Logo

Description automatically generated with medium confidence

TITLE

SUBTITLE

Marc Lefebvre

SNR: 2040052

January 2021

Master Thesis Marketing Analytics

Marketing Department

Tilburg School of Economics and Management

Tilburg University

Master Thesis Supervisor: Bobbie Krijger *Tilburg University*

Co-reader: Constant Pieters *Tilburg University*

# Management Summary

# Preface

# Table of contents

[Management Summary 2](#_Toc89013209)

[Preface 3](#_Toc89013210)

[Table of contents 4](#_Toc89013211)

[1. Introduction 6](#_Toc89013212)

[2. Theoretical framework 11](#_Toc89013213)

[2.1 Conceptual framework 11](#_Toc89013214)

[2.2 Hypotheses 11](#_Toc89013215)

[2.2.1 The main effect of expert reviews on Nr. of downloads on Steam 11](#_Toc89013216)

[2.2.2 The main effect of user reviews on Nr. of downloads on Steam 12](#_Toc89013217)

[2.2.3 The difference in effect of user- and expert reviews on Nr. of downloads on Steam 12](#_Toc89013218)

[2.2.4 The moderating effect of game mode 13](#_Toc89013219)

[2.2.5 The moderating effect of publisher 14](#_Toc89013220)

[3. Method 16](#_Toc89013221)

[3.1 Data collection 16](#_Toc89013222)

[3.2 Data processing 17](#_Toc89013223)

[3.3 Data description 18](#_Toc89013224)

[3.4 Model 19](#_Toc89013225)

[4. Results 21](#_Toc89013226)

[4.1 Assumptions of the OLS regression and model fit 21](#_Toc89013227)

[4.1.1 Heteroskedasticity 21](#_Toc89013228)

[4.1.2 Multicollinearity 21](#_Toc89013229)

[4.1.3 Model fit 22](#_Toc89013230)

[4.2 Main effects 22](#_Toc89013231)

[4.2.1 The main effect of expert reviews 23](#_Toc89013232)

[4.2.2 The main effect of user reviews 23](#_Toc89013233)

[4.2.3 The difference between expert- and user reviews 24](#_Toc89013234)

[4.3 Moderating effects 24](#_Toc89013235)

[4.3.1 The moderating effect of game mode 24](#_Toc89013236)

[4.3.2 The moderating effect of publisher 24](#_Toc89013237)

[5. Conclusions and recommendations 25](#_Toc89013238)

[5.1 Conclusion 25](#_Toc89013239)

[5.2 Discussion 25](#_Toc89013240)

[5.3 Recommendations 25](#_Toc89013241)

[5.4 Limitations and future research 25](#_Toc89013242)

[References 26](#_Toc89013243)

[Appendices 29](#_Toc89013244)

[A. GitHub repository 29](#_Toc89013245)

[B. Regression table R 30](#_Toc89013246)

# 1. Introduction

The video game industry has experienced meteoric growth. The initial forecast for the video game market revenue in 2020 was $159.3 billion (Newzoo, 2020). However, due to the increase in time spent at home amidst the COVID-19 pandemic the global video game market generated a baffling $177.8 billion in 2020 (Newzoo, 2021). In addition to a spike in revenues, the market has also seen an increase in time spent playing video games. A survey conducted in the US reported that gamers spent 45% more time playing video games amid the quarantine (Bloomberg, 2020). Moreover, one of the leading online video game retailers, Steam, hit a peak of 20.3 million concurrent players which is 11% above their previous high (The Economist, 2020). In addition to an increase in playing time, the amount of people playing video games, also known as gamers, has increased. According to the Global Games Market Report by Newzoo, the amount of gamers has already grown by 5.4% from 2020 to 2021. Of the total 3 billion gamers worldwide, 1.4 billion are active on a personal computer (PC) and 0.9 billion on a console (e.g. ps4, Wii, Xbox) (Newzoo, 2021). According to a survey conducted by brightlocal, 83% of consumers in the entertainment industry look at online reviews before making a decision (brightlocal, 2020). Also, McKinsey & Company found that there was a surge in volume of online ratings and reviews, which were 87% higher in December 2020 than in December 2019 (McKinsey, 2021). The growth of the video game industry, the surge of online reviews and the importance of online reviews in consumer decision making emphasizes the significance of understanding the influence that online reviews have on video games. Against this background, this study sets out to assess the effect of online review rating on the number of PC game downloads on Steam, and to what extent video game characteristics influence this effect.

A body of prior literature that is related to the effect of online reviews on sales has been identified, and is presented in Table 1. There are three streams of literature that relate to this study. The first stream of literature investigates the effect of expert reviews on sales related outcomes in different contexts. For instance, Basuroy et al. (2003) conducted research in the movie industry and found that positive and negative expert reviews influence box office performance. A research conducted by Cox (2014) in the video game industry had similar results, whose findings show that video games with higher expert ratings are more likely to sell a larger number of units.

The second stream of literature examines the effect of user reviews on sales related

**Table 1: Prior research related to Expert- and User Reviews and their respective effect on sales related outcomes in different contexts**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Article** | **Context** | **Expert** | **User** | **Expert vs. User** | **Dependent variable** | **Key findings** | **Method** |
| (Chevalier & Mayzlin, 2006) | Books | – | ✓ | – | Book sales | Improvement in reviews leads to increase in relative sales | Differences-in-Differences |
| (Ho-Dac et al., 2013) | Blu-ray & DVD | – | ✓ | – | Blu-ray & DVD sales | Positive reviews increase sales of weak brands, no significant impact on strong brands | Multiple linear regression |
| (Öǧüt & Onur Taş, 2012) | Hotels | – | – | ✓ | Hotel room sales | A higher user rating significantly increases online sales, no significant impact of expert (star) rating | Log-linear regression |
| (Ye et al., 2009) | Hotels | – | ✓ | – | Hotel rooms bookings | Significant relationship between online reviews and hotel room sales | Fixed log-linear regression |
| (Chintagunta et al., 2010) | Movies | – | ✓ | – | Box office earnings | Valence\* has a significant and positive impact on box office earnings | OLS regression |
| (Basuroy et al., 2003) | Movies | ✓ | – | – | Box office performance | Positive and negative expert reviews predict and influence box office performance. Furthermore, there is evidence of negativity bias | Time-series cross-section regression |
| (Tsao, 2014) | Movies | – | – | ✓ | Movie evaluations | Greater importance is attached to user reviews than expert reviews | Between-subject factorial design |
| (Zhu & Zhang, 2010) | Video Games | ✓ | – | – | Video game sales on console | Online reviews are more influential for less popular games, and if gamers have greater internet experience | Differences-in-Differences |
| (Zhu & Zhang, 2006) | Video Games | – | ✓ | – | Video game sales on console | Online consumer reviews have a significant influence on the sale of video games | Differences-in-Differences |
| (Cox, 2014) | Video Games | ✓ | – | – | Video game sales on console | Games with higher expert ratings are significantly more likely to sell a greater number of units | OLS regression |

