"Do Substitute Products Affect Seasonal Fluctuation in Box Office Revenue for Motion Picture Studios?"

Economics and Business Master Thesis 2015

Jordy Martijn van der Reijd (402990)

Supervisor: Dr. Y. Peers

Executive Summary

The motion picture industry has become stagnant in terms of total theatrical revenue earned over the past 8 years in the United States. Competitive services such as video on demand (VOD) and cable television have seen significant rise in terms of revenue and usage. Production studios are searching for ways to release big box office revenue generating movies. The goal of this research is to investigate what strategies studios can exercise in order to release profitable movies. Seeing as there is large variance in revenue generation during the different seasons, this paper will investigate what causes this fluctuation in revenue in order to gain a better understanding why some movies make more revenue than others. The research question of the thesis is therefor:

What causes the fluctuation in seasonal box office revenue?

This research will explore the fluctuations of box office revenue during different seasons based on previous research. The features analyzed in this research will relate to movie quality, which is subdivided in review ratings and star power, and movie attributes. Furthermore, this research will add to this study field by exploring the effects that substitute products have on the revenue of theatrical movies. The substitute products examined in this research are cable television and Netflix accounts. A linear regression has been performed based on these variables in order to explain the variance in box office revenue during the seasons. The results reveal that the variables movie ratings, star power, and movie attributes are significant drivers for box office revenue. The conclusion is that studios need to be aware of what variables drive box office revenue during a particular season.

Table of Content

Chapter 1 – Introduction

Four years ago, an article in GQ was placed with the title 'The Day the Movies Died'. It was an article about how in the pursuing 5 years the quality of movies would drop due to lack of stories. Furthermore, it referred to television shows attracting more audience due to a better business model. Fast forward to 2014, and there are two phenomena that are occurring in the movie industry in the United States. The first is that the industry is stagnating in yearly box office revenue since 2007. That the industry has become stagnant is not covered extensively in the media. The movie industry has increased only 7,2 percent from 2007 till 2014, equaling 1 percent per year. One article argues that theaters have become less important for movie revenues and that theater movie runs now serve as promotional activity for other media markets (Gil, 2008). The second phenomenon is that there is a quarterly box office revenue disparity. This trend has been covered more extensively in the media. The term dump month is a commonly used expression for it. It is used to describe January and February primarily because studios release movies that the studio does not believe will become big box office revenue generators. These movies are 'the stuff that barely gets promoted beyond blurbs from obscure websites and suspicious raves from local TV chefs and weathermen' (The Guardian, 2007). It comes right after the holiday season when the so-called big box office revenue generators are scheduled to debut. The general belief is that these movies are of better quality (better story, bigger stars) and therefor come out during the holiday season when consumers have more leisure time to go to the cinema. This paper will investigate what causes this seasonal disparity in box office revenue.

With all that being said, it is interesting to investigate what causes this occurrence. Is it like the GQ article said that movies have become of worse quality and/or are there substitute products for theatrical movies? The website BGR.com published in January 2015 that 'Netflix is starting to wound the movie industry'**.** The growth in consumer base of Netflix could be a cause that the movie industry is stagnating. That could indicate that variables outside of the industry are stealing revenue. On the other hand, it could be possible that quality of movies regress during different period (the so-called dump months). It is fathomable that studios decide to postpone their best movies to release during award-nomination periods. It might explain the fluctuation in seasonal revenue. Understanding what affects box office revenue is of particular interest to studios that produce theatrical movies. Do they need to adjust their revenue-generation model in order to avoid financial harm? And how does seasonal variation affect revenue? With that being said, it leads to the following problem statement:

Can box office revenue be accurately predicted by factoring in seasonal fluctuations?

By analyzing what causes the variance in box office revenue over the different seasons, it will potentially be possible to explain the stagnating movie industry in the United States. This research hypothesizes that two different phenomena are occurring. The first phenomenon is that during seasons where the revenue is low, the studios are dumping their least favorably tested movies. The term 'dump months' are primarily used to explain the January and February period as mentioned before. However, it so happens that the revenue is roughly as low during the fall period (Einav, 2007). Interestingly enough, no research has ever coined the term dump months during the period after the summer. We postulate that studios are dumping their least favorable movies during these two periods because consumers generally have less spare time than during the summer and holiday season. If studios are in fact doing this, it can explain the fluctuation in seasonality. The other phenomenon that is occurring is that there is an increased usage of substitute products. Consumers watch more television more during the first quarter of the year (Nielsen report, 2007-2014), while at the same time Netflix generates more revenue as well during the period. Seeing as this is negatively correlated to the revenue spikes in theatrical movies, it is plausible that there is a substitution effect. So this research will explore whether the quality of the movies is significantly less during some seasons, as well as investigate whether consumers are less likely to go to the movies in different periods because of substitute products like Netflix. This leads to the following research question:

What causes the fluctuation in seasonal box office revenue?

