HØGSKOLEN I OSLO
OG AKERSHUS

Kim Andreas Hermansen & Jarl Romeo Bratli

How third party reviews affect firm value

An event study in the video game industry

Masteroppgave i økonomi og administrasjon
Handelshøyskolen ved HiOA

Abstract

The purpose of this thesis is to examine the relationship between third party reviews and firm value in the video game industry. This topic has been debated in the media, but no previous research has examined the relationship. Our study aims to fill this gap in research. We find that third party reviews affect firm value. This effect is limited to good reviews, while bad reviews do not affect firm value alone. We also find that investors do not react on a single important review, but rather wait until more information is available. Our study also shows that anticipation, for a video game, plays a big part in the effect third party reviews has on firm value. Higher anticipation for a video game increases the impact third party reviews has on firm value. Our findings can improve further research that examines the effect of third party reviews on other economic variables, like revenue.

Sammendrag

Formålet med denne studien er å undersøke forholdet mellom tredjeparts anmeldelser og firmaverdi i spillindustrien. Forholdet har blitt debattert i media, men det har ikke blitt forsket på tidligere. Vår studie prøver å bidra til dette forskningstemaet. Vi finner at tredjeparts anmeldelser påvirker firmaverdi. Denne effekten er avgrenset til gode anmeldelser, mens dårlige anmeldelser ikke påvirker firmaverdi alene. Vi finner også at investorer ikke reagerer på én viktig anmeldelse, men heller venter til mer informasjon er tilgjengelig. Vår studie viser også at forventning, til et spill, har stor betydning for effekten en tredjeparts anmeldelse har på firmaverdi. Høyere forventing øker effekten av en tredjeparts anmeldelse på firmaverdi. Våre funn kan forbedre fremtidig forskning som undersøker effekten av tredjeparts anmeldelser på andre økonomiske variabler, som inntekter.

Preface

Our Master thesis marks the end of our financial studies in Business Administration, at Oslo Business School. It has been a challenging and rewarding process.

Funcom, Norway's largest video game publisher, has experienced some gigantic fluctuations in firm value over the years. In some cases have Funcom blamed third party reviews for the stock markets reaction. We found it very interesting that investors perhaps used the same third party reviews to evaluate stocks, which we used to choose video games. After some early examinations we found that the impact of third party reviews was an interesting and challenging theme for our thesis. The phenomenon was a popular topic in the media, and among industry managers and investors. There was also very little research on the subject.

Collecting data has been a very challenging part of the process. All data had to be collected manually, and we needed more data than we first expected. The exploratory nature of the thesis was also demanding, as there was very few that had examined video games before us.

We would like to thank our supervisor Helge Nordahl, for guiding us through the process and for great insight in the field of finance. We would also like to thank Stian Danielsen, editor at IGN Norway, for help in finding data sources.

Table of contents

1 Introduction

The video game industry is a relatively new industry compared to other entertainment industries. It started with “Pong” in the 1970’s, the first home video game, and has since skyrocketed in popularity. Even though the video game industry is one of the youngest within the entertainment sector, it was estimated to reach $ 91,5 billion in revenues in 2015. A 9 percent growth from $83,6 billion in 2014 (Newzoo 2015). The video game industry has now eclipsed the film and music industry. The global box office revenue reached $36,4 billion in 2014 (MPAA 2015), while the global music industry revenues was reduced to $15 billion in 2014 (Ingham 2015). In 2013 Take-Two Interactive’s ”Grand Theft Auto 5” became the fastest selling entertainment product of all time, generating $1 billion in three days (Crecente 2013). The fastest selling movie in the film industry was, by the end of 2015, ”Star Wars: The Force Awakens” generating $1 billion in 12 days (Lang 2015).

Third party reviews have become the standard of evaluating products in the entertainment industry, as well as many other industries. Due to the internet and its rapid expansion over the years, consumers can easily navigate thru a jungle of different video games, movies and TV-shows. This makes professional third party reviews an important tool for both management and investors (Banerjee 2006; Chen and Xie 2005).

Metacritic.com has become an important website as more and more third party reviews have become available. Metacritic collects and aggregates all third party reviews for an average score, called a Metacritic score. All individual reviews are weighted on how important Metacritic deems them, without disclosing this information to the public. The score ranges from 0 to 100, where a score of 100 is the best score possible. This score has been a leading benchmark in the entertainment business when evaluating how “good” a product is (Gilbert 2013; Grubb 2013b; Macdonald 2012; Schreier 2015). Metacritic scores show a census for a product, rather than just one single review or opinion. This has lead journalists, CEO’s and analysts to explain firm value changes with Metacritic scores (Banerjee 2006; Carter 2012).

The video game “Destiny” was one of the most anticipated games of 2014. It was released September 9 2014, and was the most successful launch of a new intellectual property of all time. A new intellectual property is defined as a video game that it is not a sequel or part of a franchise. The news of the successful launch was presented in a press release the day after the video game release and included retailer sale information, as sale to consumers are not available that early. However, the stock still plummeted over 9 percent the next week, resulting in a $1,5 billion reduction in the firm value. In this timeframe reviews had started to make their way on Metacritic giving it a 77 out of a 100. The reviews where published later than usual, because the game was heavily online focused, meaning reviews had to be done after the release of the game. This was not considered a good score for a game with a $ 500 million budget (Kuchera 2014).

On June 14 2011, Take-Two Interactive released “Duke Nukem”, an anticipated video game. That same day Take-Two Interactive’s stock fell 5 percent. This coincided with a Metacritic score in the low 50s out of a 100 for the product (Baker 2011). The same publisher released “Oblivion” on March 20 2006 and the stock climbed 23 percent over the next 10 days. The video game scored a 94 on Metacritic, the highest score at the time for any Xbox game. Weeks later, on April 10, Take two announced that “Oblivion” had shipped 1,7 million units, more than analysts expected before release (Banerjee 2006).

There are several examples like these in the video game industry. A new video game is reviewed, and the stock prices moves. Publishers have even gone as far as to openly blame Metacritic for negative share price movements (Carter 2012). It may seem like they are correlating, and a good or bad review score seems to cause the share price to move in the same direction. Is this always the case, and are they correlated? Some analysts do not think there is any pure correlation, but say that ratings on video games do matter on the stock price (Banerjee 2006).

If professional video game reviews affect video game revenues, an early review from a popular site should have an effect on a firm’s value, because other information is limited when a new product is released. The most reliable information is sales data, but a third party review is available earlier than any information from the company. Video game publishers are also notorious for holding back information on sales, except when the sales are record breaking. Developers and other parties in the video game production process, must sign a “non-disclosure agreements” and cannot disclose any numbers regarding sales (Wouk 2016).

Under the assumption that Metacritic scores affect sales, a third party review can give an indication of the success of a new video game. Investors will presumably use the information when making predictions in the stock market. We built our hypotheses on this assumption from financial theory and the efficient market hypothesis.

1.1 Research problem and hypotheses

In this thesis we take a closer look at the relationship between third party reviews and the firm value of video game publishers. There has not been any financial event study that we know of using the video game industry, no studies on market efficiency using this industry, and there is in general very little research of this industry. We propose the following research problem:

"Do third party reviews of video games affect firm value?"

We use multiple studies to examine this. One study to examine the release date of a video game, and use the Metacritic score to see if there is correlation between stock movements and the reception of the video game. A reception is, in this study, defined as the average third party review scores, reflected in the Metacritic score. Another study to examine the first important review that hits the market, to see if the investors react on the new information and adjust their projections, based on the review. In both cases we want to see if the consumer anticipation of a video game will have an effect on investors. If a video game is highly anticipated by potential consumers, we suggest that a third party review is more important to the investors. This means that news, good or bad, will be more impactful for more anticipated video games than less anticipated video games. We will use the number of third party reviews per video game to measure the anticipation of a video game. It is mainly video games with a high number of reviews that have been presented by the media to explain a correlation between firm value and third party reviews.

Table 1: Hypothesis description

Hypothesis	Description
H_1	A good reception for a video game will have a positive effect on firm value.
H_2	A bad reception for a video game will have a negative effect on firm value.
H_3	The first important good review for a video game will have a positive effect on firm value.
H_4	The first important bad review for a video game will have a negative effect on firm value.
H_5	A reception or third party review for an anticipated video game has a larger effect on firm value.

The research problem is tested with hypotheses defined in Table 1 and 2. Figure 1 shows our hypotheses and how third party reviews might affect firm value. The figure shows both the reception on release day and the first important review, and how anticipation might increase or decrease the effect third party reviews have on firm value.

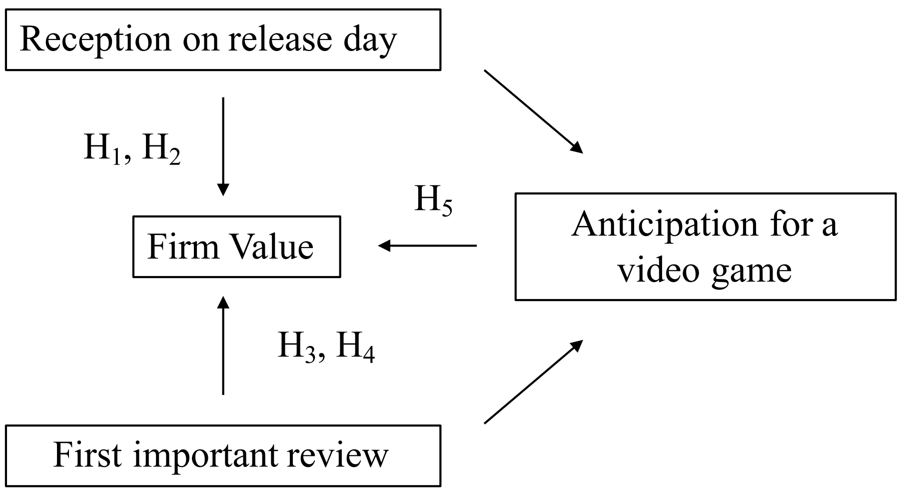


Figure 1: How third party reviews can affect firm value. The figure shows the reception on release day and the first important review, and how they can directly affect firm value. Anticipation is represented as a connection between third party reviews and firm value. Hypotheses for the different effects are shown between the different forces.

We use a one-tailed t-test to test if a good review has a positive effect on CAAR, and a bad review has a negative effect on CAAR. The null hypothesis for hypotheses H_1 and H_3 state that the cumulative average abnormal return (CAAR) is not significantly higher or lower than zero when a video game is released with good reception, or the first important good review is published. The alternative hypothesis states that a good reception or review generates a significantly higher CAAR than zero. The null hypothesis for hypotheses H_2 and H_4 states that a negative reception, or first important review, will not generate a significantly lower CAAR than zero. The alternative is that it does generate a significantly lower CAAR than zero. The null hypothesis for H_5 state that a reception or review for an anticipated video game gives the same CAAR as every other video game, while the alternative is that it gives a larger or smaller CAAR than any other video game depending on the reception or review.

Table 2: Statistical hypotheses details using one-tailed t-tests

Null hypothesis	Alternative hypothesis
H_1: Good reception, CAAR ≤ 0	H_1: Good reception, CAAR > 0
H_2: Bad reception, CAAR ≥ 0	H_2: Bad reception CAAR < 0
H_3: First good review, CAAR ≤ 0	H_3: First good review, CAAR > 0
H_4: First bad review, CAAR ≥ 0	H_4: First bad review, CAAR < 0
H_5: Anticipated CAAR = CAAR	H_5: Anticipated CAAR $\neq$ CAAR

This Master thesis consists of multiple sections. We will first present the theory and previous research on the subject, subsequently we will explain the methodology and the selection of the data sample, followed by the results, discussion and conclusion is presented.

2 Theory

The theory section uses research on market efficiency to explain the effect a third party review might have on an investor.

2.1 The Efficient Market Hypothesis

To understand how third party reviews could affect stock prices, we need to understand the stock market, and how it operates. The stock market's primary function is to allocate ownership of a firm's stock, and the stock price reflects all investors collective assessment of a firm's current and future performance (Bodie, Kane and Marcus 2014, p. 2-5). Fama (1965) suggested Efficient Market Hypothesis in an effort to explain stock price behavior. From here we will use EMH as an abbreviation for Efficient Market Hypothesis.

EMH is defined as an idea of a market where a large number of profit-maximizers are competing with each other, trying to predict future market values, and where all current information is available. This causes the full effect of any news, good or bad, to be reflected "instantaneously" in the actual stock price on average (Fama 1965). Another way to explain the theory is through a famous Wall Street joke. Two investors walk down the street, one spots a hundred-dollar bill and alerts his friend. The other investor tells him not to bother with it. If it really was a hundred-dollar bill there, someone would have picked it up already.

Malkiel (2007) described this phenomenon as a "random walk down Wall Street" where future directions cannot be predicted by using the past history. Meaning that the short run changes in a stock's price is unpredictable. Taking this to the logical extreme Malkiel suggested that a monkey throwing darts, at the stock listings, could select a portfolio that would do as good as a portfolio selected by experts. Stating that it is impossible to beat a broad index fund without taking on more risk.

Even though Fama (1965), states that news will be reflected in the price "instantaneously", it is not always the case. There has been research on both overreaction and delayed reaction on news. The same pattern found in naive undergraduates was found in professional security analysts when forecasting stock movements (De Bondt and Thaler 1990). An overreaction on price might however be adjusted by more rational investors, at a later time, to correctly reflect the news. Bernard and Thomas (1989) also found a delayed market response to information, as stock prices did not fully reflect the news "instantaneously", but rather drift to the correct price. The different changes in stock prices are shown in Figure 2.

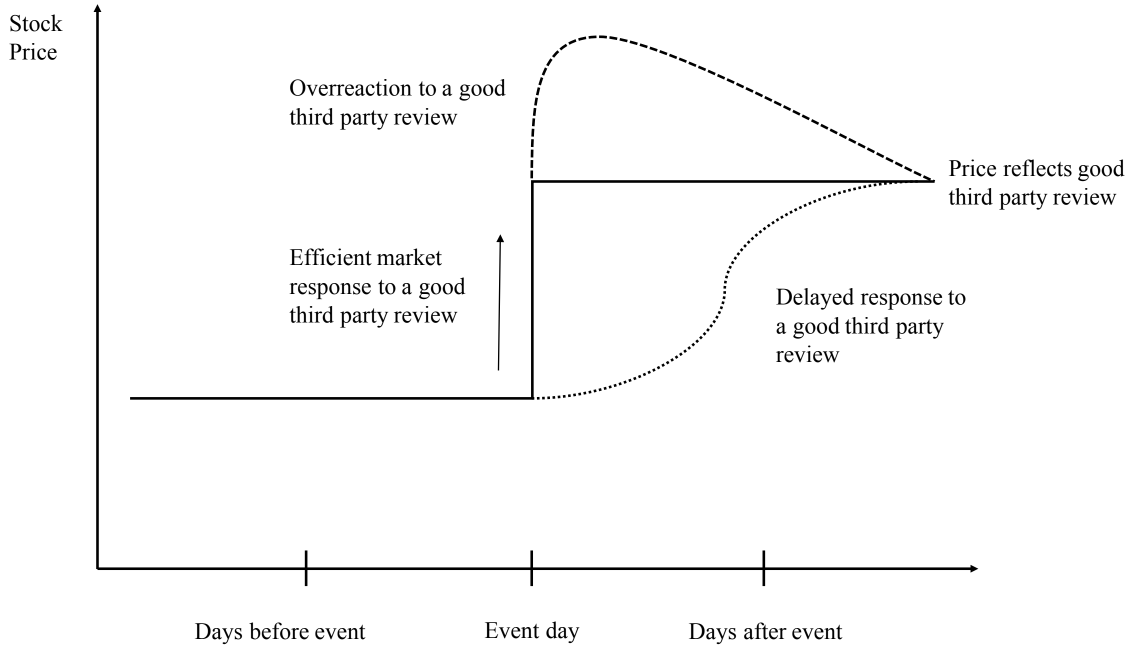


Figure 2: How third party reviews might affect stock movements in a market (Bernard and Thomas 1989; De Bondt and Thaler 1990; Fama 1965).

Malkiel and Fama (1970) suggested three forms of EMH. Weak, semi-strong and strong form, and they all reflect different amount of information in the price as seen in Figure 3.

The definition of the different forms of EMH and what type of information it includes vary between researchers. We use the most common definitions suggested by Burton and Shah (2013, p. 6-8). The semi-strong is the most commonly used form of EMH, where prices accurately summarize all publicly known information. If an investor carefully studies a company it will not give him or her an edge, because someone else has already acted on the information. Meaning that there are no "cheap" or "expensive" stocks, the current price is always the "best estimate". The weak form of EMH only takes historical share prices into account, and the price reflects only this information. This form is already embodied in the semi-strong form where all public information is known. The third form of EMH, the strong, states that prices accurately summarize all information. This includes private as well as public information, although information like this might have been illegally obtained, or legally obtained but illegal to act upon. This form might be too extreme to be an exact description of the world (Malkiel and Fama 1970).

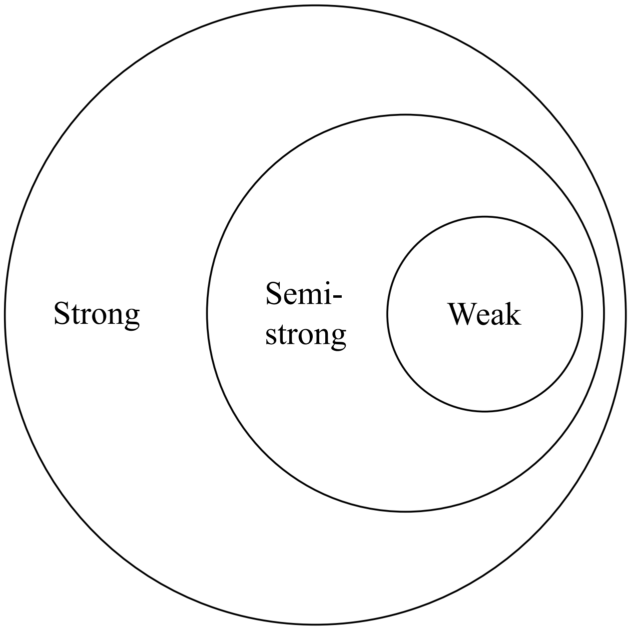


Figure 3: Amount of information reflected in stock prices (Malkiel and Fama 1970)

One of the most decisive and convincing arguments for efficient market theory is the performance of professional fund managers, compared to the performance of the market. Professionals are often compensated with strong incentives to outperform the market, and yet a large body of research (Malkiel 2003) suggest professional funds are not able to outperform index funds that buy and hold a broad stock market portfolio.