*\*Valence: Mean user rating*

outcomes in an array of different contexts, such as the hotel (Ye et al., 2009), movie (Chintagunta et al., 2010), and video game industry (Zhu & Zhang, 2006). In addition, research has been conducted on the sale of books and DVDs. For instance, Chevalier & Mayzlin (2006) found that the improvement of online book reviews leads to an increase in sales. Furthermore, research by Ho-Dac et al. (2013) on the sale of DVDs found that positive user review ratings increase sales of weak brands, however, do not have a significant effect on strong brands.

The third stream of literature is concerned with the difference between effect of expert- and user ratings on sales related outcomes in different contexts. For example, Tsao (2014) found that greater importance is attached to user reviews than expert reviews in the movie industry. Furthermore, research conducted in the hotel industry by Ögüt & Onur Tas (2012) found that a higher user rating significantly increases online sales, whilst expert rating has no significant influence.

The main focus of this study is positioned in the third stream of literature. In addition, the distinct effects of expert- and user reviews, which are the first and second stream of literature, are included. Furthermore, and contrary to prior research in the gaming industry, this study is focused on PC games instead of console games. In order to further discuss the relevance of this study, it is necessary to clarify and elaborate on the difference between PC and console gamer audiences. First, the differences in hardware (PC vs. console) will be discussed. Second, the different types of gamers will be described.

PC is notorious for its superiority in terms of computing power, which improves the graphics and speed of video gameplay. The hardware can continually be upgraded by the owner of a PC in order to keep it up-to-date. However, installing hardware upgrades require additional investment and knowledge of the PC. Also, the initial investment for a new PC is significantly higher than a new console. Although the initial investment for a console is relatively low, games have been found to be more expensive. Furthermore, consoles are built for usability, ease-of-use and no prior knowledge of the system is necessary due to the fact that hardware upgrades are not an option. An upgrade of console hardware only happens when a new console is released (Pakhrani et al., 2020; Johnson, 2021).

According to market research conducted by Clairfield International, gamers can be categorized into three different groups; hardcore gamers, casual gamers, and mass market gamers. Hardcore gamers are mainly categorized by their hefty investments in state-of-the-art gaming equipment, and the large amount of time they spend playing video games. Casual gamers play games regularly, but not as much as hardcore gamers. Moreover, casual gamers have been known to be far more price-sensitive than hardcore gamers. Mass market gamers do not spend a lot of time playing video games (Clairfield International, 2018). Although all three categories can be spread across PC and console, the amount of hardcore gamers is still skewed towards PC. According to the CEO of anzu.io (A sophisticated in-game advertising company), PC games are a form of video game that usually attracts gamers with a hardcore streak. Therefore, PC may serve as a proxy for hardcore gamers who spend more time and money on gaming. In addition, as game review ratings are most often presented on websites and PC gamers are already active on their PCs, it intuitively can be said that the threshold is lower to view review ratings on PC than on console. Moreover, on the Steam platform user review ratings are readily presented which further lowers the threshold for this segment. Also, when taking the price-sensitivity of casual gamers into consideration, it can be said that casual gamers would most likely prefer a console over a PC. All the beforementioned reasons indicate a difference in the consumer segment for PC and console gamers, and thus make the PC audience an interesting and relevant consumer group to investigate.

This study has multiple contributions to existing literature. Firstly, the main aim of is to research the difference of effect, and thus the relative importance, of expert- and user reviews on sales related outcomes in the video game industry. Prior research has investigated the difference in importance between expert- and user reviews, however, this was focused on the movie- and hotel industry (Öǧüt & Onur Taş, 2012; Tsao, 2014). This entails that this study will present new insights into the relative importance of expert- and user reviews in the video game industry.

Secondly, this study examines the effect of expert reviews on sales related outcomes for PC games. As described before, there is an evident difference between PC and console gamers. Prior studies on the influence of expert reviews in the video game industry were explicitly focused on console games (Cox, 2014; Zhu & Zhang, 2010). Therefore, this study contributes to the existing literature on the effect of expert reviews in the video game industry by adding a new dimension, namely PC games.

Finally, this study focuses on the effect of user reviews on sales related outcomes for PC games. Prior research has been conducted for this effect on the sale of books (Chevalier & Mayzlin, 2006), DVDs (Ho-Dac et al., 2013), and movies (Basuroy et al., 2003; Chintagunta et al., 2010). In terms of video games, only Zhu & Zhang (2006) researched the effect of user reviews on console game sales. Therefore, this research contributes to the existing literature by concentrating on PC games as a new dimension.

In summary, this study investigates what the effect of expert- and user review ratings is on sales related outcomes for PC games in the video game industry. To be more specific, the sales related outcome that will be used in the analysis is the number of downloads a game has on the Steam platform. In addition to the distinctive effects, the relative importance of expert- and user reviews will be examined. Furthermore, video game characteristics will be included as moderators. The video game characteristics that will be included are publisher (i.e. major- or minor publisher) and game mode (i.e. single- or multiplayer).

The findings of this research provide managers and marketers of minor- or major publishers with information regarding the relative importance of different review types (i.e. expert- and user reviews) on video games, and in specific the effect on PC games. This can lead to strategic adjustments to influence review ratings of a video game in order to improve its sales performance.

