Actions Speak Louder Than Words? Word of Mouth vs. Silence of Mouth, an application store observation

Marcia S. Merenciana 414227mm

ms.merenciana@gmail.com

Supervisor: Dr. Z. Jiang

Erasmus University, Erasmus School of Economics

MSc in Economics and Business, Marketing

22 October 2015

Keywords: word of mouth, observational learning, app store optimization, ordered probit

Preface

The copyright of the Master thesis rests with the author. The author is responsible for its contents. Erasmus School of Economics is only responsible for the educational coaching and cannot be held liable for the content.

Abstract

The impact of word of mouth and silence of mouth has been researched multiple times in the area of traditional shopping and online shopping on websites. The researches came to the conclusions that these two forms of interaction have an influence on the customer purchase of the item or service. Be that as it may, a new sort of retailing has been on the rise as of late. More and more customers are purchasing or downloading gaming applications on their mobile devices such as smartphones and tablets. These products lack tangibility just like services but aren't classified as a service. With applications becoming increasingly popular, it is starting to get more important to understand the social drivers of a successful application. This paper researches how word of mouth and silence of mouth affects application adoption decisions in mobile software distribution platforms and what the magnitude of these two social interactions is.

To research the effect of the two social interactions, the ordered probit model was employed. Ordered probit model is an ordered choice model which allows one to map an underlying continuous preference to a categorical, yet ordered dependent variable. In the case of this research, the underlying preference is based on the perceived quality of the application, and the dependent variable is the categorized number of installations. The perceived quality was determined with the social interactions word of mouth and silence of mouth. These two phenomenons are portrayed in the Play Store as the average rating and the ranking. With support of the ordered probit model, a hit-rate table was constructed, which allows us to show the model performance in terms of predictability.

This research showed that the average rating and ranking influence the adoption. Namely, the higher an application is rated, thus higher valued by the consumers, the more likely it is for an application to be in one of the higher number of installation categories. Ranking also influences the adoption as was expected based on the literature review. The higher the ranking, the more likely an application is going to be in a higher installation category.

Also, it was presumed that a higher volume of ratings would amplify the above described effects of rating and ranking. However, instead of amplifying the effect, the effect got weakened. Age of the game since its launch date, which was expected to diminish the effect of word of mouth and silence of mouth only had a diminishing effect of silence of mouth and not on word of mouth.

To conclude, the predictability of the ordered probit model was robust with a hit rate of 74%. The second dataset which consisted of the same observations but at a different time had a hit rate of 48,6% which is not robust.

1. Introduction

1.1. Application stores and their social drivers

App store, also known as an application market place, is a mobile software distribution platform which allows consumers to download applications ("a program or piece of software designed to fulfill a particular purpose") to their mobile devices. It differs from a traditional offline retail store or even from an online retail store in that the products bought, are downloadable only. Customers will never actually have a tangible product.

These distribution platforms are widely used in mobile devices such as tablets and smartphones. The two biggest mobile software distribution platforms are the Google Play Store and the App Store by Apple. As of today, these app stores are still growing. The Play Store saw an increase of 30% in application downloads in the first quarter of 2015 in comparison to the first quarter of 2014. Additionally, the revenue had an increase of 40%. Due to the big growth of the Play Store, Googles' distribution platform has 60% more application downloads than the Apple App Store. Nevertheless, Apple still generates 70% more yearly revenue from the App Store. These growth numbers give an indication of how these distribution platforms are still growing and could thus have big opportunities for the companies who offer or who would like to offer their content on these distribution platforms.

And yet, despite the tremendous growth, there are no clear strategies or blueprints to help app developers launch, grow or sustain their applications. As more than 43.000 new apps were launched in the Apple App Store in December 2014, the competition is extremely high.[1] Out of all these apps, only a few can be called a success by eventually achieving a high number of downloads. So, why are some applications successful and others not?

A first glance at the Apple App store and Google Play Store a few things are noticeable. Downloaders directly and indirectly socially interact with each other in the forms of Word of Mouth (WOM) and Silence of Mouth (SOM). WOM is in the form of reviews and ratings and SOM in the form of rankings and numbers of downloads. These types of consumer interactions are also widely used in online retailing and have proven in the past that they can influence the behavior of

1http://www.statista.com/statistics/258160/number-of-new-apps-submitted-to-the-itunes-store-per-month/

consumers. The focus of this research will therefore be on understanding the social drivers in the mobile software distribution platforms and how these social interactions are of influence on the success of application downloads.

1.1.1. Research question

How is the adoption of applications influenced by social interactions in the form of word of mouth and silence of mouth?

1.2. Academic and Managerial Relevance

The aim of this paper is to contribute to the theory and the practical market and to understand the role of social interactions the adoption of applications. This paper is academically relevant as it further explores the relationship between word of mouth and observational learning. The interrelationship between the two has been researched before, but not in the application market. To research these social interactions, an ordered probit model will be used. This model comes from the field of econometrics and therefore not widely explored in the field marketing.

In addition, this research has a managerial relevance as it furthers explores the application store optimization.

1.3. Structure of Thesis

This research follows with the second chapter, literature review, which gives a deeper insight in consumer decision making, learning and choice and builds the basis for constructing the hypotheses. Chapter three, data and method discusses all the variables used and explains the used ordered probit model. Furthermore, chapter four tests the constructed hypotheses and shows the results found. And to conclude, chapter five offers the general discussion, academic contributions and managerial implications, limitations and future research.

2. Theory

2.1. Literature review

In this paragraph general literature on consumer decision making, learning, choice and social interactions will be discussed. The goal of the general literature is to set a standard of definitions which will be used throughout the study.

2.1.1. Consumer decision making, learning and choice

Prior to a product adoption, consumers have to consider multiple factors before the eventual purchase or rejection. Trying to understand this process of product adoption, researchers have tried to connect a theory to the decision making process of consumers (Bettman, 1979; Newell, 1968; Simon, 1957; Slovic & Lichtenstein, 1971). Often it comes down to a rational process that can be predicted with the help of a model. But what has to be kept in mind is that consumers aren't rational beings but rather have versatile interactions with their environment (Holbrook & Hirschman, 1982). Additionally, the type of interactions with the environment are influenced by inherent evolutionary fundamental motives, such as making friends, finding a mate, or achieving status. These fundamental motives have to be known by the researcher to use a predictive model (Griskevicius & Kenrick, 2013). As is imaginable it might be far too complex to ever connect a fully accurate prediction model to the consumer decision making process. Consequently, still trying to link a model to the decision making process might lead to errors in predicting the consumer choice (Thaler, 1980). This does not mean that it is unnecessary to construct a predictive model. It can help guide companies in setting up a marketing tactic that increases the chances of success. However, keep in mind that a predictive model is not a perfect predictive tool.

Be that as it may, researchers have made significant contributions to uncovering the search behavior of consumers. For instance, Nelson (1974) found groundbreaking information on the differences in consumer choices between certain types of products. According to his research, products can be divided in experience goods and search goods. Experience goods are defined as a product or service where the qualities are not determined prior to purchase. These qualities can only be determined by experience, as is the case of a mobile application. Meaning that in a traditional non-online setting, consumers would have to take the doubt of the product away by gathering other user experiences before purchase, which can be a tedious job in an offline setting. Klein (1998) suggested to transform an experience good in a search good ("information for dominant product attributes can be known before purchase" (Nelson, 1974)) by giving the consumers a 'virtual experience' of the product with the help of the World Wide Web. However, a recent study (Huang, Lurie, & Mitra, 2009) proved that the distinction between the perceived ability to assess an experience products' or search products' quality beforehand disappears in online shopping. This makes it unnecessary to transform an experience good in a search good. More so, consumers spend about the same time searching for information about experience goods as for search goods. Still, there are differences between the type of searching for the two types of goods. Experience goods require a greater depth of search. And, the ability to learn from other consumers' experiences provided by the website, increases the chances of purchasing the product from that website more than for search goods. Thus, suggesting that it is best to allow consumers to conduct that greater depth of research on the online shopping location itself.

While this is true, consumers consult more sources for information than just other consumers' experiences. These sources can be divided in two types of sources. Internal sources, which come from memory and past purchase decisions, and external sources, which come from the environment (Murray, 1991). The environment may be from marketing sources or personal/impersonal interaction (Engel, Blackwell, & Miniard, 1986). This research is focusing on the last source, the environment and specifically personal and impersonal social interactions. An application has a one time adoption and therefore consumers won't be able to make an evaluation of the application from memory.

2.1.2. Social interaction

Social interaction, also defined by Godes et al. (2005) as "an action or actions that is/are taken by an individual not actively engaged in selling the product or service that impacts others' expected utility for that product or service." These actions cover any type of interaction that are up for interpretation. For example, a recommendation or an assessment of a product, or an observation such as the popularity of a good indicated by the number of users. Phenomena as such are also known as word of mouth and silence of mouth (observational learning). Word of mouth (WOM) is spoken communication as a means of transmitting information ("Word of mouth," 2015). Opposite of spoken communication is the communication by not speaking at all. In this paper it is also identified as silence of mouth, or in the general literature known as observational learning. Banerjee, a pioneer in observational learning identified the basis of herd behavior as observational learning, “people will be doing what others are doing rather than using their information” (1992). The following paragraphs will give a deeper insight in the two phenomenons.

Word of mouth

Why do consumers engage in word of mouth activities? It depends on the type of word of mouth. The motive to spread positive WOM comes from self-enhancement, altruistic and product involvement reasons, while negative WOM comes from vengeance, anxiety reduction, altruistic and advice seeking reasons (Sundaram, Mitra, & Webster, 1998). The motive to spread WOM online are slightly different, Hennig-Thurau, Gwinner, Walsh & Gremler identified concerns for others, social interaction, potential to enhance self worth and economic incentives as motivations (2004). As the motivations to engage in WOM are either positive or negative, consequently, the outcomes of such WOM may also be affected positively or negatively. Arndt (1967) was the first to research WOM as an intervening variable between purchase, the result variable and the antecedent variables. He found that WOM, when positive, can have a positive effect on the product adoption, but when negative might slow the adoption down. WOM might even be more influential than traditional marketing activities. Trusov, Bucklin & Pauwels (2009) found that WOM has a prolonged carry-over effect and a higher response rate. While this is true, not all the types of sources of word of mouth are perceived as equally influential. A distinction has to be made between strong ties (close others) and weak ties. Information that comes from strong ties are more influential than the latter (Brown & Reingen, 1987). Online word of mouth that comes from strangers who share their experiences and recommendations are weak ties as the reader often has no ties at all. This could indicate that WOM coming from an online source is less influential.

Silence of mouth

SOM, unlike WOM, has no intrinsic motivation to spread it. SOM comes from the observation of behavior of others. As explained before, the phenomenon was first explained in the academic literature by Banerjee as herd behavior. Bikhchandani, Hirshleifer, & Welch (1992) took it a step further by also identifying informational cascade as a result of observational learning which "occurs when it is optimal for an individual, having observed the actions of those ahead of him, to follow the behavior of the preceding individual without regard to his own information." The difference between herd behavior as explained by Banerjee and informational cascade is that the cascade can lead to radical changes such as cultural changes and fashion ideals. Also, informational cascade can be positive or negative where consumers could adopt or reject what is being observed.

Observational learning does not only influence behavior, a recent study suggests that it also affects quality perception. If consumers could accept an offer in consecutive order, if the first one declines, the second (and so on) will feel inclined to refuse the offer also. This happens when previous decisions are observable but the true reasons behind the decision are not. Consumers can only base the quality of a product on the choices or behavior of consumers before them. Consequently, some products are valued at a higher level even though the true quality may be lower, vice versa. For this reason, a few choices by the innovators and early adopters at the beginning of a product launch can be critical in the success of the product (Zhang, 2010). The researcher of this study, Zhang, thinks that if observational learning is combined with word of mouth, the success of the product would also be a true representation of its quality.

2.2. Hypotheses and conceptual model

Now that there is a clear idea on what online social interactions, word of mouth and silence of mouth are, hypotheses can be formed. Articles that discuss WOM and SOM in a similar setting as a mobile distribution platform will be reviewed. This will lead to the forming of hypotheses that will be tested in this research. At the end of this chapter, a conceptual model will be presented which visually show all the hypotheses. But first, a quick overview on what a mobile distribution platform is.

2.2.1. Mobile software distribution platform

App store, also known as an application market place, is a mobile software distribution platform which allows consumers to download applications (a program or software created to fulfill a particular goal) to their mobile devices. It differs from a traditional offline retail store or even from an online retail store in that the products bought, are downloadable only. Customers will never actually have a tangible product.

Distribution platforms are widely used in mobile devices such as tablets and smartphones. The two biggest mobile software distribution platforms are the Google Play Store and the App Store by Apple. As of today, these app stores are still growing. The Play Store saw an increase of 30% in app downloads in the first quarter of 2015 in comparison to the first quarter of 2014. Additionally, the revenue had an increase of 40%. Due to the big growth of the Play Store, Googles' distribution platform has 60% more application downloads than the Apple App Store. Nevertheless, Apple still generates 70% more yearly revenue from the App Store (AppAnnie, 2015a, 2015b). The exact numbers of the revenues are not published by Google, but in 2014, the revenue of the Play Store was estimated at about 4 to 5 billion US dollars (Wallenstein, 2015). In January 2015, Apple shared in a press release that the revenue is more than 10 billion US dollars for the developers (Monaghan & Neumayr, 2015). These growth numbers give an indication of how these distribution platforms are still growing and could thus have big opportunities for the companies or individuals who offer or who would like to offer their content on these distribution platforms.

Mobile application store value network

In mobile application stores are multiple roles and complex relationships at play. To the researcher's knowledge, Cuadrado & Dueñas were the first who attempted to map it (2012). The following paragraph is based on their article Mobile Application Stores: Success Factors, Existing Approaches and Future Developments.