2.2 EMH issues and anomalies

The efficient market hypothesis has never been fully accepted by many portfolio managers and researchers, because the theory implies that you can never search for undervalued or overvalued securities. The magnitude issue, selection bias issue and lucky event issue, suggested by Bodie, Kane and Marcus (2014, p. 362-363), keep this debate going.

A 0,1 percent return, above the market, would be hard to detect because of the volatility of a market, but if the portfolio is massive, the 0,1 percent would generate a lot of value for the manager. This is called the magnitude issue and tells us that even though the stocks are priced at a “fair value”, a large enough portfolio might still be able to earn enough trading profits on stocks compared to the market.

Investors that find a way to beat the market would probably not share this with the rest of the world. This creates a selection bias issue where only the findings that supports EMH, or shows negative earnings compared to the market, will be published.

In a population of different investing strategies and investors, there will always be some who outperform the market. This is basic statistics and is called the Lucky Event issue. Some people will make money with enough “bets” placed. Doubters will call it luck, the investor might call it skill, and for every big winner there might be many losers.

2.2.1 Other anomalies

There have been several studies on the efficiency in the stock market. Studies show that stocks with different price/earnings multiple have higher rates of return (Ball 1978; Basu 1977). Other studies show there are seasonal effects like the “January effect” (Haugen and Jorion 1996), where small stocks have higher return within a certain period. There has even been research results that show significant lower average returns on Mondays, also called the “Weekend effect” (French 1980). The market crash in 1987 and the Internet bubble in the late-1990s might also explain that investors are not rational.

If companies are divided into groups by size, studies show that smaller companies have a higher monthly average return (Fama and French 1993; Keim 1983). This presents a problem for the market efficiency theory as this gives investors a possibility to gain an abnormal return adjusted for risk, if the correct measure for risk is the beta from the capital pricing model (Malkiel 2003). If this is the case it is a clear indication that the market is inefficient. Company size is a far better proxy for measuring risk than beta according to Fama and French (1993). The studies are also questionable as there is no size difference in the abnormal returns from the mid-1980s to the late-1990s. Even though some anomalies and patterns have emerged through research over the years, most of the patterns cannot be exploited because of the transaction cost, and one time events are the exceptions from the rule (Malkiel 2003).

To conclude the efficient market hypothesis is widely accepted and the majority of research supports it, but an investor who is especially thorough, creative or intelligent may still outperform the market (Bodie, Kane and Marcus 2014, p. 380). Fama and Malkiel also reviewed the three forms, of EMH, and concluded that the evidence for the efficient market hypothesis was extensive and the contradicting evidence was sparse. Thus supporting EMH while also noting that "much remains to be done" (Malkiel and Fama 1970, p. 416)

2.3 EMH and event studies

"The cleanest evidence of market efficiency comes from event studies, especially event studies on daily returns" (Fama 1991, p. 1607).

Event studies are defined as a empirical financial research technique that enables you to assess the impact of a particular event on a firms stock price (Bodie, Kane and Marcus 2014, p. 359). This comes from the fact that, given a rational and efficient market, the effect of an event will be reflected in the securities price immediately (MacKinlay 1997).

There are large bulks of research done with event studies, and on average they show that the stock prices adjust quickly to new information in the market (Fama 1991). Event studies have changed over time, but the basic statistical format of event studies has remained unchanged the last thirty years (Kothari and Warner 2006), and is a great way to look at events and the way a market reacts. We explain the theory and methods, of event studies, further in the methodology section.

3 Previous research

Research on the video game industry is mostly psychological studies regarding the learning aspect of video games (Gee 2003; Mitchell and Savill-Smith 2004), and effects from violent video games (Anderson 2004; Anderson and Bushman 2001; Dietz 1998). Using research from similar industries we will see how third party reviews affect sale and firm value, and investor behavior.

3.1 Third party reviews

Third party reviews are a big part of many industries. The definition of a third party in this paper is a professional critic, often an "expert" in his or her field. The two other parties are the seller and the buyer, making the reviewer an objective party writing a subjective review, on a product or service, based on different criteria. Third party reviews are especially important in the entertainment industry. As mentioned in the introduction, managers put especially great weight on the reviews of their products. Why are they so important for the firms?

3.1.1 Third party reviews and sales

There are two ways to look at a critic. As an influencer or a predictor. Influencing consumers to buy, shifting the demand for a product, or predicting the total sales for a product. These are not opposite views and critics may be a bit of both (Eliashberg and Shugan 1997). Video game publishers often think of them as influencers. Some video game publishers has even hired former critics to play through their game, and write a confidential review, to see what they need to adjust before it is released to the public (Baker 2011). Even though publishers view them as influencers some studies show conflicting results. Eliashberg and Shugan (1997) used box office sales, close to release, and sales over the whole period a movie was screening, to find out if critics influenced the consumers early or predicted the total sale for the movie. They found no significant correlation between sales early, as trailers and other forms of marketing overshadowed the critics. Total sales, on the other hand, correlated with the movie reviews from critics. They concluded that critics might be more predictors than influencers. Another study by Reinstein and Snyder (2005) found some weak evidence that critics might be influencers, using early box office sales as well. They concluded that third party reviews are not only good mechanics to gauge uncertain products, but are themselves uncertain products. Some reviews might have more power than others in influencing consumers.

In a study by Entertainment Software Association (2015), only three percent of video game consumers reported that they are influenced by third party reviews when buying a video game. Holbrook (1999) found a positive correlation between popular appeal and expert judgment in the movie industry. This explains why reviews might not influence someone but rather predict product sales, as the consumer opinion and third party review correlate.

The studies are pretty clear in the movie industry but the studies are not as clear in the video game industry. At the 2008 DICE summit, Activision Vice President of Marketing Robin Kaminsky, stated that for every additional 5 point over 80, on Metacritic, sales could double (Orland 2014). The statement however has no statistical background and must be viewed with skepticism. Even though there are no academic studies on third party reviews and video game sales, it is still a popular topic in the industry and there have been smaller studies trying to find a correlation.

Gamasutra, a video game industry news site, reported that Susquehanna Financial Group used NPD Group sales data, and Metacritic scores to test this theory. First with 275 Playstation 2 games, a year later with 1200 Playstation 2 games. They found no correlation and could only explain 15,8 percent of the movement in game unit sales from movements in game ratings (Boyer 2006; Maragos 2005). They also looked at other variables like franchise correlation and sequels to try and catch up the other factors that might explain the findings. The research is however difficult to use as only the results are presented, not the whole research. Another flaw is that NDP Group only tracks retailers, so digital sales are not accounted for.

Another small study by Ars Technica, a software developer news site, using sales data from Steam, the leading platform for PC video games, tried to find a correlation without being successful (Orland 2014). They did find that a game is probably better of with a higher score, but nothing close to significant results. They did however use sales estimates instead of real data. The same conclusion was derived in a Master’s thesis using Swedish retailer data and Metacritic scores (Artursson 2015).

Electronic Entertainment Design and Research (EEDAR) have done multiple smaller studies in the video game industry. They calculated the average sales in 2011 and compared it to the critical reviews. 216 games rated 90 or above and the average sales over the first three months was 700,000 units. While the 80 – 90 rated games had an average of only 236,000 units. Games rated lower than this had an even lower sales average (North 2012). This is just averages however, and no correlation or causation can be derived from this. They have also studied how third party reviews affect purchase intent, finding that it does, without examining causation and studying other variables like anticipation, marketing, pricing and release timing (Sliwinski 2010).

To conclude it seems that there is no pure correlation between third party reviews and sales in the video game industry. This may be true, but using a full data sample of all video games and not taking into account variables like platform, anticipation, production budget and other factors, limits the research, as there may be a correlation on certain types of video games.

Studies on consumer reviews do show there are differences in the product group for entertainment products and especially video games. A study by Chevalier & Mayzlin (2006) show that consumer reviews on Amazon.com play a significant role in sales, while a similar study with a similar data set by Chen et. al. (2004) show no direct influence on sales. This seems to be trend according to Zhu & Zhang (2010), who studied the influence of online consumer reviews on product sales, using data from the video game industry. Their findings show that consumer reviews are more influential for less popular games, but had less influence on very popular products. They concluded that the effect of consumer reviews, on sales, differentiate within the same product group and are hard to measure without extra variables, which may explain previous conflicting research.

3.1.2 Investor behavior and prediction

For an investor it does not really matter if a third party review is a predictor or influencer, as long as it might convey information about future sales. The research indicates that sales are correlated with third party reviews in the movie industry (Eliashberg and Shugan 1997; Reddy, Swaminathan and Motley 1998; Reinstein and Snyder 2005) but no solid evidence has been presented in the video game industry. Nonetheless this does not mean investors will not react to third party reviews in the video game industry.

Video games are a type of product that is very difficult to predict future success on in advance. Every year almost 50 percent of all new products fail, making video game publishers a very uncertain venture for a potential investor (Sivadas and Dwyer 2000). Even with extensive market research new products fail at a high rate (Crawford 1977). When predicting, investors will access all information that is available, and third party reviews will be natural information sources as sales information often is limited or comes later than the reviews (Banerjee 2006).

Research show that people, when facing uncertainty, will not use the calculus of chance or the theory of prediction. They will rather rely on representativeness, predicting the outcome that appears most representative of the evidence (Kahneman and Tversky 1973). It also shows that the predication is insensitive to both the prior probability of the outcome and the reliability of the evidence. As third party reviews can predict movie box office numbers it would not be a big assumption to think investors would react to this information in the video game industry as well, before they get news about sales. This is especially true regarding video games that are non-sequels, as it will be hard to predict the success without any prior sale numbers on previous titles.

Even though investors might react positive on good reviews, they might not react negatively on bad reviews. A study by Nofsinger (2001) shows that individual investors are more likely to both buy and sell stocks because of good news, and when there is bad news they don't trade as much. Institutions however trade on both types of news.

3.1.3 Third party reviews and firm value

If third party reviews can predict sales, it would not be a grasp to say it can make investors react to the news. In our literal review, we have found very little research done on third party reviews and the effect on firm value. Some exceptions, most notable, are Chen, Liu & Zhang's (2010) contribution where they study professional movie reviews and the effect on firm value using an event study. They studied 14 news outlets and 7 movie studios, and measured the absolute and relative effect of movie reviews on firm value using the review date as the event window. Their findings showed that the absolute effect, the Metacritic score, had no significant results on firm value, but the relative effect, of one review from a previous one, had a significant effect. The absolute effect of the review is hard to measure in the movie industry as the movie studio are only represents a fraction of the revenue for the parent companies.

Another event study, by Tellis & Johnson (2007), shows that the reviewed quality, of a newly released technological products, has a strong immediate effect on abnormal returns and the firm value. They used a multitude of companies and products that had been reviewed in “The Wall Street Journal” by a specific technology writer, Walter Mossberg. They used absolute effect from the review and still found a significant abnormal return on the event day of the review showing that investors react and utilize the new information a third party review has to make new estimates for the firm value.

The study by Tellis and Johnson (2007) and Chen et al. (2012) is the only research that studies the effect a review can have on an investor. They both find a significant result that reviews do influence investors when predicting sales within the technology- and movie industry. It does not mean this is transferred to other industries, like the video game industry. It may however be an indication on how investors will react to third party reviews in other industries as well, especially as the movie industry and video game industry carry many similarities.

3.2 Consumer reviews and other forms of information

While third party reviews and consumer reviews might convey information about a product´s selling power, other tools have emerged from video game reviewers. Gamespot and IGN, two of the largest video game reviewers, have both developed Gamespot Trax and IGN GamerMetrics to measure the opinions of millions of potential video game buyers ahead of release (Banerjee 2006). Tools like this might remove some of the power a third party review has on an investor, as it is possible to measure demand, in a new way that was not possible before. However, there have yet to be done any public research on tools that measures opinions at a large scale.

4 Methodology

In this section we explain the methodology used to test our hypotheses. It is a general approach to the methodology. We explain our choices further in section 5.

4.1 Event studies

An event study is a good method to test if there is a correlation between third party reviews and a firm's value. The reviews come out, create an event, and we can see if there is a movement in the stock price, inside this event window, based on the information that is available in the market. Event studies are a tried and true form of empirical research within the financial research community. It has been refined over the last 70 years, but you cannot research events without referring to MacKinlay's (1997) review of the key features of an event study, and its applications to economics and finance. He reviews and summarizes the different methods and approaches for an event study and we will use his research as our cornerstone going forward.

4.2 Event study timeline

An event study is divided in a three-part timeline (MacKinlay 1997), the estimation window, the event window and the post event window, as shown in figure 4. This makes it easier to look at a specific event and compare it to a timeline that is excluded from the event.

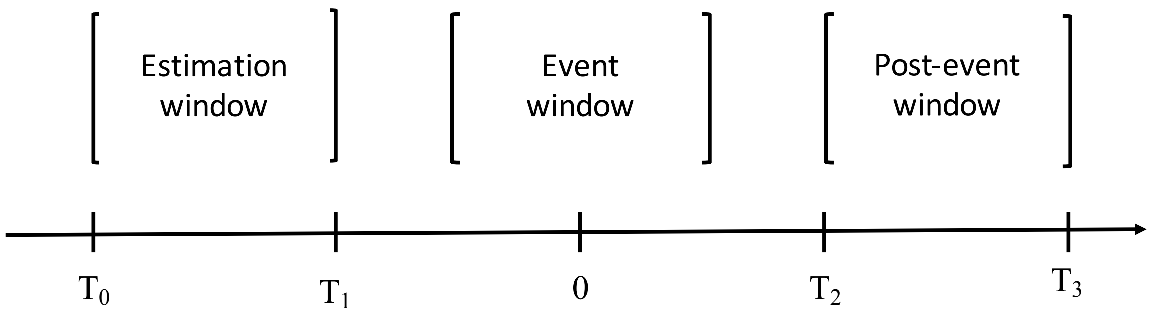


Figure 4: Timeline of an event study (MacKinlay 1997).

In a design where you separate the estimation window and the event window, the normal returns will not capture the event impact. The study will only capture the abnormal returns in the event window. On occasion, the post-event window is included in the estimation window to increase the robustness of the normal return (MacKinlay 1997).

4.3 Event and estimation window

The event window is the first to be defined in an event study. What event are we looking at? How long is the period we are examining? The event window should cover the event day we are looking at if it is a specific time, but it is quite common to include more days around the event to examine any information leakage. At least one day before and one day after the event (MacKinlay 1997). Adding more days to the window gives us the opportunity to look at the abnormal return before and after the review has been released. However, confounding events are easier to avoid if we have a smaller event window rather than a large. The estimation window on the other hand is over a large period of time before the event has occurred. Often up to a year of trading days before the event. A large estimation window sample will give a reasonable estimation for the variance so the hypothesis can be tested (MacKinlay 1997).

4.4 Abnormal return

MacKinlay (1997) defines abnormal return as the actual return of the firm minus the normal return of the firm over the event window.

$$AR_{i\tau} = R_{i\tau} - E(R_{i\tau}|X_\tau) \tag{1}$$

In equation (1), i is the firm, τ is the event date, and $AR_{i\tau}$, $R_{i\tau}$, $E(R_{i\tau}|X_\tau)$ are the abnormal, actual and normal return. When calculating the normal return, the actual return is given from the data selection, but there are multiple ways to model the normal return. The two most common models for the normal returns is constant mean return model and the market model, but there is also other forms of statistical and economic models (MacKinlay 1997). We will only explain the two most common forms here.

4.4.1 Constant mean return model

The constant mean model is the simplest of the normal return models. It assumes that the mean return of any given asset or security is constant through time (MacKinlay 1997).

$$R_{i\tau} = \mu_i + \zeta_{i\tau} \qquad E(\zeta_{i\tau}) = 0 \qquad var(\zeta_{i\tau}) = \sigma^2_{\zeta_i} \tag{2}$$

In equation (2) the normal return, $R_{i\tau}$, is calculated from μ, the mean return for asset i. Variable $\zeta_{i\tau}$ is the disturbance term and is expected to be zero. The abnormal return is then the difference between the return in the event window and the mean return calculated in the estimation window. Even though the model is very simple, the results it gives is often very similar to other models according to event studies done on stock return data (Brown and Warner 1980, 1985).

4.4.2 Market model

The market model on the other hand is a bit more complex, and uses statistics to compare the return of a stock with a market portfolio as a benchmark. Just like an investor would look at a benchmark to see if a stock, or portfolio of stocks, has performed better or worse than the market or its peers.

$$R_{i\tau} = \alpha_i + \beta_i R_{m\tau} + \varepsilon_{i\tau} \qquad E(\varepsilon_{i\tau} = 0) \qquad var(\varepsilon_{i\tau}) = \sigma^2_{\varepsilon_i} \tag{3}$$

In equation (3) $R_{i\tau}$and $R_{m\tau}$ are the returns from the stock and the market in period τ, the event study window. α, β and $\sigma^2$are the parameters of the market model and are calculated from the return of the market and the return on the stock. This is a theoretical improvement over the constant mean market model. The portion of the stocks return that is attributed to the variance of the market is efficiently removed. It may be easier to detect events when you remove this portion from the stocks return. The benefit you get from this model will depend on the market benchmark you choose and the R^2 of the market model regression. The higher R^2 you get the greater the variance reduction (MacKinlay 1997).

Brown and Warner (1980) found that the market model is a simple, well specified, and a powerful methodology when used on monthly data under a wide variety of conditions. They later confirmed the models use with daily data as well, making the market model a very good model for calculating abnormal return (Brown and Warner 1985).

4.5 Aggregating the abnormal returns

To see the cumulative effect in the event window, the abnormal return must be aggregated, both through time and through securities (MacKinlay 1997).