# 2. Theoretical framework

## 2.1 Conceptual framework

The conceptual framework for this research is presented in Figure 1. The framework consists of the main effect of expert- and user reviews on number of downloads on Steam. In addition, the main effect is moderated by video game characteristics (i.e. game mode and publisher).

**Figure 1 - Conceptual Framework**

Diagram

Description automatically generated

## 2.2 Hypotheses

### 2.2.1 The main effect of expert reviews on Nr. of downloads on Steam

In order to understand the conceptual model, it is necessary to define the variables included. Review rating is separated in two distinct variables, namely Expert rating (i.e. ratings by experts, critics or professionals in a respective industry; Cox, 2014; Chen et al., 2011) and User rating (i.e. rating by a user or consumer assessing quality of product or service; Ho-Dac et al., 2013). Furthermore, number of downloads on Steam (written as: # of downloads on Steam) represents the number of downloads for a particular game on the Steam platform.

As mentioned before, a basis of extant literature exists on the effect of expert reviews on sales related outcomes. In the movie industry, Basuroy et al. (2003) found that positive (and negative) critic reviews influence and predict box office revenue. A study conducted in the movie industry presents similar results, indicating that higher expert ratings significantly impact returns (Chen et al., 2011). In terms of the video game industry, studies by both Cox (2014) and Zhu & Zhang (2010) support these findings, offering evidence that suggest that the purchasing decision of consumers in the video games market is significantly influenced by expert review scores. However, Zhu & Zhang (2010) also propose that expert reviews are more influential for less popular games. The findings of the beforementioned studies indicate a clear direction regarding the influence of expert reviews on sales related outcomes. This study argues that this holds for PC games, the reasoning behind this is twofold. Firstly, the movie- and video game industry are both entertainment based industries and so may yield similar results. Secondly, despite the dissimilarities between PC and console gamers the general effect of expert reviews will be the similar as it is still centred around video games. This leads to the following hypothesis:

: Expert reviews positively influence the number of downloads on Steam.

### 2.2.2 The main effect of user reviews on Nr. of downloads on Steam

There is an extensive basis of literature available on the effect of user reviews on sales related outcomes. A multitude of studies in different contexts have presented similar results. For instance, studies concerning the sale of books (Chevalier & Mayzlin, 2006), new products (Cui et al., 2012), mobile phones (Decker & Trusov, 2010), and rooms in hotels (Jenq, 2019; Ye et al., 2009) all found significant relationships with user reviews. More importantly, studies conducted in entertainment based industries such as the movie industry suggest similar results. For example, Chintagunta et al. (2010) found that the mean user rating is one of the main drivers of box office performance. However, a study by Liu (2006) contradicts this stating that instead of the mean user rating the volume of user reviews offers explanatory power for box office performance. Although volume might carry more explanatory power, Duan et al. (2014) suggests that the mean user rating significantly influences the volume of user reviews, which subsequently influences box office performance. In addition to the movie industry, findings from a study by Zhu & Zhang (2006) suggest that user reviews have a significant influence on the sale of video games. The findings of the video game- and movie industry, which are both entertainment based industries, combined with the general consensus of other contexts that user reviews influence sales related outcomes leads to the following hypothesis:

: User reviews positively influence the number of downloads on Steam

### 2.2.3 The difference in effect of user- and expert reviews on Nr. of downloads on Steam

Although there is a scarce amount of extant literature regarding the difference in effect of expert- and user reviews on sales related outcomes in the video game industry, there is an array of literature from other industries available. Prior research from the hotel industry found that consumers relate more to opinions of peers, rather than those of professionals who are paid to write reviews (De Langhe et al., 2016). Ögüt & Tas (2012) had similar findings, suggesting that a higher user rating for hotels significantly increases online sales whilst an expert defined star rating does not. A study from the movie industry supports these findings, stating that potential moviegoers attach greater importance to consumer reviews than to critic reviews (Tsao, 2014). However, this is contradicted by two recent studies, which suggest that expert ratings are more influential for moviegoing decisions than user ratings as experts are critical and more consistent, and that expert reviews have a larger influence on the movie industry as a whole (Basuroy et al., 2020; Souza et al., 2019). Research on the video game industry builds upon these findings, proposing that user reviews are often highly polarized, whilst expert reviews are more balanced over time (Santos et al., 2019). The beforementioned literature leads to the impression that expert reviews are valued more, relative to user reviews in an entertainment oriented industry due to their consistency and reliability. Furthermore, Research in the movie industry suggests that infrequent moviegoers are influenced to a greater extent by user reviews, whilst frequent moviegoers are influenced to a greater extent by expert reviews (Chakravarty & Mazumdar, 2016). As mentioned in the introduction, PC gamers are often more invested in the act of playing video games relative to console gamers. Therefore, this study that PC gamers may be similarly classified as frequent moviegoers, thus placing more emphasis on the importance of expert reviews. This leads to the following hypothesis:

: Expert reviews positively influence number of downloads on Steam more than user reviews.

### 2.2.4 The moderating effect of game mode

The first moderator is game mode. Game mode is the manner in which a video game can be played. This can be multi-player, which means the video game can be played (online) by more than one player. Conversely, single-player means the video game can be played by one player (Situmeang et al., 2014; Cole & Griffiths, 2007).There is a limited amount of extant literature on the influence of game mode on the relationship between expert- and user reviews and sales related outcomes. Cox (2014) found a significant effect concerning the number of players that were able to play a video game. The more players were able to play, the more games were sold (Cox, 2014). Moreover, prior research about video games has found that mode of play predicts the time spent playing, and that there is a positive relationship between the social aspects of playing video games and time spent playing (Cole & Griffiths, 2007; Johnson et al., 2016). Furthermore, Zhu & Zhang (2010) suggest that user reviews are more important for games that are played online. Although time, as mentioned by Cole & Griffiths (2007) and Johnson et al. (2016), is not a perfect indication of preference in terms of game mode, it does give an idea towards the social aspects of gaming. When taking the social aspects and the findings of Cox (2014) into consideration, it can be said that the single-player (and thus less social) games are more focused on the quality of the game (expert opinions) than the opinion of other players. Therefore, this study formulates the following hypothesis:

: The positive effect of expert reviews on the number of downloads on Steam is stronger (weaker) when game mode is single-player (multi-player)