The three main partakers in the relationships are the application provider, the mobile application store and the application consumer. The application provider is the one that produces and supplies the application stores with applications. The provider can range from large companies to amateur app developers who make applications as a hobby. They have the option to offer their applications for free or paid, and with in-app purchases or without in-app purchases. In-app purchases means the ability to purchase additional content and/or subscriptions (Apple, 2015). At the other end of the relationship you can find the end user or the so called application consumer. As is described in figure 1, for the application provider to reach the consumer, multiple routes can be used. Social networks and search recommendation marketing allow the provider to directly interact with the end user. But in this research, the focus lies on the mediator which is the channel one has to cross to offer their content to the application consumer. The mobile application store is linked to a platform which is in the case of the Play Store, Android and in the case of the Apple App Store, iOS. The platform has to be installed on a device. For example an iPhone or a Samsung Galaxy S phone. And finally, to connect to the application store, one needs an internet network.

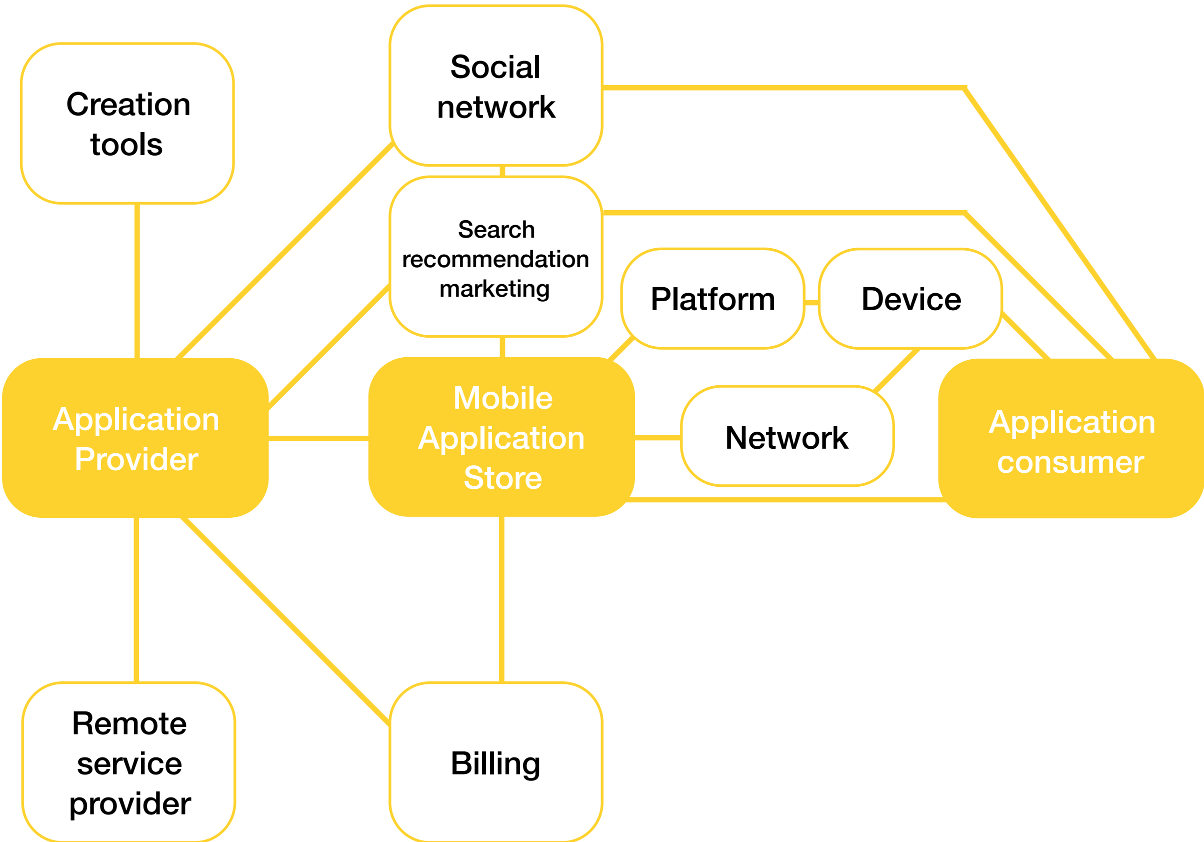


Figure 1. Mobile application store value network (Cuadrado & Dueñas, 2012)

To go back to the application store, as mentioned before, there are two major players. Apple's App Store and Google's Play Store. Between the two there are differences that also could influence the success of an application. The biggest difference is that the App Store is perceived as closed and the Play Store as open.

This means that there are less limitations in the Play Store to offer content. Due to the closed nature of the App Store, it makes it harder to offer an app but it does guarantee a certain level of quality to the end users. In the Play Store, an app is easily published, Google only steps in and eventually removes the application when multiple reports have been made by consumers. As well, due to the open nature, the Play Store is offered on a wider range of brands and devices whereas the App Store can only be accessed from Apple devices. However, with a wider range of devices, there is also a wider range of standards an application has to conform to in order to work on all these devices. Although, there might be less rules to follow to publish an app in the Play Store, due to the low barrier, consumers do encounter higher chances of facing a low quality app which, to Cuadrado & Dueñas' opinion might result in lower trust. To compensate, many applications in the Play Store are offered for free whereas the same app has to be paid for in the App Store. These are all factors that might influence the success of the app. Thus, based on this information, the data for this researched will be gathered from the Play Store. The control Apple has over the App Store might influence the natural free market flow too much in order to successfully observe the social influences.

2.2.2. Word of mouth

Senecal and Nantel (2004) researched the influence of online product recommendations on the online product choices of consumers. They found that consumer choices are influenced by online recommendations. Consumers who received a recommendation selected the recommended product twice as often in comparison to consumers who did not receive a recommendation. They also found that not all type of recommendations are as effective. A recommendation system is automatic, has no expertise of the product, or a strong tie, but it was still more influential than a recommendation from an expert. As to why this is, it still has to be researched as consumers did indicate that they perceive a recommendation system as less trustworthy. It might have to do with the overall trust of certain online shops. Size, reputation and evaluation all influence consumers' trust in online shops (Jarvenpaa, Tractinsky, & Saarinen, 1999). Other than a recommendation, consumers can also share an assessment in the form of a review and a rating (a rating summarizes a review in the form of five/four stars or a grade). Amazon and bn.com are two similar websites who allow consumers to share ratings and reviews on books bought. A research between the two showed that positive reviews and ratings resulted in a relative increase in book sales at one site. Also, negative reviews and ratings have a more powerful effect in decreasing book sales, than positive reviews and ratings in increasing book sales (Chevalier & Mayzlin, 2006).

Based on the research of Chevalier & Mayzlin (2006) it is presumed that positive word of mouth in the form of higher ratings increases the application installation.

H1 *Applications with positive word of mouth in the form of higher average ratings are more likely to be adopted.*

2.2.3. Silence of mouth

Bestseller lists, which are a visual representation of observational learning are widely used for books. An example of an influential bestsellers list is the *New York Times* bestseller list in the United States. It is found that appearing in the *New York Times* bestseller list will result in a slight increase in sales. Furthermore, this increase is even stronger for debut authors. Consequently, it can be expected that a bestseller list might decrease the variety of books published as such a list not only reflects the sales but also influences the sales. Yet, it was found that the variety of the books did not decrease and additionally, it is beneficial to books that are not bestsellers but are in the similar genres (Sorensen, 2007). The same type of research was conducted in the Apple App Store. The willingness to pay for an application decreases significantly depending on the sales rank. For example, an app in the first place is valued twice as much as an app in the second place. This steep decline in willingness to pay can be seen throughout the top ten apps. From the ranking eleven to fifty, this decline stabilizes. After the fiftieth rank, the effect of the bestseller list becomes negligible (Carare, 2012).

Unfortunately, the benefits of appearing on a bestseller list can be abused for personal gain as allegedly done by Michael Treacy and Fred Wiersema in the nineties. According to Business Week, they manipulated the *New Work Times* bestseller list by purchasing bulks of their own book, The Discipline of Market Leaders, across the United States. Resulting in their book to appear high in the list which consequently lead to their success as popular keynote speakers (Stern, 1995). Although it might be unethical, it does show the power of SOM.

According to the literature of Carare (2012) and Sorensen (2007) it is expected that higher rankings will result in higher adoption rates, leading to the following hypothesis:

H2 *Applications with positive silence of mouth in the form of a higher ranking are more likely to be adopted.*

2.2.4. Word of mouth versus silence of mouth

While the previous mentioned paragraphs focused on the online social interactions separate from each other, in the article Online Social Interactions: A Natural Experiment on Word of Mouth Versus Observational Learning (Chen, Wang, & Xie, 2011) the authors are researching the difference between the effects of the two sorts of interactions and if there is a variety in the effects over the product life cycle. Also, it was researched if there might be an interaction effect between the two. This study revealed that word of mouth and observational learning differ in their effects on sales. For example negative word of mouth is of bigger influence than positive word of mouth. In observational learning, the opposite is true, so negative observational learning has no effect while positive observational learning increases sales significantly. The effect however of the two interactions have a diminishing effect depending on the lifetime of the product. An explanation on why the effect diminishes over time might be linked to the diffusion of innovations. When a product is launched, the adoption group exist of innovators, early adopters or early majority. Eventually the application is adopted by the late majority and eventually laggards, resulting in a growth decrease. If product adoption is decreasing, the effect of WOM and SOM could decrease as well (Rogers, 2010). To conclude, they also found that there are interaction effects. That is, the impact of observational learning increases with the word of mouth volume.

Back in 2011, Chen, Wang, & Xie were the first to research these two phenomenon at the same time. Since then, no major contributions were made to the combined interactions. Cheung, Xiao, & Liu (2012) somewhat replicated the research in a beauty forum and came to the same conclusion. They did however add consumer expertise and consumer involvement to the variables and found that expertise has a negative moderating effect and involvement has a positive moderating effect.

Due to the lack of research in the field of combined WOM and SOM the opportunity came to research the two but in another online setting. Application markets show the same types of interactions to their users/consumers (appendix A). But as an application market differs in the setting, use and function than in an online shop, it would be interesting to see if similar effects as mentioned before occur.

So, besides the rating and ranking influencing the application adoption positively, the volume of ratings is expected to moderate between the independent variables and dependent variable. Based on the research by Chen et al. (2011), the more ratings there are (WOM) the stronger the effect of the ranking (SOM) will be on the application adoption. Furthermore, it is assumed that the volume of the ratings (WOM) will increase the effect of the rating (WOM):

H3a *The impact of word of mouth in the form of higher average ratings increases with the volume of word of mouth in the form of total number of ratings.*

H3b *The impact of silence of mouth in the form of a higher ranking increases with the volume of word of mouth in form of total number of ratings.*

Equally important, the variable time will play a part in the above hypotheses. Chen et al. (2011) observed that time can cause a diminishing effect on adoption. An explanation for the diminishing effect might be attributed to the diffusion of innovations (Rogers, 2010). Therefore, depending on the age of the application, all the effects above are expected to diminish.

H4a *The time since the application has been published has a diminishing moderating effect on the effect word of mouth on adoption.*

H4b *The time since the application has been published has a diminishing moderating effect on the effect of silence of mouth on adoption.*

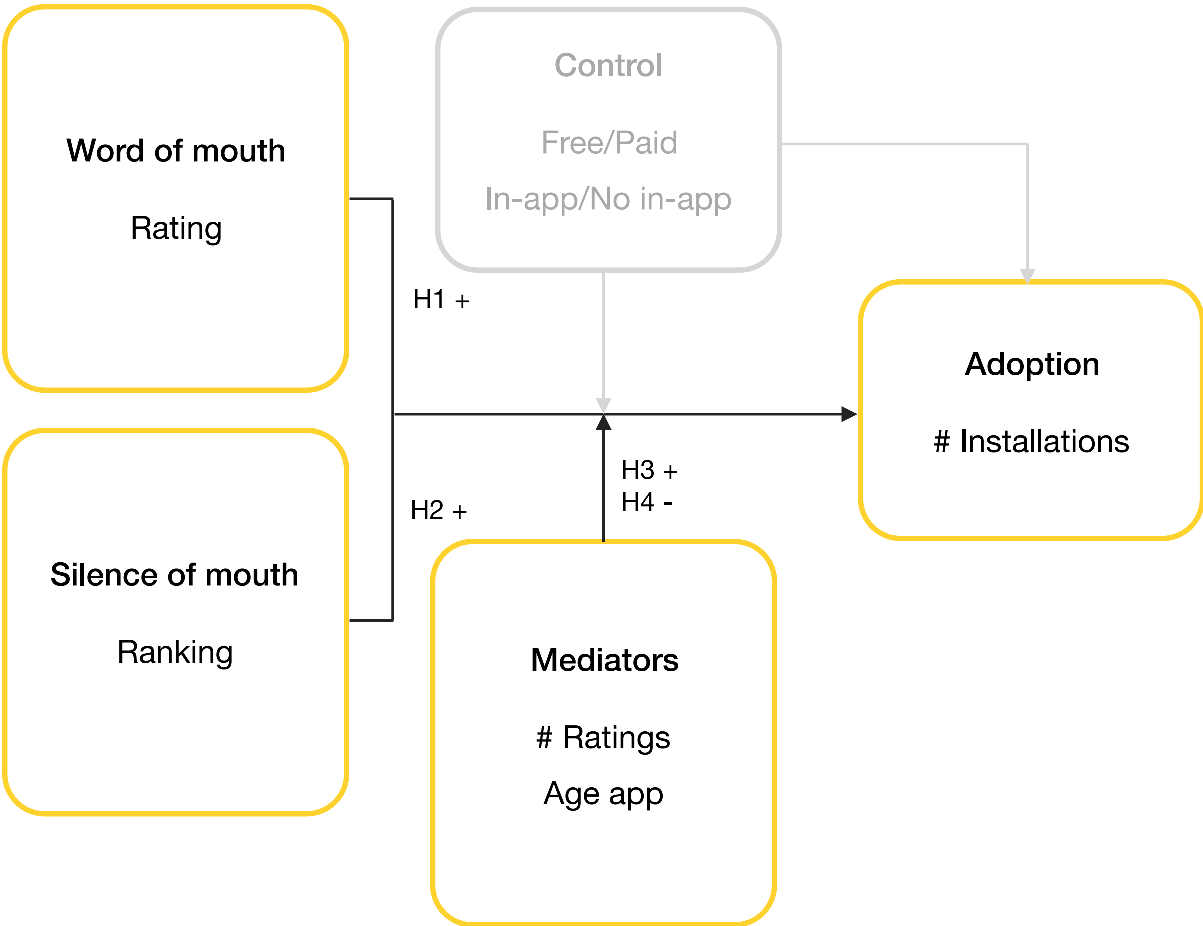


Figure 2. Conceptual model

Control

In the conceptual model, the control variables free/paid and in-app/no in-app are included as they are expected to influence the application adoption as well. It is not the aim of this research to observe the effect of payment on adoption but not including these variables in the research could lead to incomplete results. No hypotheses are formulated however, there are a few expectancies. Both these variables portray the monetizing nature of the application. The first variable, free/paid indicates if the application can be installed with or without a payment. The variable free/paid differs from the variable in-app/no in-app in that an application can be free but still require in-app payments (payments after application purchase) in order to get the complete experience of an application (table 1, monetization).