Aggregating the abnormal return, from equation (1), through securities i, gives us the Aggregated Abnormal Return (AAR) for all the securities in time period τ.

$$AAR_\tau = \frac{1}{N}\sum_{i=\tau}^{N} AR_{i\tau} \tag{4}$$

The AAR value shows how much aggregated abnormal return the stock has had, compared to the normal return it should have had in the same window, without the event. As AAR only represents one period, τ, it is needed to pool them from τ_1 to τ_2. The event window decides how many periods that are needed to test our hypothesis, using equation (5) to generate the Cumulative Average Abnormal Return (CAAR).

$$CAAR(\tau_1, \tau_2) = \sum_{\tau=\tau_1}^{\tau_2} AAR_\tau \tag{5}$$

The sample needs to be free of clustering, one event overlapping another in time, and confounding events that happen at the same time. If the abnormal return is related statistically to other events, that are common in the stock portfolio, we can conclude that the event is the cause of the abnormal return and analyze the results (Tellis and Johnson 2007). By aggregating both the abnormal return each event day, cumulative average abnormal return, you can see the event as a whole over the different securities you are testing.

4.6 Statistical properties of Abnormal Return

We have to assume normal statistical assumptions that asset returns are jointly multivariate, independent and identically distributed through time but the assumptions does generally not lead to problems as they are empirically reasonable (MacKinlay 1997).

Given null distribution and no clustering of the event windows, tests of the null hypothesis can be conducted.

Using the average variance for the estimation window given by $var(AAR_\tau)$ (MacKinlay 1997).

$$var(AAR_\tau) = \frac{1}{N^2}\sum_{i=1}^{N}\sigma_{\varepsilon_i}^2 \; . \tag{6}$$

Aggregating the results from equation (6) we get the variance for the CAAR.

$$var\big(CAAR(\tau_1, \tau_2)\big) = \sum_{\tau=\tau_1}^{\tau_2} var(AAR_\tau) \tag{7}$$

Using equation (5) and (7) a test on the hypothesis can be conducted.

$$\theta_1 = \frac{CAAR(\tau_1,\tau_2)}{var(CAAR(\tau_1,\tau_2))^{1/2}} \sim N(0,1) \tag{8}$$

4.6.1 The difference between categories

With the same equations we can also conduct tests between the good third party reviews and bad third party reviews. Using equation (5) and (7) to calculate the difference in CAAR and the standard error between the two types of third party reviews.

$$Good\ CAAR - Bad\ CAAR = CAAR(\tau_1, \tau_2)_{good} - CAAR(\tau_1, \tau_2)_{bad} \tag{9}$$

$$\sigma_{Good-Bad} = \sqrt{\frac{var\big(CAAR(\tau_1,\tau_2)\big)_{good}^2}{N_{good}} + \frac{var\big(CAAR(\tau_1,\tau_2)\big)_{bad}^2}{N_{bad}}}\sqrt{N_{event\ window}} \tag{10}$$

Equation (9) and (10) is then used to test if a good third party review gives a significant higher CAAR than a bad third party review.

$$\theta_1 = \frac{Good\ CAAR - Bad\ CAAR}{\sigma_{Good-Bad}} \tag{11}$$

Using equation (8) and (11) with the results from the event study, given the normal assumptions and that the data selected is representative of the population, a statistical conclusion can be drawn.

4.7 Cross sectional regression

When multiple hypotheses exist for the source of the abnormal return an event study will not uncover the relationships by itself. If many characteristics might explain the abnormal return a cross sectional regression is the most common approach to examine any associations (MacKinlay 1997). Ordinary Least Squares regression can be used assuming homoscedasticity and no cross-sectional correlations in the zero mean disturbance term.

$$CAR_i = \beta_0 + \beta_1 x_i + \cdots + \beta_M x_{Mi} + \varepsilon_i \tag{12}$$

Equation (12) shows the regression model where CAR_i is the i^{th} cumulative abnormal return observation and M is the characteristics being tested.

5 Data

Two event studies are used to see if third party reviews affect firm value. The first is using the release of the game and Metacritic reception, and the other is using the first important review. We perform the study with the same data sample for both studies. We first found video game publishers to study, and then extracted all video games from each publisher from Metacritic.com. As described in the introduction Metacritic has become a world leader, on third party reviews, for entertainment products in the last few years. It has a big influence on the entertainment industry, and is regarded as the leading benchmark in the entertainment industry (Macdonald 2012; Grubb 2013b; Gilbert 2013). Metacritic has been used in multiple research articles in different fields (Elberse and Anand 2007; Wiles and Danielova 2009), and most notably in an article by Chen, Liu and Zhang (2012) on how third-party reviews affect firm value in the movie industry. Metacritic has a comprehensive list of all games released by publishers since 2000. The video games have aggregated review scores, calculated by Metacritic, and information on individual reviews by third parties. We collected the video game title, release date, video game platform, and Metacritic score from Metacritic.

Information on specific titles was collected from Giantbomb, regarded as one of the most comprehensive and detailed databases for video games (Danielsen 2015). We collected information on what genre the video game was part of, if it was a sequel, a rerelease, and if it was part of a franchise from Giantbomb. Age recommendation was collected from Entertainment Software Rating Board, or ESRB. IGN and Gamespot, the two biggest third party reviewers, were used for third party review scores and review dates. Stock and index returns were collected from Thomson Reuters Eikon. A short summary of the collected data is presented in Table 3 and the data is further detailed and explained in this section.

Table 3: Data collected from all sources

Data source	Data collected
Metacritic.com	Video games, platform, release dates, score, number of reviews
Giantbomb.com	Variables to describe each video game
ESRB.com	Age recommendations
IGN.com	Review date and review score
Gamespot.com	Review date and review score
Thomson Reuters Eikon	Daily return for company stocks and indexes

Note: The table shows the data sources we used in this study and the information we extracted from each sources.

5.1 Publishers and stock returns

For our thesis, and to measure firm value changes, we needed daily stock return from the firms we are measuring, and review data from all the games released from the specific publishers. There are a multitude of video game publishers, and Metacritic.com lists over 2800 different publishers on their site (Metacritic 2016a). The publishers listed has released or co-released a game at some point in time. Many of the firms might today be bankrupt, bought by another firm or simply stopped publishing video games. Metacritic rates a publisher as major if they release at least 14 unique titles within one year, and a mid-size publisher if they release 6 to 13 unique titles within a year. In 2015 there were 9 major publishers, and 23 mid-size publishers (Dietz 2016). We used Metacritic's major and mid-size publisher lists, for the last three years, to find publishers for our study.

Not all of the publishers can be used for an event study. The companies have to be listed on a stock exchange or owned by a company listed. Many publishers are owned by parent companies that are private. We found 16 publicly traded companies that publish video games directly or through subsidiaries. Mobile video game publishers have been excluded, as they use a free-to-play business model. Publishers using a free-to-play business model are giving products away for free, and then they generate revenue on smaller transaction for playing the game. In the traditional video game business-model publishers sell complete products, to consumers, often for a one-time fee. See Appendix 1 for a full summary of all 16 publishers.

The companies for our study had to have a majority of their revenues from video game sales. Many publishers generate revenue from other sources than video games. Sony Corporation, Microsoft, Sega Sammy Holdings, Nintendo, Hasbro and Bandai Namco Holdings, are major video game publishers, but only a small percent of their revenue stem from actual video game sales[*]. This can make it hard to see any effect a review has on their firm value, as news for other segments of the business might interfere or make noise.

Regional differences are also important. Metacritic only tracks western media third party reviewers, while Koei Tecmo generate most of its revenue in Asia[*]. The markets are very different and most games are not released in the western market. Some are but they are

[*] According to the latest annual reports of the respective companies.

considered niche products and may not be dependent on third party reviews, as they do not have the same commercial appeal.

How the publisher generates the video game revenue is also important. Many generate a large portion of its revenue through subscription based video games, free-to-play video games, or other merchandise and hardware for video game arcade machines. This includes Konami, Activision Blizzard, Square Enix and Funcom*.

We also wanted the companies to be active for many years, this way we could get a good selection of events back in time. Atari, the publisher of the first video game, Pong, has not released any video games for several years, and is now a holding company for intellectual property and is mostly rereleasing video games*.

Table 4: Publishers in data sample

Company name	Stock ticker	Mkt. cap in $ (Billion)
Electronic Arts Inc.	NASDAQ: EA	$21,15
Ubisoft Entertainment SA	EPA: UBI	$3,55
Capcom Co. Ltd.	TYO: 9697	$1,65
Take Two Interactive Software Inc.	NASDAQ: TTWO	$3,33

Note: The table shows the four publishers used in this study, what stock exchange they listed and their market capitalization.

Using the selection criteria we used Electronic Arts, Ubisoft, Capcom, and Take Two Interactive in this study, as seen in Table 4. This is a good sample both geographically and in size.

5.2 Video game sample

From Metacritic we extracted 4530 video games from the selected publishers. However, not all entries are events applicable to our study. Metacritic has one entry for each platform a game has, and entries for all downloadable content and extra features. To get an exact list over unique releases that caused an event we had to remove the entries that have no newsworthy impact according to the efficient market hypothesis, and then merge the entries to unique video game releases. We removed all entries that had no Metacritic score, hand held

* According to the latest annual reports, of the respective companies.

video games, digital releases, re-releases, games that had no news coverage, and video games that were released in a period with no stock data. A summary is located in Table 5, and a detailed examination of the data we removed is given in the following sections.

Table 5: Video game sample summary

	Capcom	Ubisoft	Take two Interactive	Electronic Arts	Total
Video games	704	1538	856	1432	4530
No Metacritic score	-269	-715	-399	-674	-2057
Handheld video games	-96	-152	-59	-149	-456
Digital releases	-100	-123	-75	-149	-447
Re-releases	-79	-122	-16	-52	-269
Merging platforms	-40	-186	-159	-193	-578
No news coverage	-7	-39	-19	-16	-81
Limited stock data	-23	-2	-3	-3	-31
Total events	90	199	126	196	611

Note: The table shows the video games we extracted from Metacritic, and the different methods for removing video games that cannot be used in our sample.

5.2.1 No Metacritic score

First we removed 2057 video games, as they had no Metacritic score. An entry without a Metacritic score does not have the four reviews necessary for an average score. This is because some video games are not full games and does not have any reviews. This can be downloadable content for video games, or games that are so small they do not make any headway in the news, but still get added to Metacritic. It can also be titles that have been cancelled under development, but added on Metacritic because the publisher has announced the title.

5.2.2 Handheld video games

We also removed all handheld video game titles, 456 video games. Ubisoft development cost for the DS handheld video game system is between €0,5 and €1 million, while the development cost for a title on the major consoles is between €12 to €18 million (Boyer 2008). Take-Two Interactive has commented that development costs for their home-console titles range between $10 and $60 million, while their biggest titles exceed $60 million (Take-Two 2013). Electronic Arts often spend two or three times the amount on marketing and advertising for a game compared to the actual development cost as well (Takahashi 2009).

Information about development cost is often shrouded in mystery as investor reports and annual reports bundle all costs together. There are very few accounts of publishers talking about this, besides the excerpts used in this paper. Investors will not be interested in video games that do not have a big impact on potential earnings. If a handheld game gets a negative review it will not have the same impact as a major title for Xbox One as the production cost and potential revenue are very different. After the removal we were left with video games for Playstation, Playstation 2, Playstation 3, Playstation 4, Xbox, Xbox 360, Xbox One, Nintendo 64, Gamecube, Wii, Wii U and Dreamcast.

5.2.3 Digital releases

A video game can be published as a physical copy, as a disk-based medium, or digitally directly to the consoles. In a survey by the NPD group, a leading market research company, 74 percent of consumers still prefer a physical copy of a video game (Serrels 2014). Electronic Arts reported I 2015 that they earned \$660 million on physical sales and \$550 million on digital sales in 2014 (Beres 2015). A major video game will always be published physically and digitally, as most of the video games are still sold as physical copies, even though publishing digitally is more profitable, as publishers can use their own sale channels. There is, however, a large group of video games that only get published through digital channels. These are smaller video games, often from small independent developers. They will be reviewed, but they will not have a big impact on earnings. Similar to hand held games they have a small production cost and smaller potential earnings, because of potential sales and a sale price that is lower than a normal title. This is a good way to remove smaller games that does not have the normal risk and reward for a publisher and the possible gain for an investor. We removed 447 entries that were only released digitally and not in a physical format.

5.2.4 Re-releases

Many video games releases are re-releases of old games. It may be remakes suited for high definition televisions, or a “Game of the year edition” with all the previous released downloadable content for a game. Most of the games will have very few changes, this is because the original game was popular when it was first released, and there are low costs in releasing a video game with very little changes from the original product. The third party reviews are not defined as news as it has been reviewed earlier. There is no sale information on re-releases from any publishers, but we assume they will not have any impact on our study, and third party reviews will not bring any new information to the investors. We removed 269 entries from our data that were rereleases of old video games.

5.2.5 Merging platforms

Video games are very often divided by platform, because there can be discrepancies between them. The video games are often released at the same time on different platforms, but can also be released a year or more later than the first platform release. Most news outlets like IGN, Gamespot and Metacritic differentiate between them and give different platforms different reviews dependent on how big the discrepancies between the platforms are. One video game can have four different reviews under certain circumstances.

If a video game has multiple releases on different platforms at different dates, we only want to use the first release and first review, all other releases after this does not bring any new information to the market. If a video game has multiple platform releases, on the same day, we used an average of the Metacritic score for the different systems. This way we only have unique video game titles with the first available information for investors.

Merging the remaining 1301 entries into unique titles gave us 723 video game titles we could use to create events.

5.2.6 No news coverage

Important video games will get a lot of attention from the different news outlets, and will be reviewed by multiple third party reviewers. The publishers actively seek out this attention, sending review copies of video games to news outlets. Publishers also release small games that do not need any marketing, and will only be reviewed by a handful of third parties. The games are just not interesting for the news outlets and not expected to sell many copies. The average number of third party reviews, for a video game, is between 40 and 50 on Metacritic according to our data sample. We removed 80 titles that were reviewed by less than 10 third parties. Then we could be sure we removed any event that did not bring any relevant news to the investor about potential income or loss. Total reviews from 10 to 20 are still considered very low and the video games might not be important, but we still think they should be examined.

5.2.7 Limited stock data

We wanted a large sample covering far back in time and to be consistent. There is very little information and few video game titles before the year 2000 on Metacritic. The timeframe was set from January 1 2000 to December 31 2015, and all releases before this was removed. Capcom only had stock returns from October 26 2001 so we had to remove video games titles before this date including the estimation window timeframe. 31 titles were removed because of the two issues. At the end we end up with 611 video games.

5.3 News category type

Based on how an investor can look at the news we divided the Metacritic reception in three different news types; good, mixed and bad. The difference between a good, mixed and bad reception is not universal. Metacritic uses two different scales when referring to their scores, see Table 6, one general and one more specific (Metacritic 2016b).

Table 6: Metacritic score summary

Reception	General Meaning of Score	Movies, TV & Music	Video Games
Good reception	Universal Acclaim	81 - 100	90 - 100
	Generally Favorable Reviews	61 - 80	75 - 89
Mixed reception	Mixed or Average Reviews	40 - 60	50 - 74
Bad reception	Generally Unfavorable Reviews	20 - 39	20 - 49
	Overwhelming Dislike	0 - 19	0 - 19

Note: The table is a representation on how Metacritic.com shows their ratings for different entertainment products.

For video games a critical score of 75 to 100 is considered a good score, 50 to 74 a mixed score, and 0 to 49 a bad score. They are however divided in five different receptions.

This does not fully represent the way companies and probably investors look at the scores. Obsidian Entertainment, a well known and successful developer, disclosed that a bonus payment would be credited by Bethesda, their publisher, if they managed to get a Metacritic score of 85 or higher (Gilbert 2012). Nintendo, another publisher also emphasize on 85 and above through a marketing campaign showing all the “great video games” they had released. No video game under 85 was used in the campaign. A non disclosed publisher has even set 95 as minimum limit to pay out royalties to a developer (Schreier 2015). Funcom, a Norwegian developer, even blamed Metacritic for a stock price slide. They referred to their score of 72 as a low score (Carter 2012). There is clearly a mismatch between the ratings on Metacritic and what is considered good and bad in the industry.

Another problem is the actual reviewing process of a video games. Video games are often defined as art (Stuart 2014), like movies, books or music. This might be true as defining art itself is a controversy (Adajian 2012). Video games however has one key difference in its reviews compared to the other mediums of entertainment. They are judged on technology. A book would not be scored on what kind of paper it was printed on, and a movie can get an Academy Award even if it is shot in black and white. Third party reviews for video games often factor in how the graphics are rendered and how the video games are coded. This is not present in books, movies, or music. This creates a problem in that a sequel to a game often will be a better game and get a higher score, although this isn’t the case for all video games. Making sequels upon sequels will saturate the market and scores will eventually drop without big shifts in the recipe (Fischberg 2012). It will however be true for many video games and inflate the scores.

Metacritic has taken this into account and made the scores different from other media, but we argue that it is not adequate enough. IGN has a scoring average of 75 in our sample, and it is hard to believe that investors expect 75 to be a general good score for a video game, when the average is that high. With what we know of the mentioned publishers, on the score they want to achieve or promote, and the research we did when collecting reviews, we use a scale that has more extreme values compared to Metacritic ’s scale. See Table 7 for a summary.

Table 7: Event study score summary

Reception/review	Review Score
Good	85-100
Mixed	70-84
Bad	0-69

Note: Our scoring scale for third party reviews

We suggest that this will give a better scale on how an investor will interpret third party reviews, and give us an absolute good and absolute bad review scoring scale.

5.4 Video games summary

Our sample ended up with 611 video games and Metacritic scores, as seen in Table 8. This is a very good sample and all the released video games for the publishers in the timeframe, excluding the type of games mentioned in section 5.2.