Furthermore, based on the literature this study argues that multi-player (online) games are more likely to have a stronger influence on the relationship between user reviews and the number of downloads on Steam than single-player games. The reasoning behind this is twofold. Firstly, multi-player games are played by a large community of gamers and thus user (gamer) opinions are regarded highly. Secondly, user reviews have been found to have a significant impact on the performance of games that are played online (Zhu & Zhang, 2010). This leads to the following hypotheses:

: The positive effect of user reviews on the number of downloads on Steam is stronger (weaker) when game mode is multi-player (single-player)

### 2.2.5 The moderating effect of publisher

The second and final moderator is publisher. Publisher is the company that publishes the video game to the market for sale to the general public (Cox, 2014). In context of this study, a major publisher is a company with an annual revenue above $50 million and a minor publisher is a company with an annual revenue below $50 million. Once again, a limited amount of extant research on the video game industry is available. Prior research conducted on the movie industry found that movie stars and budgets moderated the impact of expert reviews on movies (Basuroy et al., 2003). Moreover, Souza et al. (2019) found that the influence of expert reviews on blockbusters is null, whilst the influence on small release films is large. The findings in these academic sources lead to the following hypothesis:

: The positive effect of expert reviews on the number of downloads on Steam is stronger (weaker) when the game is released by a minor publisher (major publisher).

Furthermore, a study conducted by Zhu & Zhang (2010) found that online user reviews were more influential for less popular games. In addition, Ho-Dac et al. (2013) supports this stating that online user reviews increases the sale of brands with low brand equity. The beforementioned findings lead to the following hypothesis:

: The positive effect of user reviews on the number of downloads on Steam is stronger (weaker) when the game is released by a minor publisher (major publisher).

# 3. Method

In this chapter the data collection, processing, description and statistical method will be discussed. The goal of this research is to assess what the effect of expert- and user review rating is on the number of PC game downloads on Steam, and to what extent video game characteristics influence this effect. In order to assess this effect, data on expert- and user ratings, video game characteristics (publisher and game mode), and control variables (release date and price) has to be collected. The method that will be used to assess this effect is OLS regression. All files that are mentioned in this chapter can be found in this study’s respective GitHub repository (See Appendix A).

## 3.1 Data collection

To test the conceptual framework, data for all included variables has to be collected. A detailed description of all variables is provided in Table 2. There are three sources from which the data is collected, each source will be discussed in the following paragraphs.

Firstly, data is collected from Kaggle.com, which is a website that offers its users the ability to share and collaborate on datasets. From this website a rich Steam store dataset is downloaded. The dataset is named *Steam Store Games (Clean dataset)*, and is developed and shared by Nik Davis. The dataset contains data from 27,075 game titles released between 1998 and 2019. Data on user reviews (x2), game mode (m1), publisher (m2), release date (c1, c2, c3), and price (c4) is available in this dataset.

Secondly, expert review (x1) data is collected from steamspy.com, which is a Steam stats service which automatically gathers data from Steam user profiles using a Web API. To collect data from steamspy a web scraper is developed. This web scraper uses the game IDs provided in the Kaggle dataset to create a list of unique URLs which are used to scrape all necessary web pages.

Finally, the number of downloads per game on Steam (y) is collected from SteamDB, which is a third-party website that gives insight into the Steam platform and its database. Similar to the last steamspy scraper, the collection method for SteamDB will involve a web scraper. However, in order to correctly scrape this website, the web scraper detects a temporary ban and waits until it is able to continue scraping in addition to using the game IDs to create unique URLs.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Measurement** | **Description** | **Data source** |
| # Downloads on Steam (y) | Ratio | Number of game owners on Steam | SteamDB |
| Expert review (x1) | Ratio | Average expert review rating per game | steamspy |
| User review (x2) | Ratio | Average user review rating per game | Kaggle |
| Game mode (m1) | Nominal | Dummy variable: 0 for multi-player, 1 for single-player | Kaggle |
| Publisher (m2) | Nominal | Dummy variable: 0 for minor publisher, 1 for major publisher | Kaggle |
| Release date old (c1) | Nominal | Dummy variable: 0 for > 2010, 1 for <=2010 | Kaggle |
| Release date mid (c2) | Nominal | Dummy variable: 0 for < 2011 & 2016 <, 1 for >=2011 & <= 2016 | Kaggle |
| Release date new (c3) | Nominal | Dummy variable: 0 for < 2017, 1 for >= 2017 | Kaggle |
| Price (c4) | Nominal | Dummy variable: 0 for paid, 1 for free | Kaggle |

**Table 2: Detailed description of variables**

## 3.2 Data processing

Before the collected data can be analysed, it is necessary to process the data with the goal of creating one final, clean dataset consisting of all the variables in the conceptual framework. The data cleaning process can be broken down into three succinct phases. These phases will be elaborated on further in the following paragraphs.

Phase one is aimed at cleaning the readily available Kaggle dataset. After the dataset is loaded, the first course of action is to remove all games with no playing time. This step reduces the total number of observations from 27,075 to 6,170. After this is done, the total amount of negative and positive user reviews are summed in a new column. All games with less than 1000 total user reviews are removed. The reasoning behind this is that as user review rating is a percentage, a game with a low amount of total reviews may skew the results. This further reduces the total amount of observations to 2,366. Next any unnecessary columns are removed. Finally, all game IDs of the remaining observations are extracted. These IDs are used in the steamspy and SteamDB web scraper.

Phase two is focused on importing, merging and cleaning the steamspy web scraper results. After the necessary datasets are loaded and imported, all datasets are merged. Next all expert review rows with NA values are removed, reducing the total number of observations to 1,203. This is followed by the removal of all unnecessary columns and categories. After this, the expert rating variable is changed to a percentage, and both expert rating and user rating are rounded to two decimals in order to match in terms of notation. Finally, a dummy variable is created for game mode (m1).