	NO IN-APP	IN-APP
FREE	Free, no in-app	Free, in-app
PAID	Paid, no in-app	Paid, no in-app

Table 1: Monetization

3. Data and Method

3.1. Data

3.1.1. Data collection

Before the actual data collection, multiple factors had to be considered to select the right data. When looking at the main page of the Play Store, six downloadable categories are portrayed, apps, games, films, music, books and kiosk. Within the apps category, the first category shown is games. So games is one of the most noticeable categories in the Play Store. The prominence of the gaming category might be attributed as 24% of smartphone users indicated that they use their smartphone mainly for gaming.[2] As gaming being one of the most downloaded applications in the Play Store it was decided to gather the data on these applications. Within the gaming category, users can navigate within different kinds of gaming categories. No information could be found on what the most popular gaming category in application stores is. Therefore, the puzzle category is selected at random as the subject of analysis.

Although the Play Store does portray all the data that is necessary for this research, the data was collected from a database for convenience reasons. This database is called AppAnnie.com, a website that offers analytical services and data for application providers. During data collection an aim was made to collect data on 500 applications, 125 applications in every category as seen in table 1, monetization. By gathering equal amounts of data per category an optimal control for the two control variables could be achieved. However, AppAnnie allows users to see the top 500 ranked applications per category only. In the top 500 of paid applications, only 88 had in-app purchases, thus resulting in observation reduction. In addition, observations had to be removed due to incorrect or missing data. The remaining 457 observations are distributed by monetization as portrayed in table 2.

[2] https://www.consumerbarometer.com/en/insights/?countryCode=NL

	NO IN-APP	IN-APP	TOTAL
FREE	123	121	244
PAID	125	88	213
TOTAL	248	209	457

Table 2: Free/paid * No in-app/in-app purchases cross-tabulation 8/7/15

The data is a snapshot of the United States Play Store on 8 June 2015 and 16 September 2015 (for distribution of the data collected on 16 September please refer to table 3 in appendix B). The gathering itself was conducted in the second and third week of June 2015 and the third and fourth week of September 2015.[3] The researched country is the United States because in the first quarter of 2015, the U.S. was the country with the most Google Play installations (AppAnnie, 2015a). A limitation of AppAnnie is that all the variables provided are global with an exception of the ranking which is only portrayed per country. Hence, the ranking might not be a true representation of silence of mouth. Why ranking was still used as an indicator of silence of mouth will be explained in data description. Nonetheless, ranking is the same information which is shown to the users of the Play Store. The United States, being the biggest market, battles this limitation the most.

3.1.2. Data description

The variables were already briefly noted in the conceptual model. In this paragraph an in-depth description of the variables will be given. First, the list of variables and the short description can be found in table 4.

[3]An explanation on why a second dataset was collected can be found in paragraph 3.2 Method.

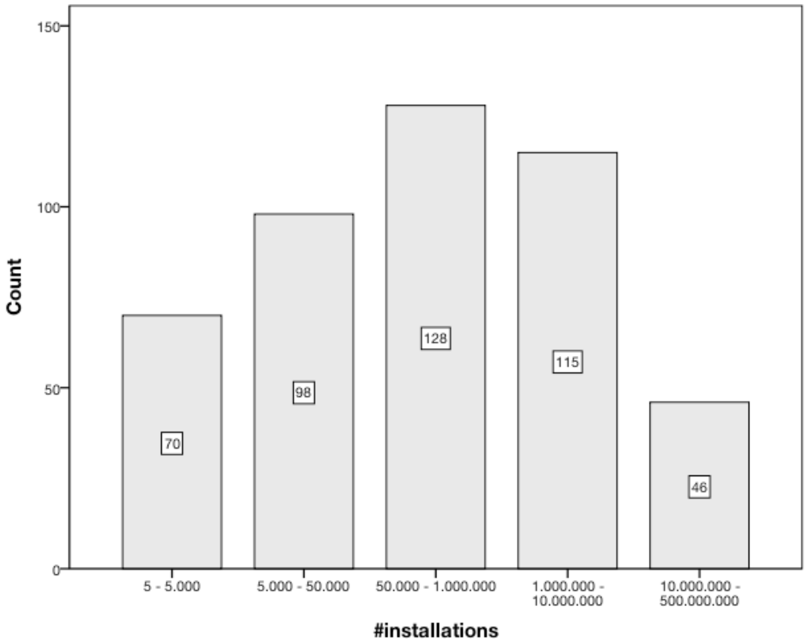
VARIABLE NAME	DEFINITION	MEASURE
INSTALLS	Number of installations, ordinal ranked in five categories. 1 = 5 - 5.000, 2 = 5.000 - 50.000, 3 = 50.000 - 1.000.000, 4 = 1.000.000 - 10.000.000, 5 = 10.000.000 - 500.000.000.	Ordinal
INSTALLS = I	Four thresholds in between the installation categories.	
RATING	Average rating rounded to the nearest tenth with 5 being the most favorable and 1 being the least favorable.	Ratio/Scale
RANKING	Ranking with 1 being the most favorable.	Ratio/Scale
SUM_RATING	Total number of ratings.	Ratio/Scale
AGE	Age in days	Ratio/Scale
FREE_PAID	Free or paid app. 0 = free, 1 = paid.	Nominal
IN_APP	In-app or no in-app purchases. 0 = no in-app, 1 = in-app.	Nominal

Table 4: List of variables used

INSTALLS

INSTALLS, the dependent variable, is indicated by the number of installations an application has. The number of installations is also a proxy for adoption. Namely, the higher the number, the higher the adoption. Consequently, this also leads to a restriction of the data. It is only portrayed on an aggregate level. Only the total adoption is visible. Thus, only the choices of people who adopted are being observed, and not the choice of people who do not adopt. Due to this considerable restriction, the selection of a potential analytical model are limited. The model that was eventually used is explained in 3.2. Method.

Another restriction of this variable is that Google, and thus AppAnnie, provides users with the number of installations in an ordinal categorized nature. It is uncertain how many times the application has been downloaded exactly. The dependent variable has a total of 16 installation categories in the observed data with in some categories only one observation. Due to the low amount of observations per category, goodness-of-fit of the necessary models could be insignificant as a result. In order to prevent this, the dependent variable has been collapsed to five categories, making sure to have sufficient observations per category. In addition, the categories have been recoded to 1, 2, 3, 4 and 5 so SPSS (the analytical software used in this research) knows how to order them. The distribution of the five categories is displayed in graph 1: Distribution INSTALLS.



Graph 1: Distribution INSTALLS

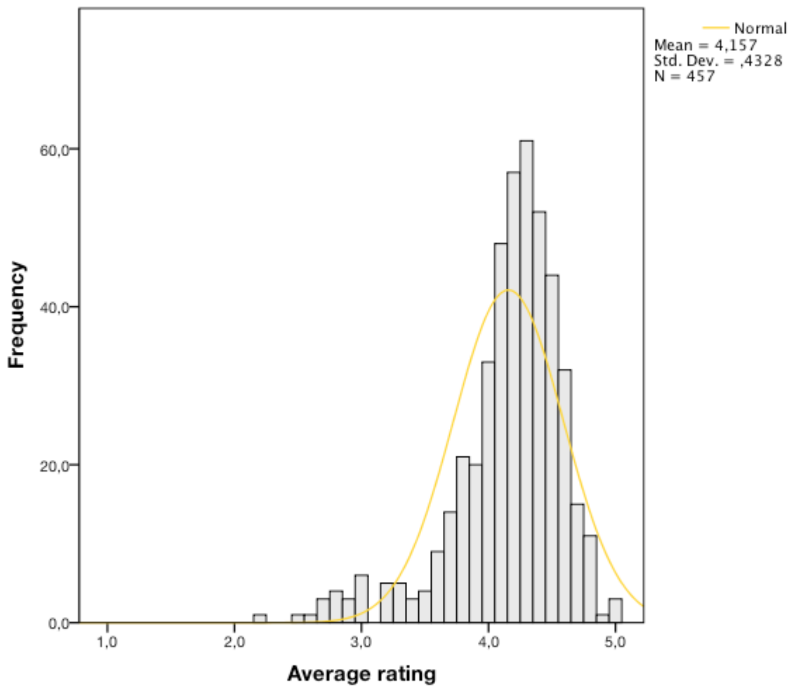
Please refer to table 5: number of installations, in appendix C for a detailed look of the dependent variable for both datasets.

RATING

The rating of the applications will be used as an indicator of WOM. Reviews are a more detailed form of online WOM than ratings and would have been the most suitable variable to measure the interaction. Reviews are constructed of text where the valence would have to be measured with the help of text mining. However, in light of this research and the limited time, it has been decided to use ratings as an indicator of WOM. In the case of the Google Play Store, ratings can be given from one star up to five stars with five stars being the most favorable (appendix A for a visualization). Application consumers are only allowed to rate an application after downloading the app. Not all the consumers are actively approached to leave a rating. Some providers ask their consumers to leave a rating and others not. This means that leaving a rating can either be intrinsic or extrinsic motivated. No data will be collected on which apps in the data actively approach their consumers, meaning no distinction can be made if the ratings come from an intrinsic or extrinsic motive. As can be seen in appendix A, consumers can see the average rating and

the distribution of the five stars. To see the exact number of votes per star, users have to click on the ratings. After clicking on the ratings, consumers will also be able to read more reviews than the three shown on the home page of the application.
The collected data consists of the cumulative votes and the average rating rounded to the nearest tenth on a ratio level of measurement.

Of all the applications gathered, the average rating is 4,157 with a light negative skew, see graph 2: histogram average rating. This average rating is quite stable across all the installation categories (graph 3 in appendix D). As a possible result, average rating could be of insignificant effect on the number of installations.



Graph 2: Histogram average rating

RANKING

Based on the literature review, the ranking will be used as an indicator of SOM. It is often assumed that the ranking is a true representation of the number of downloads indicating an application's success. However, other factors also play a role when it comes to ranking. To illustrate, applications with 100.000 downloads could be in the top ten while applications with 1.000.000 installations could be somewhere below the fiftieth ranking. A second possible theory is that ranking is constructed of growth. It would explain why an application with 100.000 installations could be in the top ten. Truly, this is all speculation. Unfortunately, Google does not share information on the algorithm they use to construct the ranking, but it has been assumed that multiple factors play a role. These factors are considered to be number of installations, number of un-installations, application quality, repetitiveness of interaction with the application, number of ratings, quality of reviews, country, keywords used, shares on social media and the number of backlinks (Butters, 2014). These speculations make it uncertain if ranking is a true representation of SOM. One research by Walz (2015) tried to deconstruct the algorithm to help application developers with ASO (app store optimization[4]). By conducting five studies, Walz was able to attach weights to a selection of the above mentioned factors and see which one affects ranking the most. The studies came to the following conclusion for the Play Store:

Total number of ratings > Number of installations > Rating > Growth trends

Now, according to these weights, ranking could be both WOM and SOM as it combines the two in one variable. The first factor, total number of ratings, is in this research used as a quantification of word of mouth as it quantifies the opinions given by the users. However, it can be argued that the number of ratings is also silence of mouth as it shows the behavior of others. In addition, the second factor, number of installations, can be used as an indicator of SOM as it shows the popularity of an application without reasons as to why the application is high in demand. High numbers of installations could, according to the theory of Banerjee, result in herd type of behavior. Moreover, the same arguments can be used for rating and growth trends. The reason that ranking is used as SOM in this research is that the ranking does not show the users why an app is ranked as it is. This is proved by the fact that nobody, other than Google, knows how ranking is constructed. To understand why this is observational learning, it is important to refer back to the definition of informational cascade by Bikchandani et al.(1992): “it occurs when it is optimal for an individual, having observed the actions of those ahead of him, to follow the behavior of the preceding individual without regard to his own information.” Informational cascade is a result of observational learning. By just observing the actions of others (popularity of an app), consumers can copy that behavior (download the app), while ignoring their own motives.

[4] https://moz.com/blog/app-store-seo-the-inbound-marketers-guide-to-mobile

SUM_RATING

The variable total number of ratings is constructed by adding up the total votes per star per application. The sum of the ratings per application make it possible to research the effect of the volume of ratings on WOM, SOM and adoption. Before running the analysis, a few assumptions can be made about the volume of ratings. If an application has an average of five stars but only three votes, the reliability of that rating could be perceived as low. So, although the average rating is valid, it is unreliable and could have less of an impact on adoption.

A first look at the data already confirms these views. SUM_RATING shows a significant correlation with RANKING, FREE_PAID and IN_APP. The correlations with the first two variables are negative. Put another way, when the ranking decreases (increase in number) the volume of rating decreases as well. FREE_PAID and IN_APP are binary variables so the interpretation is different. But if FREE_PAID has a value of 1 (= paid), an application will have less ratings. In case of in-app payments, when the variable has a value of 1 (= no in-app) an application will have more ratings. These correlations are important to take note of as they will play an important role in the interactions in chapter 4: Result. For a correlation matrix of all the independent variables, please refer to table 6 in appendix E.