Table 8: Release and Metacritic summary

		Reception			Metacritic score summary			
Company	Total	Good	Mixed	Bad	Mean	Min	Max	SD
Capcom	90	11 (12%)	35 (39%)	44 (49%)	70	38	96	12
Electronic Arts	196	42 (21%)	102 (52%)	52 (27%)	75	38	95	11
Ubisoft	199	26 (13%)	83 (42%)	90 (45%)	69	23	98	13
Take-two Interactive	126	27 (22%)	57 (45%)	42 (33%)	74	37	92	14
All publishers	611	106 (17%)	277 (45%)	228 (37%)	72	23	98	13

Note: The table shows how the different video game receptions are divided between the publishers in our sample, and simple statistics on the Metacritic score for each publisher.

5.5 Third party reviews

Collecting the video games, their release dates and their Metacritic score, is enough for studying the firm value around a video game release. For studying the third party review effect directly we needed reviews for all the video games. After finding the 12 most read video game news outlets on the internet (Ebizmba 2016), that review games, we did a small study on third party reviews for Ubisoft and Capcom. We excluded normal newspapers and magazines as very few newspapers review games, and game magazines are issued, at most once a month. See Appendix 2 for a summary of all the third party reviewers.

IGN and Gamespot is by far the largest news outlets and third party reviewers of video games. At Alexa.com, a global web analytics provider, they rank at 385 and 820 (2016), where 1 is the most visited website in the world, counting unique users over a 3 month average. Metacritic is at rank 1721 in comparison. The third most popular third party reviewer is Gamesradar at rank 5637. It is estimated that IGN and Gamespot have 20,5 million and 15 million unique monthly visitors (Ebizmba 2016). They have also been active the longest, they both have the most reviewed games in our period, and their reviews are on average earlier, relative to the release, then most of the others. A search result for a video game on Google.com will present IGN's rating of the game next to the Metacritic score. They are also the only third party review sites referenced by stock market analysts, as an information source (Baker 2011; Banerjee 2006). IGN and Gamespot are also weighted more than the average reviewer according to an unpublished study (Parkin 2013). Metacritic was however quick to point out that the study was flawed. It would not be difficult to believe the two biggest third party reviewers are weighted more than the average, as this is practice of Metacritic.

According to research it may seem like people are influenced differently by the different third party reviewers (Reinstein and Snyder 2005), this should also extend to investors. By this we reason that IGN and Gamespot are probably the most influential third party reviewers in this industry. That is why we only use their rating to measure the effect on firm value.

We gathered 1155 reviews for our 611 video games. Dates and scores were gathered directly from Gamespot and IGN, and we recalculated the scores to the Metacritic standard. By using the first available review from either IGN or Gamespot we can see if it has an effect on firm value. See Table 9 for a summary.

Table 9: Summary of third party reviewers

A: Activity summary

	Number of reviews			
	Total	Before release	On release	After release
IGN	590	184	118	288
Gamespot	565	98	107	360
First of the two	600	219	120	261

B: Review summary

	Review score				Review relative to release date (days)				
	Avg.	Min	Max	SD	Avg.	Min	Max	Median	SD
IGN	75	20	100	16	4	-20	125	0	12
Gamespot	71	19	100	15	6	-39	78	2	12
First of the two	74	20	100	15	2	-39	125	0	10

Note: The table shows how many reviews IGN and Gamespot have in our sample, and how they are scored. "First of the two" is the first review published by either, using an average if it happened on the same day.

Comparing the scores, we found a high correlation between the first review published by IGN or Gamespot, and the Metacritic score. The first review score correlated 88 percent with the Metacritic score and there is only a small difference on average as seen in Table 10. This means that the first reviews closely reflect the Metacritic score as well.

Table 10: The difference between Metacritic and first review from IGN or Gamespot

Average	Min	Max	Median	STD
-1,58	-22	29	-2	7,26

Note: The table shows simple statistical information for the difference in scores between the Metacritic score and first review score from either IGN or Gamespot.

5.6 Event study modeling

There are no studies on third party reviews and video games, and modeling of the study needs to be discussed. Specifically answering the following questions:

- How many days do we include in our event and estimation windows?
- How do we model the abnormal returns?
- How do we measure anticipation for a video game?
- How do we handle clusters?

We did two event studies to see the effect on firm value, so we will discuss modeling for both studies, one for the reception of the video game and one for the first important review.

5.6.1 Event window for reception of a video game

Third party reviews are always centered on the release date. The critics want to have the review published as early as possible as it will get the most amount of attention from readers in the time frame around the video game release. Reviewing games is a longer process than reviewing movies and other software reviews, as it can be time consuming playing a video game. It varies, but when it takes 90 hours (Ingenito 2014) just to play through the video game the reviews will be spread out over the release date in a larger degree than with movie reviews. Depending on when the critic was able to get a hold of the video game. See Figure 5 for an example of reviews relative to release of two video games from publishers Electronic Arts and Ubisoft.

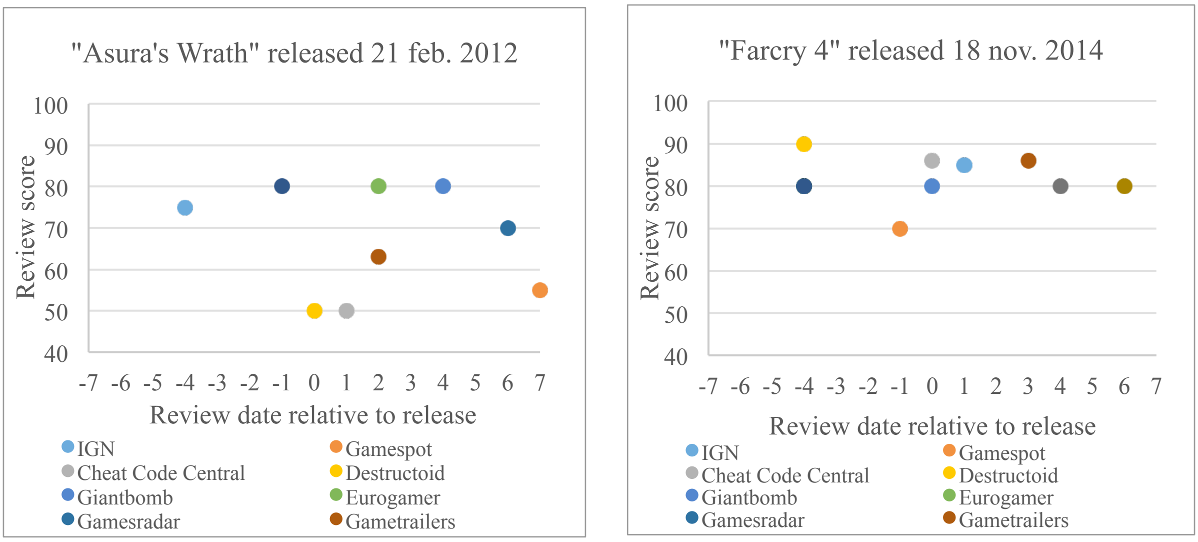


Figure 5: Video game examples for reviews and score relative to release. This figure shows when and what review score two video games received from different third party reviewers, in a period of seven days before and after release.

Based on our data sample, most of the third party reviews from IGN and Gamespot is published one week before and after the release date of a video game as seen in Figure 6. There are still some reviews in week two and three, but other sources of information should be available to investors by then. The NDP Group release monthly sales information on video games between the tenth and twentieth the following month (Blundon and Lindemann 2010; Dunning 2014; Grubb 2013a, 2015, 2016; Shea 2016; Hatfield 2007). To make sure no other sales information was released within the event window we used a 15-day event window studying the release dates, measuring seven days before and after the event. This way we also minimize the risk of additional noise and clustering of our data sample (MacKinlay 1997).

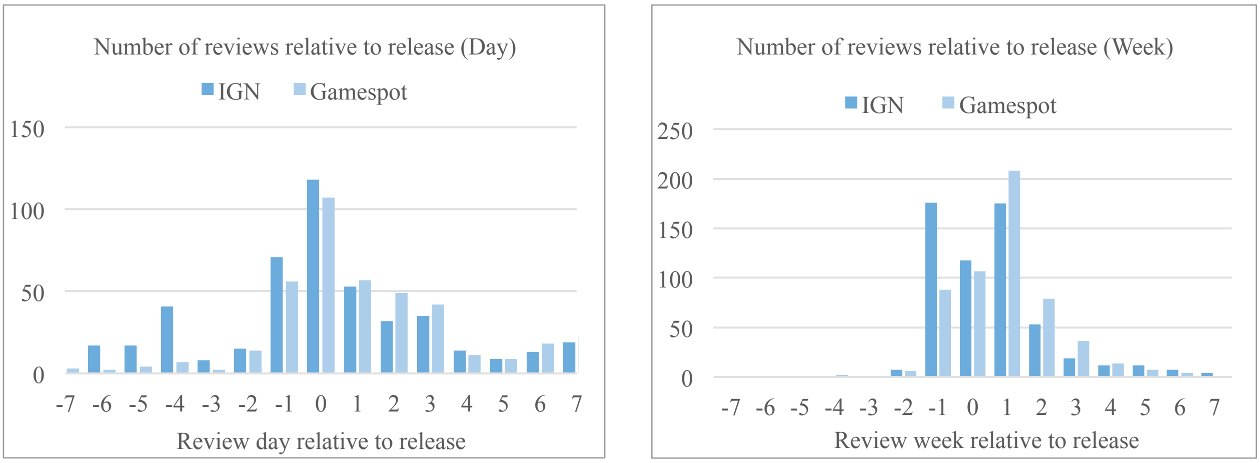


Figure 6: Distribution of reviews relative to release. This figure shows the amount of third party reviews published relative to release day, showed daily and weekly.

5.6.2 Event window for first important review

In our study on how the first important reviews affect firm value we use a 9-day event window, one day before and seven days after the event. A smaller window we captured the effect with minimal noise from other confounding events. We only use one day before, as an investor does not anticipate the review, and the changes in firm value will happen after the review date.

5.6.3 Estimation window

Our event studies used a 250 days estimation window for each event. 250 days is a full year of trading days at an average stock exchange. It was calculated eight days before the event, to ensure that the event did not contaminate the estimation period. This was done to both studies. Using 250 days ensures that the estimation for normal return is accurate, even though other video games were released from the same publisher in the estimation windows. Figure 7 shows the timeline for both our event studies.

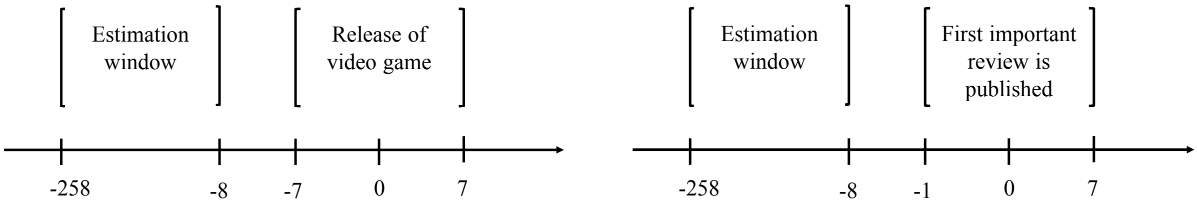


Figure 7: Time lines used in our event studies. This figure shows the two different time lines used for reception on release day and first important review.

5.6.4 Abnormal return model

As the market model is the most common model (MacKinlay 1997), to measure abnormal return, and presents a possible improvement over the constant mean model, it will be our main model for both studies.

The video game publishers are often volatile companies where the stock movements can be huge even though the market is stable, and there are no indexes with just video game publishers or other similar companies. We used the most common benchmarks for the different stock exchanges. S&P 500 for the NASDAQ based companies, Nikkei 225 for Capcom and CAC 40 for Ubisoft. They are the common benchmarks for all companies at their respective stock exchange.

We used daily stock return to give us the opportunity to examine the precise reactions of investors to third party reviews. Daily stock return data was collected from Thomson Reuters Eikon and all the stock returns are adjusted for stock splits and dividends. If the event day was on a day the stock market was closed we measured the abnormal return using the next day it was open, as traders can only react to information when it is possible to trade stocks.

5.6.5 Anticipation for a video game

We suggest that news about more anticipated games should get a larger response from an investor. Games that have been in production for a long time and had a large production budget will be more anticipated by a consumer and be more interesting to an investor. How do we measure the anticipation for a video game?

Metacritic track how many times a video game have been reviewed by a third party. We suggest that this number is a good proxy to measure anticipation. The more anticipated a game is, the more third party reviews it will get. The simple reason behind this is that third party reviewers generate most of their revenue from website traffic. If a video game is highly anticipated by potential consumers the third party reviewer will review the video game to attract traffic to their web site. This will generate a high amount of third party reviews for the most anticipated video games. Number of third party reviews on Metacritic can then be a proxy for a consumer's anticipation for a video game. If a game has a low amount of third party reviews we suggest that it will not be interesting for an investor.

5.6.6 Clustering and modeling for reception of video games

In the 611 video game release events we have gathered there are a lot of clusters, where one event overlaps another. Like for instance video games that are released on the same day or within the event window of another video game from the same publisher. Removing all clustering events is not an optimal solution as the product difference within the sample is large (Zhu and Zhang 2010). This will remove many events we believe is important to the investors.

Table 11: Number of third party reviews per video game

Avg.	Min	Max	Median	SD
38	10	98	32	21

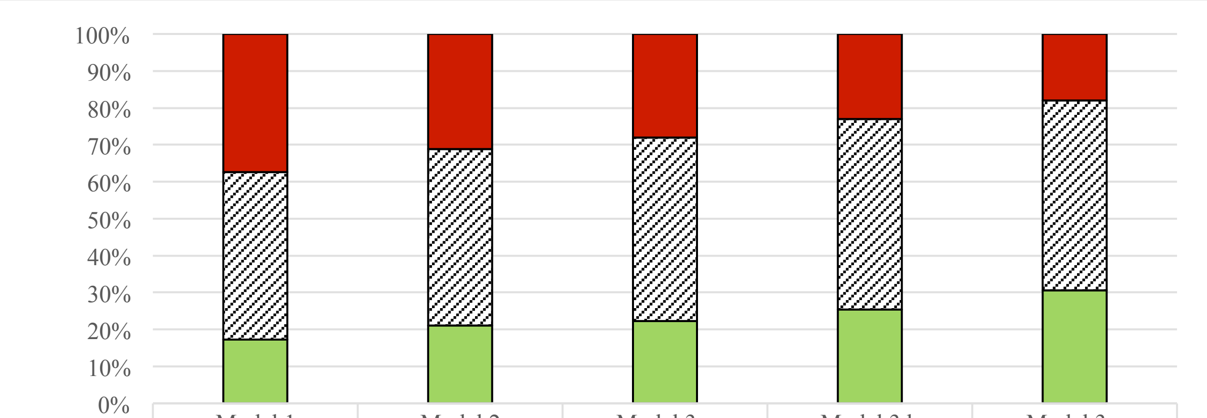
Note: The table shows simple statistical information about number of third party reviews received per video game.

We propose multiple models to address this issue, by using our proxy for anticipation. The average number of third party reviews, in our data sample, is 38, see Table 11, but it is not always the case that games under the average is less important for investors. We did multiple event studies by removing first a set number of reviews under a limit, and then removed the remaining clusters to see if anticipation is important for an investor. We assume, in the different models, that investors are impartial to video games that get fewer than a specific amount of third party reviews and should not be in the event study, as they are non-events. See Table 12 and Figure 8 for a full summary of the models. This also gives us different samples for the different models.

Table 12: Release day reception models (-7,+7 days between each event)

	Review number intervals	Description
Model 1	[10,98]	All video games, none removed.
Model 2	[10,98]	Removed all clusters
Model 3.a-c	[20-40,98]	Removed events under 20-40 reviews, then removed clusters.

Note: The review number intervals show the range between the lowest and the highest number of reviews per game within each model.



	Model 1	Model 2	Model 3.a	Model 3.b	Model 3.c
Bad	228	56	48	36	26
Mixed	277	86	85	81	74
Good	106	38	38	40	44

Figure 8: Number of events for release model 1-3.c. This figure shows total number of events and percentage share within each category.

5.6.7 Clustering and modeling for first important review

We followed the same principal for the first important review as we did with reception of a video game, removing video games that have a low amount of third party reviews before removing clusters. This ensures us that we measure third party reviews that will be of interest for the investor, and we still withhold a good amount of events, ranging over the different models. We also categorized all the reviews on the relative review date, meaning we want to see if there is a difference if the review is published before or after the release of the video game. We suggest that the third party reviews that come out before the release of the video game have more information for an investor than a review that comes out after the video game has been released. See Table 13, Table 14, Figure 9 and Figure 10 for a full summary of the models.

Table 13: First important review before and after the release of the video game (-1,+7 days between each event)

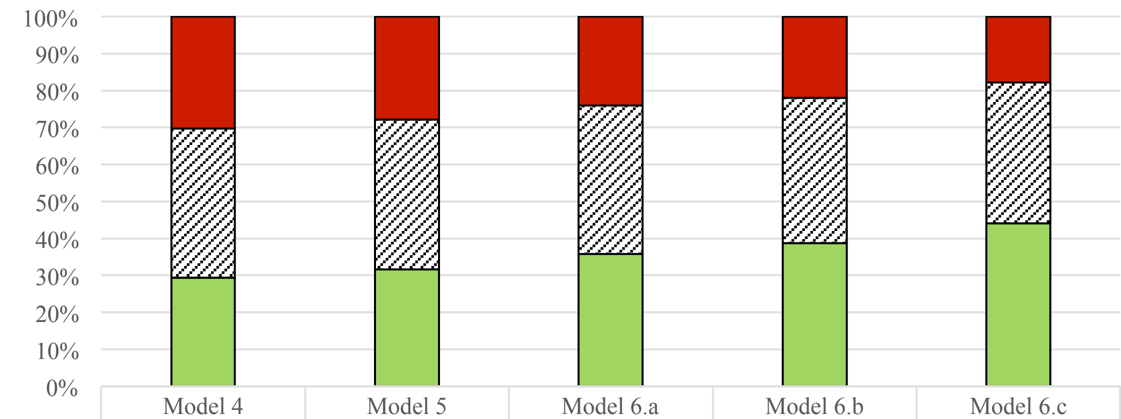
	Review number intervals	Description
Model 4	[10,98]	All video games, none removed.
Model 5	[10,98]	Removed all clusters.
Model 6.a-c	[20-40,98]	Removed events under 20-40 reviews, then removed clusters.

