BI Norwegian Business School - campus Oslo

# GRA 19703

Master Thesis

Thesis Master of Science

Electronic word of mouth and new releases: An empirical study in video games industry


Navn: Yunman Jiang


Start: 15.01.2020 09.00

Finish: 01.09.2020 12.00

# Electronic word of mouth and new releases: An empirical study in video games industry

Programme:
Master of Science in Business Analytics

Supervisor:
Rutger Daniel van Oest

# Table of Contents

# Abstract

In today's world, product iterations continue to accelerate, new product introduction becomes more and more frequent. Marketers are also utilizing digital marketing tools to realize a more efficient conversion from marketing efforts to sales performance. Content marketing and the corresponding electronic word-of-mouth (eWOM) are examples that digital marketing tools used more and more frequently today.

This research aims to uncover the effect of prerelease eWOM on new products' short-term and long-term sales performance and to discover the moderator roles of free sample and company awareness.

I was indented to build a dynamic regression model that can use prerelease eWOM data to predict future sales performance. However, there are too many blocks to collect and process open data sources to a usable panel data.

I collect the analysis data through Steam in the video games industry. The present research uses linear regression analysis and moderation analysis to solve the research question on the prerelease eWOM effect and the moderator roles.

The analysis result shows that audience reaction volume significantly affects long-term and short-term sales performance. Free sample and company awareness can negatively influence the prerelease eWOM effect. The analysis model also suggests that the impact of using prerelease publicity (from 0 to 1) may be very different from increasing the number of prerelease publicity (an increase from 1), and the effect of using prerelease publicity may be negative. Therefore, driving the audience reaction volume is the most crucial factor for a successful prerelease digital campaign.

The research results on the moderator roles of free sample and company awareness give different companies using different product pricing models trailed suggestions on making the prerelease marketing strategy. In general, products not offering free samples or from unwell-known companies should leverage the prerelease eWOM effect from focusing on increasing audience reaction volume.

# 1. Introduction

Since 2000, digital media platforms have revolutionized marketing, offering new ways to reach, inform, engage, sell to, learn about, and provide service to customers. From a macro-level thematic discussion of digital, social media, and mobile marketing (DSMM) between 2000 and 2015, three themes are identified emerging, which are DSMM as a facilitator of individual expression, DSMM as a decision support tool, and DSMM as a market intelligence source (Lamberton and Stephen 2016).

Researchers explore the way to observe, analyze, and predict consumer behavior through the internet. Companies also try to gain knowledge about their customers through the internet. With the emerging of social media platforms and the booming of user-generated content (UGC) around 2010, marketers realized the value of social media as a marketing tool. They started to manipulate electronic word-of-mouth (eWOM) for marketing purposes. Both the consumers and firms are benefited from the viral transmission and content creation (Toubia and Stephen, 2013).

With a deeper understanding of eWOM, companies and marketers are moving forward to embrace the digital era. According to The State of Content Marketing 2019 survey by SEMrush, 91% of organizations worldwide use content marketing. Content marketing is also commonly used in introducing new products to the market.

It is important for marketers to know the effectiveness of marketing efforts and to understand the factors that may impact efficacy. With a good understanding of factors that influence the effectiveness, marketers can better utilize the marketing tools to achieve the business goals.

Many articles attempted to demonstrate the value of UGC or social media based WOM. Trusov (2010) linked review rating dynamics to subsequent ratings and product sales; Tirunillai and Tellis (2012) related UGC with firm's stock performance.

However, researches on the effect of prerelease eWOM on new product's sales performance are still limited.

Prerelease buzz data can use to predict new product performance (Xiong and Bharadwaj, 2014). The prerelease buzz in Xiong's research is represented by online search volume, the online blog, and online forum postings, and the opening sales represent the product performance.

Another research compares the impact from pre-launch and post-launch publicity and advertising on both new product opening sales and half-year sales, demonstrating that pre-launch publicity is more effective than advertising (Burmester et al., 2015). Only publicity and advertising volume from the company included in the analysis, reviews, and previews from customers are not included in Burmester's research.

In terms of the reviews and previews from customers, there is a positive impact of social media "like" volume in the prerelease on box office (Ding et al., 2016). But pre-releasing marketing activity is not included in Ding's analysis.

*Table 1 Comparison of previous research and current research on eWOM and product performance*

| **Previous research** | **Pre-releasing period** | **Performance** |
|---|---|---|
| Xiong and Bharadwaj, 2014 | online search volume, online blog and online forum postings | Opening sales |
| Burmester, 2015 | Magazine press release and advertisings | Sales in a half year |
| Ding et al., 2016 | social media "like" volume | box office |
| Nguyen & Chaudhuri, 2019 | eWOM volume, sentiment | Sales |
| **The present study** | Company release and audience reactions volume | Opening sales and long-term sales |

Even fewer researches studied the moderator of prerelease eWOM effect on new product performance. Nguyen & Chaudhuri did a study explore the moderator in 2019; they examined the moderator effect of eWOM channels, announcement time, communication richness, and branding.

Companies must have a clear idea of whether they should do any prerelease marketing before the new product launching? How many marketing efforts should they put before a new product launching? How will marketing efforts become sales performance? And how to maximize the sales performance brought by the prerelease marketing efforts?

Therefore, the effect of prerelease eWOM on sales performance after releasing; and moderators of this effect are important research questions.

This research wants to address the following two questions:

1. Does pre-releasing eWOM has a positive impact on both short term and long term sales performance after release?

2. Which factors influence the pre-releasing eWOM effects?

The present research will consider both the official prerelease announcement made by companies and the coproduction content form consumers to explore the prerelease eWOM's effect on new product performance in both the short-term and long-term. The present research will also examine the moderator for this effect.

To sum up, this study contributes to the literature in three ways. First, extend the previous research in the digital marketing context. Second, answer the question that should or should not make a prerelease announcement through empirical data analysis. Third, find the moderator roles that influence the effect of prerelease eWOM on product performance.

The empirical study results show that audience reaction volume to the company's prerelease announcement (publicities) has a positive effect on product sales performance both short-term on the opening sales and long-term for the product lifetime sales. And with the audience reaction volume growth, the positive impact would become more prominent. However, prerelease publicities that cannot raise audience attention to have a good audience reaction volume may harm product sales performance, both short-term and long-term.

Offering free sample of the new product and company awareness has moderation influence on the prerelease eWOM effect. Not having free sample would strengthen the eWOM impact on sales performance in the short-term; not from a well-known company would also enhance the short-term eWOM effect. While in the long-term, the two moderators influence the effects from prerelease publicities and audience reactions differently.

In practice, the recent study's empirical analysis results can be used to help companies build their new product launching marketing strategy. Based on the present study's conclusion, companies can make the new product launching marketing plan combined with their pricing strategy. Because the results disclose the mechanism of how prerelease publicities and eWOM works for generate sales performance and how free sample or pricing strategy and company awareness can influence the mechanism.

## 2. Theory and hypothesis development

According to the network coproduction model, electronic word of mouth (eWOM) is formed by the marketing-mix elements released by the marketer and coproducing the marketing message by consumers (Kozinets, 2010). It is important to include both the marketer side and the consumer side when considering the impact of prereleasing activities on new product performance.

In the present research's conceptual framework construction, I use prerelease publicity and audience reactions together to represent prerelease eWOM.

Prerelease publicity refers to the press release, blog article, introduction video, or even social media posts made by the company that aims to release a limited message to the consumers about the upcoming new release. Prerelease publicity represents the marketing-mix released by marketers.

Audience reaction refers to the review, discussion, likes, recommendation, and recreation of the consumers. It can be many different formats, such as likes, comments, reposts, blog articles, forum topics, and videos. Audience reaction shows the extend of consumers' interest in the publicity made by marketers, and audience reaction amplifies and coproduct the marketing message.

Prerelease publicity and audience reaction together can capture the trend of prerelease WOM to some extent. In the present research, prerelease WOM refers to the combination of prerelease publicity and audience reaction.

Short-term performance refers to the sales performance of the new release at the end of the launcing week. Corresponding long-term performance refers to the performance of the product throughout its lifecycle.

I also include Moderators that will influence the pre-relase WOM effect on product sales performance. Consumers are looking for eWOM to eliminate uncertainty and risk of purchase behavior (Reimer & Benkenstein, 2016). Therefore, factors that bring uncertainty and risk to customer purchases would have moderator impact on the pre-releasee eWOM effect.

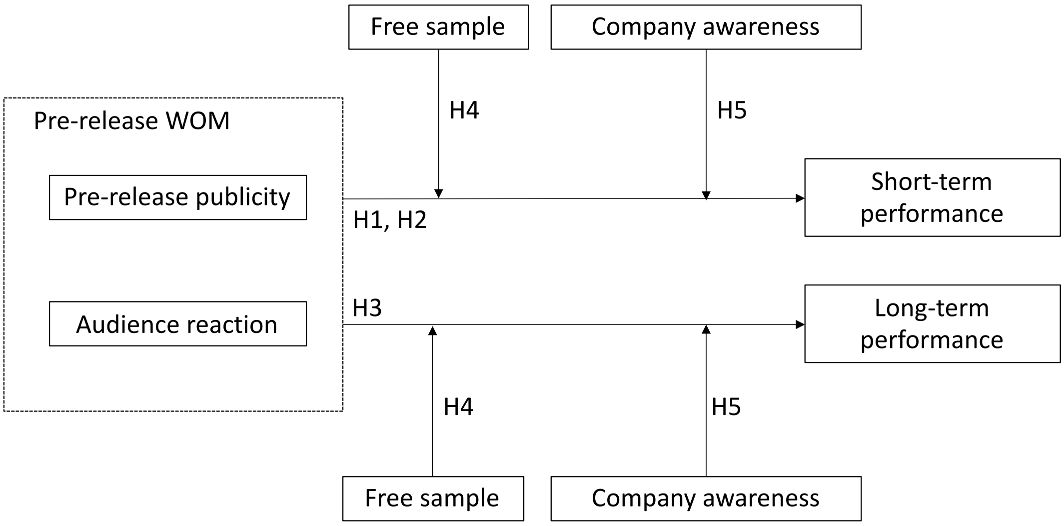



*Figure 1 Conceptual framework*

The increase of brand name exposure can promote a favorable consumer attitude towards the brand (Janiszewski 1993). Prerelease publicity can increase brand exposure and awareness, which would promote consumer's favorable and perception of the product.