AGE

AppAnnie provides information on when applications were launched in the Play Store. With this information, age in number of days can be constructed and used as an indicator of time, or age of the application. On average, puzzle games in the top 500 are 619 days old (table 7: descriptive statistics in appendix F). However, the average age does differ across the installation categories. Categories two and five contain the oldest applications with an average of 799 and 862 (graph 4 in appendix G). The remaining categories are on average at least 200 days younger. As this is not what was expected, it is difficult to determine what this will mean for the analysis. Although the age is oddly distributed across the installation categories, there is still a correlation between AGE and the other independent variables. Rating has a positive correlation of 0,130 meaning that the higher the age, the higher the rating. The positive correlation can also be seen with FREE_PAID (0,325). Hence, older applications have a higher portion of paid applications.

FREE_PAID and IN_APP

The two control variables are both linked to the monetization used by application developers. A short description of both variables has already been given in the past few paragraphs. The only thing to add to that is that as they are control variables, they will be added to every test conducted. Although these two variables and their possible effects are not an objective of this research, if they do influence the number of installations significantly, it will be shortly discussed.

3.2. Method

As mentioned in the data description, the Play Store provides users with the number of installations. These number of installations showcase an underlying preference based on the perceived quality of the application. This means that the higher number of downloads an application has the higher the perceived quality of the application must be. Back in paragraph 2.1.2. Silence of mouth, a research by Zhang (2010) suggested that when WOM and SOM are combined, quality can be truly represented.

So even though scales are observed, these scales underly a preference which is based on a higher perceived application quality. The higher the number of downloads, the higher the perceived quality of the application is. However, the differences between these scales do not make any sense in a way that the categories and their differences are not of equal size. The only aspect that does make sense is that it goes from a low number of downloads to a high number of downloads.

Therefore, the model used for this research is the ordered probit model which is an ordered choice model. The ordered probit model allows you to map an underlying attitude or preference to an ordered outcome which is observed. The ordered nature of the dependent variable is also the reason a multinomial regression is not sufficient for this research. In this case, the observed outcome is the number of installations, divided in the ordered categories. The underlying preference, also known as a latent variable, is the perceived quality of the application. The perceived quality is not directly observed but rather shines through the number of installations and the normal distribution of the error term.

For each application it is hypothesized that there is a continuous underlying perceived quality that underlies the number of installations. This latent variable will be denoted as y_i*, which is unobservable. What can be observed is the crossing of the thresholds, labelled as μ_j, between the categories. This crossing happens when y_i* surpasses a certain value. Furthermore, as there are five categories in the number of installations, only four intercepts are identified. To calculate y_i* the following function, which is constructed of observed and unobserved variables, will be used:

(1) $y_i* = x'_i\beta + u_i$. u_i denotes the error term.

In this research $x'_i\beta$ is

(2) $$\begin{aligned}&\beta_0 + \beta_1 x_{RATING} + \beta_2 x_{RANKING} + \beta_3 x_{SUM\_RATING} + \beta_4 x_{AGE} + \beta_5 x_{FREE\_PAID} + \beta_6 x_{IN\_APP}\\&+\beta_7 x_{RATING \bullet SUM\_RATING} + \beta_8 x_{RANKING \bullet SUM\_RATING} + \beta_9 x_{RATING \bullet AGE} + \beta_{10} x_{RANKING \bullet AGE} + u_i\end{aligned}$$

After calculating the y_i*, it can be determined if the value crosses a threshold. The rule for crossing a threshold and determining the outcome category:

(3) $y_i = j$ if $\mu_{j-1} < y_i* \le \mu_j$

However, it is hard to predict in exactly what category an app belongs to based on all the predictors. Therefore, the ordered probit allows you to calculate the probability that it belongs to in a certain category. The probability is calculated with the function:

(4) $P_{ij} = P(y_i = j) = P(\mu_{j-1} < y_i* \le \mu_j) = \phi(\mu_j - x'_i\beta) - \phi(\mu_{j-1} - x'_i\beta)$[5]

Note that all the probabilities should add up to 1 (= 100%). After calculating the probabilities, the category with the highest probability is selected as the most probable installation category.

To test if this ordered probit model actually works, a hit rate table is used. A hit rate table is a cross-tabulation which crosses the expected installation category with the actual installation category. When the hit rate table achieves a hit rate of at least 50%, the model used for this research can be classified as robust. As a second test, the hit rate table will also be used on the dataset collected on 16 September. The coefficients of dataset one are used to calculate the expected outcome values

[5]The equations 3 and 4 are expressed in more detail in appendix G

of dataset two and will be crossed with the actual values of the second dataset. When the second dataset is also robust, it would mean that the coefficients can be transferred to other real life datasets.

4. Results

All the hypotheses are tested based on the (significant) coefficients. The coefficients in an ordered probit can not be interpreted as in a normal regression. Instead, statements like ‘in the case of a unit increase, the application is more or less likely to be in a higher or lower category outcome.’ will be used. In this chapter, there will be two rounds of testing the hypotheses. During the first round an obstacle was met. The last two interactions added to the model shifted the data dramatically. As these interactions are all insignificant, it reduces the predictability of the model tremendously. Therefore a second round of hypotheses testing is conducted. Instead of just starting over and and not mentioning this in the research, the differences between the two models will be shown. It is advised to read the first round of hypotheses testing in detail in addition to the second round.

4.1. Hypotheses testing first round

The hypotheses are tested with the parameter estimates in table 8.

		ESTIMATE	STD. ERROR	WALD	DF	SIG.	95% CONFIDENCE INTERVAL	
							LOWER BOUND	UPPER BOUND
THRESHOLD	[INSTALLS = 1]	-3,634	0,928	15,319	1	**0,000**	-5,454	-1,814
	[INSTALLS = 2]	-1,891	0,913	4,288	1	**0,038**	-3,681	-0,101
	[INSTALLS = 3]	0,375	0,906	0,171	1	0,679	-1,401	2,151
	[INSTALLS = 4]	3,706	0,965	14,739	1	**0,000**	1,814	5,598
LOCATION	RATING	0,140	0,212	0,434	1	0,510	-0,276	0,555
	RANK	-0,009	0,002	31,050	1	**0,000**	-0,012	-0,006
	SUM_RATING	7,134E-05	1,594E-05	20,039	1	**0,000**	4,011E-05	0,000
	AGE	0,000	0,001	0,150	1	0,699	-0,002	0,003
	FREE_PAID	-3,989	0,262	231,119	1	**0,000**	-4,503	-3,475
	IN_APP	0,097	0,129	0,564	1	0,452	-0,156	0,351
	RATING * SUM_RATING	-1,467E-05	3,406E-06	18,544	1	**0,000**	-2,134E-05	-7,991E-06
	RANK * SUM_RATING	3,094E-08	1,514E-08	4,178	1	**0,041**	1,271E-09	6,062E-08
	RATING * AGE	0,000	0,000	1,082	1	0,298	0,000	0,001
	RANK * AGE	1,187E-06	1,719E-06	0,477	1	0,490	-2,183E-06	4,557E-06

Table 8: Parameter Estimates

4.1.1. H1

H1: Applications with positive word of mouth in the form of higher average ratings are more likely to be adopted.

In order to test the first hypothesis, the following equation is constructed: $x'_i\,\beta = \beta_0 + \beta_1 x_{RATING} + u_i$. Given the insignificance (0,510>0,05) of the rating (0,140) the hypothesis is rejected. The average rating of an application is unlikely to influence the number of installations. In order to explain why rating is insignificant, the variables are added to the model in a hierarchical fashion (appendix I). In the fourth block, after SUM_RATING and AGE are added, RATING becomes insignificant. Based on this information, it becomes clear why RATING is no longer significant. The SUM_RATING, as partially expected during data description, takes over the effect and acts as a mediator. Partially expected, because it was presumed that SUM_RATING would only act as a moderator instead of a mediator.

4.1.2. H2

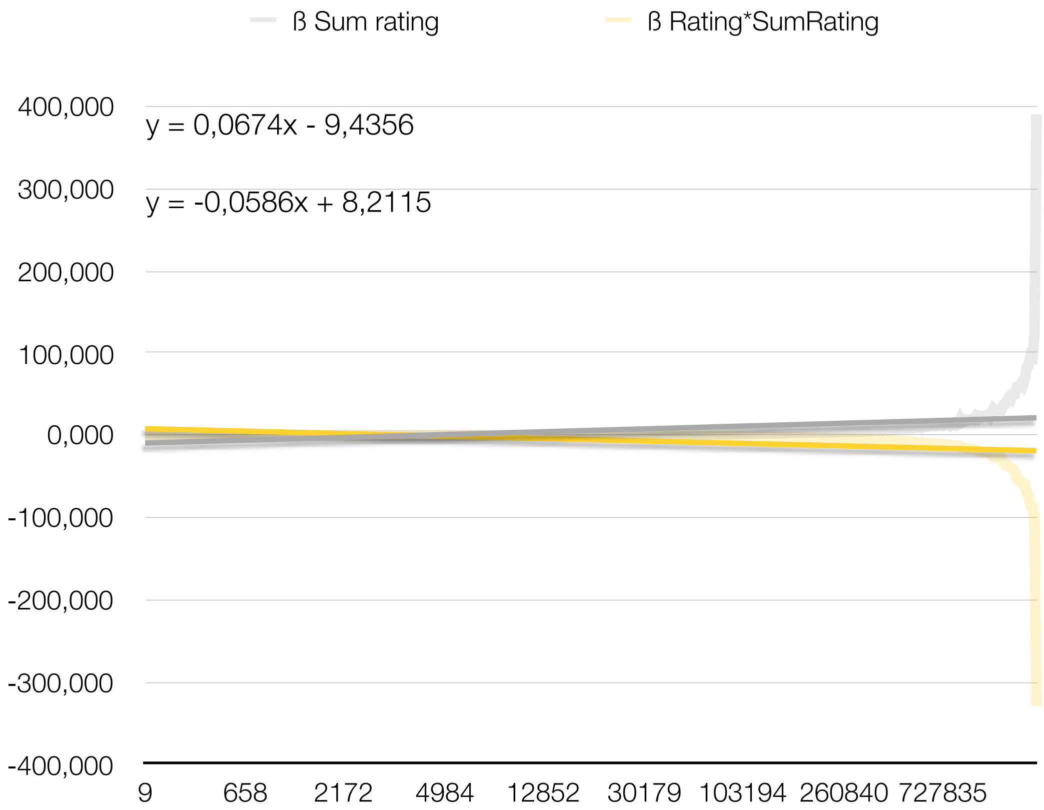
H2: Applications with positive silence of mouth in the form of a higher ranking are more likely to be adopted.

In order to test the second hypothesis, the following equation is constructed: $x'_i\,\beta = \beta_0 + \beta_2 x_{RANKING} + u_i$. Unlike rating, ranking does have a significant (0,000<0,05) effect on the number of installations. As the coefficient is negative (-0,009) it can be stated that, the lower the ranking (higher number), the less likely the application is going to be in a higher installation category. As ranking is stated per country, but the remaining data is worldwide, it was uncertain if ranking was going to have an effect at all. This uncertainty proved to be wrong. Therefore, hypothesis 2 is supported.

4.1.3. H3a and H3b

H3a: The impact of word of mouth in the form of higher average ratings increases with the volume of word of mouth in the form of total number of ratings.

In order to test this hypothesis, the following equation is constructed: $x'_i\,\beta = \beta_0 + \beta_1 x_{RATING} + \beta_3 x_{SUM\_RATING} + \beta_7 x_{RATING \cdot SUM\_RATING} + u_i$. As mentioned before, SUM_RATING does act as a mediator between RATING and INSTALLS. So on its own, the average rating has no significant effect or impact at all, but when it is added in the model in the form of an interaction with the total number of ratings, a negative (-1,467E-05) significant (0,000<0,05) effect is observed. This effect is also known as a mediated interaction (VanderWeele, 2014). In other words, the average rating an application has, becomes more relevant with the volume of the ratings. What is important to note is that the main effect of SUM_RATING has a positive (7,134E-05) significant (0,000<0,05) effect, which is stronger than the main effect. These numbers are suggesting that the average rating is important in combination with a high volume of ratings, but not as important as the volume of the ratings on its own. In this case, actions speak slightly harder than the volume of the words. To illustrate this better, a graph is used:

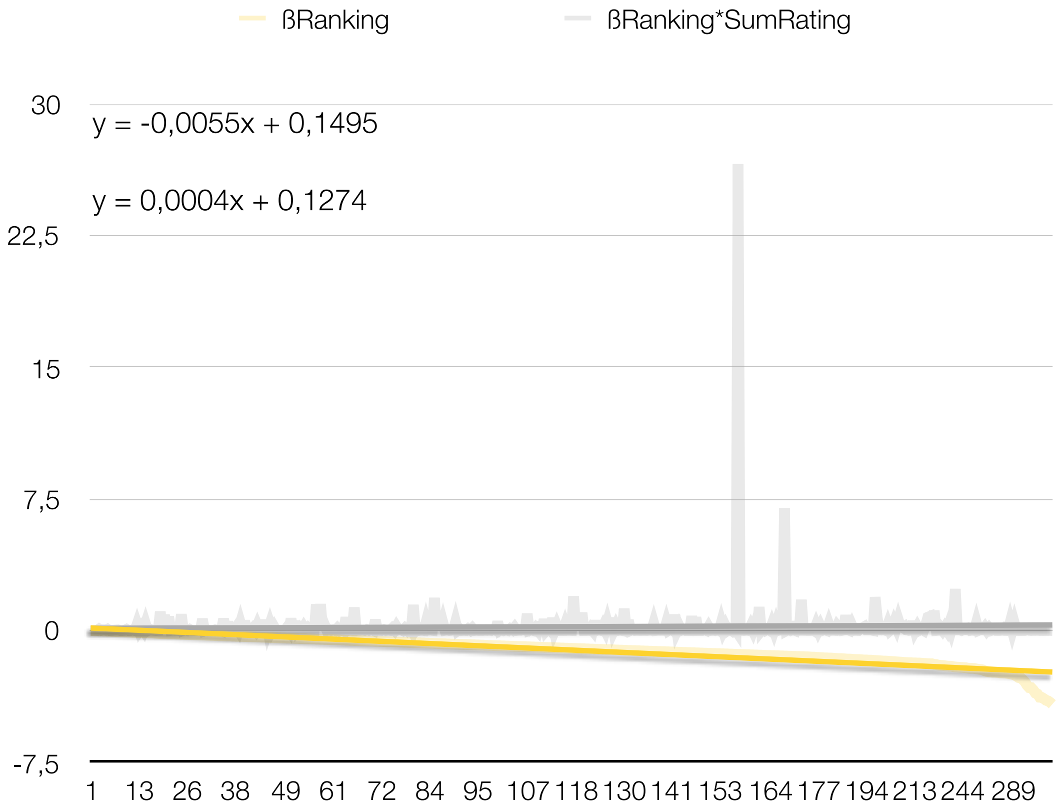


Graph 5: SUM_RATING, RATING*SUM_RATING

The values of SUM_RATING of each application in the first dataset have been ordered from low to high. The accompanied betas and the betas of RATING*SUM_RATING are the values presented in the graph. To see which effect is stronger, linear regression lines have been added. Although barely visible due to the size of the graph, the regression line of SUM_RATING has a slightly stronger gradient. This is confirmed by the equation, because SUM_RATING has a relatively higher coefficient. Although, it does take some time for SUM_RATING to be more important due to the relatively higher constant of the variable. However, as a higher average rating has no significant effect, hypothesis 3a is rejected.