Note: The review number intervals show the range between the lowest and the highest number of reviews per game within each model.

Table 14: First important review before release of the video game (-1,+7 days between each event)

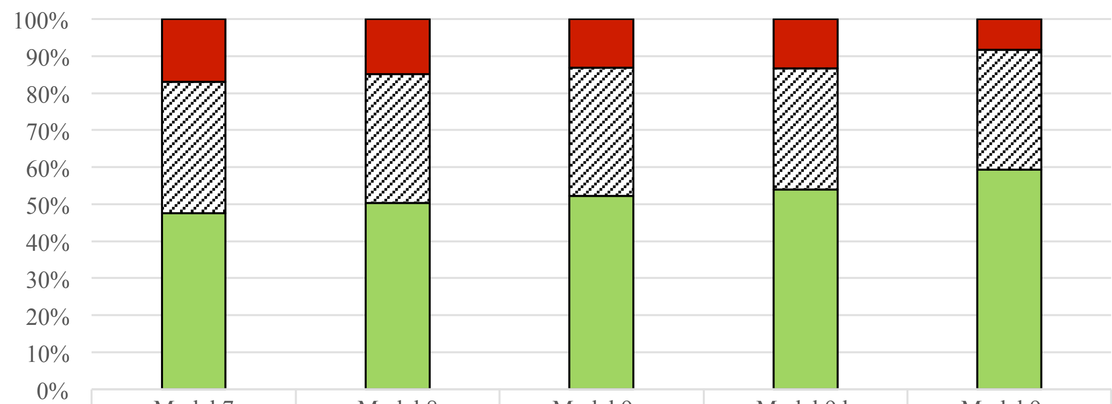
	Review number intervals	Description
Model 7	[10,98]	All video games, none removed.
Model 8	[10,98]	Removed all clusters.
Model 9.a-c	[20-40,98]	Removed events under 20-40 reviews, then removed clusters.

Note: The review number intervals show the range between the lowest and the highest number of reviews per game within each model.



	Model 4	Model 5	Model 6.a	Model 6.b	Model 6.c
Bad	182	88	70	54	36
Mixed	242	128	117	96	77
Good	176	100	104	95	89

Figure 9: The number of events for model 4-6.c. This figure shows total number of events and precentage share within each category.



	Model 7	Model 8	Model 9.a	Model 9.b	Model 9.c
Bad	37	20	18	17	9
Mixed	78	47	48	42	35
Good	104	68	72	69	64

Figure 10: The number of events for model 7-9.c. This figure shows total number of events and precentage share within each category.

5.7 Cross-sectional regression

We suggest that Metacritic score and anticipation in the form of the total number of reviews per video game has an effect of the abnormal return. Our different models explained in section 5.6.5 and 5.6.6 will examine this to some degree. The common approach to examine this is with a cross-sectional regression of the abnormal return and the other variables of interest (MacKinlay 1997). Our regression is defined in equation (13) with META as the Metacritic score and NUMBER as the total number of reviews a video game has.

$$CAR_i = \beta_0 + \beta_1 META_i + \beta_2 NUMBER_i + \beta_3 META_i \times NUMBER_i + \ldots + \beta_M X_{Mi} + \varepsilon_i \quad (13)$$

META and NUMBER are multiplied to see the effect total number of reviews have on the Metacritic score relative to the abnormal return. The three variables constitute our base model. Variable X_{Mi}represents control variables and all public available variables that may influence an investor and the reaction on the stock market. Variables include publisher of the video game, the season the video game was released in and if it was released on multiple platforms. We also included standard video game description variables such as genre, ESRB rating, if it is a sequel or part of a franchise at release day. A full summary of all variables and data source is shown in Table 15.

Table 15: Description of variables used in cross sectional regression

Variable	Description	Source
META	Metacritic score of the video game	Metacritic.com
NUMBER	The total number of reviews per video game	Metacritic.com
PUBLISHER dummies	The company that published the video game (e.g. Capcom, Ubisoft, Take-Two, EA)	Metacritic.com
PLATFORM	Whether a video game is released on one or multiple video game consoles.	Metacritic.com
SEQUEL	Whether a video game is sequel to another video game	Giantbomb.com
FRANCHISE	Whether a video game is part of a franchise at release	Giantbomb.com
GENRE dummies	Dummies for type of genre (e.g. Action, Sports)	Giantbomb.com
ESRB dummies	Dummies for ESRB ratings (e.g. Mature, Teen)	ESRB.com
SEASON dummies	Dummies for the season a video game is released (e.g. Spring, Summer, Fall)	Metacritic.com

Note: The different dummy variables are mentioned in the description. Observations that do not have one of the mentioned dummy variables are calculated as a base value and included in the intercept.

6 Empirical results

In this section we present the results of our studies. We examine the effect of the reception for a video game and the effect the first important review has on firm value. A cross-sectional regression is also presented to explain how independent variables, like age recommendation and genre, explain the cumulative abnormal return.

Good and Bad category are tested with one-tailed t-test in all our models, based on the assumption that a good third party review generate a positive abnormal return, $CAAR_{good} > 0$, and a bad third party review generates a negative abnormal return, $CAAR_{bad} < 0$. As expected, the Mixed category shows no CAAR significantly different from zero. Results for mixed third party reviews are given in Appendix 3.

We also present the CAAR and t-test value for the difference between the good and bad category. The t-test value is calculated from a one-tailed t-test under the assumption that good category reviews generates a positive abnormal return compared to a bad category review, $CAAR_{good} - CAAR_{bad} > 0$. Model specific standard errors are presented in Appendix 5.

6.1 Event study on video game release day reception

Using the event study methodology and our data sample, of 611 events we modeled five different event studies as mentioned in section 5.6.6. Table 16, 17 and 18 presents the cumulative average abnormal return and the T-test value of the different models. Model 1 uses the full sample of 106 good receptions without removing any clusters. We find no values significantly higher than zero, as there is clustering of events. Still abnormal returns are positive in the Good category and negative in the Bad category, which is interesting when measuring 611 events that cluster.

Model 2 is a general model, as the sample is not adjusted for review numbers, only cluster free. Model 3.a-c is adjusted for number of total reviews, to exclude events that is possibly, and probably, not important to an investor. There are also two different CAAR values. One measures the CAAR from seven days before and after the event, and the other measures the CAAR from one day before the event to seven days after the event. Both CAAR's are measured from the same sample in a 15-day event window.

Table 16: Good reception on release day, 9-day and 15-day event window (-7 and -1,7)

Model	Review interval	N	CAAR (-7,7)		CAAR (-1,7)	
Model 1	[10,98]	106	0,76%	(0,74)	0,48%	(0,61)
Model 2	[10,98]	38	2,17%	(1,21)	2,44%**	(1,75)
Model 3.a	[20,98]	38	2,23%	(1,25)	2,42%**	(1,75)
Model 3.b	[30,98]	40	2,43%*	(1,44)	2,54%**	(1,95)
Model 3.c	[40,98]	44	3,01%**	(2,04)	2,45%**	(2,14)

Note: The table shows cumulative average abnormal return (CAAR) for companies that have received a good reception for a video game. CAAR is measured in event windows ranging from 7 (1) days before release to 7 days after release. Each models sample varies because of the different cluster elimination methods as explained in section 5.6.6. T-stats are in parentheses. (*) means significant on a 90% confidential level, (**) are significant on a 95% level and (***) are significant on a 99% level. For model specific σ see appendix 5.

Examining the Good reception CAAR values we see that all values are positive when a video game gets a good reception. Not all the CAAR values are significantly higher than zero when measuring the full window of fifteen days. When we remove the less anticipated video games and then remove the clusters we gain a higher CAAR that is also significantly higher than zero. The more anticipated video games are, the larger and more significant the CAAR value is in the larger event window. The CAAR values that are cumulated from one day before to seven days after the release day is significant in all the models that are cluster free. This means that good third party reviews does affect firm value. It also indicates that a good reception for more anticipated video games increase the abnormal return.

Table 17: Bad reception on release day, 15-day event window (-7,7)

Model	Review interval	N	CAAR (-7,7)		CAAR (-1,7)	
Model 1	[10,98]	228	-0,58%	(-0,83)	-0,45%	(-0,83)
Model 2	[10,98]	56	-0,71%	(-0,52)	-0,17%	(-0,16)
Model 3.a	[20,98]	48	-0,49%	(-0,34)	-0,93%	(-0,84)
Model 3.b	[30,98]	36	-1,44%	(-0,96)	-2,10%**	(-1,82)
Model 3.c	[40,98]	26	-0,67%	(-0,40)	-1,67%	(-1,27)

Note: The table shows cumulative average abnormal return (CAAR) for companies that have received a bad reception for a video game. CAAR is measured in event windows ranging from 7 (1) days before release to 7 days after release. Each models sample varies because of the different cluster elimination methods as explained in section 5.6.6. T-stats are in parentheses. (*) means significant on a 90% confidential level, (**) are significant on a 95% level and (***) are significant on a 99% level. For model specific σ see Appendix 5.

A bad category reception have a negative CAAR in all our models, and in both the event windows we have measured, but a majority of the models and event windows are not significantly lower than zero. The t-test value is however more significant when we remove less anticipated video games. Model, 3.b has a significant lower CAAR than 0, and Model 3.c has to few events to correctly measure any change. It does indicate though that a bad reception affects firm value when the video games are more anticipated. However, we cannot say that a bad reception has a significant effect on firm value.

Table 18: Difference between Good and Bad reception, 15-day event window (-7,7)

Model	Review interval	N	Difference (-7,7)		Difference (-1,7)	
Model 1	[10,98]	334	1,34%	(1,08)	0,93%	(0,98)
Model 2	[10,98]	94	2,88%	(1,28)	2,61%*	(1,49)
Model 3.a	[20,98]	86	2,72%	(1,19)	3,35%**	(1,90)
Model 3.b	[30,98]	76	3,87%**	(1,72)	4,65%***	(2,66)
Model 3.c	[40,98]	70	3,68%*	(1,63)	4,12%**	(2,36)

Note: The table shows the difference between cumulative average abnormal return (CAAR) for category good and bad receptions. CAAR is measured in event windows ranging from 7 (1) days before release to 7 days after release. Each models sample varies because of the different cluster elimination methods as explained in section 5.6.6. T-stats are in parentheses. (*) means significant on a 90% confidential level, (**) are significant on a 95% level and (***) are significant on a 99% level. For model specific σ see Appendix 5.

The difference between a good reception and a bad reception is also significant in all models that are free of clusters. The good reception CAAR's is significantly higher than a bad reception CAAR's. As we remove less anticipated video games, the CAAR is higher and more significant.

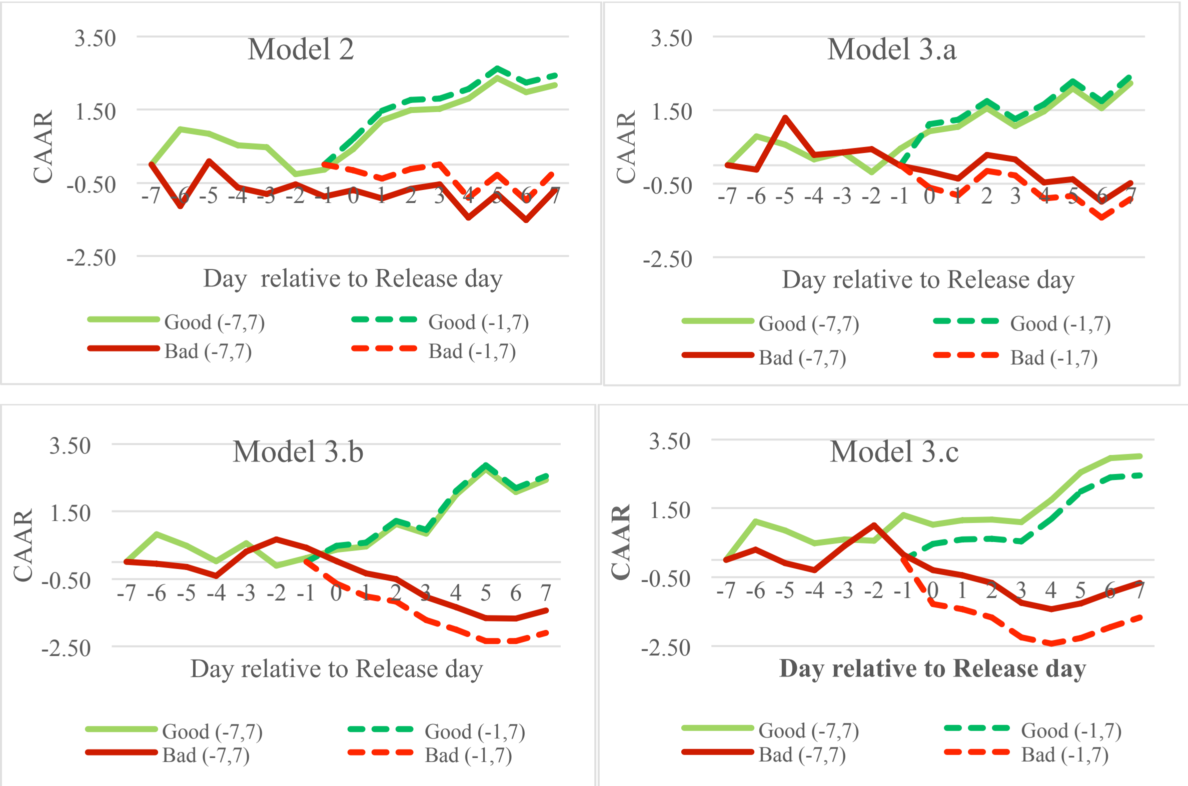


Figure 11: Daily CAAR over 15-day event window. This figure shows daily changes in CAAR for models without clusters, starting 7(1) day(s) before release, and ending 7 days after release

Figure 11 shows how the average abnormal return is aggregated during the event window. We can see a clear difference between a good and bad reception. According to Figure 11 there seems to be a weak reaction from investors before the release of the video game. There are many reviews that are available before the release of the video game. This means that an indication of the Metacritic score is also often available before the release, but investors seem to react to the reception closer to the release of the video game. This means that we can reduce our event window to one day before the release and seven days after the release. A shorter event window lets us include more observations in our sample for all the models, because of less clustering. We still use the same principals to model new samples, but we use a shorter event window. The new models include the same video games that we have in model 2 and 3.a-c, however because the event window is now shorter we can add events that previously overlapped another event within a few days. Table 19, 20 and 21 shows the new models and the respective CAAR and T-test values.

Table 19: Good reception on release day, 9 day event window (-1,7)

Model	Review interval	N	CAAR (-1,7)	
Model 10	[10,98]	65	1,33%*	(1,32)
Model 11.a	[20,98]	66	0,95%	(0,97)
Model 11.b	[30,98]	64	1,66%**	(1,69)
Model 11.c	[40,98]	62	1,73%**	(1,80)

Note: The table shows cumulative average abnormal return (CAAR) for companies that have received a good reception for a video game. CAAR is measured in an event window 1 day before release to 7 days after release. Each models sample varies because of the different cluster elimination methods as explained in section 5.6.6. T-stats are in parentheses. (*) means significant on a 90% confidential level, (**) are significant on a 95% level and (***) are significant on a 99% level. For model specific σ see appendix 5.

In Table 19 the CAAR is now reduced to 1,33 percent in Model 10. A reduction from 2,44 percent in model 2, but it is still significantly higher than zero. The CAAR for model 11.a, b and c, have all been reduced compared to model 3.a, b and c, but 11.b and 11.c are still significantly higher than zero. So even with more events the cumulative abnormal return is still higher than zero for a good reception video game. We also see the same effect of anticipation now. As we remove less anticipated video games, and add more anticipated video games, both the CAAR and the t-test value increase.

Table 20: Bad reception on release day, 9 day event window (-1,7)

Model	Review interval	N	CAAR (-1,7)	
Model 10	[10,98]	92	-0,08%	(-0,09)
Model 11.a	[20,98]	75	-0,58%	(-0,66)
Model 11.b	[30,98]	54	-1,70%**	(-1,76)
Model 11.c	[40,98]	34	-1,22%	(-1,00)

Note: The table shows cumulative average abnormal return (CAAR) for companies that have received a bad reception for a video game. CAAR is measured in an event window 1 day before release to 7 days after release. Each models sample varies because of the different cluster elimination methods as explained in section 5.6.6. T-stats are in parentheses. (*) means significant on a 90% confidential level, (**) are significant on a 95% level and (***) are significant on a 99% level. For model specific σ see appendix 5.

The bad reception depicted in Table 20 shows the same results as Table 17. There is only one CAAR that is significantly lower than zero, Model 11.b. The CAAR and t-test values are also lower when we remove less anticipated video games. We still only have one model that is significantly lower than zero and we cannot say that a bad reception for a video game affects firm value. Our results indicate tough that a more anticipated video game has a larger and more significant effect on firm value.

Table 21: Difference between Good and Bad reception, 9 day event window (-1,7)

Model	Review interval	N	Difference (-1,7)	
Model 10	[10,98]	157	1,40%	(1,07)
Model 11.a	[20,98]	141	1,53%	(1,16)
Model 11.b	[30,98]	118	3,36%***	(2,44)
Model 11.c	[40,98]	96	2,95%**	(1,90)

Note: The table shows the difference between cumulative average abnormal return (CAAR) for category good and bad receptions. CAAR is measured in an event window 1 day before release to 7 days after release. Each models sample varies because of the different cluster elimination methods as explained in section 5.6.6. T-stats are in parentheses. (*) means significant on a 90% confidential level, (**) are significant on a 95% level and (***) are significant on a 99% level. For model specific σ see appendix 5.

Table 21 shows that the CAAR for a good reception is now not significantly higher than a bad reception for Model 10 and Model 11.a. As we remove more and more of the less anticipated video games we get a CAAR that is significantly higher for a good reception than a bad reception.

Examining the CAAR during the event window, in Figure 12, we see that the abnormal return aggregate slowly from one day before the release to seven days after the release.

