.278	-0.587*	0.238
Pseudo R^2	0.864		0.888	
AIC	9,277.671		4,259.590	
BIC	9,374.523		4,341.127	

*Table 2. Regression results (*p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001)*

The higher an app's upfront price, the lower the daily demand that is generated independent of the bandwagon effect. Reducing an app's price by one $US improves the logarithm of the app's demand by 0.011. Figure 1 shows the effect of price changes on the demand of an app that is downloaded 1,000 times a day for 2.99 Euro (dashed lines in Figure 1).

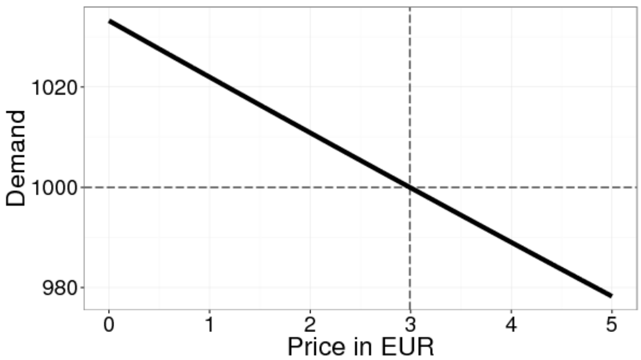

Figure 1. Effect of Price Changes (based on an app with 1,000 downloads a day at $US 2.99).

We found a significant effect of the number of apps i' that are similar to i (i.e., number of substitutes) on the demand of i. The demand of apps with upfront price benefits from more substitutes whereas the reversed is true for apps with IAP option[2]. An app's substitutes have been identified based on the app's description and independent of its pricing strategy. Substitutes of apps with IAP option can hence have an upfront price. We found an average price of $US 2.81 of the substitutes of apps with upfront price and an average price of only $US 0.93 of the substitutes of apps with IAP option. We

[2] The confidence interval of the estimate for the number of substitutes does not include 0 for both the model for apps with upfront price and the model for apps with IAP option.

also found that on average 83% of the substitutes of apps with IAP option do not have an upfront price whereas only 0.01% of the substitutes of apps with an upfront price do not have an upfront price. Apps with an upfront price hence benefit from more competitors that typically have a high price whereas apps with IAP option and no upfront price do benefit from less competitors that typically also offer their app without an upfront price. The average upfront price of the substitutes consequently has a positive effect on the demand of those apps that are offered with an upfront price.

Description length only influences the demand of apps without upfront price but IAP option. We hence must reject H1a but can support H1b. Apps with a neutral description attract more customers than apps with a positive description in the case of an upfront price model. If 1,000 consumers by an app with upfront price that is positively written, there will be approximately 26 additional consumers for this app if the description will be formulated in a neutral tone. H2 is thus partially supported. Readability has been found to not significantly influence an app's demand. H3 is thus rejected. Formality and objectivity have been found to influence app demand as proposed for apps with IAP option. H4 and H5 are hence partially supported.

6 Discussion

We interestingly found a significant negative effect of the polarity of an app's description on the demand of apps with upfront price. Since only 11.1% of apps with upfront price have a negative polarity, this finding does not suggest writing negative app descriptions but not writing too positive descriptions.

Description length, formality and objectivity have been found to have a significant but rather small effect on the demand of apps with IAP option. Providers of apps with IAP option should write rather short app descriptions in a formal and subjective style. Since these apps do not have an upfront price, consumers do not need lengthy and highly informative descriptions to decide about whether to download the app. A more formal style helps consumers to easier understand app descriptions because not much (additional) contextual information is necessary. Several apps are rather hedonic than utilitarian goods. A subjective description better transports the spirit of using the app and more confidentially expresses that the provider of the apps itself is an experienced user of the app.

Our analysis also revealed that the average price of apps that can be categorized as substitutes to a particular app has a positive influence on the demand of the particular app. This is somewhat unexpected, because substitutes are typically characterized by a negative cross elasticity of price on demand (Krugman and Wells 2013). The effect of price on demand is significantly positive ($p < 0.001$) when not controlling for clusters of substitutable apps. This indicates that the cheapest apps with upfront price have rather a low demand whereas more expensier apps have a significantly higher demand. Thus, the higher the price of the substitutes of an app with upfront price, the higher also demand of the particular app.