H3b: The impact of silence of mouth in the form of a higher ranking increases with the volume of word of mouth in form of total number of ratings.

In order to test this hypothesis, the following equation is constructed: $x'_i \beta = \beta_0 + \beta_2 x_{RANKING} + \beta_3 x_{SUM\_RATING} + \beta_8 x_{RANKING \cdot SUM\_RATING} + u_i$. Although the interaction effect is small (3,094E-08), the impact of ranking does significantly (0,041<0,05) increase with the volume of word of mouth. But there is a difference between the effect of ranking on its own and ranking in an interaction. Ranking on its own has a negative effect, meaning the lower the ranking the less likely an application is going to be in a higher installation category. The interaction however has a, although small, positive effect. This interaction effect, decreases the effect of the ranking. Also for this hypothesis, it is better to illustrate it visually:



Graph 6: RANKING, RANKING*SUM_RATING

The values of RANKING have been ordered from low to high. The accompanied betas and the betas of RANKING*SUM_RATING are presented in the graph. So yes, the ranking has a stronger effect on the number of installations, but interaction effect reduces the influence of the main effect. Hypothesis 3b is rejected.

4.1.4. H4a and H4b

H4a: The time since the application has been published has a diminishing moderating effect on the effect word of mouth on adoption.

H4b: The time since the application has been published has a diminishing moderating effect on the effect of silence of mouth on adoption.

In order to test the hypothesis, the following equations are constructed: $x'_i \beta = \beta_0 + \beta_1 x_{RATING} + \beta_4 x_{AGE} + \beta_9 x_{RATING \bullet AGE} + u_i$.

$x'_i \beta = \beta_0 + \beta_2 x_{RANKING} + \beta_4 x_{AGE} + \beta_{10} x_{RANKING \bullet AGE} + u_i$. First the main effect AGE. AGE has been added to the parameter estimates in block 4 and had a positive, significant effect. In block, 7 when the interaction RATING*AGE is added, the main effect becomes insignificant. As a matter of fact, the interaction is also insignificant. As a result, the effect of age is removed overall. AGE remains insignificant when the second, also insignificant, interaction is added to the model. When the process of adding the interactions is turned around, so first AGE*RANKING, the main effect AGE does remain significant. This phenomenon suggests there is something wrong in the order of how the variables have been added to the model and that the interactions are incorrect.

Why the interaction proved to be insignificant can be attributed to the research method. Instead of making interactions, the coefficients of ranking and average rating of the first dataset should be compared to the coefficients of the second dataset. Accordingly, the hypotheses should exclude moderation and only mention diminishing.

4.1.5. Modifications

Based on this information, a different set of parameter estimates should be used for this research. Instead of using all the information in block 8, the information in block 6 should be used. AGE, can still be in the data, but should act as a control variable only. The reason the above described problem occurred and was not noticed, is caused by the order the variables have been added to the model and the interactions being incorrect. In case of an ordered probit, variables can only be manually added to the model in a hierarchical fashion. The variables are added based on the literature review. The information in the literature review was not

sufficient to predict this problem of occurring. A new conceptual model can be found in figure 3 in appendix J.
In order to test the diminishing effect, it is not age that should be compared but the difference in time. Because data was collected on two separate occasions, the hypotheses can be tested using another method.
The second round of hypotheses testing will be used to write the general discussion.

4.2. Hypotheses testing second round

The second round is not as broadly discussed as the first round. Many observations still stand but the values of the betas changed. Only rating saw a considerable change. The hypotheses are tested with the help of the parameter estimates in the table 10 below, also block 6 in I:

		ESTIMATE	STD. ERROR	WALD	DF	SIG.	95% CONFIDENCE INTERVAL	
							LOWER BOUND	UPPER BOUND
THRESHOLD	[INSTALLS = 1]	-2,841	0,643	19,520	1	**0,000**	-4,102	-1,581
	[INSTALLS = 2]	-1,102	0,627	3,084	1	0,079	-2,331	0,128
	[INSTALLS = 3]	1,157	0,625	3,426	1	0,064	-0,068	2,382
	[INSTALLS = 4]	4,489	0,706	40,470	1	**0,000**	3,106	5,872
LOCATION	RATING	0,300	0,150	4,020	1	**0,045**	0,007	0,593
	RANK	-0,008	0,001	73,546	1	**0,000**	-0,010	-0,006
	SUM_RATING	7,149E-05	1,583E-05	20,389	1	**0,000**	4,046E-05	0,000
	AGE	0,002	0,000	109,398	1	**0,000**	0,002	0,002
	FREE_PAID	-3,973	0,259	235,507	1	**0,000**	-4,480	-3,465
	IN_APP	0,108	0,128	0,718	1	0,397	-0,142	0,359
	RATING * SUM_RATING	-1,469E-05	3,385E-06	18,832	1	**0,000**	-2,133E-05	-8,056E-06
	RANK * SUM_RATING	3,032E-08	1,499E-08	4,094	1	**0,043**	9,497E-10	5,970E-08

Table 10: Parameter estimates

4.2.1. H1

H1: Applications with positive word of mouth in the form of higher average ratings are more likely to be adopted.

In order to test the first hypothesis, the following equation is constructed: $x'_i \beta = \beta_0 + \beta_1 x_{RATING} + u_i$. In this set of parameter estimates, rating does have a significant (0,045<0,05) positive (0,300) effect. Thus, in case of a unit increase in the average rating, the application is more likely to be in a higher installation category. In the first round of hypotheses testing it was mentioned that rating became insignificant when SUM_RATING was added. However, as soon as the interaction RATING*SUM_RATING is added, the average rating becomes significant again. The reason this wasn't noticed in the first round was due to the insignificance of RATING when the AGE interactions were added. It still doesn't change the fact that SUM_RATING acts as a mediator without the interactions. Hypothesis 1 is supported.

4.2.2. H2

H2: Applications with positive silence of mouth in the form of a higher ranking are more likely to be adopted.

In order to test the second hypothesis, the following equation is constructed: $x'_i \beta = \beta_0 + \beta_2 x_{RANKING} + u_i$. The coefficient of RANKING is still of significant (0,000<0,05) negative (-0,008) effect. That is, in case of a unit increase in the ranking, an application is more likely to be in a lower installation category. Hypothesis 2 is supported.

4.2.3. H3a and H3b

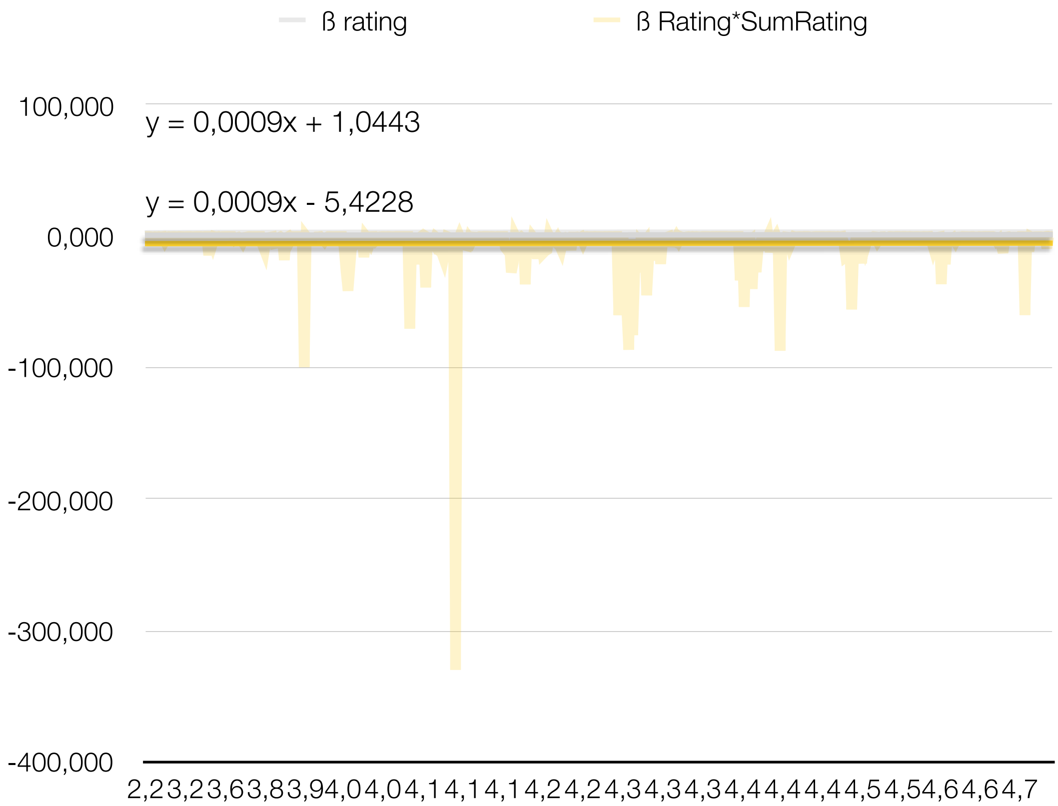
H3a: The impact of word of mouth in the form of higher average ratings increases with the volume of word of mouth in the form of total number of ratings.

H3b: The impact of silence of mouth in the form of a higher ranking increases with the volume of word of mouth in form of total number of ratings.

In order to test these hypotheses, the following equations are constructed:

$x'_i \beta = \beta_0 + \beta_1 x_{RATING} + \beta_3 x_{SUM\_RATING} + \beta_7 x_{RATING \cdot SUM\_RATING} + u_i$.

$x'_i \beta = \beta_0 + \beta_2 x_{RANKING} + \beta_3 x_{SUM\_RATING} + \beta_8 x_{RANKING \cdot SUM\_RATING} + u_i$. As SUM_RATING acted as a mediator before the interactions were added, the first interaction can still be classified as a mediated interaction. The betas of the interactions are still significant. What changed in comparison to the first round of hypothesis testing is that rating is significant. The conclusion of the first round was that the positive influential power of SUM_RATING has a stronger effect than the negative power of RATING*SUM_RATING. But now that rating is significant, the positive effect of RATING also decreases the influence of RATING*SUM_RATING:



Graph 7: RATING, RATING*SUM_RATING

The values of RATING have been ordered from low to high. The accompanied betas and the betas of RATING*SUM_RATING are presented in the graph. The regression lines in this graph are also hard to see due to an outlier (an application called Paperama), but the equation explains it all. The slope of both variables are the same. However, the positive constant of the rating opposes the negative effect slightly. So, the impact of word of mouth in the form of higher average ratings does not increase with the volume of WOM. It does the opposite. Hypothesis 3a is rejected.

The coefficients of the second hypothesis changed slightly but RANKING still has a negative effect and RANK*SUM_RATING a positive effect. The impact of SOM

(RANKING) decreases with the volume of WOM (RANK*SUM_RATING). Hypothesis 3b is rejected.

4.2.4. H4a and H4b

H4a: The time since the application has been published has a diminishing effect on the effect word of mouth on adoption.

H4b: The time since the application has been published has a diminishing effect on the effect of silence of mouth on adoption.

In order to test these hypotheses, the coefficients of rating and ranking of the first dataset should be compared to the second dataset. A change in coefficients demonstrates the effect of time. The variable AGE does not demonstrate this effect as it compares the age of one application to another application. Instead it should compare the age of the same set of applications on different time points.

		ESTIMATE	STD. ERROR	WALD	DF	SIG.	95% CONFIDENCE INTERVAL	
							LOWER BOUND	UPPER BOUND
THRESHOLD	[INSTALLS = 1]	-3,186	0,917	12,074	1	0,001	-4,984	-1,389
	[INSTALLS = 2]	-1,496	0,912	2,694	1	0,101	-3,283	0,290
	[INSTALLS = 3]	0,882	0,893	0,977	1	0,323	-0,867	2,632
	[INSTALLS = 4]	4,226	0,933	20,532	1	0,000	2,398	6,055
LOCATION	RATING	0,347	0,206	2,842	1	0,092	-0,056	0,751
	RANK	-0,006	0,001	63,381	1	0,000	-0,007	-0,004
	SUM_RATING	8,307E-05	1,989E-05	17,447	1	0,000	4,409E-05	0,000
	AGE	0,002	0,000	66,987	1	0,000	0,001	0,002
	FREE_PAID	-4,215	0,328	165,449	1	0,000	-4,858	-3,573
	IN_APP	-0,139	0,138	1,013	1	0,314	-0,410	0,132
	RATING * SUM_RATING	-1,722E-05	4,409E-06	15,251	1	0,000	-2,586E-05	-8,576E-06
	RANK * SUM_RATING	1,821E-08	9,799E-09	3,453	1	0,063	-9,970E-10	3,742E-08

Table 11: Parameter estimates

RATING changed from a coefficient of 0,300 to 0,347, indicating that the effect of average rating increased. However, this proved to be insignificant. Why rating is no longer significant could have been caused by the decrease in the number of observations. Either way, hypothesis H4a is rejected.