Phase three seeks to import, merge and clean the SteamDB web scraper results. In addition, a dummy variable for publisher (m2) is created. The first course of action is to create a dataset of all unique publishers. This dataset can subsequently be used to code the dummy variable by systematically searching for the annual revenue per publisher online[[1]](#footnote-1). When this is completed, the months and days of release date are removed in order to merely depict the year of release. This is followed by the creation of dummy variables for the control variables release date (c1, c2, c3) and price (c2). The dummy variables for release date are computed as follows, c1 covers all games prior to 2011, c2 covers all games from 2011 to 2016, and c3 covers all games from after 2016. The dummy variable for price is computed to determine if a game is offered for free or not. Finally, publisher and the SteamDB results are merged into one final, clean dataset.

## 3.3 Data description

After the data is processed, the final dataset consists of 1,203 video game titles. When observing the descriptive statistics in Table 3, the data suggests that the mean of both expert- and user reviews are high. Past research has found similar patterns in review ratings for video games on Xbox and PS2 (Zhu & Zhang, 2010) and book reviews on Amazon.com (Chevalier & Mayzlin, 2006). In terms of the moderators, there is a significantly higher amount of single-player games present in the dataset than multi-player games. Furthermore, there is a higher amount of minor publishers in the dataset than major publishers.

**Table 3: Descriptive Statistics**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **N** | **0** | **1** | **Min** | **Max** | **Mean** | **SD** |
| **Dependent variable** |  |  |  |  |  |  |  |
| # Downloads on Steam (y) | 1203 |  |  | 140000 | 80500000 | 3030000 | 3990000 |
| **Independent variables** |  |  |  |  |  |  |  |
| Expert review (x1) | 1203 |  |  | 0.23 | 0.96 | 0.762 | 0.0966 |
| User review (x2) | 1203 |  |  | 0.13 | 0.99 | 0.829 | 0.126 |
| **Moderators** |  |  |  |  |  |  |  |
| Game mode (m1) | 1203 | 502 | 701 | 0 | 1 | 0.583 | 0.493 |
| Publisher (m2) | 1203 | 701 | 502 | 0 | 1 | 0.417 | 0.493 |
| **Control variables** |  |  |  |  |  |  |  |
| Release date old (c1) | 1203 | 986 | 217 | 0 | 1 | 0.180 | 0.385 |
| Release date mid (c2) | 1203 | 468 | 735 | 0 | 1 | 0.611 | 0.488 |
| Release date new (c3) | 1203 | 952 | 251 | 0 | 1 | 0.209 | 0.407 |
| Price (c4) | 1203 | 1120 | 83 | 0 | 1 | 0.069 | 0.254 |

When taking a closer look at the descriptive statistics for game mode presented in Table 4, it is notable that the mean user review rating for single-player games is significantly higher than multi-player games. In addition, single-player games have a lower mean when it comes to downloads in comparison to multi-player games. The summary statistics of publisher also present notable insights. The data suggests that minor publishers receive a significantly higher mean user review rating than major publishers. However, major publishers receive significantly higher mean expert review ratings than minor publishers. Intuitively, this can be explained by the origin of the games. Minor publishers may release niche, indie games with a lower production value that are enjoyed by a relatively small audience and thus receive most of their attention from users (gamers). On the contrary, there are many eyeballs aimed at the production of games that stem from major publishers. These publishers have a large amount of resources, and thus their games would have a higher production value resulting in higher quality games.

**Table 4: Descriptive Statistics Dummy Variables**

|  |  |  |
| --- | --- | --- |
| **A: Summary Statistics for dummy variable game mode** | | |
| **Variable** | **Multi-player (0)** | **Single-player (1)** |
| **# Downloads on Steam (y)** |  |  |
| Mean (SD) | 359000 (5560000) | 2620000 (2190000) |
| **Expert review (x1)** |  |  |
| Mean (SD) | 0.762 (0.0981) | 0.762 (0.0957) |
| **User review (x2)** |  |  |
| Mean (SD) | 0.797 (0.141) | 0.852 (0.110) |
| **B: Summary statistics for dummy variable publisher** | | |
| **Variable** | **Minor publisher (0)** | **Major publisher (1)** |
| **# Downloads on Steam (y)** |  |  |
| Mean (SD) | 2580000 (2300000) | 3650000 (5480000) |
| **Expert review (x1)** |  |  |
| Mean (SD) | 0.748 (0.0967) | 0.781 (0.0933) |
| **User review (x2)** |  |  |
| Mean (SD) | 0.843 (0.119) | 0.810 (0.134) |

## 3.4 Model

The statistical analysis that will be performed in this research is an OLS regression. The regression analysis includes a number of variables that could theoretically explain variation in the number of downloads on Steam. This is due to the fact that these variables

are largely consistent or comparable with variables included in the analysis of for instance the sale of console video games in the US (Cox, 2014), and the impact of reviews on movie revenues (Basuroy et al., 2020; Chintagunta et al., 2010). However, the main difference is that the beforementioned studies account for time (longitudinal data) whilst this study does not (cross-sectional data). Because the main aim of this study is to explain the variation of the y-variable as robustly as possible with the variation of the x-variables, only one model will be developed in order to capture as much variation as possible. In this model, the dependent variable will be taken as a logarithm in order to respond to skewness caused by the large size of the dependent variable. In addition, the independent variables are mean-centered in order to improve interpretation. Mean-centering does not change the computational accuracy, sampling accuracy of main effects, moderating effects, nor the R2 (Echambadi & Hess, 2007). The functional form of the model developed for this study is as follows:

The unit of analysis is a video game (i). The number of downloads on Steam for a video game i in this model is depicted on the left-side of the equation. All relevant coefficients and variables are depicted on the right side of the equation. As discussed earlier in this chapter, there are: two independent variables & , two moderators which are represented as four variables in the equation due to their interaction with each independent variable (, , and six control variables when including the two moderators distinctively as additional control variables (.

# 4. Results

In this chapter the results of the beforementioned model will be discussed. Firstly, the model will be tested for problems with heteroskedasticity and multicollinearity, followed by a paragraph concerning the model fit. Secondly, the main effects and the corresponding hypotheses for these effects will be analysed. Finally, the moderating effects will be discussed in addition to their corresponding hypotheses.