Moreover, the product awareness triggered in the pre-releasing period would result in a prerelease "shadow diffusion". Customers may decide to buy the new product at any time before it is available, the demand would accumulate until the product officially releases (Goldenberg et al., 2007; Peres et al., 2009).

Prerelease publicities would increase the product's awareness and consumer interest in the new release. Still, the accumulated demand for the new product can only convert to sales performance after release. Therefore, the increased number of prerelease publicity would lead to a better new release performance in the first week after releasing.

**H1.** The number of prerelease publicities is positively related to release week performance.

Previous studies already show that eWOM positively affects product sales (Trusov, 2010).

For a product that only has limited information available (e.g., the product that hasn't available on the market), the amount of WOM buzz (e.g., mentions and discussions) can be perceived as an indicator of high quality and popularity (Godes and Mayzlin 2004).

The audience reaction, such as like and discuss, contributes to the amount of WOM buzz of the product, that would not only be perceived as high quality and popularity, but also reinforce itself through social impact (Banerjee, 1992). Therefore, audience reactions may have a positive effect on new releases performance.

**H2.** The number of audience reactions on the pre-releasing period has a positive effect on the release week performance.

New releases' long-term performance refers to the performance of the product throughout its lifecycle.

Researches show that new products in industries with short life cycles, such as movies, video games, and new technology devices, usually experience a sales peak when first released and then followed by sales decreasing (Burmester et al. 2015). And For products with this innovation diffusion pattern, it is important to use prerelease marketing activities to drive overall sales (Elberse & Eliashberg, 2003).

For products with the innovation diffusion pattern, prerelease eWOM has the same effect on product long-term performance as on product short-term performance.

Ordinary products have a bell-shape diffusion pattern that "expert" opinions trigger customers to get interested and follow the products (Bass, 1969). Prerelease eWOM can not only be the "expert" opinion but also trigger more after release "expert" opinion through social impact (Banerjee, 1992).

For ordinary products, prerelease eWOM also has an impact on product long-term performance.

**H3.** Prerelease eWOM has a positive effect on product long-term performance.

Many products are offering free samples to potential consumers. For digital products, free sample is involved in the product pricing model as "freemium". A lot of forms of free sample are offering to potential customers today. Free samples can be offered to allow consumers resolving uncertainty before purchase (Wu et al., 2018).

If there is no free sample available, how can consumers resolve uncertainty? Online reviews can reduce uncertainty about service or product quality (Reimer & Benkenstein, 2016).

Therefore, offering free sample or not can be a moderator of prerelease eWOM effect on sales performance. Because not offering free samples leads to consumers relying more on the prerelease eWOM to resolve uncertainty and risk on the newly released products.

**H4.** If the product not offering free sample, the effect of prerelease eWOM on release week performance would be strengthened.

Similar to free sample, company awareness is also helping to reduce consumers' concern about the quality of the new products.

Research on the hotel industry find that eWOM volume (reviews) does not affect the revenue growth of branded chain hotels and a positive effect on revenue growth for not-branded chain hotels (Raguseo et al., 2017).

Consumers are more likely to rely more on prerelease eWOM to eliminate uncertainty and risks for making purchase decisions from unknown companies.

**H5.** If the product is not from a well-known company, the effect of prerelease eWOM would be strengthened.

# 3. Modeling Approach

*3.1. Model challenges*

The model specification considers three major challenges.

The first challenge is the possible different situations from two kinds of product release scenarios. One is no publicities before the official release; the second is employing publicities before the official release.

In order to build a linear regression model estimating the effect of prerelease eWOM on new releases performance, and also consider the possibility that no publicities before the official release behaves differently from employing publicities in the prerelease stage. I introduce a dummy variable that distinguishes these two scenarios (Event = 1 if having prerelease publicities, 0 otherwise). I add this dummy indicator into the regression model to give control of the two different scenarios.

Secondly, we have the company's prerelease publicities and the audience reactions contribute together to the prerelease e-WOM but need to research the moderator effect on prerelease e-WOM. In order to investigate the moderator effect of the free sample (Free) and company size (Major_mid_publisher), I used four interactions to capture the moderator effect. The four interactions are the interaction between the free sample and the number of publicities, the interaction between the free sample and the number of audience reaction, the interaction between company size and the number of publicities, and the interactions between company size and the number of audience reaction.

The third challenge is how to model the long-term effect and short-term effect differently. According to market expansion effects (Beck, 2007), the good performance on the release period brought by the prerelease eWOM affects the long-term effect. For example, the product gains short-term success when it is officially released; the post-release WOM produced by those newly acquired customers would influence others. However, in the present research, I want to focus on the prerelease effect on long-term product performance. Therefore, I made two parallel models to capture the impact of prerelease eWOM on short-term performance and long-term performance, respectively.

*3.2. Model specification*

For the sales performance of product $i$ in the release week:

$$Y\_release_i = \beta_0 + \beta_1 D_i + \beta_2 X_{1i} + \beta_3 X_{2i} + \beta_4 Z_{1i} + \beta_5 Z_{2i}$$
$$+\beta_6 Z_{1i} \cdot X_{1i} + \beta_7 Z_{1i} \cdot X_{2i}$$
$$+\beta_8 Z_{2i} \cdot X_{1i} + \beta_9 Z_{2i} \cdot X_{2i} + \varepsilon$$

Where

$Y\_release_i$ = sales performance of product $i$ in the release week;

$D_i$ = 1 if product $i$ has prerelease publicities, 0 otherwise;

$X_{1i}$ = volume of prerelease publicities of product $i$;

$X_{2i}$ = volume of audience reaction of product $i$;

$Z_{1i}$ = 1 if product $i$ offering free sample, 0 otherwise;

$Z_{2i}$ = 1 if product $i$ is introduced by a well-known company, 0 otherwise.

For the sales performance of product $i$ in the long-term:

$$Y\_lifetime_i = \beta_0 + \beta_1 D_i + \beta_2 X_{1i} + \beta_3 X_{2i} + \beta_4 Z_{1i} + \beta_5 Z_{2i}$$
$$+\beta_6 Z_{1i} \cdot X_{1i} + \beta_7 Z_{1i} \cdot X_{2i}$$
$$+\beta_8 Z_{2i} \cdot X_{1i} + \beta_9 Z_{2i} \cdot X_{2i} + \varepsilon$$

Where

$Y\_lifetime_i$ = the aggregated sales performance of product $i$ at a point that near to the end of its growth;

$D_i$ = 1 if product $i$ has prerelease publicities, 0 otherwise;

$X_{1i}$ = volume of prerelease publicities of product $i$;

$X_{2i}$ = volume of audience reaction of product $i$;

$Z_{1i}$ = 1 if product $i$ offering free sample, 0 otherwise;

$Z_{2i}$ = 1 if product $i$ is introduced by a well-known company, 0 otherwise.

# 4. Empirical Analyses and Results

### *4.1. Data Analysis Method*

This research use regression analysis to test hypothesis 1, the relationship between pre-prelease publicity and new releases performance, and hypothesis 2, the relationship between audience reaction and new releases performance.

To test hypothesis 3, I will make a parallel regression and moderation analysis, only changing the independent variable from short-term product performance to long-term product performance.

Moderation analysis through the moderator interaction terms in the same regression models will be used to test hypotheses 4 and hypothesis 5, to determine if the free sample and company awareness have a significant moderator effect in both the short-term and long-term.

For all the hypotheses, I will use the p-value approach to decide whether to accept or reject the hypotheses I made in the previous section.

In the way of the analysis, I will also evaluate the potential multicollinearity by the correlation of each two variables.

### *4.2. Data and Measures*

The empirical study is conducted in the video game industry for two reasons.

First, the video games industry is digital. With the development of internet technology, more and more video games are sold online through digital access; consumers can buy the game and play online through the game publisher's official website or video games platform, such as Steam.

Compared with products sold mostly offline, digital marketing channels are more crucial for the video game industry to communicate with their consumers. And there are fewer traditional offline marketing activities in the video games industry.

Therefore, doing the empirical study in the video games industry to some extent can eliminate the effect from traditional marketing activities, such as magazine advertising, press release, and television advertising, which are not included in this present research model.

Second, the lifecycle of products in the video game industry is relatively short.

In this present research, I am going to study the effect of prerelease eWOM on both short-term and long-term product performance. The long-term product performance refers to the aggregate sales performance of the product in its lifecycle.

For traditional products, the demand and product adoption follow the classic diffusion theory and shows a bell-shape diffusion pattern (Bass,1969). The demand and product adoption increase gradually to a certain turning point and then gradually decline.

In contrast, entertainment products, such as movies, books, music, and video games, follow a prerelease "shadow diffusion" and shows an exponentially declining pattern (Goldenberg et al., 2007). The demand and product adoption for entertainment products often experience a peak immediately after the launch and then followed by a strong decline.

Therefore, it is easier to capture the lifetime sales performance of an entertainment product. And it is much more feasible to collect the long-term performance data and model the prerelease eWOM effect on long-term sales performance in the video games industry.

*4.2.1. Data collection*

I focus on video games on Steam. Steam is the biggest PC video game platform in the world. Steam doesn't divide the district markets very strict so that the company releases and gamers reviews on Steam are globally available. Therefore, the data I collect for the present study does not focus on a specific geographic district.

I collected all the video games published in 2019 on Steam from Steamspy (https://steamspy.com/year/2019), in total 8,064 titles released in 2019. There are 3,427 games in these games with lower than 1,000 owners who own the game in their library. Those small titles are normally produced and published by independent game studios who are rarely investing in marketing activities; some even do not have an official website. Therefore, in this research, these small titles can be considered as an extreme value. Those titles are excluded from the research.