The coefficient of ranking changed from -0,008 to -0,006, indicating that the effect of ranking diminished. Hypothesis H4b is supported.

4.2.5. Control variables

The control variables play an important role in the predictability of the ordered probit model. With a significant (0,000<0,05) effect of 0,002, when AGE increases with a day, an application is more likely to be in a higher installation category.

When consumers need to make a preliminary payment to use an application, with a significant (0,000<0,05) effect of -3,973, the likelihood of an application to be in a lower installation category increases. In-app or no in-app payments do not have a significant effect on the installation categories.

4.2.6. Most influential variable

It is hard to determine which variable has the biggest influence on the number of installations. In this dataset al the coefficients and their accompanying x values were added up and divided by the total number of applications in de dataset. SUM_RATING had the highest absolute value and the biggest influence on the number of installations. Of course this is different per application and even maybe per dataset.

4.3. Predictive power

The coefficients presented were sufficient to test the hypotheses. Now it is important to test how well the coefficients act as a predictor for the success of an application. All the significant variables of block 6, were inserted to produce the model. In the form of an equation it looks like the following:

$$x'_i \beta = \beta_0 + \beta_1 x_{RATING} + \beta_2 x_{RANKING} + \beta_3 x_{SUM\_RATING} + \beta_4 x_{AGE} + \beta_5 x_{FREE\_PAID} + \beta_7 x_{RATING \cdot SUM\_RATING} + \beta_8 x_{RANKING \cdot SUM\_RATING} + u_i$$

Along with this equation, equations 1, 3 and 4 of chapter 3 were used to predict the most likely installation category. The category with the highest likelihood is placed in a hit rate table together with the actual category.

		PREDICTED				
		1	2	3	4	5
ACTUAL	1	56	13	1	0	0
	2	14	66	18	0	0
	3	1	25	89	13	0
	4	0	0	22	89	4
	5	0	0	0	8	38
				Total hit	338	
				Hit Rate	0,739606	

Table 12: Hit rate table 1

338 applications of 457 applications in total were predicted correctly producing a hit rate of 0,74. To illustrate, the model is capable of predicting the success of an application in 74% of the cases. When a model is able to predict more than 50%, it is deemed robust.

To test if the model and its coefficients can be transferred to other datasets a second dataset was collected on 16 September 2015. The data comes from the exact same applications but approximately two months later. Some applications disappeared from the Play Store either by being removed by Google, withdrawal of the application by the application provider or disappearing from the top 500. Consequently, only 391 observations remained to be subjected to the research. The betas of the first dataset were used along the x values of the second dataset to predict the outcome category. These predicted outcome categories were placed in the hit rate table below with the actual outcome categories.

		PREDICTED				
		1	2	3	4	5
ACTUAL	1	12	34	6	0	0
	2	24	38	22	0	0
	3	18	25	26	26	0
	4	2	1	18	78	7
	5	0	0	0	18	36
				Total hit	190	
				Hit Rate	0,485934	

Table 13: Hit rate table 2

190 from the 391 categories were predicted correctly producing a hit rate of 0,486. In other words, the model was able to predict 48,6% of the applications' success correctly. This results in the model not being able to translate to other datasets and not being deemed robust. An explanation of why the model seems not to be able to be transferred to other data might lie in the removal of the applications that disappeared from the top 500. Increased success has been documented more than a decrease in success.

5. Conclusion

5.1. General discussion

How is the adoption of applications influenced by social interactions in the form of word of mouth and silence of mouth?

To quote this research: “keep in mind that a predictive model is not a perfect predictive tool.” This is also the case of the model used in this research. Knowing this beforehand, it also wasn’t the aim to construct a perfect predictive model. Instead, “It can help guide companies in setting up a marketing tactic that increases the chances of success.” And one can increase its chances of success by having an insight in how social interactions in the form of word of mouth and silence of mouth influence the adoption of applications in an application store. The results show that the two social interactions influence the adoption but not in the form as was expected.

To research the effect of the two social interactions, the ordered probit model was employed. Ordered probit model is an ordered choice model which allows one to map an underlying continuous preference to a categorical, yet ordered dependent variable. In the case of this research, the underlying preference is based on the perceived quality of the application and the dependent variable is the number of installations. The perceived quality was determined with the social interactions word of mouth and silence of mouth. These two phenomenons are portrayed in the Play Store as the average rating and the ranking. With support of the ordered probit model, a hit-rate table was constructed, which allows us to show the model performance in terms of predictability.

The basis of this research was supported and build upon multiple researches (Carare, 2012; Chevalier & Mayzlin, 2006; Sorensen, 2007) which showed that the two social interactions, separate from each other, influence adoption. A 2011 research by Chen et al., showed that the impact of word of mouth and silence of mouth can get amplified by the volume of word of mouth, resulting in an interaction effect and that the main effects get diminished over time. However, the findings of this research are not completely similar as the past researches.

The average rating and ranking do influence the adoption as was expected. Namely, the higher an application is rated, thus higher valued by the consumers, the more likely it is for an application to be in one of the higher number of installation categories. Ranking, although at first unclear if it was a true representation of silence of mouth, also influences the adoption as expected. The higher the ranking, the more likely an application is going to be in a higher installation category.
However, there were also unexpected effects. It was presumed that a higher volume of ratings would amplify the above described effects of rating and ranking. However, instead of amplifying the effect, the effect got weakened. In addition, age of the game since its launch date only had a diminishing effect on silence of mouth and not on word of mouth.
The predictability of the ordered probit model was robust with a hit rate of 74%. The second dataset which consisted of the same observations but at a different time had a hit rate of 48,6% which is not robust.

These results lead to the discussion if "actions actually do speak louder than words". In case of the Play Store, the answer to this question is not black or white. The variables average rating and ranking were not the biggest influencers of adoption. Instead, the total numbers of ratings, when high, influenced the number of installations the most. So do actions speak louder than words? Not really. Words are still important, especially when they are quantified. And quantity does not mean much without words.

5.2. Academic contribution and managerial implications

The aim of this paper was to contribute to the theory and the practical market and to understand the role of social interactions on the adoption of applications. The paper is academically relevant as it further explores the relationship between word of mouth and observational learning. This interrelationship has mainly been researched in web shops and specifically in the digital camera product group, and an online forum. To date it has not been researched in the upcoming and fast growing software distribution platform market. The findings in this research partly support past papers. But the findings also provide food for thought on the effect of time or age on word of mouth and silence of mouth. What makes this research unique and therefore adds to academic researches is the use of an ordered probit model.

This research adds to the field of marketing and specifically to the app store optimization theories. It is advised that application developers more actively approach their users to leave a rating as the number of ratings affects the number of installations the most. Whether an application has in-app or no in-app purchases proved to be of no significant effect on the adoption. Although the parameters of this ordered probit model can not be translated to other datasets, an ordered probit model can still help developers to predict which social interaction to push in order to grown their installation numbers. Unfortunately, the algorithm of ranking is still unknown, but ranking is only an indicator of success instead of a mean to success.

5.3. Limitations and directions for future research

This research had multiple limitations. First the variable ranking. To start, every variable used in the research was global, only ranking is per country. Rankings of other countries also have an influence on the number of installations but were not used in this research. In addition, ranking is constructed of multiple unknown factors. The algorithm Google uses is still a subject of debate.
Second, the adoption of the applications is publicly available on an aggregate level. This means that only the people who adopt are being observed, resulting in the considerations of consumers who decide not to install to be unknown.
Finally, Google does not disclose the exact number of installations, but only in a categorized form. To make matters more complicated, the categories are all dissimilar. Exact numbers of installation of the applications would have made it easier to predict the influential effect of the different variables.

Further research could dig deeper in the motivation to leave a rating. Ratings can come from intrinsic and extrinsic motivations. If one motivation is preferred over the other, marketing activities could be focused on attracting those ratings. For example, in the dataset used there is one application called Paperama. Paperama has a significantly higher amount of ratings in comparison to other applications. Even more ratings than applications with ten times more installations. Despite this, in the two month time slot between data collection, their ranking decreased with 45.

Also, Zhang (2010) showed in her research that the choice by the first consumer affects the choices of the consumers thereafter. The research by Zhang was in the donor kidney market. This leads to the question, how much influence does the first rating have on an application. It would be an interesting research subject as it portrays a completely different market than the donor kidney market.

Moreover, price, which has a significant effect on the installation success, was added to the research in the form of a binary variable. Instead of adding it in a binary fashion, it could be added in a continuous order to see what the optimal price setting strategy is.

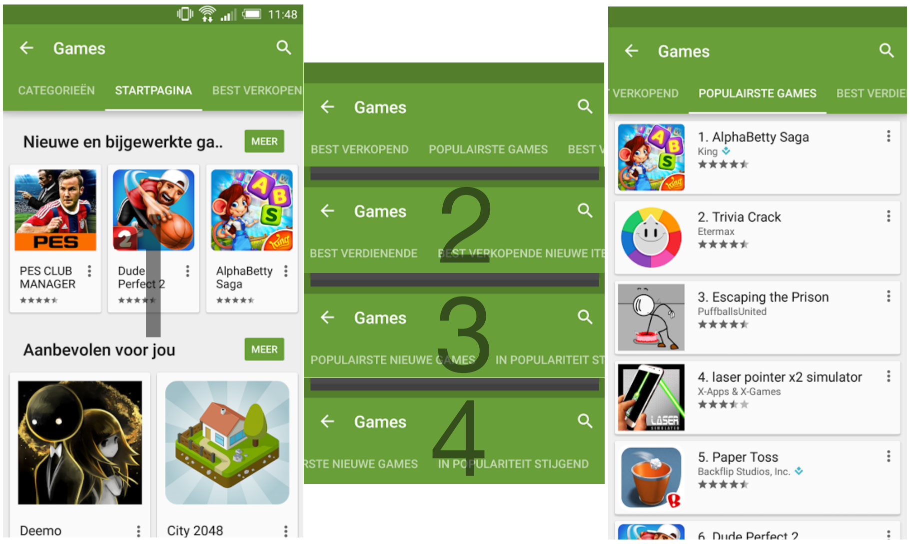
Furthermore, ranking was the reason for the hardest limitations to overcome. Another research could include the rankings of other countries. Not all the rankings are as important to growth as each country represents only a fraction of the total market. However, this is only relevant for applications which adopt a global strategy.

And finally, rating was used to indicate word of mouth. The reviews would have been a better alternative but were not used due to the limited time. Further research could use text mining in order to research the social interaction in the application store.

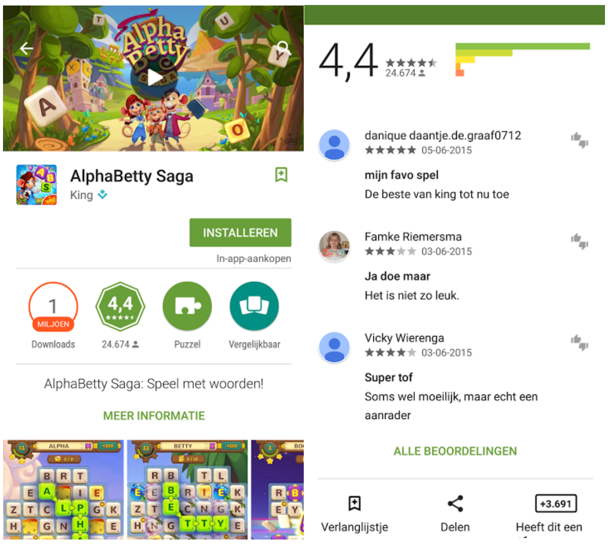
Appendix

A.

Home page of the gaming section in the Google Play Store. On top in the Play Store, sub menus can be selected: (1) Subcategories (categorieën), home page (startpagina), top selling games (best verkopende), (2) top games (populairste games), (3) top grossing games (best verdienende), top selling new games (best verkopende nieuwe items), (4) top new games (populairste nieuwe games), rising in popularity (in populariteit stijgende).



When an app is opened, information on the number of downloads, number of ratings can be viewed and read.



B.