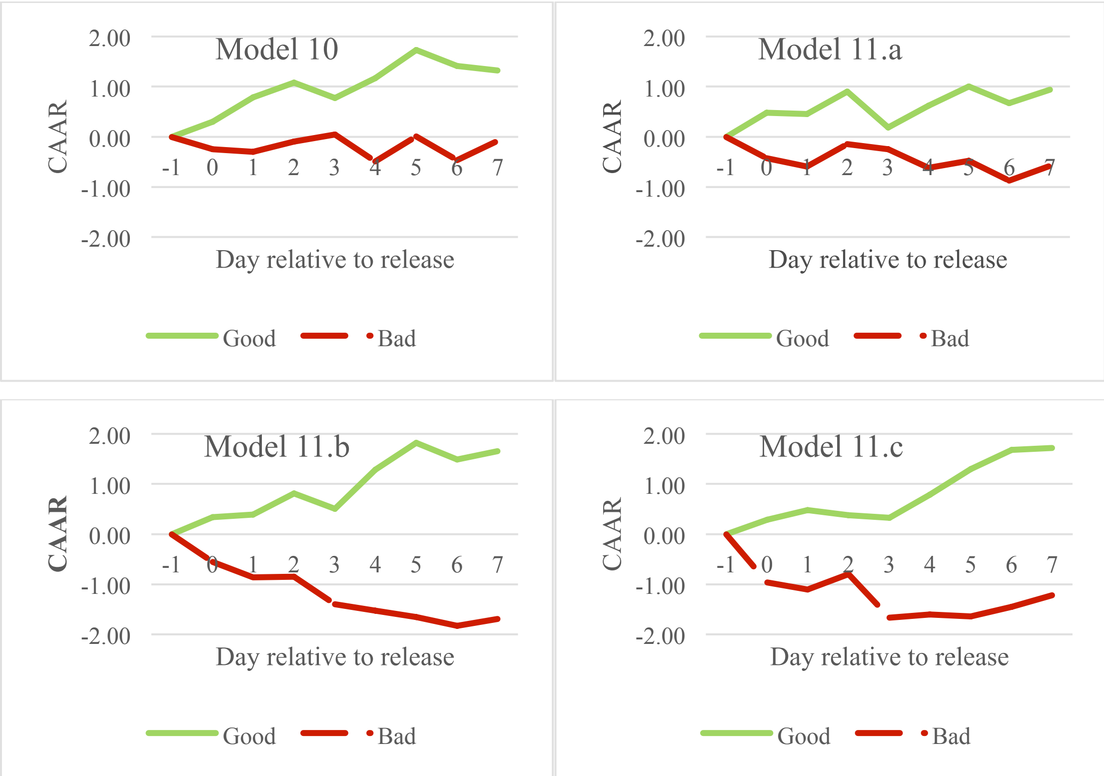


Figure 12: Daily CAAR over a 9-day event window. This figure shows the daily change in CAAR for model 10 and 11, starting 1 day before release, and ending 7 days after release.

To summarize, we see that a good reception affects firm value positively and the CAAR is significantly higher than zero. A bad reception gives a negative CAAR value, but it is not significantly lower than zero. Both types of reception seem to be affected by the anticipation of a video game, giving more significant results as we remove less anticipated video games.

6.2 Event study on first important review

We started with a sample of 600 reviews that was the first review from IGN or Gamespot. Using the models explained in section 5.7.7, to remove clusters, we examined the effect on abnormal return one day before the review and seven days after the review, to see if there was any effect on the firm value. The results are presented in two tables based on the day the review is published relative to the release for the video game. Table 22 includes first important reviews published both before and after the release. Table 23 only includes reviews published before the release of the video game.

Table 22: Event study on first review from important third party reviewer (Before, on and after the release of video game)

		Good			Bad				
Model	Review interval	N	CAAR (-1,7)		N	CAAR (-1,7)		Difference between Good and Bad	
Model 4	[10,98]	176	-0,62%	(-0,97)	182	-0,20%	(-0,33)	-0,42%	(-0,48)
Model 5	[10,98]	100	0,21%	(0,25)	88	0,23%	(0,27)	-0,02%	(-0,02)
Model 6.a	[20,98]	104	-0,01%	(-0,02)	70	0,67%	(0,76)	-0,69%	(-0,57)
Model 6.b	[30,98]	95	0,27%	(0,33)	54	0,29%	(0,30)	-0,02%	(-0,01)
Model 6.c	[40,98]	89	0,03%	(0,04)	36	-0,65%	(-0,57)	0,68%	(0,49)

Note: The table shows cumulative average abnormal return (CAAR) for companies that have received a good/bad review for a video game. CAAR is measured in an event window 1 day before release to 7 days after first review. Each models sample varies because of the different cluster elimination methods as explained in section 5.6.7. T-stats are in parentheses. (*) means significant on a 90% confidential level, (**) are significant on a 95% level and (***) are significant on a 99% level. For model specific σ see appendix 5.

There are no significant results in Table 22. The CAAR values are also not representative of reception of the video game, as found in section 6.1. Good category CAAR that is negative and bad category CAAR that is positive. The sample includes reviews that come out after the release of the game. There should be other information sources investor can use if the review is not early relative to the release, and reviews might be published the same time as confounding events. Hence, the models that only study reviews that are published before the release of the video game are more important.

Table 23: Event study on first review from important third party reviewer (Before the release of video game)

		Good			Bad				
Model	Review interval	N	CAAR (-1,7)		N	CAAR (-1,7)		Difference between Good and Bad	
Model 7	[10,98]	104	0,33%	(0,42)	37	-0,78%	(-0,59)	1,10%	(0,73)
Model 8	[10,98]	68	0,60%	(0,66)	20	-2,34%*	(-1,45)	2,94%*	(1,58)
Model 9.a	[20,98]	72	0,58%	(0,63)	18	-1,48%	(-0,89)	2,05%	(1,08)
Model 9.b	[30,98]	69	0,79%	(0,87)	17	-0,56%	(-0,33)	1,35%	(0,71)
Model 9.c	[40,98]	64	0,86%	(0,94)	9	-3,20%*	(-1,60)	4,07%**	(1,84)

Note: The table shows the difference between cumulative average abnormal return (CAAR) for category good and bad reviews. CAAR is measured in an event window 1 day before release to 7 days after first review. Each models sample varies because of the different cluster elimination methods as explained in section 5.6.7. T-stats are in parentheses. (*) means significant on a 90% confidential level, (**) are significant on a 95% level and (***) are significant on a 99% level. For model specific σ see appendix 5.

Examining only reviews that are published before the release, shown in Table 23, gives us different results. The Good category CAAR is no significantly larger than zero, but two of the bad category models are significantly lower than zero. The sample size used in bad category for examining reviews published before release of the video game is however very small. The power of the study is so small that we cannot reject our null hypothesis that state that bad first important reviews affect firm value.

The categories are now positive for good review CAAR's and negative on bad review CAAR's. As we remove less anticipated video games the t-test value also increases in the good category. Anticipation seems have an effect on first important review as well.

6.3 Cross-sectional regression

A cross-sectional regression gives us the opportunity to see the interactions between multiple variables in our study, and to explain the variance in the cumulative abnormal return in our events. We use the sample from Model 10 to do our regression. Model 10 has 296 observations, it is free of clusters, events are not removed because of anticipation and it has a good category CAAR significantly larger than zero. In an event study we cannot use the Metacritic score directly to test an events effect on firm value, we had to use a range for good and bad. In a regression on our cumulative abnormal return we can use the Metacritic score directly instead of categories.

We did two cross-sectional regressions, one base model,

$$CAR_i = \beta_0 + \beta_1 META_i + \beta_2 NUMBER_i + \beta_3 META_i \times NUMBER_i + \varepsilon_i \tag{14}$$

and on full model including all variables represented by X_{Mi}, presented in section 5.7.

$$CAR_i = \beta_0 + \beta_1 META_i + \beta_2 NUMBER_i + \beta_3 META_i \times NUMBER_i + \ldots + \beta_M X_{Mi} + \varepsilon_i \tag{15}$$

A summary of the observations for the variables in the base model is presented in Table 24, while a summary of the observations for the variables in the full model is presented in Appendix 4. The results for both regressions are found in Table 25.

Table 24: Variables summary base model

Variable	Min	1st Q	Median	Mean	3rd Q	Max
META	23	67	76	74	84	98
NUMBER	10	23	42	42,4	59	98

Note: The table shows simple statistical information for variable Metacritic score and number of reviews.

Table 25: Cross-sectional regression results

Variable	Base Model		Full Model	
(INTERCEPT)	7,4001	(1,49)	8,0985	(1,50)
META	-0,1049	(-1,56)	-0,1031	(-1,43)
NUMBER	-0,2696**	(-2,29)	-0,2691**	(-2,13)**
NUMBER × META	0,0036**	(2,41)	0,0037**	(2,34)**
EA			0,2249	(0,16)
UBISOFT			-0,7489	(-0,47)
CAPCOM			0,8094	(0,49)
PLATFORM			-0,7261	(-0,74)
SEQUEL			0,1420	(0,12)
FRANCHISE			-0,6634	(-0,48)
ACTION			-0,5806	(-0,49)
SPORTS			-0,8792	(-0,65)
MATURE			-0,7110	(-0,55)
TEEN			0,3216	(0,26)
SPRING			0,7184	(0,58)
SUMMER			0,5887	(0,44)
FALL			0,1140	(0,09)
N	296		296	
R^2	0,02464		0,03868	
R^2 adjusted	0,01462		-0,01644	

Note: The dependent variable is CAR in percentage. T-stats are in parentheses. (*) means significant on a 90% confidential level, (**) are significant on a 95% level and (***) are significant on a 99% level.

The number of reviews (NUMBER), and the number of reviews multiplied with the Metacritic score (NUMBER*META) are both significant. They are significant in both the models and the Metacritic score (META), although failed to be significant, has a very high t-test value in both models. None of the other variables explain the variance of the cumulative abnormal return very well. The base model explains the variance more efficient than the full model according to the adjusted R^2, because the full model includes to many variables that does not help explaining the CAR. This means that the base model is a better explanatory model, because it explains the dependent variable more efficient than the full model.

We also examined the Metacritic score and the number of reviews alone to see if they are significant on their own. The results show that they cannot explain the variance in the abnormal return alone, as seen in Appendix 4.

$$CAR_i = 7{,}4001 - 0{,}1049 \times META_i - 0{,}2696 \times NUMBER_i + 0{,}0036 \times META_i \times NUMBER_i \quad (16)$$

Equation (16) is calculated from the results in the base model regression to show how the different variables explain the CAR. Even though both the coefficient for META and NUMBER are negative it does not mean that a high Metacritic score and review number gives negative abnormal return, as the positive coefficient for the Metacritic score multiplied with the number of reviews explain most of the variance.

$$CAR_{meta\ 98} = -2{,}879 + 0{,}083 \times NUMBER \quad (17)$$

$$CAR_{meta\ 23} = 4{,}9876 - 0{,}1869 \times NUMBER \quad (18)$$

We used numeric values for the Metacritic score 98 and 23 in equation (16) to construct two linear functions (17) and (18). These are then used to explain the cumulative abnormal return in our observations. Our sample, as seen in Table 24, has a maximum Metacritic score of 98 and a minimum score of 23, while the minimum number of reviews is 10 and the maximum is 98. Using the minimum and maximum Metacritic score as a constant we can see how the number of reviews affects the cumulative abnormal return in Figure 13.

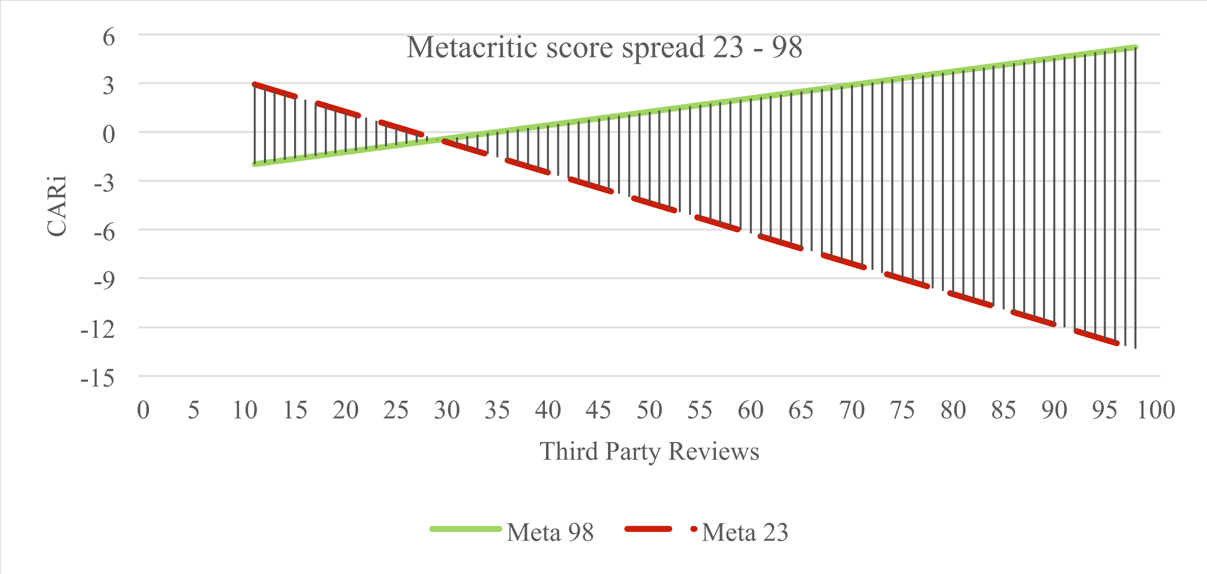


Figure 13: Metacritic score and number of reviews effect on cumulative abnormal return. This figure shows estimated CAR for video games with a Metacritic score of 23 and 98, relative to number of reviews.

Figure 13 shows that a high and low Metacritic score cross after 28 reviews and then start to go in the direction one would assume from the results in the event study. Our model is not strong enough to explain the variations in the abnormal return with a low amount of observations. It clearly shows however that the number of third party reviews plays a significant role in explaining the abnormal return. Both Metacritic score and the number of third party reviews significantly affect the slope of the abnormal return.

Using a Metacritic score spread from our third party reception categories, detailed in Table 7, a good reception has a Metacritic score between 85 and 98. A bad reception has a Metacritic score between 23 and 69. Using the same method as explained in Equation (17) and (18) we can see the differences between the categories as well. Calculating the highest and lowest score within each category.

Figure 14: Good and bad Metacritic reception and number of reviews effect on abnormal return. This figure shows estimated CAR for video games in category good and bad relative to number of reviews.

In Figure 14 the bad category has a steeper slope than the good category, and there is higher variation in the bad category compared to our good category. This is also indicated in our event study where the CAAR for the bad category was not significantly lower than zero, while the good category CAAR was significantly higher than zero.

The regression results show that a third party reception for an anticipated video game has a larger effect on firm value, than a less anticipated video game. We reject the null hypothesis for H5, stating that they are the same.

6.4 Robustness checks

We have analyzed multiple different models and observations in each event study to increase the robustness of the results. This was a natural solution to the clusters, but also help us feel more confident in our results.

Video game publisher stocks are very volatile, so using a mean model was also considered when calculating the abnormal return. We used the market model in our study, but used the constant mean model, explained in section 4.4.1, to see how robust the results are. A summary of the results using the constant mean model can be found in Table 26. We examined the constant mean return on Model 2 and 3 with a 15 day event window and compared them to their counterpart in section 6.1.

Table 26: The difference between abnormal return models

Model	Description	CAAR Good	P-value	CAAR Bad	P-value
	Market model	2,17 %	0,117	-0,71 %	0,303
Model 2	Mean model	2,34 %	0,069	-1,20 %	0,221
	Difference	0,17 %	-0,048	-0,49 %	-0,082
	Market model	2,23 %	0,109	-0,49 %	0,366
Model 3.a	Mean model	1,86 %	0,176	-1,49 %	0,178
	Difference	-0,37 %	0,067	-1,00 %	-0,188
	Market model	2,43 %	0,079	-1,44 %	0,171
Model 3.b	Mean model	2,85 %	0,069	-2,25 %	0,099
	Difference	0,41 %	-0,01	-0,82 %	-0,073
	Market model	3,00 %	0,024	-0,67 %	0,348
Model 3.c	Mean model	3,25 %	0,028	-2,38 %	0,117
	Difference	0,25 %	0,004	-1,71 %	-0,231

Note: The table shows the difference between our results, using the market-model and constant mean model, when calculating the cumulative average abnormal returns (CAAR).

Examining the difference between the CAAR and p-values we see that using the constant mean model, to calculate abnormal return, gives more significant results. The Good category CAAR is higher and the Bad category CAAR is lower using the constant mean value. The p-value is also reduced across all models. The results strengthen our findings even further.

6.5 Summary of hypotheses

Summarizing our results, in light of our hypotheses, a good reception affects firm value as shown in section 6.1, so we reject our null hypothesis for H_1. In the same section a bad reception gave a negative CAAR, but it was not significantly lower than zero, so we accept our null hypothesis for H_2. We found that the impact of a bad reception is higher for an anticipated video game, but without removing less anticipated video games, the null hypothesis for a bad category reception still holds.

In section 6.2, we found that the impact of the first important review was not significantly higher or lower than zero in any of our event studies. When we only tested reviews that were published before the release and we removed less anticipated video games from good reviews, the CAAR increased and became more significant, but not to a degree where we could say the CAAR is significantly higher than zero. We keep both null hypotheses for H_3 and H_4 and reject the alternative.

In our event studies in section 6.1 to 6.2 and the cross-sectional regression in section 6.3, our results showed that the third party reviews for anticipated video games has more effect on the firm value than less anticipated video games. Our event studies showed that anticipated third party reviews had a greater impact on the abnormal return, and the cross sectional regression showed a slope, for Metacritic score and anticipation, significantly different than zero. Based on the results we reject the null hypothesis for H_5. Table 27 shows a summary of the null hypotheses.