I tried to use random sampling methods to choose 300 titles from the remaining 4,637 titles released in 2019 to form the research video game dataset. However, even exclude those small titles with lower than 1,000 owners, there are still many titles that are hard to find the relevant dependent variables data, the time-based owners' number. So, the final set of video game titles we used are the top 300 released in 2019, in order to collect the essential data for dependent variables.

After collecting the dependent variable data, I continue to collect the independent variables data (prerelease publicities and audience reactions representing prerelease eWOM). The independent variables data are coded one by one for the 300 observations in my dataset through Steam's video games platform. I check the publicity events for all the 300 observations, count the events number in the half-year before the official release, and collect the audience reactions ("rate up" and "discuss") volume for the ten latest events before the official release for the 300 observations. For those observations with less than 10 prerelease events, I collect the audience reactions volume data for all the prerelease events that observation has.

The moderator variables data is processed from two different sources; one is the dataset of the 8,064 titles released in 2019 from Steamspy, the other is external sources for the major and mid-size game publishers. I use python to process the data and extract the moderator variables data. For the moderator "free sample", I use the game price form Steamspy dataset. If the price is 0, is 1, otherwise, is 0. For the moderator "company awareness", I use the publisher data in the Steamspy dataset, combined with an external list of the major and mid-size publishers. If the publisher is in the list of the major and mid-size publisher list, is 1, otherwise, is 0.

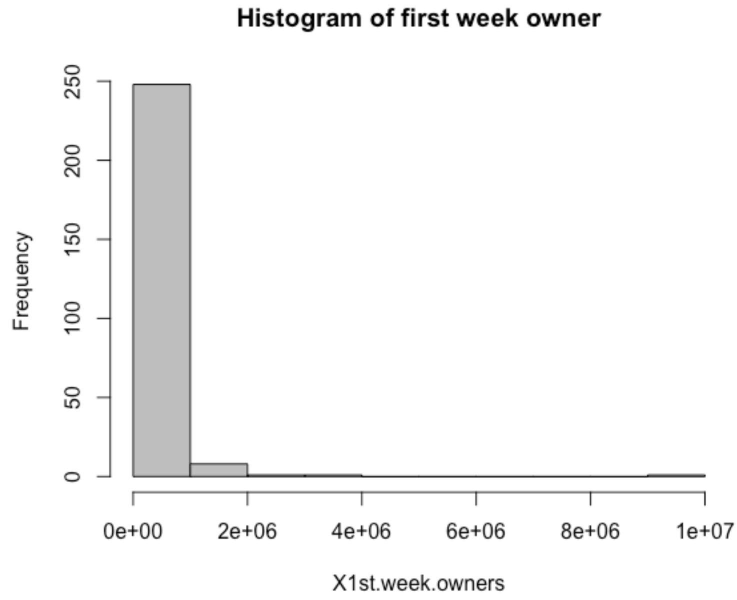
I also collect data of several control variables that may affect the video games sales performance both short-term and long-term by manipulating the Steamspy dataset. There are three dummy control variables, which are "early access", "external developer", and "indie". I also collect the meta score data representing the game quality. However, there is not enough data for this variable, so it is not used in the final analysis model.

### *4.2.2. Dependent variable*

I use the total owner number to capture the product performance. In the video games industry, there are mainly two types of pricing model, one is freemium, that gamers can play the game for free, but there are in-game advertisings or in-game purchases that can bring revenue for the video game company; the other is premium, that gamers must pay a certain amount to buy the game before playing it. The owner number is directly related to the revenue and sales performance of both freemium and premium pricing models. Therefore, I use the total owner number as a proxy for product sales performance, both short-term and long-term.

The short-term effect model's dependent variable is "owners" volume at the end of release week (on the 7th day after the official release). The long-term effect model's dependent variable is "owners" volume at the end of the 26th week after release (on the 182nd day after the official release). According to the diffusion pattern of experience products with a relatively short lifecycle, half-year owners' numbers can be used to proxy the product's long-term performance (Burmester et al., 2015).

Among those 300 titles, 259 observations having owner data at the end of the first week after release (short-term performance) are used to estimate the short-term effect regression model; and 285 observations having owner data at the end of 26th week after release (long-term performance) are used to estimate the long-term effect regression model.



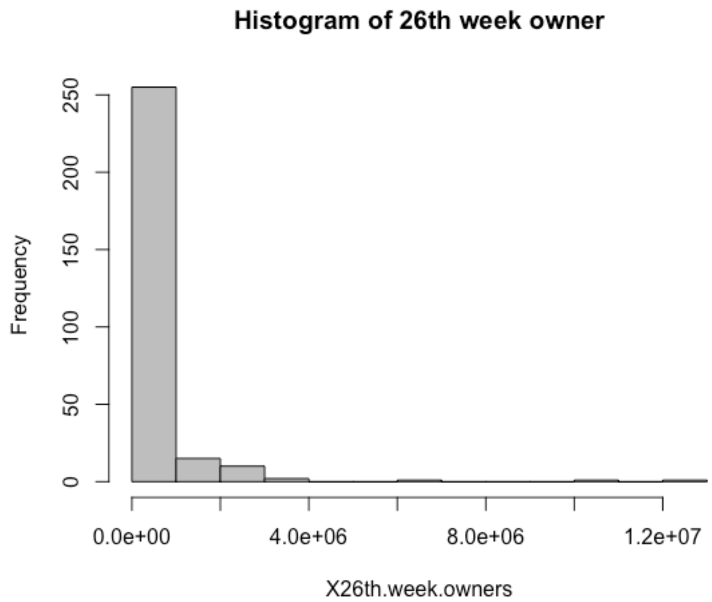




*Figure 2 Histogram of dependent variables*

The owner data for both short-term and long-term performance is very skewed. Because I will use the data estimating linear regression models, the dependent variables must be normally distributed. I consider technics such as logarithmic transformation to achieve a normal distribution. And the owner data here is count data that starts from 0, which means if I use logarithmic transformation, the owner number value 0 will become NA. Therefore, I use log(n+1) transformation to make the data points have a more symmetric distribution and keep data as much as possible. The log(n+1) transformation is used in both the short-term model for the owner data at the end of the first week after release and long-term model for the owner data at the end of the 26th week after release.

*Table 2 Descriptive statistics of dependent variable*

| | Obs. | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | Std. |
|---|---|---|---|---|---|---|---|---|
| **1st week owner** | 259 | 0 | 27,500 | 65,000 | 223,762 | 170,000 | 9,208,000 | 685,048.9 |
| **26th week owner** | 285 | 571 | 104,000 | 166,000 | 472,591 | 347,000 | 12,935,000 | 1,172,743 |

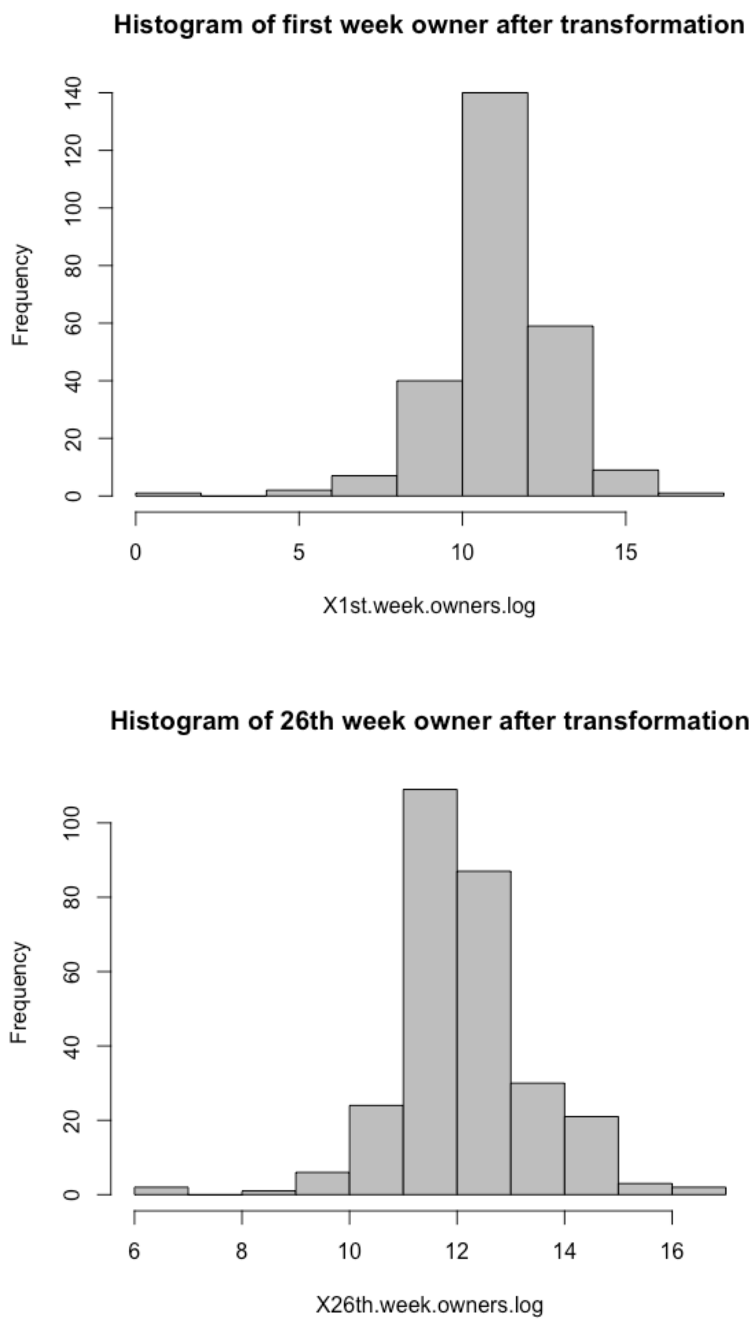



*Figure 3 Logarithm of dependent variables*

*4.2.3. Independent variables*

In the conceptual framework, prerelease eWOM contains two parts, the volume of prerelease publicities and the volume of prerelease audience reactions. Corresponding with the conceptual framework, the independent variables assumed to have the main effect on short-term and long-term product performance contain two variables and a dummy variable to indicate the two scenarios of doing prerelease publicities and not doing pre-releasing publicities.