	NO IN-APP	IN-APP	TOTAL
FREE	94	106	200
PAID	116	75	191
TOTAL	210	181	391

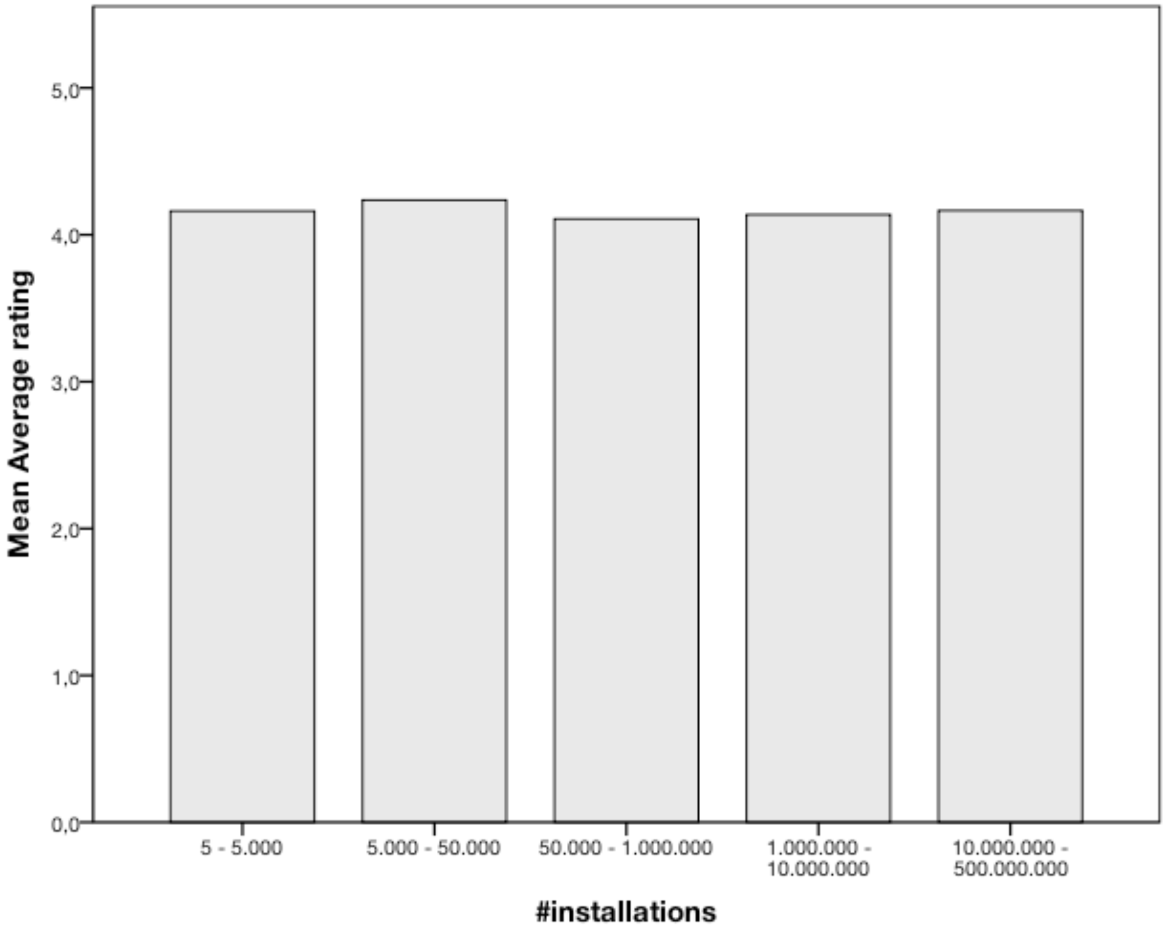
Table 3: Free/paid * No in-app/in-app purchases cross-tabulation 16/9/15

C.

OLD CATEGORIES	NEW CATEGORIES	TOTAL SAMPLE 8/7/15	TOTAL SAMPLE 16/9/15
5 - 10	5 - 5000	70	52
10 - 50	5000 - 50000	98	84
50 - 100	50000 - 1000000	128	95
100 - 500	1000000 - 10000000	115	106
500 - 1000	10000000 - 500000000	46	54
1000 - 5000			
5000 - 10000			
10000 - 50000			
50000 - 100000			
100000 - 500000			
500000 - 1000000			
1000000 - 5000000			
5000000 - 10000000			
10000000 - 50000000			
50000000 - 100000000			
100000000 - 500000000			
TOTAL		457	391

Table 5: Number of installations

D.



Graph 3: Average rating per installation category

E.

		AGE IN DAYS	RANKING	AVERAGE RATING	TOTAL # OF RATINGS	FREE OR PAID	IN-APP OR NO IN-APP
AGE IN DAYS	Pearson Correlation	1	-0,037	,130**	0,023	,325**	-0,050
	Sig. (2-tailed)		0,436	0,006	0,630	0,000	0,284
	N	457	457	457	457	457	457
RANKING	Pearson Correlation	-0,037	1	-,166**	-,127**	-0,017	0,057
	Sig. (2-tailed)	0,436		0,000	0,007	0,713	0,225
	N	457	457	457	457	457	457
AVERAGE RATING	Pearson Correlation	,130**	-,166**	1	0,043	,223**	,174**
	Sig. (2-tailed)	0,006	0,000		0,364	0,000	0,000
	N	457	457	457	457	457	457
TOTAL # OF RATINGS	Pearson Correlation	0,023	-,127**	0,043	1	-,230**	,121**
	Sig. (2-tailed)	0,630	0,007	0,364		0,000	0,009
	N	457	457	457	457	457	457
FREE OR PAID	Pearson Correlation	,325**	-0,017	,223**	-,230**	1	-0,083
	Sig. (2-tailed)	0,000	0,713	0,000	0,000		0,077
	N	457	457	457	457	457	457
IN-APP OR NO IN-APP	Pearson Correlation	-0,050	0,057	,174**	,121**	-0,083	1
	Sig. (2-tailed)	0,284	0,225	0,000	0,009	0,077	
	N	457	457	457	457	457	457

**** CORRELATION IS SIGNIFICANT AT THE 0.01 LEVEL (2-TAILED).**

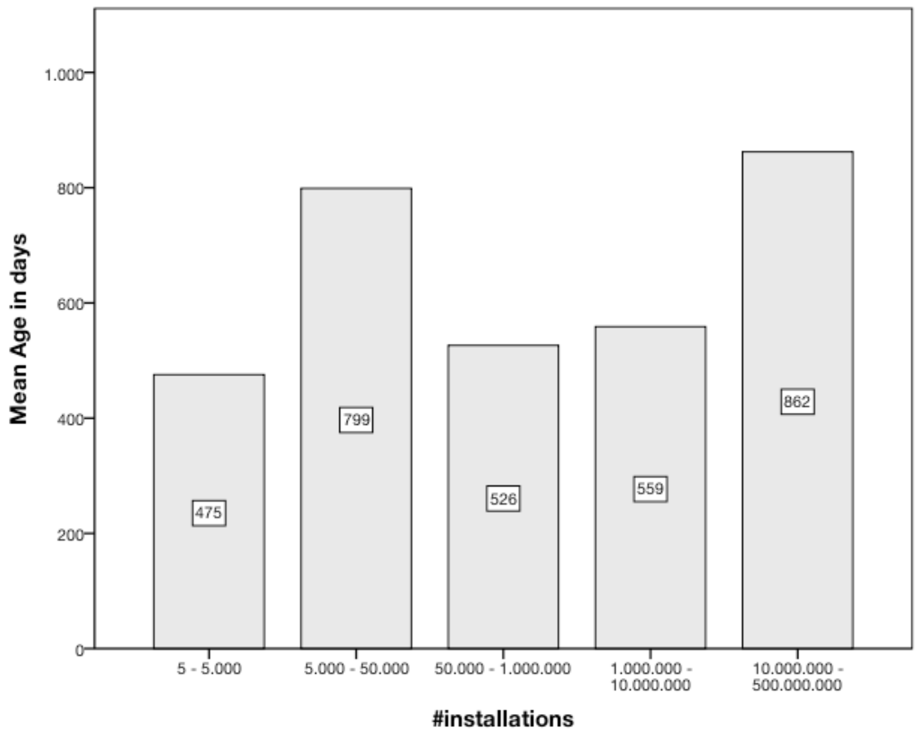
Table 6: Correlation matrix

F.

	N	MINIMUM	MAXIMUM	MEAN	STD. DEVIATION
AGE IN DAYS	457	1	1825	619,02	442,500
RANKING	457	1	486	128,47	89,799
AVERAGE RATING	457	2,2	5,0	4,157	0,4328
TOTAL # OF RATINGS	457	2	5474464	84102,14	318082,362
FREE OR PAID	457	0	1	0,47	0,499
IN-APP OR NO IN-APP	457	0	1	0,46	0,499
#INSTALLATIONS	457	1	5	2,93	1,217
VALID N (LISTWISE)	457				

Table 7: Descriptive statistics

G.



Graph 4: Average age per installation category

H.

Threshold equation:

$y_i = 1$ if $-\infty < y_i^* \le \mu_1$

$y_i = 2$ if $\mu_1 < y_i^* \le \mu_2$

$y_i = 3$ if $\mu_2 < y_i^* \le \mu_3$

$y_i = 4$ if $\mu_3 < y_i^* \le \mu_4$

$y_i = 5$ if $\mu_4 < y_i^* < \infty$

Probability equations:

$$P(y_i = 1) = P(y_i^* \le \mu_1) = \phi(\mu_1 - x'_i \beta)$$

$$P(y_i = 2) = P(\mu_1 < y_i^* \le \mu_2) = P(y_i^* \le \mu_2) - P(y_i^* \le \mu_1) = \phi(\mu_2 - x'_i \beta) - \phi(\mu_1 - x'_i \beta)$$

$$P(y_i = 3) = P(\mu_2 < y_i^* \le \mu_3) = P(y_i^* \le \mu_3) - P(y_i^* \le \mu_2) = \phi(\mu_3 - x'_i \beta) - \phi(\mu_2 - x'_i \beta)$$

$$P(y_i = 4) = P(\mu_3 < y_i^* \le \mu_4) = P(y_i^* \le \mu_4) - P(y_i^* \le \mu_3) = \phi(\mu_4 - x'_i \beta) - \phi(\mu_3 - x'_i \beta)$$

$$P(y_i = 5) = 1 - P(y_i = 1) - P(y_i = 2) - P(y_i = 3) - P(y_i = 4) = 1 - P(y_i^* \le \mu_4) = 1 - \phi(\mu_4 - x'_i \beta)$$

I.

			ESTIMATE	STD. ERROR	WALD	DF	SIG.	95% CONFIDENCE INTERVAL	
								LOWER BOUND	UPPER BOUND
1	THRESHOLD	[INSTALLS = 1]	-0,820	0,516	2,525	1	0,112	-1,831	0,191
		[INSTALLS = 2]	0,339	0,514	0,433	1	0,510	-0,670	1,347
		[INSTALLS = 3]	1,839	0,514	12,781	1	**0,000**	0,831	2,847
		[INSTALLS = 4]	3,128	0,524	35,611	1	**0,000**	2,101	4,155
	LOCATION	RATING	0,546	0,127	18,398	1	**0,000**	0,297	0,796
		FREE_PAID	-2,684	0,161	278,556	1	**0,000**	-2,999	-2,369
		IN_APP	-0,002	0,105	0,000	1	0,988	-0,208	0,205
2	THRESHOLD	[INSTALLS = 1]	-2,688	0,566	22,532	1	**0,000**	-3,798	-1,578
		[INSTALLS = 2]	-1,317	0,558	5,564	1	**0,018**	-2,411	-0,223
		[INSTALLS = 3]	0,363	0,550	0,434	1	0,510	-0,716	1,442
		[INSTALLS = 4]	1,792	0,556	10,375	1	**0,001**	0,702	2,882
	LOCATION	RATING	0,403	0,132	9,305	1	**0,002**	0,144	0,662

			ESTIMATE	STD. ERROR	WALD	DF	SIG.	95% CONFIDENCE INTERVAL	
		RANK	-0,007	0,001	88,547	1	**0,000**	-0,008	-0,005
		FREE_PAID	-3,054	0,177	297,767	1	**0,000**	-3,401	-2,707
		IN_APP	0,044	0,110	0,162	1	0,687	-0,171	0,260
3	**THRESHOLD**	[INSTALLS = 1]	-3,008	0,000		1		-3,008	-3,008
		[INSTALLS = 2]	-1,689	0,000		1		-1,689	-1,689
		[INSTALLS = 3]	0,047	0,000		1		0,047	0,047
		[INSTALLS = 4]	2,037	0,000		1		2,037	2,037
	LOCATION	RATING	0,171	0,000		1		0,171	0,171
		RANK	-0,005	0,000		1		-0,005	-0,005
		SUM_RATING	6,942E-06	0,000		1		6,942E-06	6,942E-06
		FREE_PAID	-2,595	0,000		1		-2,595	-2,595
		IN_APP	-0,035	0,000		1		-0,035	-0,035
4	**THRESHOLD**	[INSTALLS = 1]	-3,536	0,628	31,665	1	**0,000**	-4,768	-2,305
		[INSTALLS = 2]	-1,746	0,610	8,180	1	**0,004**	-2,942	-0,549
		[INSTALLS = 3]	0,499	0,600	0,690	1	0,406	-0,678	1,675
		[INSTALLS = 4]	3,491	0,646	29,181	1	**0,000**	2,225	4,758
	LOCATION	RATING	0,160	0,145	1,203	1	0,273	-0,126	0,445
		RANK	-0,007	0,001	76,602	1	**0,000**	-0,009	-0,006
		SUM_RATING	7,830E-06	1,239E-06	39,948	1	**0,000**	5,402E-06	1,026E-05
		AGE	0,002	0,000	131,651	1	**0,000**	0,002	0,002
		FREE_PAID	-4,238	0,261	263,395	1	**0,000**	-4,750	-3,726
		IN_APP	0,034	0,124	0,076	1	0,783	-0,209	0,278
5	**THRESHOLD**	[INSTALLS = 1]	-2,686	0,643	17,440	1	**0,000**	-3,947	-1,426
		[INSTALLS = 2]	-0,912	0,627	2,114	1	0,146	-2,141	0,317
		[INSTALLS = 3]	1,456	0,628	5,369	1	**0,020**	0,225	2,688
		[INSTALLS = 4]	5,500	0,767	51,478	1	**0,000**	3,998	7,003
	LOCATION	RATING	0,332	0,151	4,841	1	**0,028**	0,036	0,627
		RANK	-0,007	0,001	72,968	1	**0,000**	-0,009	-0,006
		SUM_RATING	0,000	1,933E-05	35,946	1	**0,000**	7,799E-05	0,000
		AGE	0,002	0,000	112,474	1	**0,000**	0,002	0,002
		FREE_PAID	-4,065	0,270	226,887	1	**0,000**	-4,594	-3,536
		IN_APP	0,060	0,130	0,217	1	0,641	-0,194	0,314
		RATING * SUM_RATING	-2,330E-05	4,088E-06	32,490	1	**0,000**	-3,131E-05	-1,529E-05