6.1 Research Implications

Our study contributes in two aspects to recent research on the determinants of mobile app demand. First, we demonstrate the effects of app descriptions on app demand. More specifically, we propose that linguistic characteristics of app descriptions influence demand. In an analysis of mobile apps for Apple's iOS, we show that a description's polarity influences the demand of apps with an upfront price whereas a description's length, formality, and objectivity influences the demand of apps without an upfront price but an IAP option. Although recent research has investigated the impact of several factors such as price, customer reviews and an app's file size (Ghose and Han 2014; Lee and Raghu 2014; Liu et al. 2014), this study is, to the best of our knowledge, the first that investigates the effects of app descriptions.

Second, we used app descriptions to identify substitutes. Recent research has considered apps being in the same category as substitutes to each other (Ghose and Han 2014). Since app categories typically consist of thousands of apps that are often very different[3], categories are not a good proxy for app substitutes. We used an approach (multiscale bootstrap resampling, Shimodaira 2004) that identifies clusters of apps having similar app descriptions. This approach allows identifying substitutes of an app independent of product categories and based on only app features that are typically explained in app descriptions. Based on the identified substitutes we were furthermore able to integrate the number of substitutes and the average price of substitutes in our econometric model.

6.2 Managerial Implications

Our results offer interesting insights for app providers. App providers should carefully write descriptions for their apps and in case of apps with upfront price avoid a too positive tone. Description length does not play a role for consumers' download decisions. When consumers have to pay for an app, they are willing to read also rather not easy readable and lengthy descriptions in order to diagnose the app's quality prior to purchase.

This is not the case with apps without an upfront price but an IAP option. Consumers of these apps have the possibility to judge an app's quality by using the app without the risk of loosing money. Descriptions for apps without an upfront price but an IAP option should be rather short and written in a formal and subjective style.

We further found that competition (i.e., the number of substitutes) helps apps with an upfront price to attract more consumers but decreases the demand of apps without an upfront price. App providers can use app descriptions in order to identify their competitors and in combination with the demand estimation approach by Garg and Telang (2013) to evaluate their competitors' market share.

6.3 Limitations

Our study has some limitations that provide avenues for further research. Since we have no indication about the consumers' real usage of app descriptions, we can not be sure that all consumers read app descriptions before making a download decision. Several consumers might hence use other information sources than the description to judge an app's quality and to decide whether to download the app. For example, some new games might just be downloaded because of their names or consumers might download an app of which they heard from their friends. However, we have no indication that such consumers vary systematically across our analyzed apps. We rather can assume that such consumers have equally influenced the demand of the top 200 apps. Future research should analyze to what extent consumers read and adopt app descriptions in order to get further insights in the impact of app descriptions on app demand.

We collected data only for iPhone apps. Tablets such as Apple's iPad have larger screens and make it much easier to read especially lengthy app descriptions. Investigating the impact of app descriptions on the demand of tablet apps might hence provide an interesting avenue for further research.

An app's readibility might be especially of importance for younger people having a rather low educational level so far. Since we only have an approximation of the total demand, we are not able to test our hypotheses for specific age groups.

The demand of apps might be also determined by some unobservable factors such as the sorting and filtering algorithms used on app stores. The higher an app is ranked in an app store, the more likely consumers will buy this app. How an app's ranking position will be determined is, however, strictly confidential.

---

[3] The app category "Music" in Apple's iTunes App Store, for example, consists of apps for playing music, learning a music instrument, calibrating music instruments, editing music files, and recording music.

References

Carare, O. (2012). The impact of bestseller rank on demand: Evidence from the app market. *International Economic Review* 53 (3), 717–742.

Chintagunta, P.K, Gopinath, S., and Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science* 29 (5), 944–957.

comScore (2016). comScore Reports December 2015 U.S. Smartphone Subscriber Market Share, http://www.comscore.com/Insights/Market-Rankings/comScore-Reports-December-2015-US-Smartphone-Subscriber-Market-Share.

De Reuver, M. and Bouwman, H. (2015). Dealing with self-report bias in mobile internet acceptance and usage studies. *Information & Management* 52 (3), 287–294.

Dimoka, A., Hong, Y. and Pavlou, P.A. (2012). On product uncertainty in online markets: Theory and evidence. *MIS Quarterly* 36 (2), 395–426.

Garg, R. and Telang, R. (2013). Inferring app demand from publicly available data. *MIS Quarterly*, 37 (4), 1253–1264.

Ghose, A. and Han, S.P. (2014). Estimating demand for mobile applications in the new economy. *Management Science* 60 (6), 1470–1488.