The first independent variable is the prerelease publicity volume. I collect the number of "events" in the half-year before release. The events here are the official publicity from the company on Steam; it can be an article with pictures and videos; it also allows people to share on social networking such as Facebook and Twitter.

Although it would be better if I also collect data on the different types of events and investigate how the different type influences the effect, I do not have enough computer capacity and coding capacity to collect the relevant data.

In the dataset, only 121 games have prerelease events. Among those games, the number of prerelease events ranges from 1-68; the median value is 8. The data is heavily skewed left. Logarithm may also be considered. But in this case, it is not necessary because I want to interpret the effect from the increasing number of prerelease publicity instead of the elasticity of pre-releasing publicity; therefore, I do not logarithm transform the event number variable.

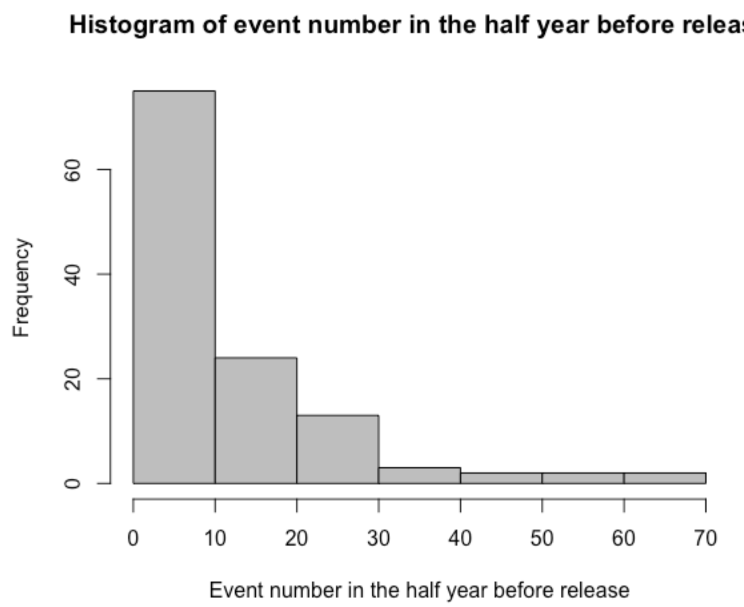



*Figure 4 Histogram of pre-releasing events number and logarithm*

The second variable is audience reactions. This variable represents how consumers interact with the company's pre-releasing publicities. In the present study, I collect the audience's reaction volume from the ten latest events before releasing each title. For each event, I collect each event's date, the "like" volume of each event, and the "discuss" volume of each event.

Because each observation has a different number of events before release, I calculate the minimum, maximum, mean, and standard deviation of "like" and "discuss" volume for each title's events to make it easier to compare and model.

*Table 3 Descriptive statistics of like volume and discussion volume*

| | **Obs.** | **Min.** | **1st Q.** | **Median** | **Mean** | **3rd Q.** | **Max.** | **Std.** |
|---|---|---|---|---|---|---|---|---|
| **Like min** | 121 | 0 | 19 | 56 | 103.7 | 117 | 1613 | 208.63 |
| **Like max** | 121 | 6 | 54 | 134 | 339.4 | 280 | 9149 | 1147.28 |
| **Like mean** | 121 | 2.8 | 32 | 93.6 | 108.9 | 178.4 | 3757.5 | 443.28 |
| **Like std** | 121 | 0 | 4.77 | 19.95 | 82.37 | 46.49 | 3120.71 | 352.65 |
| **Discussion min** | 121 | 0 | 0 | 6 | 19.06 | 20 | 248 | 39.71 |
| **Discussion max** | 121 | 0 | 15 | 38 | 88.14 | 92 | 1144 | 141.82 |
| **Discussion mean** | 121 | 0 | 6.3 | 21.78 | 43.38 | 48.5 | 406.25 | 66.66 |
| **Discussion std** | 121 | 0 | 2.66 | 7.1 | 23.28 | 24.44 | 295.98 | 39.96 |

The correlation between "like" volume and "discuss" volume is very high. The correlation coefficient of like mean volume and discussion mean volume for each game title among all events they have is 0.80. Therefore, I should avoid model both "like" and "discussion" in one model.

Both like and discuss are the audience's reactions to the pre-releasing events. The definition of electronic word of mouth (eWOM) contains all information communicated through the Internet, including reviews, tweets, blog posts, likes, pins, images, and videos (Rosario et al., 2016). I construct a variable sum up the like mean volume and discuss mean volume to represent the eWOM volume in this research. In this way, I can avoid multilinearity and capture both "like" and "discuss" in the meanwhile.

Audience reactions mean volume data is also very skewed. Logarithm transformed audience reactions mean volume can be considered if I can interpret the audience reaction elasticity. However, it is not very essential to transforming the audience reactions mean volume. Because it is an independent variable, and the unsymmetrical data distribution of independent variables would not have a severe bad influence on the linear regression model quality. We will need more judgment to decide on this problem.

*Table 4 Descriptive statistics of audience reaction mean*

| | **Min** | **1st Qu.** | **Median** | **Mean** | **3rd Qu.** | **Max** |
|---|---|---|---|---|---|---|
| **Audience reactions mean** | 3.0 | 48.5 | 109.2 | 224.3 | 237.2 | 4,147.5 |

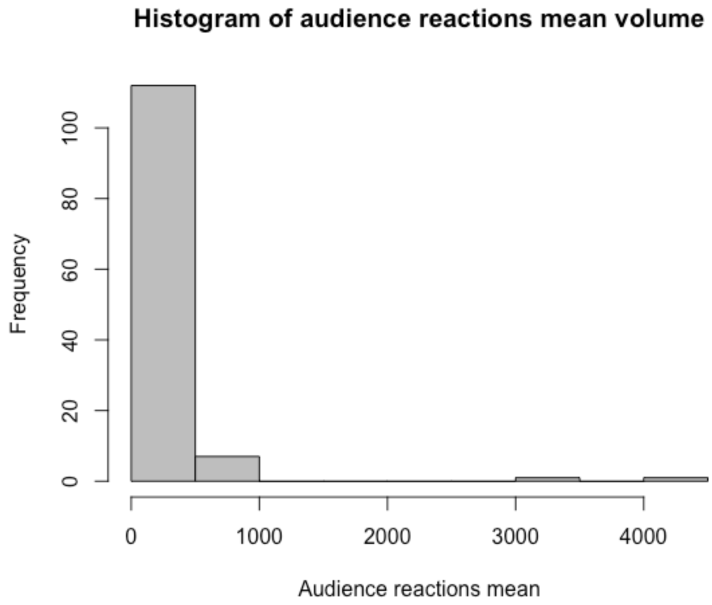



*Figure 5 Histogram of the audience reactions mean volume*

As mentioned before, only 121 games of the 300 have prerelease events in my dataset. For those 179 games did not do prerelease publicities, they do not have audience reactions as well. Therefore, there are two different scenarios, and it must be considered in the model in order to have a better estimation. Aimed for that, I introduced a dummy variable to indicate the two different scenarios.

Considered the possibility of multicollinearity, I calculate the correlation coefficient and coefficient of multiple determination (R-squared) between audience reactions mean volume and event dummy and between log transformed audience reactions mean volume and event dummy.

*Table 5 Coefficient of Multiple Determination for audience reactions and log transformed audience reaction*

| event dummy & | Audience reactions | Log(n+1) audience reactions |
|---|---|---|
| **Cor** | 0.30 | 0.88 |
| **R-squared** | 0.09 | 0.77 |

According to the estimated coefficient of multiple determination, log transformed audience reaction volume has a 0.88 correlation with the variable event dummy. The estimated R-squared using one variable to predict another is also high as 0.77. Therefore, to avoid multilinearity, I choose to use the simple audience reactions mean volume in the modeling.

Except for the audience reactions to each event, I collect the date of each event and trying to include the timing in the model. Since the observations in the dataset are released during the year of 2019 on a different date, I processed the event's date to how many days before release (e.g., "-8" represent to 8 days before the official release date). Therefore, this variable is comparable with the observations.

*Table 6 Descriptive statistics of the date of latest ten events*

| Event | Obs. | Min. | $1^{st}$ Q. | Median | Mean | $3^{rd}$ Q. | Max. | Std. |
|---|---|---|---|---|---|---|---|---|
| **1** | 121 | -150 | -9 | -6 | -14.47 | -2 | -1 | 25.96 |
| **2** | 106 | -176 | -30 | -14 | -25.73 | -7 | -2 | 29.39 |
| **3** | 96 | -172 | -55 | -27 | -42.02 | -13 | -3 | 40.66 |
| **4** | 87 | -171 | -57.5 | -39 | -48.48 | -17 | -5 | 41.07 |
| **5** | 80 | -178 | -82.5 | -49 | -61.66 | -28.75 | -6 | 45.10 |
| **6** | 74 | -181 | -95 | -55 | -68.51 | -32.25 | -7 | 47.79 |
| **7** | 66 | -175 | -104 | -64.5 | -73 | -35 | -8 | 46.05 |
| **8** | 62 | -181 | -111.75 | -75.5 | -81.34 | -39.5 | -12 | 47.39 |
| **9** | 52 | -174 | -118.25 | -81.5 | -82.96 | -46.75 | -16 | 43.23 |
| **10** | 48 | -180 | -133 | -88 | -92.19 | -51.25 | -20 | 48.08 |

For those observations that have more than ten events in the half-year before release, I also calculated the events interval, which is how often it has an event. 75% of observations have an interval day range from 5.1 days to 13.6 days; the mean value is 9.4 days.