## 4.1 Assumptions of the OLS regression and model fit

### 4.1.1 Heteroskedasticity

The statistical analysis used to test the conceptual framework is OLS regression. The analysis assumes that the variance of the error term is constant and unrelated to the independent variables, also known as homoskedasticity. If the error term is heteroskedastic, the results might have incorrect standard errors. In this analysis, it is expected that the error term might be heteroskedastic due to the high mean of expert- and user review ratings. This entails that the majority of the ratings will be clustered at one side of the graph. In order to check for heteroskedasticity white’s test is performed. As predicted, heteroskedasticity is present in the model as the p-value of white’s test falls below 0.05. In order to correct this, Huber-White robust standard errors are used. These are standard errors that are heteroskedasticity-robust, and enable us to interpret the results correctly.

### 4.1.2 Multicollinearity

Moreover, the analysis also assumes that there is no correlation between independent variables in the OLS regression model. If there is correlation, the independent variables are not independent which can cause problems for the model fit and interpretation of results. In order to identify if there is multicollinearity present in the model, the variance inflation factor (VIF) is calculated for each variable (See Table 5). All but one of the variables have a VIF close to 1, which indicates that there is hardly any case of multicollinearity present in this model. The reason there is no VIF in Table 5 for release date new (c3) is because this variable instigated an ‘aliased coefficients’ error in R. This means that this variable is linearly dependent on others. In order to fix this issue, release date new (c3) will be removed due to its collinearity issues. This will ensure that the model fit and interpretation are not affected by multicollinearity.

|  |  |
| --- | --- |
| **Variable** | **VIF** |
| Expert review (x1) | 1.55 |
| User review (x2) | 1.59 |
| Game mode (m1) | 1.13 |
| Publisher (m2) | 1.20 |
| Release date old (c1) | 1.63 |
| Release date mid (c2) | 1.56 |
| Release date new (c3) | - |
| Price (c4) | 1.07 |

**Table 5: VIF values per variable**

### 4.1.3 Model fit

Two statistics are used to evaluate model fit, namely R2 and the F-test (See Table 6). There are differing opinions regarding what is an acceptable R2. A study from 1992 recommends that an R2 value should be equal to or greater than 0.10 in order for the variance explained of a particular construct to be deemed adequate (Falk & Miller, 1992). A study from 1998 recommends a higher R2, arguing that an R2 of 0.19 is the lowest acceptable level (Chin, 1998). Finally, Cohen (1988) suggests that R2 values that are above 0.13 are adequate. The OLS regression in this study has an R2 of 0.197, which is acceptable according to all three scholarly sources. In terms of the F-test, if the F-statistic is at least 3.95 or higher the model stands a mere 1% chance of being wrong (Archdeacon, 1994). The OLS regression in this study has a F-statistic of 26.61 which is well over 3.95 and thus acceptable. Therefore, the model in this study can be deemed a well-fitted regression model capable of predicting values close to the observed data values.

## 4.2 Main effects

Before delving into the results of the main effects, it is important to mention how the coefficients in an OLS regression table with a logarithmic dependent variable are interpreted. The unconditional expected mean of the intercept (number of owners on Steam) is 14.566. The geometric mean is the exponent of the unconditional expected mean, which is exp(14.566) = 2,118,036. This entails that the intercept for amount of owners is 2,118,036 per video game. In order to get the percentage change in owners per independent variable, the exponent of each β is taken. This depicts the percentage change in the dependent variable for a one unit increase of the independent variable in question. Furthermore, it is necessary to mention the importance of the control variables. In addition to the interpretation, it is necessary to briefly mention the relevance of the added control variables. The added control variables are all significant, and thus enhance the internal validity of the model by limiting the influence of unaccounted extraneous variables. This helps with to establish robust and interpretable relationships between the included variables.

### 4.2.1 The main effect of expert reviews

In support of , the results show that expert ratings positively affect the number of downloads on Steam (β = 0.154, p = 0.000). This means that when expert rating increases by one unit (0.01 or 1%), the number of downloads on Steam increases by 16.65%, ceteris paribus. This aligns with literature from the movie- (Basuroy et al., 2003; Chen et al., 2011) and video game industry (Cox, 2014; Zhu & Zhang, 2010), which also found that box office revenues and the sale of video games on console are significantly impacted by expert ratings. There are no clear-cut contradictions of these findings in literature from reputable journals.

|  |  |  |  |
| --- | --- | --- | --- |
| **OLS** | **β** |  | **SE** |
| (Intercept) | 14.566\*\*\* | (0.000) | 0.024 |
| Expert rating | 0.154\*\*\* | (0.000) | 0.026 |
| User rating | 0.124\*\*\* | (0.000) | 0.030 |
| User rating x Game mode | -0.066\* | (0.023) | 0.029 |
| Expert rating x Game mode | 0.034 | (0.169) | 0.025 |
| User rating x Publisher | 0.025 | (0.381) | 0.028 |
| Expert rating x Publisher | 0.070\*\* | (0.008) | 0.027 |
| Publisher | 0.080\*\* | (0.001) | 0.024 |
| Game mode | -0.081\*\*\* | (0.001) | 0.023 |
| Release old | 0.172\*\*\* | (0.000) | 0.027 |
| Release mid | 0.233\*\*\* | (0.000) | 0.029 |
| Price | 0.095\*\*\* | (0.000) | 0.025 |
|  |  |  |  |
| Num. Obs | 1203 |  |  |
| F-statistic | 26.61 |  |  |
| R2 | 0.197 |  |  |
| R2 Adj. | 0.190 |  |  |
| Std. Errors. | Heteroskedasticity-robust | | |
| + p < 0.1, \* p < 0.05, \*\* < 0.01, \*\*\* p < 0.001 | | | |

**Table 6: OLS regression output**

### 4.2.2 The main effect of user reviews

In line with , the results show that user rating positively affects the number of downloads on Steam (β = 0.124, p = 0.000). This means that when user rating increases by one unit (0.01 or 1%), the number of downloads on Steam increases by 13.20%, ceteris paribus. This effect confirms findings in existing literature, which ascertained that the sale of books (Chevalier & Mayzlin, 2006), movies (Chintagunta et al., 2010) and video games on console (Zhu & Zhang, 2006) found a significant positive relationship with user reviews. Contradicting is Liu (2006), who found that the volume is of greater importance than the rating.