Table 27 Null hypotheses summary

Hypothesis	Description	Null hypothesis
H_1	A good reception for a video game has a positive effect on firm value.	Rejected
H_2	A bad reception for a video game has a negative effect on firm value.	Accepted
H_3	The first important good review for a video has a positive effect on firm value.	Accepted
H_4	The first important bad review for a video game has a negative effect on firm value.	Accepted
H_5	A reception or third party review for an anticipated video game has a larger effect on firm value.	Rejected

7 Discussion

7.1 Hypotheses 1 and 2: Reception effect on the firm value

Reviewing the hypotheses, using the tables and figures described in section 6, we see that third party reviews for video games affects firm value. We found that when a video game gets a Metacritic score of 85 or higher, the CAAR is significantly higher than zero, meaning that good third party reviews affect firm value. A bad third party review and the difference between a good and bad third party review does not necessarily affect firm value alone. The results in section 6.1 where not significant for bad third party reviews, but when we removed less anticipated video games, we found more significant results for both. This indicates that less anticipated video games are smaller investments for the publishing companies and has a more limited down side when it comes to firm value.

There can be multiple explanations to why a bad category third party reception does not give a negative CAAR significantly lower than zero. A bad category reception might not affect sales for a video game just because a good reception gives better sales. Especially sequels and yearly titles, like sports games, might not sell any worse just because it gets a bad reception, but it might sell more if it gets a good reception. The effect can also be explained by research on investor behavior (Nofsinger 2001). Investors seem to trade more on good news compared to bad news and this extends to good and bad receptions of a game. Investor can be more willing to “ride it out” if a video game gets a bad reception, waiting for an upside in news or wait for the actual sales information to make the final prediction. While on the other hand, if the third party reviews are very good, an investor immediately react on the information.

It seems to be a “winner takes it all” mentality among the investors. We see a significantly higher CAAR when a video game gets a good reception, but everything other than a good reception does not affect the firm value significantly. This investor response is not necessarily irrational. Electronic Entertainment Design and Research found that the average sales for video games that got a Metacritic score of 90 and above sold three times as much as video games that got a score of 80 to 89 (North 2012). The study was just an average measurement on sales data and leaves a lot of questions regarding the spread of the video games and the sales data. It was also just for video games released in one year. On the other hand, it might point to a “winner takes it all” scenario in the industry where highly rated video games sell more when it gets a very good reception, but does not necessarily reduce sales if the reception is bad.

7.2 Hypotheses 3 and 4: First important review effect on firm value

Examining the first important review we found no correlation between firm value and first important third party review. It does not matter if the third party review is published before or after the video game release, or if it is good or bad review. Investors are not reacting on one single third party review, but rather multiple third party reviews as shown in the results for the reception of a video game.

The results are in line with the study by Chen, Liu and Zhang (2012). They found that a third party review's absolute score did not affect firm value in the movie industry. However, when they studied one review relative to the one before, they did find a correlation.

Our findings also differ from other research that found a correlation between important third party reviews and firm value. Tellis and Johnson (2007) found that firm value correlate with third party reviews that is considered important in the consumer technology industry. IGN and Gamespot are arguably the two most important third party review sites for video games. We found no correlation between a single review and firm value, even though we used reviews that should be of importance to investors, using a proxy for anticipation. However, the studies are different in one aspect. Tellis and Johnson (2007) used one single critic, and the publisher, The Wall Street Journal, is highly regarded by the financial community. When only one professional critic is handling reviews for a paper, there is limited time to review products. This ensures that only highly anticipated products, and probably products that are always interesting for the financial community, are reviewed. This would be a natural assumption, because the publisher would want to maximize the readership, which is aimed at the financial community.

Our results do raise some questions. The study, on reception of a video game release, showed that third party reviews do affect firm value. A good review will give an abnormal return significantly higher than zero, and in section 5.5 we showed that there is little difference between the first important review and the Metacritic score. The correlations between the scores are 88 percent, and the average difference in critic scores are less than two points out of a hundred. Why are not investors reacting to an early important review if it includes the similar information as the Metacritic score?

Reinstein and Snyder (2005) found that third party reviews can be considered uncertain products and have different impact on consumers. This can extend to investors as well. Our study indicates that Metacritic might be viewed as a more influential third party reviewer than any ordinary critic, like IGN and Gamespot, even though Metacritic do not review any products themselves. The Metacritic score is unique and the information is valuable for the investor, according to our study on reception in section 6.1.

The lack of research on the industry also presents a challenge for the investors. There is no definitive research that third party reviews affect sales or other variables, and this causes an information uncertainty. Investors are not making any prediction based on a single review, but when the total reception and the Metacritic score is published, closer to release, the power of multiple third party reviews makes the information more assuring for the investors. This explains why there is little change in firm value, as shown in Figure 11 in section 6.1, until the game is released and the change in firm value is cumulating slow and late in the event window. The more third party reviews that give the same information, the more valuable the information is for the investor.

There is also some uncertainty regarding a single review for an investor. A review score of 85 is considered good, but it does not guarantee a good Metacritic score. Other third party reviews might give a lower score reducing the reception to mixed, or even bad. This can explain why we do not see the same results for single reviews as we do for the reception of the video game. A score of a 95 or 100 gives a higher probability that the Metacritic score also reflect a good score, but our study is not strong enough to measure the extreme values this problem presents, and more observations are required. The difference between the first important review and the Metacritic score in section 5.5 was small, but there were still some differences that can lead to investor uncertainty.

7.3 Hypothesis 5: Anticipation for a video game

Our cross-sectional regression and event studies showed that the Metacritic score does not explain the change in abnormal return alone. We found that anticipation plays an important role in how Metacritic affects firm value. The more anticipated the video game is, the more third party reviews will affect firm value. This might explain why other unpublished studies have failed to find a correlation between sales and third party reviews.

Our study shows that investors believe that the third party review effect on sales are enhanced or reduced dependent on how anticipated the video game is. This is a reasonable assumption from the investors, as the more news coverage a video game gets, odds are that it has a large production budget and a mass consumer appeal. A larger production budget because the publisher wants to market the product as much as possible, and a mass consumer appeal because it is the best interest of the news outlet to get a higher readership. A higher production budget makes the third party review effect stronger, as the possible downside of a failed product is larger, and a mass consumer appeal makes the possible upside for a highly rated product larger.

Chen, Liu and Zhang (2012) showed that there are fundamental product differences in video games. Consumer reviews do not affect video games sales alone. First when video games are attributed other variables, they affect sales. Chen et al. (2012) used popularity as a variable to show a correlation, just as we used anticipation in our study. This might be the reason why there has been no correlation found between third party reviews and sales.

8 Implications

8.1 Managerial implications

This is the first study that finds a correlation between third party reviews and firm value in the video game industry. Our study shows that third party reviews is a good way to attach a value to a video game when other evaluations are not possible, like sales information. Calculating return on investment requires historical information, and it is difficult for mangers to assess the success or failure, of a product, before this information is available. The stock prices reflect all investors combined knowledge predicting future earnings for a product. This gives managers a new tool to assess the success of a new product before the first economical information on sales is available. Managers have actively been reviewing developers track record when assessing new possibilities and funding of new projects (Schreier 2015). This practice should be continued to create shareholder value.

This study is also important to the investor body as it shows that third party reviews are a good source of information to evaluate future prospects. The results from our event study on reception, in section 6.1, show that less anticipated video games has a smaller downside for investors than more anticipated video games. In the models where we included less anticipated video games, the upside of a good category score is significantly higher than zero, but the downside is not significantly lower than zero. As we remove the less anticipated video games from the good category, the impact of third party reviews only increase a little, while the impact of bad category third party reviews is increased more. The increased downside can also be seen in the cross sectional regression in section 6.3, and in Figure 13 and 14. The Bad category regression slope is a lot steeper for a bad category reception than for a good category reception.

Our results show that the market is efficient, according to EMH, and investors incorporate the new information into the stock price. The general belief in the market seems to be that a good reception affects the publishers positively, as it does for managers at the publishers. According to our results, the general response from investors is slow. It seems to be a delayed reaction as many third party reviews are available to investors at release day. The delayed reaction is in line with event study research, showing that the stock price “drifts” to the correct price, rather than “instantaneously” increase or decrease (Bernard and Thomas 1989).

The slow reaction might be a case of investor uncertainty. They believe third party reviews affect sales, but do not know how much, as there is little research on this subject. The uncertainty is also reflected in that investors do not react on a single early important review that is highly correlated with the Metacritic score.

8.2 Research implications and future research

Relative to its importance, there is little research done in the video game industry, and researchers should inspect the economical mechanics of this industry. It is the fastest growing industry in the entertainment sector and should be examined by the financial markets as well. Our study supplements the small body of research done on this industry, but also contributes to the research examining the effects of third party review on firm value. We found a significant abnormal return in the third party reception of a video game, but this in large part only correlated to good third party reviews. Bad third party reviews did not have the same impact on firm value. The different effects should be studied in future research, and we think examining the third party reviews effect on sales is the most adequate way to find an answer to this question. Our results indicate that it should be possible to find a correlation between sales and third party reviews.

Previous unpublished studies in the video game industry have focused on the relationship a third party review can have on sales, without any variables to explain the differences between the video games. (Artursson 2015; Boyer 2006; Maragos 2005; Orland 2014). All the studies failed to find any correlation. Even though no one has found a correlation between sales and third party reviews, the debate is still on going. The assumption that there is a correlation between the two is rooted within the minds of managers and investors alike. Multiple third party reviewers have even decided not to report a numerical score in their reviews because of the believed effect they have on the economy of the industry (Welsh 2015). Our study shows that it is important to add anticipation as a variable, to explain the power of the effect a third party review can have on firm value, and this should extend to sales as well. Our proxy for anticipation, the number of third party reviews, showed to be a good and available variable to use, and should be incorporated in future research regarding sales and third party reviews.

This industry is constantly evolving, and the way investors use third party reviews might have changed over the years. Other types of information have become available for investors and third party reviews might have lost some of the informational power. Our study is not strong enough to examine this effect, as more observations are required within in each period. Examining the changes between time periods can give some insight to this subject.

9 Conclusion

Our thesis has examined the effect third party reviews have on abnormal return in the video game industry. A sample of four publishers, 611 video games, and 600 reviews has been used to examine the abnormal return for stocks over a 15-year period. We carried out an event study following MacKinlay's (1997) methodology to examine the effect a good and bad third party review has on firm value, using both important single reviews and the total reception for a video game with the Metacritic score. We also performed a cross-sectional regression on the abnormal return, to examine the mechanics between abnormal return and third party reviews.

Our study is the first to find a correlation between third party reviews and firm value in the video game industry. We found that third party reviews affect firm value, but only high scoring reviews. In general, low scoring third party reviews do not affect firm value, but we found that the anticipation of a video game has an important role in how the third party reviews affect firm value. Anticipation increases the effect on firm value of both good and bad third party reviews. We found that product difference between video games is an important factor, and anticipation has a large effect on the power of third party reviews. Our results show that investors react to the total reception of the video games, but not on single reviews with the same information, even if the information is available before the release of the video game. The change in firm value is consistent with the efficient market hypothesis and is affected by the news of third party reviews, and is also late and drifts to its new price rather than change instantaneously. Both observations can be attributed to investor uncertainty, as the research regarding sales and third party reviews are limited.

10 References

Adajian, Thomas. 2012. "The Definition of Art." (Winter 2012) Accessed 24 April 2016. http://plato.stanford.edu/archives/win2012/entries/art-definition/.

Alexa. 2016. "Site Overview." Alexa Accessed March 30 2016. http://www.alexa.com/siteinfo.

Anderson, Craig A. 2004. "An update on the effects of playing violent video games." *Journal of adolescence* 27 (1): 113-122.

Anderson, Craig A and Brad J Bushman. 2001. "Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: A meta-analytic review of the scientific literature." *Psychological science* 12 (5): 353-359.

Artursson, Erik. 2015. "Online Ratings – who decides what games you buy?" Master thesis, Ekonomihögskolan, Lunds Universitet. http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=5469912&fileOId=5469919.

Baker, Liana B. 2011. "Shares of video game companies swing on reviews." Reuters Accessed January 16 2016. http://www.reuters.com/article/us-videogame-reviews-idUSTRE78F52320110916.

Ball, Ray. 1978. "Anomalies in relationships between securities' yields and yield-surrogates." *Journal of Financial Economics* 6 (2): 103-126.

Banerjee, Scott. 2006. "Video game reviews presage stock moves." Accessed January 15 2016. http://www.marketwatch.com/story/video-game-reviews-presage-stock-moves.

Basu, Sanjoy. 1977. "Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis." *The journal of Finance* 32 (3): 663-682.

Beres, Damon. 2015. "Stop Buying Physical Video Games Already!". Huffington Post Accessed March 29 2016. http://www.huffingtonpost.com/entry/digital-video-game-sales_us_5654946be4b0879a5b0c7fd6.

Bernard, Victor L and Jacob K Thomas. 1989. "Post-earnings-announcement drift: delayed price response or risk premium?" *Journal of Accounting research*: 1-36.

Blundon, Matthew and Jon Lindemann. 2010. "NPD Sales Results for August 2010." NintendoWorldReport Accessed April 28 2016.

http://www.nintendoworldreport.com/news/24003/npd-sales-results-for-august-2010.

Bodie, Zvi, Alex Kane and Alan J Marcus. 2014. *Investments*. 10th Global Edition ed. Berkshire: McGraw-Hill Education.

Boyer, Brandon. 2006. "Survey: Game Score-to-Sale Theory Again Disproven." Gamasutra Accessed April 24 2006. http://gamasutra.com/view/news/101876/Survey_Game_ScoretoSale_Theory_Again_Disproven.php.

———. 2008. "Ubisoft: Wii To Rule Them All, Microsoft/Sony Battle Split In U.S/Europe." Gamasutra Accessed March 29 2016. http://www.gamasutra.com/php-bin/news_index.php?story=18389.

Brown, Stephen J and Jerold B Warner. 1980. "Measuring security price performance." *Journal of financial economics* 8 (3): 205-258.

———. 1985. "Using daily stock returns: The case of event studies." *Journal of financial economics* 14 (1): 3-31.

Burton, Edwin and Sunit Shah. 2013. *Behavioral Finance: Understanding the Social, Cognitive, and Economic Debates*. Wiley Finance. New York: New York : Wiley.

Carter, Grey. 2012. "Funcom Blames MetaCritic For Share Price Drop." Escapist Magazine Accessed April 27 2016. http://www.escapistmagazine.com/news/view/119015-Funcom-Blames-MetaCritic-For-Share-Price-Drop.

Chen, Pei-Yu, Shin-yi Wu and Jungsun Yoon. 2004. "The impact of online recommendations and consumer feedback on sales." *ICIS 2004 Proceedings*: 58.

Chen, Yubo, Yong Liu and Jurui Zhang. 2012. "When do third-party product reviews affect firm value and what can firms do? The case of media critics and professional movie reviews." *Journal of Marketing* 76 (2): 116-134.

Chen, Yubo and Jinhong Xie. 2005. "Third-party product review and firm marketing strategy." *Marketing Science* 24 (2): 218-240.

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

Crawford, C Merle. 1977. "Marketing research and the new product failure rate." *The Journal of Marketing*: 51-61.

Crecente, Brian. 2013. "Grand Theft Auto 5 hits $1 billion in three days, sets new record." Polygon, 20.09.2013 Accessed February 29 2016.

http://www.polygon.com/2013/9/20/4752380/grand-theft-auto-5-hits-1-billion-in-three-days-sets-new-record.

Danielsen, Stian. 2015. Masteroppgave. edited by Kim Hermansen.

De Bondt, Werner FM and Richard H Thaler. 1990. "Do security analysts overreact?" *The American Economic Review*: 52-57.

Dietz, Jason. 2016. "Metacritic's 6th Annual Game Publisher Rankings." Metacritic Accessed March 27 2016. http://www.metacritic.com/feature/game-publisher-rankings-for-2015-releases.

Dietz, Tracy L. 1998. "An examination of violence and gender role portrayals in video games: Implications for gender socialization and aggressive behavior." *Sex roles* 38 (5-6): 425-442.

Dunning, Jason. 2014. "March 2014 NPD: PS4 is #1 Yet Again, Titanfall Beats inFamous: Second Son." PlayStationLifestyle Accessed April 28 2016. http://www.playstationlifestyle.net/2014/04/17/march-2014-npd-ps4-is-1-yet-again-titanfall-beats-infamous-second-son/.

Ebizmba. 2016. "Top 15 Most Popular Video Game Websites | March 2016." Ebizmba Accessed March 30 2016. http://www.ebizmba.com/articles/video-game-websites.

Elberse, Anita and Bharat Anand. 2007. "The effectiveness of pre-release advertising for motion pictures: An empirical investigation using a simulated market." *Information Economics and Policy* 19 (3): 319-343.

Eliashberg, Jehoshua and Steven M Shugan. 1997. "Film critics: Influencers or predictors?" *The Journal of Marketing*: 68-78.

ESA. 2015. *Essential facts about the computer and video game industry*: Entertainment Software Association. March 31. http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf.

Fama, Eugene F. 1965. "Random walks in stock market prices." *Financial analysts journal* 21 (5): 55-59.

———. 1991. "Efficient capital markets: II." *The journal of finance* 46 (5): 1575-1617.

Fama, Eugene F and Kenneth R French. 1993. "Common risk factors in the returns on stocks and bonds." *Journal of financial economics* 33 (1): 3-56.

Fischberg, Jason. 2012. "State of the Video Game Union: Sequel Saturation." Examiner Accessed April 15 2016. http://www.examiner.com/article/state-of-the-video-game-union-sequel-saturation.

French, Kenneth R. 1980. "Stock returns and the weekend effect." *Journal of financial economics* 8 (1): 55-69.

Gee, James Paul. 2003. "What video games have to teach us about learning and literacy." *Computers in Entertainment (CIE)* 1 (1): 20-20.

Gilbert, Ben. 2012. "Obsidian missed Fallout: New Vegas Metacritic bonus by one point." Engadget Accessed April 13 2012. http://www.engadget.com/2012/03/15/obsidian-missed-fallout-new-vegas-metacritic-bonus-by-one-point/.

Gilbert, Henry. 2013. "Michael Bay proves you should be happy about Metacritics influence." Gamesradar Accessed March 30 2016. http://www.gamesradar.com/games-vs-movies-editorial/.