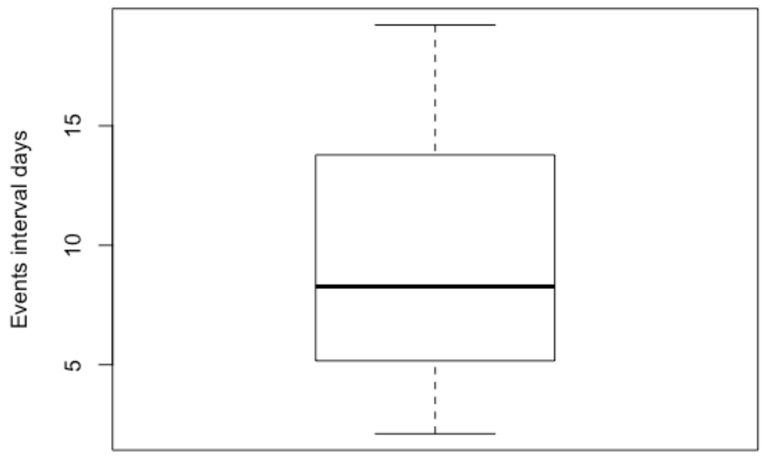



*Figure 6 Boxplot of events interval days*

However, the timing and events intervals are not including in the final analysis model. Because only 48 observations have all of those data, the observations volume is not big enough to make a meaningful analysis model for this present research question.

### *4.2.4. Moderators*

For each observation, characteristics related to the game itself are collected, such as the game's price, the publisher and developer, and the game genre.

Freemium and premium are the two main types of video games pricing model. The freemium model allows consumers to try out the product before paying, which can also be seen as a free sample. But the freemium model aims to increase revenue either in sales and subscription of digital products and services or through advertising revenue.

In this study's dataset, most games are below 5 USD, and nearly 1/3 of the games are free (94 titles in 300 titles). Therefore, a binary variable that indicates a game title is free or not is constructed and will be further investigated in the model.

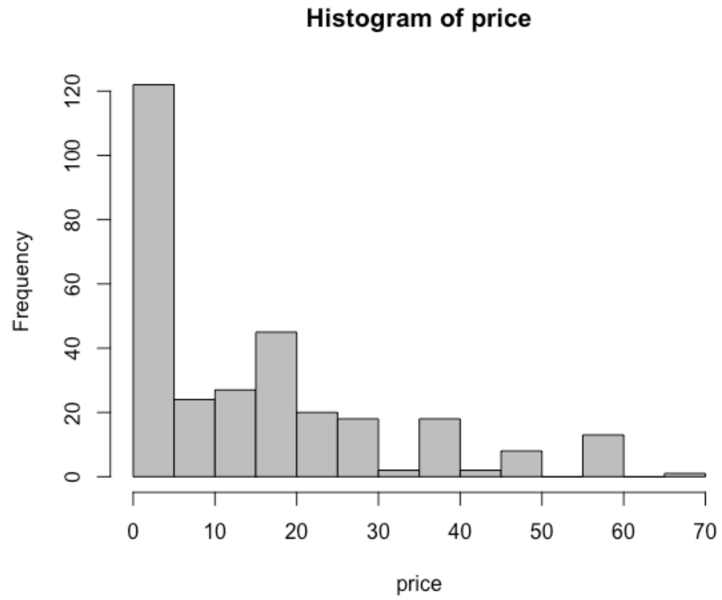



*Figure 7 Histogram of price*

To investigate the company awareness's moderation effect, I look into the publisher and developer of each observation.

I use the size of the company as a proxy for the company awareness. I construct a binary variable to describe the size of the publishing company. If a major or mid-sized publisher publishes the game, company awareness value is 1; otherwise, it is 0.

According to data on Wikipedia and Metacritic, I use table 6 to determine if a publisher is a major or mid-size publisher. There are only 16 titles published by major publishers in the dataset; the mid-size publisher publishes 58 titles.

*Table 7 Major publishers and mid-size publishers*[1]

| | |
|---|---|
| Major publishers | Tencent Games, Sony Interactive Entertainment, Apple, Microsoft, Activision Blizzard, NetEase, Google, EA, Nintendo, Bandai Namco, Ubisoft |
| Mid-size publisher | Telltale Games, Paradox Interactive, Capcom, Take-Two Interactive, Sega, Zen Studios, Devolver Digital, Konami, Slitherine Strategies, NIS America, Warner Bros. Interactive Entertainment, Koei Tecmo, Atlus, 505 Games, Aksys Games, Deep Silver, Focus Home Interactive, Valve, Gameloft, Bethesda |

#### *4.2.5. Control variable*

In the data collection stage, I tried to collect relevant control variables as much as possible. The control variables I collected are "external developers", "indie", "early access", and "meta score". In the final analysis model, the meta score is not used due to not enough data points for this variable.

Despite the four variables, I also looked into game genres. But in the dataset, most games have genre labels such as "action" and "role-play games". So, I didn't look more in-depth into this characteristic, since there is limited variety.

The first control variable is the external developer. I construct a binary variable that describes if a game is published by its developer. Because there exists a business model that some companies are focusing on video game publishing, and some are focusing on developing, these two kinds of companies cooperate together to complete the value chain. Many companies also do the whole value chain that publishes their own developed games. In the dataset, there are nearly half (147) games are developed and published by different companies.

The second control variable is indie. "indie" is also a genre label in the data I collected. Indie games are the games developed by a small studio that doesn't have a big budget and is published through easy-access internet channels. This type of game is more niche and followed by a group of loyal audiences. Therefore, I also made a binary variable to describe if a game title is an indie game. If the observation

[1] Video game publisher. (2020, August 31). Retrieved August 31, 2020, from https://en.wikipedia.org/wiki/Video_game_publisher
Metacritic's 10th Annual Game Publisher Rankings. (n.d.). Retrieved August 31, 2020, from https://www.metacritic.com/feature/game-publisher-rankings-for-2019-releases

is an indie game, the value for an indie is 1, otherwise it is 0. There are 128 titles of the 300 are indie game.

The third control variable is early access. I construct a binary variable of "early access", if the observation allows a limited number of users to buy and play before it is publicly released, the value for early access is 1, otherwise it is 0. Early access would allow the game to have reviewed before it releases from gamers. Moreover, it is a way for small game companies to test their game and their ability to host gamers. There are 130 games that have early access. For those games that have early access, I also collect the total days of early access. The range of early access days is very big, from 10-1,726 days.

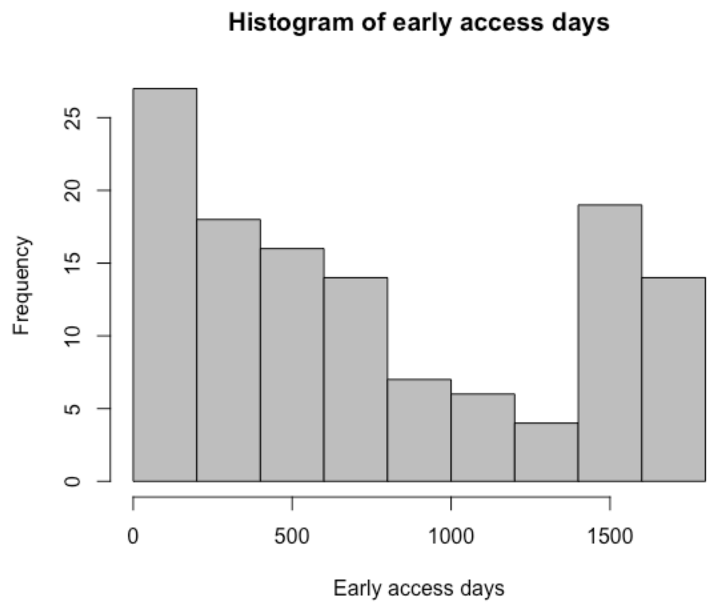



*Figure 8 Histogram of early access days*

Variable related to the game quality also been collected. In this research, I use meta-score (0-100) from Steam to represent the game quality. But only 93 titles have meta-score data. Therefore, it doesn't include in the model to keep the amount of data as much as possible.

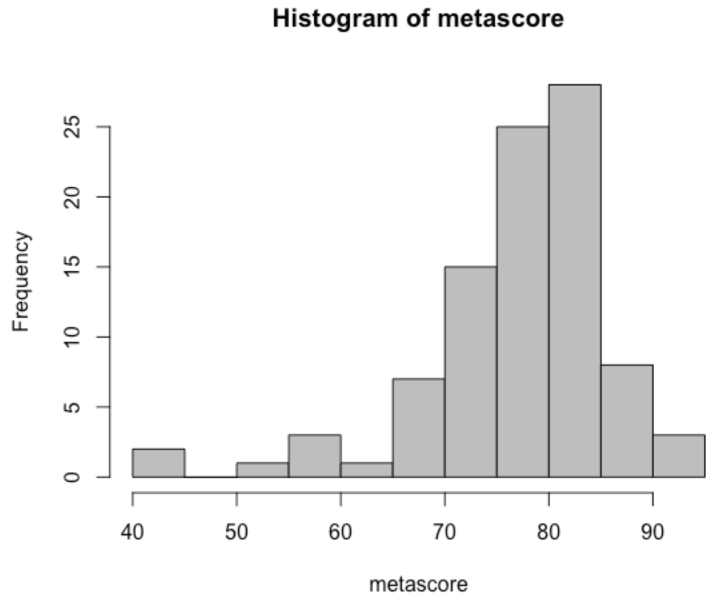



*Figure 9 Histogram of game title meta-score*

### *4.4. Results*

The empirical analysis built two regression models according to the conceptual model indicated in the part 3 modeling approach. And because of the skewed data point distribution on the dependent variable, owner volume, for both short-term and long-term, I use log(n+1) transformation of the independent variables in both short-term and long-term models. Here is the result of the regression coefficients estimation for the two regression models.