### 4.2.3 The difference between expert- and user reviews

The results of the regression analysis are also in accordance with , showing that expert ratings (16.60%) positively influence the number of downloads on steam more than user reviews (13.20%). These results fall in line with literature from the movie (Basuroy et al., 2020; Souza et al., 2019) and video game industry (Santos et al., 2019), which found that expert ratings were valued more highly and seen as more trustworthy in comparison with user reviews. Contradictory literature regarding this hypothesized relationship is present for the hotel (De Langhe et al., 2016; Öǧüt & Onur Taş, 2012) and movie industry (Tsao, 2014), which found that consumers relate more to peers than professionals.

## 4.3 Moderating effects

### 4.3.1 The moderating effect of game mode

is not supported, as the influence of game mode on the effect of expert rating on the number of downloads on Steam is not significant in the results (β = 0.034, p = 0.169). **WHY?**

In contrast to the prior hypothesis, is supported. Game mode has a significant influence on the effect of user rating on the number of downloads on Steam (β = -0.066, p = 0.023). These results show that the game single-player (dummy = 1) the slope decreases, and vice versa. To be precise, when a game is single-player and the user rating increases by one unit, the number of downloads on Steam decreases by 6.39%, ceteris paribus. There is no concrete literature in line with these findings. The main intuition was built upon that a large community of gamers play multi-player games together, and thus value each others opinions highly.

### 4.3.2 The moderating effect of publisher

The results show contradictory findings for . The positive effect of expert reviews on the number of downloads on Steam is stronger when the game is released by a major publisher (β = 0.070, p = 0.008). This shows that if the game is released by a major publisher (dummy = 1) the slope increases, and vice versa. When a major publisher releases a game, and the expert rating increases by one unit, the number of downloads on Steam increases by 7.25%, ceteris paribus. Literature from the movie industry contradicts these findings, stating that the influence of expert reviews on blockbusters is not significant (Basuroy et al., 2003; Souza et al., 2019).

The results from the regression analysis do not support . This is due to the fact that the moderating effect of publisher on the relationship between user ratings and the number of downloads on Steam is not significant (β = 0.025, p = 0.381). **WHY?**

# 5. Conclusions and recommendations

## 5.1 Conclusion

This study investigates the effect of expert- and user review ratings on the number of downloads on Steam, and to what extent video game characteristics influence this effect. e

|  |  |
| --- | --- |
| **Hypotheses** | **Outcome** |
| : Expert reviews positively influence the number of downloads on Steam. | Supported |
| : User reviews positively influence the number of downloads on Steam | Supported |
| : Expert reviews positively influence number of downloads on Steam more than user reviews | Supported |
| : The positive effect of expert reviews on the number of downloads on Steam is stronger (weaker) when game mode is single-player (multi-player) | Not supported |
| : The positive effect of user reviews on the number of downloads on Steam is stronger (weaker) when game mode is multi-player (single-player) | Supported |
| : The positive effect of expert reviews on the number of downloads on Steam is stronger (weaker) when the game is released by a minor publisher (major publisher) | Not supported |
| : The positive effect of user reviews on the number of downloads on Steam is stronger (weaker) when the game is released by a minor publisher (major publisher) | Not supported |

## 5.2 Discussion

## 5.3 Recommendations

Academic and managerial

## 5.4 Limitations and future research

# References

**BIBLIOGRAPHY WORD**

Archdeacon, T. J. (1994). *Correlation and Regression Analysis.* Wisconsin: University of Wisconsin Press.

Bloomberg. (2020, December 1). *Global gaming market expected to exceed $256 billion by 2025*. Retrieved from Bloomberg: https://www.bloomberg.com/press-releases/2020-12-01/global-gaming-market-expected-to-exceed-256-billion-by-2025

brightlocal. (2020, December 9). *Local Consumer Review Survey 2020*. Retrieved from brightlocal: https://www.brightlocal.com/research/local-consumer-review-survey/

Clairfield International. (2018). *Gaming Industry - Facts, Figures and Trends.* Geneva: Clairfield International.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2 ed.). United States of America: Lawrence Erlbaum Associates.

Johnson, B. (2021, June 24). *PC vs Console Gaming: A War On Multiple Fronts: 2020 Edition*. Retrieved from WePC: https://www.wepc.com/tips/pc-vs-console-gaming/

McKinsey. (2021, August 12). *Five-star growth: Using online ratings to design better products*. Retrieved from McKinsey: https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/five-star-growth-using-online-ratings-to-design-better-products

Newzoo. (2017, January 25). *What Is Core Gaming Anyway? How the Hardware Habits of PC Gamers Are Breaking Down Stereotypes*. Retrieved from Newzoo: https://newzoo.com/insights/articles/the-hardware-habits-of-pc-gamers-and-the-defintion-of-core/

Newzoo. (2020, May 8). *The World’s 2.7 Billion Gamers Will Spend $159.3 Billion on Games in 2020; The Market Will Surpass $200 Billion by 2023*. Retrieved from Newzoo: https://newzoo.com/insights/articles/newzoo-games-market-numbers-revenues-and-audience-2020-2023/

Newzoo. (2021). *Global Games Market Report: The VR & Metaverse Edition.* Amsterdam: Newzoo.

Newzoo. (2021, May 8). *Global Games Market to Generate $175.8 Billion in 2021; Despite a Slight Decline, the Market Is on Track to Surpass $200 Billion in 2023*. Retrieved from Newzoo: https://newzoo.com/insights/articles/global-games-market-to-generate-175-8-billion-in-2021-despite-a-slight-decline-the-market-is-on-track-to-surpass-200-billion-in-2023/

The Economist. (2020, March 19). *The rise and rise of video games*. Retrieved from The Economist: https://www.economist.com/prospero/2020/03/19/the-rise-and-rise-of-video-games

Zhu, F., & Zhang, X. (2006). The Influence of Online Consumer Reviews on the Demand for Experience Goods: The Case of Video Games. *Proceedings of the 27th International Conference on Information Systems*, 367-381.