Grubb, Jeff. 2013a. "August 2013 NPD: Retail sales up year-over-year for first time since October 2011." Venturebeat.com Accessed April 28 2016. http://venturebeat.com/2013/09/12/august-2013-npd-retail-sales-up-year-over-year-for-first-time-since-october-2011/.

———. 2013b. "Metacritic works: Why the review-aggregation site is important for the average consumer." VentureBeat Accessed March 30 2016. http://venturebeat.com/2013/08/07/metacritic-works-why-the-review-aggregation-site-is-important-for-the-average-consumer/.

———. 2015. "December 2015 NPD: Call of Duty, Star Wars, Fallout end the year strong." Venturebeat Accessed April 28 2016. http://venturebeat.com/2016/01/14/december-2015-npd-call-of-duty-star-wars-fallout-end-the-year-strong/.

———. 2016. "March 2016 NPD: Ubisoft, Nintendo shake up sales chart with The Division, Zelda." Venturebeat Accessed April 28 2014. http://venturebeat.com/2016/04/14/march-2016-npd-division-zelda/.

Hatfield, Daemon. 2007. "Npd spree: August Sales." IGN Accessed April 28 2016. http://www.ign.com/articles/2007/09/14/npd-spree-august-sales.

Haugen, Robert A and Philippe Jorion. 1996. "The January effect: Still there after all these years." *Financial Analysts Journal* 52 (1): 27-31.

Holbrook, Morris B. 1999. "Popular appeal versus expert judgments of motion pictures." *Journal of consumer research* 26 (2): 144-155.

Ingenito, Vince. 2014. "Dragon Age: Inquisition Review." IGN Accessed April 28 2016. http://www.ign.com/articles/2014/11/11/dragon-age-inquisition-review.

Ingham, Tim. 2015. "Global record industry income drops below $15 billion for first time in decades." Music Business Worldwide Accessed February 25 2016. http://www.musicbusinessworldwide.com/global-record-industry-income-drops-below-15bn-for-first-time-in-history/.

Kahneman, Daniel and Amos Tversky. 1973. "On the psychology of prediction." *Psychological review* 80 (4): 237.

Keim, Donald B. 1983. "Size-related anomalies and stock return seasonality: Further empirical evidence." *Journal of Financial Economics* 12 (1): 13-32.

Kothari, SP and Jerold B Warner. 2006. "Econometrics of Event Studies." *Working Paper, Center for Corporate Governance, Tuck School of Business at Dartmouth University*.

Kuchera, Ben. 2014. "Destiny is a $500 million game? Yes, and that's not insane." Polygon Accessed April 14 2016. http://www.polygon.com/2014/6/13/5807752/destiny-500-million-budget-activision-bungie.

Lang, Brent. 2015. "Box Office: 'Star Wars' Crosses $1 Billion Globally at Record Pace." Variety, 27.12.2015 Accessed February 29 2016. http://variety.com/2015/film/box-office/star-wars-box-office-christmas-daddys-home-point-break-1201668001/.

Macdonald, Keza. 2012. "Is metacritic ruining the industry?". IGN Accessed March 3 2016. http://www.ign.com/articles/2012/07/16/is-metacritic-ruining-the-games-industry.

MacKinlay, A. Craig. 1997. "Event Studies in Economics and Finance." *Journal of Economic Literature* 35 (1): 13-39.

Malkiel, Burton G. 2003. "The efficient market hypothesis and its critics." *The Journal of Economic Perspectives* 17 (1): 59-82.

———. 2007. *A random walk down Wall Street: The time-tested strategy for successful investing*: Norton & Company, Incorporated, W. W.

Malkiel, Burton G and Eugene F Fama. 1970. "Efficient capital markets: A review of theory and empirical work." *The journal of Finance* 25 (2): 383-417.

Maragos, Nich. 2005. "Survey: Game Score-to-Sale Theory Disproven." Gamasutra Accessed April 24 2016. http://www.gamasutra.com/php-bin/news_index.php?story=7453.

Metacritic. 2016a. "All video game publishers." Metacritic Accessed March 27 2016. http://www.metacritic.com/browse/games/company/popular?num_items=100.

———. 2016b. "How We Calculate Our Scores: The Long FAQ." Metacritic Accessed April 14 2016. http://www.metacritic.com/about-metascores.

Mitchell, Alice and Carol Savill-Smith. 2004. "The use of computer and video games for learning: A review of the literature." *Learning and Skills Development Agency*.

MPAA. 2015. *Theatrical Statistics Summary 2014*: MPAA. March 11. http://www.mpaa.org/wp-content/uploads/2015/03/MPAA-Theatrical-Market-Statistics-2014.pdf.

Newzoo. 2015. "Global games market will grow 9.4% to $91.5BN in 2015." Newzoo, 22.04.2015 Accessed February 2 2016. https://newzoo.com/insights/articles/global-games-market-will-grow-9-4-to-91-5bn-in-2015/.

Nofsinger, John R. 2001. "The impact of public information on investors." *Journal of Banking & Finance* 25 (7): 1339-1366.

North, Dale. 2012. "GDC: How important review scores are to game sales." Destructoid Accessed February 25 2016. http://www.destructoid.com/gdc-how-important-review-scores-are-to-game-sales-223570.phtml.

Orland, Kyle. 2014. "Steam Gauge: Do strong reviews lead to stronger sales on Steam?". ArsTechnica Accessed April 23 2016. http://arstechnica.com/gaming/2014/04/steam-gauge-do-strong-reviews-lead-to-stronger-sales-on-steam/.

Parkin, Simon. 2013. "Full Sail study attempts to shed light on Metacritic's weighting system." Gamasutra Accessed May 4 2016. http://gamasutra.com/view/news/189448/Full_Sail_study_attempts_to_shed_light_on_Metacritics_weighting_system.php.

Reddy, Srinivas K, Vanitha Swaminathan and Carol M Motley. 1998. "Exploring the determinants of Broadway show success." *Journal of Marketing Research*: 370-383.

Reinstein, David A and Christopher M Snyder. 2005. "The influence of expert reviews on consumer demand for experience goods: A case study of movie critics*." *The journal of industrial economics* 53 (1): 27-51.

Schreier, Jason. 2015. "Metacritic Matters: How Review Scores Hurt Video Games." Kotaku Accessed February 23 2016. http://kotaku.com/metacritic-matters-how-review-scores-hurt-video-games-472462218.

Serrels, Mark. 2014. "Gamers Still Prefer Physical Copies Of Video Games." Kotaku Accessed March 29 2016. http://www.kotaku.com.au/2014/05/gamers-still-prefer-physical-copies-of-video-games/.

Shea, Brian. 2016. "February 2016 NPD: PlayStation 4 Leads Hardware Sales, Far Cry Primal Tops Software Sales." Gameinformer Accessed April 28 2016. http://www.gameinformer.com/b/news/archive/2016/03/10/february-2016-npd-results.aspx.

Sivadas, Eugene and F Robert Dwyer. 2000. "An examination of organizational factors influencing new product success in internal and alliance-based processes." *Journal of marketing* 64 (1): 31-49.

Sliwinski, Alexander. 2010. "EEDAR/SMU study: review scores affect perceived quality, purchase intent." Engadget Accessed April 25 2016. http://www.engadget.com/2010/07/06/eedar-smu-study-review-scores-affect-perceived-quality-purchas/.

Stuart, Keith. 2014. "Video games and art: why does the media get it so wrong?". The Guardian Accessed April 8 2016. https://www.theguardian.com/technology/gamesblog/2014/jan/08/video-games-art-and-the-shock-of-the-new.

Takahashi, Dean. 2009. "EA's chief creative officer describes game industry's re-engineering." VentureBeat Accessed March 29 2016. http://venturebeat.com/2009/08/26/eas-chief-creative-officer-describes-game-industrys-re-engineering/.

Take-Two. 2013. *Take-Two Interactive 2012 Annual Report*: Take-Two Interactive. http://phx.corporate-ir.net/External.File?item=UGFyZW50SUQ9NDc1NzM2fENoaWxkSUQ9NTA4MDQ4fFR5cGU9MQ==&t=1.

Tellis, Gerard J and Joseph Johnson. 2007. "The value of quality." *Marketing Science* 26 (6): 758-773.

Welsh, Oli. 2015. "Eurogamer has dropped review scores." Eurogamer Accessed May 21 2016. http://www.eurogamer.net/articles/2015-02-10-eurogamer-has-dropped-review-scores.

Wiles, Michael A and Anna Danielova. 2009. "The worth of product placement in successful films: An event study analysis." *Journal of Marketing* 73 (4): 44-63.

Wouk, Kristofer. 2016. "Jonathan Blow's The Witness earned more in one week than his previous game did in a year." Accessed April 14 2016.

http://www.digitaltrends.com/gaming/the-witness-outperforms-braid-in-under-a-week/.

Zhu, Feng and Xiaoquan Zhang. 2010. "Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics." *Journal of marketing* 74 (2): 133-148.

11 Appendix

Appendix 1: Publisher Summary[*]

Company name	Stock ticker	Mkt. cap. in $ (Billion)	Included in study?
Microsoft Corporation	NASDAQ: MSFT	$397,10	No, not the main revenue form
Sony Corporation	NYSE:SNE	$31,68	No, not the main revenue form
Activision Blizzard Inc.	NASDAQ: ATVI	$25,91	No, subscription based revenue
Electronic Arts Inc.	NASDAQ: EA	$21,15	Yes
Nintendo Co., Ltd	TYO: 7974	$18,87	No, not the main revenue form
Hasbro, Inc.	NASDAQ: HAS	$10,72	No, not the main revenue form
Bandai Namco Holdings Inc.	TYO: 7832	$4,49	No, not the main revenue form
Konami Holdings Corp	TYO: 9766	$4,44	No, not the main revenue form
Ubisoft Entertainment SA	EPA: UBI	$3,55	Yes
Square Enix Holdings Co. Ltd.	TYO: 9684	$3,35	No, subscription based revenue
Take Two Interactive Software Inc.	NASDAQ: TTWO	$3,33	Yes
Sega Sammy Holdings Inc.	TYO: 6460	$2,80	No, not the main revenue form
Koei Tecmo Holdings Co Ltd	TYO: 3635	$1,78	No,Asian markets prioritized
Capcom Co. Ltd.	TYO: 9697	$1,65	Yes
Atari SA	EPA: ATA	$0,03	No, no recent games published
Funcom N.V.	FUNCOM.OL	$0,03	No, subscription based revenue

[*] Mkt. Cap. Extracted from www.google.com/finance - Accessed March 30 2016

Appendix 2: Summary of third party reviewers (Ubisoft and Capcom)

A: Activity summary

	Web traffic (Alexa.com)		Number of reviews				
	Rank World	Rank US	Total	Before release	On release	After release	First review
IGN	385	171	279	74	51	154	01/09/00
Gamespot	820	402	269	55	48	166	28/08/00
Cheat Code Central	12852	7016	127	8	14	105	19/08/05
Destructoid	6870	2453	64	14	13	37	27/10/07
Giantbomb	4652	1739	43	5	13	25	18/03/08
Eurogamer	5315	3933	155	33	23	99	25/02/03
Gamesradar	4637	2826	82	25	17	40	23/03/04
Gametrailers	31109	26615	85	19	15	51	31/08/06
Gamezone	29248	12466	176	14	5	157	15/10/01
Gamerevolution	22732	8395	157	16	12	129	01/09/00
The Escapist	9527	3702	28	5	7	16	01/04/10
Videogamer	23370	15027	103	33	18	52	13/03/06
All			1568	301	236	1031	

B: Review summary

	Review score				Review relative to release date (days)				
	Avg	Min	Max	SD	Avg	Min	Max	Median	SD
IGN	72	20	98	17	5	-20	77	1	12
Gamespot	68	19	96	16	6	-39	77	2	12
Cheat Code Central	76	30	96	16	12	-27	198	4	28
Destructoid	72	25	95	19	5	-40	35	2	11
Giantbomb	71	20	100	21	4	-1	21	1	6
Eurogamer	67	20	100	18	21	-47	200	7	38
Gamesradar	72	20	100	18	8	-61	139	0	24
Gametrailers	77	40	94	13	5	-13	52	1	11
Gamezone	74	25	97	15	16	-55	122	13	18
Gamerevolution	62	0	100	23	14	-29	123	10	19
The Escapist	76	30	100	17	7	-13	29	4	10
Videogamer	70	20	90	15	11	-47	130	1	28
All	71	0	100	18	9	-61	200	3	21

Appendix 3: Model summary of Category mixed

			Mixed				
Model	Review interval	Cluster	Event window	N	CAAR	T-test	P-value
Model 1	[10,98]	0	7,7	277	-0,21	-0,34	0,74
Model 2	[10,98]	[-7, 7]	7,7	86	0,47	0,43	0,67
Model 3.a	[20,98]	[-7, 7]	7,7	85	-0,33	-0,30	0,76
Model 3.b	[30,98]	[-7, 7]	7,7	81	-0,31	-0,29	0,77
Model 3.c	[40,98]	[-7, 7]	7,7	74	-1,47	-1,37	0,17
Model 1	[10,98]	0	-1,7	277	-1,00	-2,05	0,04
Model 2	[10,98]	[-7, 7]	-1,7	86	0,22	0,26	0,80
Model 3.a	[20,98]	[-7, 7]	-1,7	85	-0,41	-0,48	0,63
Model 3.b	[30,98]	[-7, 7]	-1,7	81	-0,77	-0,93	0,35
Model 3.c	[40,98]	[-7, 7]	-1,7	74	-1,14	-1,37	0,17
Model 4	[10,98]	0	-1,7	242	0,13	0,26	0,80
Model 5	[10,98]	[-1,7]	-1,7	128	0,76	1,13	0,26
Model 6.a	[20,98]	[-1,7]	-1,7	117	1,08	1,59	0,11
Model 6.b	[30,98]	[-1,7]	-1,7	96	-0,47	-0,67	0,51
Model 6.c	[40,98]	[-1,7]	-1,7	77	-0,77	-0,99	0,33
Model 7	[10,98]	0	-1,7	78	-0,01	-0,01	0,99
Model 8	[10,98]	[-1,7]	-1,7	47	0,64	0,63	0,53
Model 9.a	[20,98]	[-1,7]	-1,7	48	1,23	1,21	0,23
Model 9.b	[30,98]	[-1,7]	-1,7	42	0,61	0,59	0,56
Model 9.c	[40,98]	[-1,7]	-1,7	35	0,07	0,06	0,95
Model 10	[10,98]	[-1, 7]	-1,7	139	-0,84	-1,27	0,21
Model 11.a	[20,98]	[-1, 7]	-1,7	140	-0,48	-0,75	0,45
Model 11.b	[30,98]	[-1, 7]	-1,7	130	-0,91	-1,42	0,16
Model 11.c	[40,98]	[-1, 7]	-1,7	107	-0,98	-1,41	0,16

Appendix 4: Summary full model variables, Meta and Number regression

Full Model

Variable	Min	1st Q	Median	Mean	3rd Q	Max
META	23	67	76	74	84	98
NUMBER	10	23	42	42,4	59	98
PUBLISHER						
EA	0	0	0	0,3209	1	1
UBISOFT	0	0	0	0,25	0,25	1
CAPCOM	0	0	0	0,1926	0	1
PLATFORM	0	0	1	0,5236	1	1
SEQUEL	0	0	0	0,4257	1	1
FRANCHISE	0	1	1	0,7973	1	1
ESRB						
MATURE	0	0	0	0,3446	1	1
TEEN	0	0	0	0,2568	1	1
SEASON						
SPRING	0	0	0	0,2669	1	1
SUMMER	0	0	0	0,2095	0	1
FALL	0	0	0	0,2905	1	1

Meta Regression

Coefficients	Estimate	Std.Error	t-value	P-value
(Intercept)	-2.97265	2.49069	-1.194	0.234
META	0.03838	0.03317	1.157	0.248

Number Regression

Coefficients	Estimate	Std.Error	t-value	P-value
(Intercept)	-0.82393	0.92146	-0.894	0.372
NUMBER	0.01632	0.01929	0.846	0.398

Appendix 5: Model summary of standard error σ

	σ			σ * event window			
Model	Good	Mixed	Bad	Good	Mixed	Bad	Event window
Model 1	0,26	0,16	0,18	1,02	0,63	0,70	15
Model 2	0,46	0,28	0,35	1,80	1,09	1,36	15
Model 3.a	0,46	0,28	0,37	1,78	1,08	1,42	15
Model 3.b	0,44	0,27	0,39	1,69	1,06	1,49	15
Model 3.c	0,38	0,28	0,44	1,47	1,08	1,70	15
Model 1	0,26	0,16	0,18	0,79	0,49	0,54	9
Model 2	0,46	0,28	0,35	1,39	0,84	1,06	9
Model 3.a	0,46	0,28	0,37	1,38	0,84	1,10	9
Model 3.b	0,44	0,27	0,39	1,31	0,82	1,16	9
Model 3.c	0,38	0,28	0,44	1,14	0,83	1,32	9
Model 4	0,21	0,17	0,20	0,63	0,51	0,60	9
Model 5	0,28	0,22	0,29	0,83	0,67	0,86	9
Model 6.a	0,28	0,23	0,30	0,83	0,68	0,89	9
Model 6.b	0,27	0,24	0,31	0,81	0,71	0,94	9
Model 6.c	0,27	0,26	0,38	0,81	0,78	1,15	9
Model 7	0,26	0,29	0,44	0,77	0,86	1,31	9
Model 8	0,31	0,34	0,54	0,92	1,02	1,61	9
Model 9.a	0,31	0,34	0,55	0,92	1,02	1,66	9
Model 9.b	0,31	0,34	0,56	0,92	1,03	1,68	9
Model 9.c	0,31	0,38	0,67	0,92	1,14	2,00	9
Model 10	0,33	0,22	0,28	1,00	0,66	0,85	9
Model 11.a	0,33	0,22	0,30	0,98	0,65	0,89	9
Model 11.b	0,33	0,21	0,32	0,98	0,64	0,96	9
Model 11.c	0,32	0,23	0,41	0,96	0,69	1,22	9