			ESTIMATE	STD. ERROR	WALD	DF	SIG.	95% CONFIDENCE INTERVAL	
6	THRESHOLD	[INSTALLS = 1]	-2,841	0,643	19,520	1	**0,000**	-4,102	-1,581
		[INSTALLS = 2]	-1,102	0,627	3,084	1	0,079	-2,331	0,128
		[INSTALLS = 3]	1,157	0,625	3,426	1	0,064	-0,068	2,382
		[INSTALLS = 4]	4,489	0,706	40,470	1	**0,000**	3,106	5,872
	LOCATION	RATING	0,300	0,150	4,020	1	**0,045**	0,007	0,593
		RANK	-0,008	0,001	73,546	1	**0,000**	-0,010	-0,006
		SUM_RATING	7,149E-05	1,583E-05	20,389	1	**0,000**	4,046E-05	0,000
		AGE	0,002	0,000	109,398	1	**0,000**	0,002	0,002
		FREE_PAID	-3,973	0,259	235,507	1	**0,000**	-4,480	-3,465
		IN_APP	0,108	0,128	0,718	1	0,397	-0,142	0,359
		RATING * SUM_RATING	-1,469E-05	3,385E-06	18,832	1	**0,000**	-2,133E-05	-8,056E-06
		RANK * SUM_RATING	3,032E-08	1,499E-08	4,094	1	**0,043**	9,497E-10	5,970E-08
7	THRESHOLD	[INSTALLS = 1]	-3,467	0,895	14,997	1	**0,000**	-5,222	-1,712
		[INSTALLS = 2]	-1,713	0,873	3,848	1	0,050	-3,425	-0,001
		[INSTALLS = 3]	0,537	0,875	0,377	1	0,539	-1,178	2,252
		[INSTALLS = 4]	3,863	0,936	17,034	1	**0,000**	2,028	5,697
	LOCATION	RATING	0,150	0,211	0,503	1	0,478	-0,264	0,563
		RANK	-0,008	0,001	73,889	1	**0,000**	-0,010	-0,006
		SUM_RATING	7,172E-05	1,593E-05	20,282	1	**0,000**	4,051E-05	0,000
		AGE	0,001	0,001	0,323	1	0,570	-0,002	0,003
		FREE_PAID	-3,967	0,259	234,903	1	**0,000**	-4,474	-3,459
		IN_APP	0,111	0,128	0,755	1	0,385	-0,139	0,361
		RATING * SUM_RATING	-1,474E-05	3,405E-06	18,742	1	**0,000**	-2,142E-05	-8,068E-06
		RANK * SUM_RATING	3,088E-08	1,510E-08	4,185	1	**0,041**	1,295E-09	6,047E-08
		RATING * AGE	0,000	0,000	0,985	1	0,321	0,000	0,001
8	THRESHOLD	[INSTALLS = 1]	-3,634	0,928	15,319	1	**0,000**	-5,454	-1,814
		[INSTALLS = 2]	-1,891	0,913	4,288	1	**0,038**	-3,681	-0,101
		[INSTALLS = 3]	0,375	0,906	0,171	1	0,679	-1,401	2,151
		[INSTALLS = 4]	3,706	0,965	14,739	1	**0,000**	1,814	5,598
	LOCATION	RATING	0,140	0,212	0,434	1	0,510	-0,276	0,555

	ESTIMATE	STD. ERROR	WALD	DF	SIG.	95% CONFIDENCE INTERVAL	
RANK	-0,009	0,002	31,050	1	**0,000**	-0,012	-0,006
SUM_RATING	7,134E-05	1,594E-05	20,039	1	**0,000**	4,011E-05	0,000
AGE	0,000	0,001	0,150	1	0,699	-0,002	0,003
FREE_PAID	-3,989	0,262	231,119	1	**0,000**	-4,503	-3,475
IN_APP	0,097	0,129	0,564	1	0,452	-0,156	0,351
RATING * SUM_RATING	-1,467E-05	3,406E-06	18,544	1	**0,000**	-2,134E-05	-7,991E-06
RANK * SUM_RATING	3,094E-08	1,514E-08	4,178	1	**0,041**	1,271E-09	6,062E-08
RATING * AGE	0,000	0,000	1,082	1	0,298	0,000	0,001
RANK * AGE	1,187E-06	1,719E-06	0,477	1	0,490	-2,183E-06	4,557E-06

Table 9: parameter estimates in blocks

J.

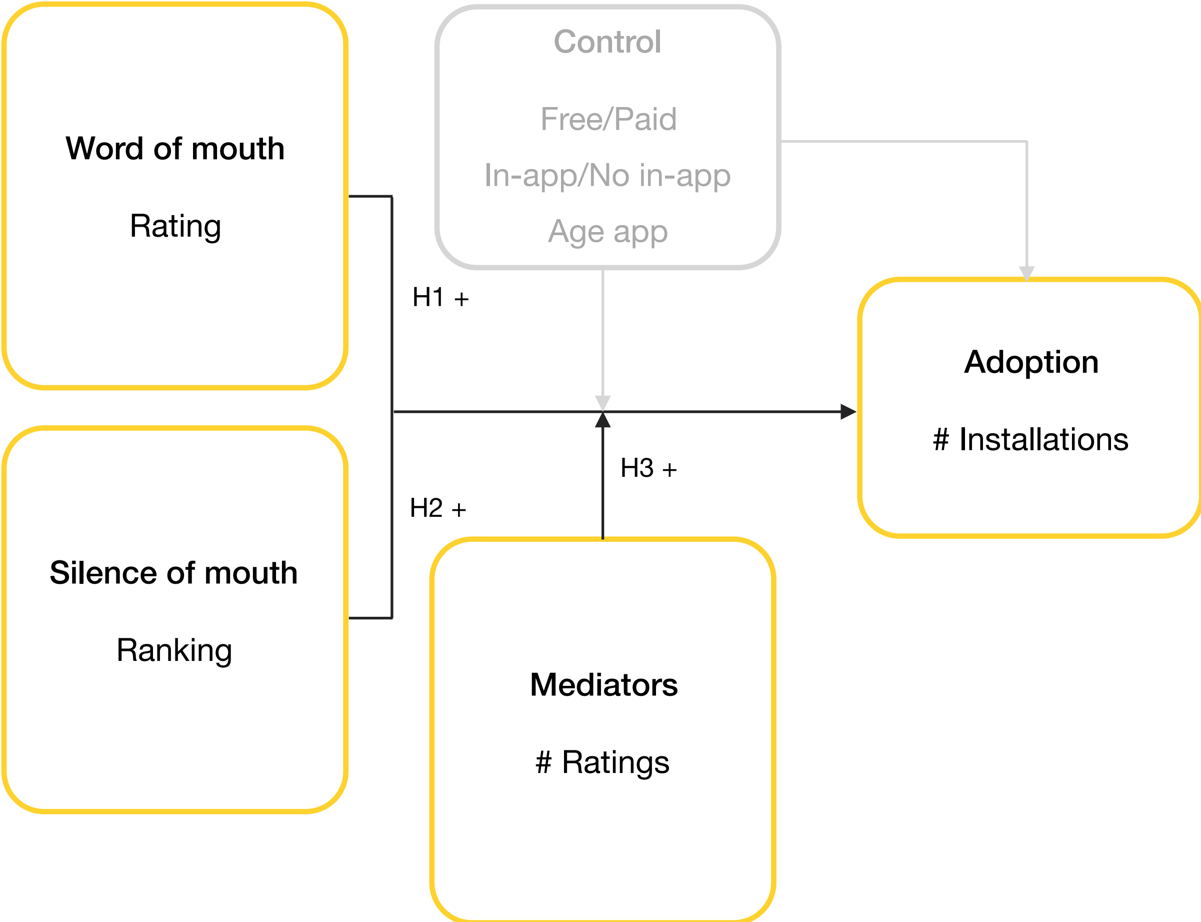


Figure 3. New conceptual model

Bibliography

AppAnnie. (2015a). *App Annie Index Market Q1 2015*. Retrieved from https://s3.amazonaws.com/files.appannie.com/reports/App-Annie-Index-Market-Q1-2015.pdf?mkt_tok=3RkMMJWWfF9wsRokua3PZKXonjHpfsX87u0rUaSzIMI%2F0ER3fOvrPUfGjI4DSsBkI%2BSLDwEYGJlv6SgFSrfEMbFl3rgPUxU%3D

AppAnnie. (2015b). *The State of Play: A Look at the Growth of Google Play*. Retrieved from https://s3.amazonaws.com/files.appannie.com/reports/The-State-of-Play-Google-IO-2015-EN.pdf?mkt_tok=3RkMMJWWfF9wsRokvqXJZKXonjHpfsX87u0rUaSzIMI%2F0ER3fOvrPUfGjI4CTMtgI%2BSLDwEYGJlv6SgFSrfEMbFl3rgPUxU%3D

Apple. (2015). Make in-app purchases. Retrieved from https://support.apple.com/en-us/HT202023

Arndt, J. (1967). Role of product-related conversations in the diffusion of a new product. *Journal of marketing research*, 291-295.

Banerjee, A. V. (1992). A Simple-Model of Herd Behavior. *Quarterly Journal of Economics, 107*(3), 797-817.

Bettman, J. R. (1979). An information processing theory of consumer choice. *Journal of Marketing*.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 992-1026.

Brown, J. J., & Reingen, P. H. (1987). Social Ties and Word-of-Mouth Referral Behavior. *Journal of consumer research, 14*(3), 350-362.

Butters, K. (2014). How Does Google Play Rank Mobile Apps. Retrieved from https://getappcase.com/blog/app-stores/how-does-google-play-rank-mobile-apps

Carare, O. (2012). The Impact of Bestseller Rank on Demand: Evidence from the App Market. *International Economic Review, 53*(3), 717-742.

Chen, Y., Wang, Q., & Xie, J. (2011). Online social interactions: A natural experiment on word of mouth versus observational learning. *Journal of marketing research, 48*(2), 238-254.

Cheung, C. M. K., Xiao, B., & Liu, I. L. B. (2012). The Impact of Observational Learning and Electronic Word of Mouth on Consumer Purchase Decisions: The Moderating Role of Consumer Expertise and Consumer Involvement. 3228-3237.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research, 43*(3), 345-354.

Cuadrado, F., & Dueñas, J. C. (2012). Mobile application stores: success factors, existing approaches, and future developments. *Communications Magazine, IEEE, 50*(11), 160-167.

Engel, J. F., Blackwell, R. D., & Miniard, P. W. (1986). *Consumer behavior* (fifth edition ed.): New York: The Dryden Press.

Field, A. (2013). Discovering statistics using IBM SPSS statistics: Sage.

Godes, D., Mayzlin, D., Chen, Y., Das, S., Dellarocas, C., Pfeiffer, B., . . . Verlegh, P. (2005). The firm's management of social interactions. *Marketing Letters, 16*(3-4), 415-428.

Greene, W. H. (2003). *Econometric analysis*: Pearson Education India.

Griskevicius, V., & Kenrick, D. T. (2013). Fundamental motives: How evolutionary needs influence consumer behavior. *Journal of Consumer Psychology, 23*(3), 372-386.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? *Journal of interactive marketing, 18*(1), 38-52.

Holbrook, M. B., & Hirschman, E. C. (1982). The Experiential Aspects of Consumption - Consumer Fantasies, Feelings, and Fun. *Journal of consumer research, 9*(2), 132-140.

Huang, P., Lurie, N. H., & Mitra, S. (2009). Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods. *Journal of Marketing, 73*(2), 55-69.

Jarvenpaa, S. L., Tractinsky, N., & Saarinen, L. (1999). Consumer trust in an internet store: a cross-cultural validation. *Journal of Computer-Mediated Communication, 5*(2), 0-0.

Klein, L. R. (1998). Evaluating the potential of interactive media through a new lens: Search versus experience goods. *Journal of business research, 41*(3), 195-203.

Monaghan, C., & Neumayr, T. (2015). App Store Rings in 2015 with New Records [Press release]. Retrieved from https://www.apple.com/pr/library/2015/01/08App-Store-Rings-in-2015-with-New-Records.html

Murray, K. B. (1991). A Test of Services Marketing Theory - Consumer Information Acquisition Activities. *Journal of Marketing, 55*(1), 10-25.

Nelson, P. (1974). Advertising as information. *The journal of political economy*, 729-754.

Newell, A. (1968). Judgment and its representation: An introduction. *Formal representation of human judgment. New York: Wiley*.

Rogers, E. M. (2010). Diffusion of innovations: Simon and Schuster.

Senecal, S., & Nantel, J. (2004). The influence of online product recommendations on consumers' online choices. *Journal of Retailing, 80*(2), 159-169.

Simon, H. A. (1957). Models of man; social and rational.

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational behavior and human performance, 6*(6), 649-744.

Sorensen. (2007). Bestseller lists and product variety. *Journal of Industrial Economics, 55*(4), 715-738.

Stern, W. (1995). Did Dirty Tricks Create a Best Seller? *Business Week, 7*, 195-196.

Sundaram, D. S., Mitra, K., & Webster, C. (1998). Word-of-mouth communications: A motivational analysis. *Advances in consumer research, 25*(1), 527-531.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization, 1*(1), 39-60.

Trusov, M., Bucklin, R. E., & Pauwels, K. (2009). Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site. *Journal of Marketing, 73*(5), 90-102.

VanderWeele, T. J. (2014). A unification of mediation and interaction: a 4-way decomposition. *Epidemiology, 25*(5), 749-761.

Wallenstein, A. (2015). How Google Play Is Closing The App Gap With Apple. Retrieved from http://variety.com/2015/digital/features/google-play-apple-itunes-app-store-1201449326/

Walz, A. (2015). Deconstructing the App Store Rankings Formula with a Little Mad Science. Retrieved from https://moz.com/blog/app-store-rankings-formula-deconstructed-in-5-mad-science-experiments

Word of mouth. (2015). Oxford University Press.

Zhang, J. (2010). The Sound of Silence: Observational Learning in the U.S. Kidney Market. *Marketing Science, 29*(2), 315-335.