*Table 8 Regression coefficients*

| | Hypothesis | Model 1 1st week | Model 2 26th week |
|---|---|---|---|
| Intercept ($\boldsymbol{\beta_0}$) | | 10.1652*** | 11.7703*** |
| **Main effect variables** | | | |
| Event dummy ($\boldsymbol{\beta_1}$) | H1, H2, H3 | -0.2078 | -0.2287 |
| Event volume ($\boldsymbol{\beta_2}$) | H1, H3 | 0.0409* | -0.0112 |
| Audience reaction mean volume ($\boldsymbol{\beta_3}$) | H2, H3 | 0.0042*** | 0.0035*** |
| Free sample ($\boldsymbol{\beta_4}$) | | 1.0501*** | 0.4746** |
| Company awareness ($\boldsymbol{\beta_5}$) | | 0.6668* | 0.4947** |
| **Moderator** | | | |
| Event volume × Free sample ($\boldsymbol{\beta_6}$) | H4 | 0.0086 | 0.0431** |
| Audience reaction mean volume × Free sample ($\boldsymbol{\beta_7}$) | H4 | -0.0034** | -0.0024** |
| Event volume × Company awareness ($\boldsymbol{\beta_8}$) | H5 | -0.0013 | 0.0299. |
| Audience reaction mean volume × Company awareness ($\boldsymbol{\beta_9}$) | H5 | -0.0032** | -0.0028*** |
| $\boldsymbol{R^2}$ | | 0.2 | 0.1868 |
| $\boldsymbol{R^2_{adj.}}$ | | 0.1705 | 0.1596 |
| **p-value** | | 1.111e-08 | 6.802e-09 |
| **Observations** | | 245 | 270 |

*Note*: . $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

I use Model 1 (short-term performance model, using owners volume at the end of 1st week as the independent variable) to test H1, H2, H4 and H5, because Model 1 captures the main effects from event volume and audience reactions, and captures the moderation effects from interactions between eWOM and free sample and between eWOM and company awareness in the short-term.

I use Model 2 (long-term performance model, using owners volume at the end of 26th week as the independent variable) to test H3-H5, because Model 2 captures the main effects from prerelease eWOM and captures the moderation effects from interactions between eWOM and free sample and between eWOM and company awareness in the long-term.

The $\boldsymbol{R^2}$ of Model 1 and Model 2 are 0.2 and 0.18, respectively, indicating the level of model fit. Both Model 1 and Model 2 are statistically significant, as the p-value of the two models are smaller than 0.05.

*4.4.1. Prerelease publicities and new release performance*

Hypothesis 1 is tested through the main effect variable "event volume" ($\beta_2$) in Model 1.

The p-value of event volume is 0.015, which is smaller than 0.05. Therefore, the estimated coefficient is statistically significant. The estimated coefficient of event volume is 0.0409, which means holding other variables constant. One unit increase in prerelease event volume leads to an approximate 4.09% increase in the owner's volume at the end of the first week after release.

According to the analysis, we can accept the hypothesis that prerelease publicities have a positive effect on the short-term product performance.

*4.4.2. Audience reactions to prerelease publicities and new release performance*

Hypothesis 2 tests through the main effect variable "audience reactions mean volume" ($\beta_3$) in Model 1.

The p-value of audience reactions mean volume is 6.28e-05, smaller than 0.05. Therefore, the estimated coefficient is statistically significant. The estimated coefficient of audience reactions mean volume is 0.0042, which means holding other variables constant, one unit increase in audience reaction mean volume of events in the prerelease period leads to approximate 0.42% increase in the owner's volume at the end of first week after release.

Although 0.42% increasement sounds like a very small effect, the average audience reactions mean for the games who having prerelease publicities is 224.3. Giving other variables unchanged, Audience reactions mean volume of 224.3 can bring 94.2% increasement to the owner's volume at the end of first week after release.

According to the analysis, we can accept the hypothesis that audience reactions of prerelease publicities have positive effect on the short-term product performance.

*Table 9 Descriptive statistics of audience reactions mean volume for observations having prerelease publicities*

| | **Min** | **1st Qu.** | **Median** | **Mean** | **3rd Qu.** | **Max** |
|---|---|---|---|---|---|---|
| **Audience reactions mean** | 3.0 | 48.5 | 109.2 | 224.3 | 237.2 | 4,147.5 |

*4.4.3. Event dummy*

Except event volume and audience reactions volume, I introduced a dummy variable "event dummy" ($\beta_1$) to capture the possible different effect from no prerelease publicity to 1 prerelease publicity and other publicities volume increasement.

The p-value of event dummy is 0.44, which is not statistically significant. However, the absolute value of estimated coefficient for event dummy (-0.2078) is material bigger than the estimated coefficient of event volume (0.0409).

The model estimates doing prerelease publicities can decrease 20.78% owner's volume at the end of first week after release, holding other variables constant. But every 1 increasement of prerelease publicities after wards can increase 4.09% owner's volume at the end of first week after release.

For example, if one product would do 6 prerelease publicities, then the effect from the 6 prerelease publicities itself is -20.78% + 6×4.09% = 3.76%. Therefore, giving other variables constant, do 6 prerelease publicities would lead to 3.76% increasement of owner's volume at the end of first week after release.

Event dummy also connected with another variable; audience reactions mean volume. If event dummy is 0, then audience reactions mean is also 0, because no reactions can happen if no prerelease publicities exists.

For example, if one product would do prerelease publicities, and can record 500 average audience reactions per event. Given other variables constant, the prerelease publicities can lead to -20.78% + 500×0.42% = 189.22% increasement of owner's volume at the end of first week after release.

However, comparing with not doing prerelease publicities, doing prerelease publicities can decrease 20.78% owner's volume at the end of first week after release. If the publicities are not attractive enough to generate audience reactions or the publicities volume is very limited, doing publicities can even bring negative effects.

*4.4.4. Prerelease eWOM and long-term product performance*

Hypothesis 3 is tested through main effect variable "event dummy" ($\beta_1$), "event volume" ($\beta_2$), and "audience reactions volume" ($\beta_3$) in Model 2.

In Model 2, both event dummy and event volume have p-value bigger than 0.05, not significant statistically. And the estimated coefficients of event dummy and event volume are -0.2287 and -0.0112 respectively. In other word, prerelease publicity (event) from 0 to 1 would lead to 22.87% decrease on owner's volume at the end of 26$^{th}$ week (long-term performance); prerelease publicities (event) one unit increase from 1 would lead to 1.12% decrease on owner's volume at the end of 26$^{th}$ week (long-term performance).

The p-value of audience reactions mean volume in Model 2 is 2.27e-06, smaller than 0.05, statistically significant. The estimated coefficient of audience reactions mean volume is 0.0035, which means one unit increase in audience reactions mean volume of the pre-releasing publicities leads to 0.35% increase in owner's volume at the end of 26$^{th}$ week (long-term performance).

Among event dummy, event volume and audience reactions mean volume, only audience reactions mean volume is estimated to have positive effect on the products long-term performance.

If a product is going to have $n$ prerelease publicities, and the average audience reaction volume of the $n$ prerelease publicities is m, given other variables constant, the effect of the prerelease eWOM on product's long-term sales performance increase or decrease percentage is:

$$-22.87\% + (-1.12\%)n + 0.35\%m$$

Therefore, the direction of the prerelease eWOM effect on product's long-term performance is relying on the level of audience reaction. If the audience reaction is at a very high level, the effect is positive; if the audience reaction is at a very low level, the effect is negative.

*4.4.5. Free sample and effect of prerelease eWOM*

Hypothesis 4 is tested through moderator interaction "Event volume × Free sample" ($\beta_6$) and moderator interaction "Audience reaction mean volume × Free sample" ($\beta_7$) in Model 1 for its short-term effect and in Model 2 for its long-term effect.

*1. Short-term effect of moderator free sample*

The short-term moderator effect is evaluated through Model 1.

The p-value of the interaction between event volume and free sample is 0.7, not statistically significant. However, the p-value of the interaction between audience reaction mean volume and free sample is 0.001, which is smaller than 0.05. The estimated coefficient of the interaction between audience reaction mean volume and free sample is statistically significant.

Estimated coefficient of the interaction between audience reaction mean volume and free sample is -0.0034. Which means holding other variables constant, when the product offering free sample, one unit increase in audience reaction mean volume leads to 0.42% - 0.34% = 0.08% increase in the owner's volume at the end of first week after release (short-term performance); when not offering free sample, one unit increase in audience reaction mean volume leads to 0.42% increase in the owner's volume.

Because 0.08% is much smaller than 0.42%, the prerelease eWOM effect on short-term new release sales performance would decrease if the product offering free sample; the prerelease eWOM effect on short-term new release sales performance would increase if the product does not offer free sample.

According to the analysis, we can accept the hypothesis that not offering free sample can lead to a stronger prerelease eWOM effect on short-term new release performance.

*2. Long-term effect of moderator free sample*

The long-term moderator effect is evaluated through Model 2.

The p-value of the interaction between event volume and free sample is 0.007, smaller than 0.05, statistically significant. The p-value of the interaction between audience reaction mean volume and free sample is 0.001, also statistically significant. The estimated coefficients of both the interaction between event volume and free sample and the interaction between audience reaction mean volume and free sample are statistically significant.

Estimated coefficient of the interaction between event volume and free sample is 0.0431. Which means holding other variables constant, when the product offering free sample, one unit increase in event volume leads to -1.12% + 4.31% = 3.19%

increase in the owner's volume at the end of 26th week after release (long-term performance); when not offering free sample, one unit increase in event volume leads to 1.12% decrease in the owner's volume.

Estimated coefficient of the interaction between audience reaction mean volume and free sample is -0.0024. Which means holding other variables constant, when the product offering free sample, one unit increase in audience reaction mean volume leads to 0.35% - 0.24% = 0.11% increase in the owner's volume at the end of 26th week after release (long-term performance); when not offering free sample, one unit increase in audience reaction mean volume leads to 0.35% increase in the owner's volume.

Moderator free sample influence the effect of prerelease eWOM on long-term product performance in two ways. First, it strengthens the prerelease eWOM effect from the event volume (publicities); second, it weakens the prerelease eWOM effect from the audience reactions volume.

#### *4.4.6. Company awareness and effect of prerelease eWOM*

Hypothesis 5 is tested through moderator interaction "Event volume $\times$ $company\ awareness$" ($\beta_8$) and moderator interaction "Audience reaction mean volume × company awareness" ($\beta_9$) in Model 1 for its short-term effect and in Model 2 for its long-term effect.

*1. Short-term effect of moderator company awareness*

The short-term moderator effect is evaluated through Model 1.