**BIBLIOGRAPHY MENDELEY**

Basuroy, S., Abraham Ravid, S., Gretz, R. T., & Allen, B. J. (2020). Is everybody an expert? An investigation into the impact of professional versus user reviews on movie revenues. *Journal of Cultural Economics*, *44*(1), 57–96. https://doi.org/10.1007/s10824-019-09350-7

Basuroy, S., Chatterjee, S., & Abraham Ravid, S. (2003). How Critical are Critical Reviews? The Box Office Effects of Film Critics, Star Power, and Budgets. *Journal of Marketing*, *67*(4), 103–117. https://doi.org/10.1509/jmkg.67.4.103.18692

Chakravarty, A., & Mazumdar, T. (2016). The Differential Effects of Online Word-of-Mouth and Critics ’ Reviews on Pre-release Movie Evaluation. *Journal of Consumer Behaviour*, *42*(3), 185–197. https://doi.org/10.1016/j.intmar.2010.04.001

Chen, Y., Liu, Y., & Zhang, J. (2011). When Do Third-Party Product Reviews Affect Firm Value and What Can Firms Do ? The Case of Media. *Journal of Marketing*, *75*(September), 116–134. https://doi.org/10.1509/jm.09.0034

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. https://doi.org/10.1509/jmkr.43.3.345

Chin, W. W. (1998). The partial least squares approach for structural equation modeling. *Modern Methods for Business Research*, *January 1998*, 295–336.

Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science*, *29*(5), 944–957. https://doi.org/10.1287/mksc.1100.0572

Cole, H., & Griffiths, M. D. (2007). *Social Interactions in Massively Multiplayer Online Role-Playing Gamers*. *10*(4), 575–583. https://doi.org/10.1089/cpb.2007.9988

Cox, J. (2014). What makes a blockbuster video game? An empirical analysis of US sales data. *Managerial and Decision Economics*, *35*(3), 189–198. https://doi.org/10.1002/mde.2608

Cui, G., Lui, H. K., & Guo, X. (2012). The effect of online consumer reviews on new product sales. *International Journal of Electronic Commerce*, *17*(1), 39–58. https://doi.org/10.2753/JEC1086-4415170102

De Langhe, B., Fernbach, P. M., & Lichtenstein, D. R. (2016). Navigating by the Stars : Investigating the Actual and Perceived Validity of Online User Ratings. *Journal of Consumer Research*, *42*, 817–833. https://doi.org/10.1093/jcr/ucv047

Decker, R., & Trusov, M. (2010). Estimating aggregate consumer preferences from online product reviews. *International Journal of Research in Marketing*, *27*(4), 293–307. https://doi.org/10.1016/j.ijresmar.2010.09.001

Echambadi, R., & Hess, J. (2007). Mean-Centering Does Not Alleviate Collinearity Problems in Moderated Multiple Regression Models. *Marketing Science*, *26*(3), 438–445. https://doi.org/10.1287/mksc.1060.0263

Falk, R. F., & Miller, N. B. (1992). A Primer for Soft Modeling. *The University of Akron Press*, *April*, 80. http://books.google.com/books/about/A\_Primer\_for\_Soft\_Modeling.html?id=3CFrQgAACAAJ

Ho-Dac, N. N., Carson, S. J., & Moore, W. L. (2013). The effects of positive and negative online customer reviews: Do brand strength and category maturity matter? *Journal of Marketing*, *77*(6), 37–53. https://doi.org/10.1509/jm.11.0011

Jenq, S.-N. (2019). Exploring the Impact of a Hotel’s Rating and Number of Reviews through Online Transactions. *OALib*, *06*(05), 1–15. https://doi.org/10.4236/oalib.1105401

Johnson, D., Gardner, J., & Sweetser, P. (2016). Computers in Human Behavior Motivations for videogame play : Predictors of time spent playing. *Computers in Human Behavior*, *63*, 805–812. https://doi.org/10.1016/j.chb.2016.06.028

Öǧüt, H., & Onur Taş, B. K. (2012). The influence of internet customer reviews on the online sales and prices in hotel industry. *Service Industries Journal*, *32*(2), 197–214. https://doi.org/10.1080/02642069.2010.529436

Santos, T., Lemmerich, F., Strohmaier, M., & Helic, D. (2019). What’s in a review: Discrepancies between expert and amateur reviews of video games on Metacritic. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW). https://doi.org/10.1145/3359242

Souza, T. L. D., Nishijima, M., & Fava, A. C. P. (2019). Do consumer and expert reviews affect the length of time a film is kept on screens in the USA? *Journal of Cultural Economics*, *43*(1), 145–171. https://doi.org/10.1007/s10824-018-9332-6

Tsao, W. (2014). Which type of online review is more persuasive ? The influence of consumer reviews and critic ratings on moviegoers. *Electronic Commerce Research*, *57*, 559–583. https://doi.org/10.1007/s10660-014-9160-5

Ye, Q., Law, R., & Gu, B. (2009). The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management*, *28*(1), 180–182. https://doi.org/10.1016/j.ijhm.2008.06.011

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on Sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, *74*(2), 133–148. https://doi.org/10.1509/jmkg.74.2.133

# Appendices

## A. GitHub repository

Overview of the of the directory structure and files:

├── README.md  
├── makefile  
├──.gitignore  
├── data   
│      ├── d\_publishers.csv

│      ├── df\_merged2.csv  
│      ├── game\_id.csv

│      ├── game\_url.csv  
│      ├── metascore.csv

│      ├── publishers.csv  
│      ├── results.csv  
│      └── steam.csv  
├── gen  
│      ├── analysis  
│      │      └─ output  
│      │          └── regression\_table.png

│      └─── data-preparation  
│            └─ output  
│                └── df\_clean.csv  
└── src  
       ├── collect  
       │      ├── SteamDB.ipynb  
      │      ├── kaggle.R  
      │      └── steamspy.ipynb  
      ├── preparation  
       │      ├── summary\_stats.R

     │      ├── clean\_1.R

    │      ├── clean\_2.R

      │      └── clean\_3.R  
       ├── analysis  
       │      └── regression.R  
       └── .DS\_Store

## B. Regression table R

1. Websites: zoominfo.com, owler.com, growjo.com, games-stats.com, dnb.com, rocketreach.co, vginsights.com. [↑](#footnote-ref-1)