Ghose, A. and Ipeirotis, P.G. (2011). Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE Transactions on Knowledge and Data Engineering* 23 (10), 1498–1512.

GSMA and NTT DOCOMO. (2013). Children's use of mobile phones: An international comparison 2012. Technical report.

Heylighen, F. and Dewaele, J.-M. (2002). Variation in the contextuality of language: An empirical measure. *Foundations of Science* 7 (3), 293–340.

Hu, M. and Liu, B. (2004). Mining opinion features in customer reviews. In *AAAI'04 Proceedings of the 19th National Conference on Artificial Intelligence*, 755–760.

Khalid, H., Shihab, E., Nagappan, M., and Hassan, A.E. (2015). What do mobile app users complain about? *IEEE Software* 32 (3), 70–77.

Kostoff, R.N. and Geisler, E. (1999). Strategic management and implementation of textual data mining in government organizations. *Technology Analysis & Strategic Management* 11(4), 493–525.

Krugman, P. and Wells, R. (2013). *Microeconomics*, 3rd edition, Worth Publishers.

Lee, G. and Raghu, T.S. (2014). Determinants of mobile apps' success: Evidence from the app store market. *Journal of Management Information Systems* 31 (2), 133–169.

Lee, H.-H. and Ma, Y.J. (2012). Consumer perceptions of online consumer product and service reviews: Focusing on information processing confidence and susceptibility to peer influence. *Journal of Research in Interactive Marketing* 6 (2), 110–132.

Liu, C.Z., Au, Y.A. and Choi, H.S. (2014). Effects of freemium strategy in the mobile app market: An empirical study of google play. *Journal of Management Information Systems* 31(3), 326–354.

Ludwig, S., de Ruyter, K., Friedman, M., Brüggen, E.C., Wetzels, M. and Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing* 77 (1), 87–103.

Mudambi, S.M. and Schuff, D. (2010). What makes a helpful online review? a study of customer reviews on amazon.com. *MIS Quarterly* 34 (1), 185–200.

Netzer, O. and Srinivasan, V. (2011). Adaptive self-explication of multiattribute preferences. *Journal of Marketing Research* 48 (1), 140–156.

Rinker, T.W. (2013) *qdap: Quantitative Discourse Analysis Package*. University at Buffalo/SUNY, Buffalo, New York.

Salo, M. and Frank, L. (forthcoming). User behaviours after critical mobile application incidents: the relationship with situational context. *Information Systems Journal*.

Salo, M., Olsson, T., Makkonen, M. and Frank, L. (2013). User perspective on the adoption of mobile augmented reality based applications. In Lee, I., editor, *Strategy, Adoption, and Competitive Advantage of Mobile Services in the Global Economy,* pp. 165–188. IGI Global, Hershey.

Salo, M., Olsson, T., Makkonen, M., Hautamäki, A., and Frank, L. (2013). Consumer value of camera-based mobile interaction with the real world. *Pervasive and Mobile Computing* 9 (2), 258–268.

Sawyer, A.G., Laran, J. and Xu, J. (2008). The readability of marketing journals: Are award-winning articles better written? *Journal of Marketing* 72 (1), 108–117.

Scholz, M. and Dorner, V. (2013). The recipe for the perfect review? an investigation into the determinants of review helpfulness. *Business & Information Systems Engineering* 5 (3), 141–151.

Scholz, M., Dorner, V., Landherr, A., and Probst, F. (2013). Awareness, interest, and purchase: The effects of user- and marketer-generated content on purchase decision processes. *Proceedings of the 34th International Conference on Information Systems*.

Scholz, M., Franz, M., and Hinz, O. (2016). The Ambiguous Identifier Clustering Technique – A Method for Unobserved Product Heterogeneity in Online Transaction Data. *Electronic Markets* 26 (2), 143–156.

Shimodaira, H. (2004). Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling. *The Annals of Statistics* 32 (6), 2616–2641.

Sonnier, G.P., McAlister, L., and Rutz, O.J. (2011). A dynamic model of the effect of onlien communications on firm sales. *Marketing Science* 30 (4), 702–716.

Srivastava, S. and Misra, M. (2016). Assessing and forecasting technology dynamics in smartphones: a TFDEA approach. *Technology Analysis & Strategic Management* 28 (7), 783–797.

Statista (2017). Number of apps available in leading app stores as of March 2017. https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/

Tan, S.J. (1999). Strategies for reducing consumers' risk aversion in Internet shopping. *Journal of Consumer Marketing* 15 (2), 163–180.