The p-value of the interaction between event volume and company awareness is 0.95, not statistically significant. However, the p-value of the interaction between audience reaction mean volume and company awareness is 0.003, which is smaller than 0.05. The estimated coefficient of the interaction between audience reaction mean volume and company awareness is statistically significant.

Estimated coefficient of the interaction between audience reaction mean volume and company awareness is -0.0032. Which means holding other variables constant, when the product is from a well-known company, one unit increase in audience reaction mean volume leads to 0.42% - 0.32% = 0.1% increase in the owner's volume at the end of first week after release (short-term performance); when not from a well-known company, one unit increase in audience reaction mean volume leads to 0.42% increase in the owner's volume.

Because 0.1% is smaller than 0.42%, the prerelease eWOM effect on short-term new release sales performance would decrease if the product from a well-known company.

According to the analysis, we can accept the hypothesis that low company awareness leads to a stronger prerelease eWOM effect on short-term new release performance.

*2. The long-term effect of moderator company awareness*

The long-term moderator effect is evaluated through Model 2.

The p-value of the interaction between event volume and company awareness is 0.054, a little big bigger than 0.05, not statistically significant but important. The p-value of the interaction between audience reaction mean volume and company awareness is 0.003, smaller than 0.05, statistically significant.

Estimated coefficient of the interaction between event volume and company awareness is 0.0299. Which means holding other variables constant, when the product is from a well-known company, one unit increase in event volume leads to -1.12% + 2.99% = 1.87% increase in the owner's volume at the end of $26^{th}$ week after release (long-term performance); when not from well-known company, one unit increase in event volume leads to 1.12% decrease in the owner's volume.

Release from a well-known company can lead to a stronger effect of prerelease publicities on long-term sales performance.

Estimated coefficient of the interaction between audience reaction mean volume and company awareness is -0.0028. Which means holding other variables constant, when the product from a well-known company, one unit increase in audience reaction mean volume leads to 0.35% - 0.28% = 0.07% increase in the owner's volume at the end of $26^{th}$ week after release (long-term performance); when not from well-known company, one unit increase in audience reaction mean volume leads to 0.35% increase in the owner's volume.

Release from a not well-known company can lead to a stronger effect of prerelease audience reactions on long-term sales performance.

Moderator company awareness influence the effect of prerelease eWOM on long-term product performance in two ways. First, it strengthens the prerelease eWOM effect from the event volume (publicities); second, it weakens the prerelease eWOM effect from the audience reactions volume.

### *4.4.7. Robustness checks*

To check the robustness of my findings, I estimate alternative model adding the control variables mentioned in the data and measures sector. The control variables are "external developers" ($\beta_{10}$), "indie" ($\beta_{11}$) and "early access" ($\beta_{12}$).

As shown in Table 10 Robustness checks, the results of the new Model 3 (short-term performance) and Model 4 (long-term performance) are highly similar to the results in the original model. Despite a little bit higher or lower coefficient, the results are almost the same for including control variables and excluding control variables.

Because all of the control variables are not significant in both short-term and long-term performance model, except external developer, who is statistically important for the long-term performance ($p = 0.08$). I did not include control variables in the analysis model as shown in Table 8 Regression coefficients.

*Table 10 Robustness checks*

| | Hypothesis | Model 3 1st week | Model 4 26th week |
|---|---|---|---|
| Intercept ($\boldsymbol{\beta_0}$) | | 10.1504*** | 11.8122*** |
| **Main effect variables** | | | |
| Event dummy ($\boldsymbol{\beta_1}$) | H1, H2, H3 | -0.1703 | -0.2877 |
| Event volume ($\boldsymbol{\beta_2}$) | H1, H3 | 0.0406* | -0.0144 |
| Audience reaction mean volume ($\boldsymbol{\beta_3}$) | H2, H3 | 0.0040*** | 0.0034*** |
| Free sample ($\boldsymbol{\beta_4}$) | | 1.0312*** | 0.3860* |
| Company awareness ($\boldsymbol{\beta_5}$) | | 0.6677* | 0.5662** |
| **Moderator** | | | |
| Event volume × Free sample ($\boldsymbol{\beta_6}$) | H4 | 0.0071 | 0.0453** |
| Audience reaction mean volume × Free sample ($\boldsymbol{\beta_7}$) | H4 | -0.0033** | -0.0023* |
| Event volume × Company awareness ($\boldsymbol{\beta_8}$) | H5 | -0.0004 | 0.0321 |
| Audience reaction mean volume × Company awareness ($\boldsymbol{\beta_9}$) | H5 | -0.0031** | -0.0027*** |
| **Controller** | | | |
| External developers ($\boldsymbol{\beta_{10}}$) | | -0.1689 | -0.2752. |
| Indie ($\boldsymbol{\beta_{11}}$) | | 0.1913 | 0.1248 |
| Early access ($\boldsymbol{\beta_{12}}$) | | 0.0264 | 0.1988 |
| $\boldsymbol{R^2}$ | | 0.2043 | 0.202 |
| $\boldsymbol{R^2_{adj.}}$ | | 0.1647 | 0.166 |
| p-value | | 1.099e-07 | 1.401e-08 |
| Observations | | 242 | 266 |

*Note*: . p < 0.1; * p < 0.05; ** p < 0.01; *** p < 0.001

Additionally, I use the correlation matrix to check multilinearity for the analysis model.

From Table 11 Correlation matrix of the variables in analysis model, we can see that only the correlation between event volume and event dummy is relatively high (0.56), but still acceptable. Because event dummy is an indicator of event volume equals 0 or not equals 0, and the correlation still smaller than 0.8.

Therefore, the multilinearity of the independent variables is avoided in the analysis model. The empirical analysis results are trustworthy.

*Table 11 Correlation matrix of the variables in analysis model*

| | **Event dummy** | **Event volume** | **Audience reactions** | **Free** | **Company awareness** |
|---|---|---|---|---|---|
| **Event dummy** | 1 | 0.56 | 0.32 | -0.09 | -0.03 |
| **Event volume** | 0.56 | 1 | 0.20 | -0.08 | -0.002 |
| **Audience reactions** | 0.32 | 0.20 | 1 | -0.10 | 0.07 |
| **Free** | -0.09 | -0.08 | -0.10 | 1 | -0.28 |
| **Company awareness** | -0.03 | -0.002 | 0.07 | -0.28 | 1 |

# 5. Discussion and Conclusions

### *5.1. Research implications*

The present research studied the effect of pre-releasing eWOM on new preleases products sales performance in both the short-term and long-term and studied the moderator roles of free sample and company awareness.

The analysis results of empirical study in the video games industry show that audience reaction volume to company prerelease announcement has a positive effect on both short-term and long-term product sales performance. However, if announce the new releases product before the official launch should be considered differently because the analysis results show that given other variables constant, having prerelease events from 0 to 1 will decrease 16.69% owners in the short-term (proxy of sales performance), and reduce 21.75% owners in long-term. However, good performance on audience reaction volume can bring a decrease back to an impressive percentage increase.

The analysis results also show that not offering free samples and new releases from an unwell-known company will strengthen the short-term prerelease eWOM effect. In the long-term, moderators (free sample and company awareness) influence the effects of prerelease publicity volume and audience reaction volume differently.

The present research contributes to the studies on "prerelease eWOM and new product performance" in two main areas.

One is in the main effect area. The present study constructs eWOM by combining prerelease announcement publicities from the company side and the cocreation of eWOM from audience reactions, based on the eWOM network coproduction model (Kozinets, 2010). Therefore, it is possible to research the effect of prerelease eWOM not only from 1 to n, but also from 0 to 1. Although the effect from 0 to 1 does not appear statistically significant in my analysis model, the estimated coefficient is big enough to raise our attention.

Second is the moderator effect area. The present study researches the moderator roles of free sample and company awareness in this topic for the first time.

Firstly, the present research extends the previous study by investigating the moderator role of free sample, a variable studied as the main effect in previous research. Li, Jain and Kanna (2019) found that free samples of the entire content can be very effective in increasing revenues. The present study innovatively studying the moderator role of free sample, therefore build a connection between the pre-launch marketing strategy and product pricing strategy.

Secondly, the present research extends the previous study by investigating the moderator role of company awareness. Although company awareness was being investigated as a moderator of the eWOM effect on product sales (Lopez & Sicilia, 2013), it is the first time investigating the moderator role of company awareness to the prerelease eWOM effect.

### *5.2. Managerial implications*

Based on the finding through empirical study and analysis results, this research can offer four managerial implications.

First, using prerelease announcement publicity carefully.

Prerelease publicities not always have a positive effect on the new product sales performance. From the analysis result, given other variables constant, having prerelease events from 0 to 1 will decrease 16.69% owners in the short-term (proxy of sales performance), and reduce 21.75% in the long-term. However, the audience reaction volume from prerelease publicities can eliminate the negative effect and bring impressive positive effect if the volume is big enough.

The energy and budget of producing and marketing a new product are limited; if the input can hardly ensure high-quality prerelease publicity and a good audience reaction, the prerelease publicity would likely lead to even worse product performance.

Therefore, using or not using a prerelease announcement is depending on the specific situation. Not all products are suitable to have prerelease announcement publicities. For example, for the product with limited labor capacity and budget, it is better to concentrate on product production and find other marketing activity solutions such as post-release events because the poor performance of prerelease publicities would negatively influence the ultimate sales performance.

Second, focus on the quality and performance of prerelease marketing activities comparing with the frequency of prerelease marketing activities.

According to the analysis results, the average audience reaction volume to the prerelease publicities positively affects both short-term and long-term product sales performance. However, the frequency of prerelease publicities has a negative effect on long-term product sales performance.

Moreover, the improvement room for the audience reactions is unlimited to some extent, and high WOM buzz volume also increases the likelihood that people contribute more to the buzz through social impact (Banerjee, 1992).

Therefore, companies should consider more on improving the audience reactions, instead of simply increasing the frequency of prerelease publicity to achieve high exposure.

Third, digital products that use premium pricing models should pay more attention to prerelease marketing activities and generate eWOM as much as possible to leverage the new products' opening sales and lifetime sales performance.

According to the findings of free sample moderator roles in influencing pre-releasing eWOM effect on new product performance, using a premium pricing model that does not allow free trials strengthens the prerelease eWOM effect because consumers are more likely to seek eWOM to eliminate uncertainty and decrease the risk of purchasing products that cannot try before making a purchase.

Similarly, for traditional products that do not offer free sample in any form, Prerelease eWOM is also more relevant to the sale performance of that kind of product than those offering free samples. For those products, eWOM is good leverage to generate new product sales in both the opening and long-term.

Fourth, Small companies that do not have a good awareness among consumers can benefit from the prerelease eWOM more.

Because of the analysis results, unwell-known companies strengthen the prerelease eWOM effect on sales performance through audience reactions.

Therefore, these companies should keep closer communication with customers, leverage the eWOM to sales performance, and eventually increase company value and company awareness.

### *5.3. limitations and future research*

At the data collecting stage of the present research, I meet a lot of drawbacks that make me unable to realize some research ideas I had. The drawbacks in the data collection stage limited the variables been considered in the present research.

First, this research is limited in the volume dimension of eWOM and does not include other dimensions, such as eWOM valence and eWOM variance, because I do not have enough capacity to collect and process valence and variance data.

Second, some variables I mentioned in the thesis did not include in the final analysis modeling, such as the meta score of each game title, because of un-enough data points. The ability to control for some important variables is not enough in this research analysis. And the reason for this limitation is also by lack of data.

In future research in this area, panel data can be used to construct a dynamic regression model to capture the effect of timing.

The use of digital marketing channels to make an announcement before the products' official release is common in today's digital marketing environment. Companies know the power of eWOM and prerelease publicities and keen to find a way to measure the effect bring by the marketing efforts and want to know when the best timing is for making an announcement.

Through construct a dynamic model on panel data, the model can also be used as a tool to predict the new product performance through prerelease eWOM, which can bring unlimited value for companies because a lot of business decisions can be made based on the future demand prediction.

Future studies can also consider other factors, such as the different types of publicities and how can the types influence the effect of the prerelease eWOM effect. For example, can divide the type by different digital channels or by different forms (article, video, picture, podcast, or more), and research the influence of the different types.

Last but not least, research on the factors that influence the eWOM volume may also be an interesting future research direction. Previous research already demonstrates that the eWOM volume significantly influences product sales performance (Trusov, 2010). The present study also finds that prerelease eWOM volume significantly influences the new product sales performance.

But for practitioners, improving the eWOM volume is more relevant to their daily work, and can bring more direct value.

Therefore, it is also important for the academic to research on the mechanism of eWOM volume accumulation and find out the practical and useful methods for increasing eWOM volume.

## Acknowledgements

This article was completed under the guidance of Professor Oest. Thanks to my thesis supervisor Professor Oest. In the whole process of thesis topic selection, research method determination, and final completion, Professor Oest always responsibly put forward many valuable opinions and suggestions for me and helped me develop research ideas and sort out the research framework. Give me deep and detailed guidance. On the occasion of the completion of this paper, I would like to express my heartfelt thanks to Professor Oest.

The authors also thank her friends Penglong Li, Lian Zhang, Xifan Cui, and Yizhen Sun for their encourage and support during the hard time of COVID-19.

# References

Ainslie, Andrew, Drèze, Xavier, & Zufryden, Fred. (2005). Modeling Movie Life Cycles and Market Share. Marketing Science (Providence, R.I.), 24(3), 508-517.

Alexa B. Burmester, Jan U. Becker, Harald J. van Heerde, Michel Clement. (2015). The impact of pre- and post-launch publicity and advertising on new product sales. International Journal of Research in Marketing, 32(4), 408-417. doi:10.1016/j.ijresmar.2015.05.005.

Babić Rosario, A., Sotgiu, F., De Valck, K., & Bijmolt, T. H. A. (2016). The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors. Journal of Marketing Research, 53(3), 297–318. https://doi.org/10.1509/jmr.14.0380

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

Bass, F. M. (1969). A New Product Growth for Model Consumer Durables. Management Science, 15(5), 215-227.

Beck, Jonathan (2007). The Sales Effect of Word of Mouth: A Model for Creative Goods and Estimates for Novels. Journal of Cultural Economics, 31 (1), 5–23.

Casidy, R., & Wymer, W. (2015). The impact of brand strength on satisfaction, loyalty and WOM: An empirical examination in the higher education sector. Journal of Brand Management, 22(2), 117-135. doi:http://dx.doi.org.ezproxy.library.bi.no/10.1057/bm.2015.6

Ding, Chao, Cheng, Hsing Kenneth, Duan, Yang, & Jin, Yong. (2017). The power of the "like" button: The impact of social media on box office. Decision Support Systems, 94, 77-84.

Elberse, Anita, & Eliashberg, Jehoshua. (2003). Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures. Marketing Science (Providence, R.I.), 22(3), 329-354.

Godes, David, & Mayzlin, Dina. (2004). Using Online Conversations to Study Word-of-Mouth Communication. Marketing Science (Providence, R.I.), 23(4), 545-560.

Goldenberg, Jacob, Libai, Barak, Moldovan, Sarit, & Muller, Eitan. (2007). The NPV of bad news. International Journal of Research in Marketing, 24(3), 186-200.

Guiyang Xiong, & Sundar Bharadwaj. (2014). Prerelease Buzz Evolution Patterns and New Product Performance. Marketing Science, 33(3), 401-421.

Janiszewski, C. (1993). Preattentive Mere Exposure Effects. The Journal of Consumer Research, 20(3), 376-392.

Kozinets, R. V., De Valck, K., Wojnicki, A. C., & Wilner, S. J. S. (2010). Networked Narratives: Understanding Word-of-Mouth Marketing in Online Communities. Journal of Marketing, 74(2), 71–89. https://doi.org/10.1509/jm.74.2.71

Lamberton, Cait, & Stephen, Andrew T. (2016). A Thematic Exploration of Digital, Social Media, and Mobile Marketing: Research Evolution from 2000 to 2015 and an Agenda for Future Inquiry. Journal of Marketing, 80(6), 146-172.

Li, H. (Alice), Jain, S., & Kannan, P. K. (2019). Optimal Design of Free Samples for Digital Products and Services. Journal of Marketing Research, 56(3), 419–438. https://doi.org/10.1177/0022243718823169

Li, Hongshuang, Jain, Sanjay, & Kannan, P.K. (2019). Optimal Design of Free Samples for Digital Products and Services. Journal of Marketing Research, 56(3), 419-438.

Lopez, Manuela, & Sicilia, Maria. (2013). How WOM marketing contributes to new product adoption. European Journal of Marketing, 47(7), 1089-1114.

Mäntymäki, M., Hyrynsalmi, S., & Koskenvoima, A. (2019). How do small and medium-sized game companies use analytics? an attention-based view of game analytics. Information Systems Frontiers, , 1-16. doi:http://dx.doi.org.ezproxy.library.bi.no/10.1007/s10796-019-09913-1

Misner, Ivan R. (1999), The World's Best Known Marketing Secret: Building Your Business with Word-of-Mouth Marketing, 2d ed. Austin: Bard Press.

Nguyen, Hang T, & Chaudhuri, Malika. (2019). Making new products go viral and succeed. International Journal of Research in Marketing, 36(1), 39-62.

Peres, Renana, Muller, Eitan, & Mahajan, Vijay. (2010). Innovation diffusion and new product growth models: A critical review and research directions. International Journal of Research in Marketing, 27(2), 91-106.

Raguseo, Elisabetta, & Vitari, Claudio. (2017). The Effect of Brand on the Impact of e-WOM on Hotels' Financial Performance. International Journal of Electronic Commerce, 21(2), 249-269.

Reimer, Thomas, & Benkenstein, Martin. (2016). When good WOM hurts and bad WOM gains: The effect of untrustworthy online reviews. Journal of Business Research, 69(12), 5993-6001.

Schatzel, Kim, & Calantone, Roger. (2006). Creating market anticipation: An exploratory examination of the effect of preannouncement behavior on a new product's launch. Journal of the Academy of Marketing Science, 34(3), 357-366.

SEMrush. (December 10, 2019). Use of content marketing in organizations worldwide as of June 2019 [Graph]. In Statista. Retrieved August 28, 2020, from https://www-statista-com.ezproxy.library.bi.no/statistics/251434/use-of-content-marketing-worldwide/

Sonnier, G., McAlister, L., & Rutz, O. (2011). A Dynamic Model of the Effect of Online Communications on Firm Sales. Marketing Science, 30(4), 702-716. Retrieved August 9, 2020, from www.jstor.org/stable/23012020

Tirunillai, Seshadri, & Tellis, Gerard J. (2012). Does Chatter Really Matter? Dynamics of User-Generated Content and Stock Performance. Marketing Science (Providence, R.I.), 31(2), 198-215.

Toubia, Olivier, & Stephen, Andrew T. (2013). Intrinsic vs. Image-Related Utility in Social Media: Why Do People Contribute Content to Twitter? Marketing Science (Providence, R.I.), 32(3), 368-392.

Trusov, Michael, Bodapati, Anand V, & Bucklin, Randolph E. (2010). Determining Influential Users in Internet Social Networks. Journal of Marketing Research, 47(4), 643-658.

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. https://doi.org/10.1509/jmkg.73.5.90

Wang, Y., Li, M., Feng, H. et al. Optimal sequential releasing strategy for software products in the presence of word-of-mouth and requirements uncertainty. Inf Technol Manag 20, 153–174 (2019). https://doi-org.ezproxy.library.bi.no/10.1007/s10799-018-0296-1

Wu, Lingli, Deng, Shiming, & Jiang, Xuan. (2018). Sampling and pricing strategy under competition. Omega, 80, 192-208.

Xiao Liu, Param Vir Singh, and Kannan Srinivasan. (2016). A Structured Analysis of Unstructured Big Data by Leveraging Cloud Computing. Marketing Science, 35(3), 363-388. doi:10.1287/mksc.2015.0972.