If there is significant evidence that the quality of movies is deteriorating and/or consumers are replacing theatrical movies with video on demand services or television shows, then it is logical that the movie industry has been stagnant over the past 7 years. So this leads to the objective of this paper. The objective of this paper is to investigate what causes the differences in theatrical movie revenue during the seasons in the United States by comparing different variables in the motion picture industry. The first aspect that will be investigated is if the movie quality differs during the seasons. The idea behind the 'dump months' is that movies that are of lesser quality come out during the winter period after the first weekend of January. Movies that debut after that period are not eligible for the Oscar awards in February that year, so movie distributors generally do not release their best movies during this period. The second aspect of this research investigates how substitution products like cable television and Netflix affects box office revenue. There is seasonality in Netflix earnings. The summer period is the lowest in terms of revenue generation, while the winter period is the highest. This paper reasons that consumers prefer spending their budget on theatrical movies instead of a Netflix account because there are better quality movies on display in the theaters. The belief is that movies of better quality are released during the summer. That causes consumers to switch their time watching cable television and Netflix and will instead visit theatrical movies. The last aspect of this research will investigate how movie specific attributes such as production budget affects box office revenue. The belief is that movies of better quality are displayed in more theaters across the country. That will consequently lead to more revenue. The same goes for production budget of a movie. The higher the budget, the more revenue a movie will make. This can also be a predictor why there is such a fluctuation in box office revenue during the different seasons. If movies during the dump months were to be displayed in more theaters, then it could potentially lead to higher revenue.

If these variables are predictors for the seasonality, then movie distributors can develop a strategy based on these variables in order to determine when to release their movie as to create more revenue. Another option is to know if they need to allocate more of their budget to a certain movie in a specific season depending on whether investing more will result in increased revenue. This paper will show what strategies studios can exploit during certain periods in order to maintain a competitive advantage over competitors.

The outline of this paper is described next. In order to answer the aforementioned research question, this paper will first describe the motion picture industry in more detail. The hypotheses are developed based on the theoretical background afterwards, followed by the variables that will represent the hypotheses. After this, the methodology section will describe the models used in this research and argue why the models are appropriate to use. Following the methodology, the data section will provide explanation on how the data was gathered, as well as some descriptive statistics on the variables used. An initial analysis will then be used to determine whether the linear model is correct or whether moderators or mediators should be added. After these analyses, the results section will interpret the results found in the data. Finally, there is a conclusion provided as well as the managerial implications and future research to improve on this paper.

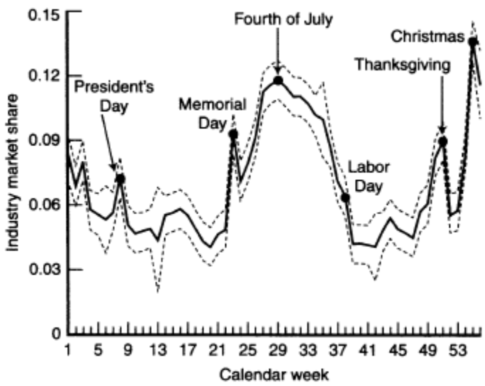
Chapter 2 – Theory

2.1 Literature Review

There has been a lot of research into the motion picture industry over the years. A lot of that research has been directed at what causes a theatrical movie to generate box office revenue (Terry et al., 2005; Hennig-Thurau et al., 2007; Albert, 1998). It is especially interesting for producers and distributors alike to know beforehand how much revenue a movie will generate. Most of that research split the motion picture industry in three main players; namely producers, distributors, and exhibitors (Gill, 2008; Einav, 2003). The producers basically can be split in two groups. On the one side there are small independent movie producers who have a small budget. On the other side are the producers that have been hired by large studios like *20TH CENTURY FOX*, *Warner Bros.*, or *Universal* that give the producers millions of dollars to spend as a movie budget. The distributors also consist fundamentally of two sides. There are the small independent distributors that try and get the movie to be screened in a theater. Opposite are the large studios that have a nation wide distribution. The last part of the industry comprises out of exhibitors, which are movie visitors.

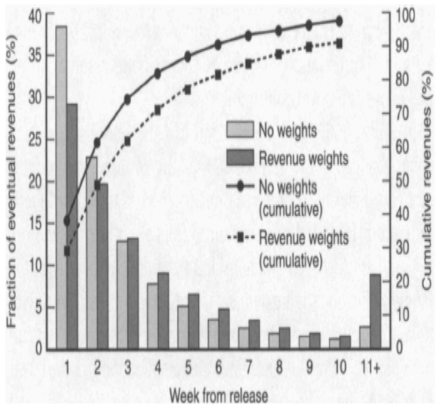
The movie industry in the United States sees very large variance in box office revenue during the seasons. Einav (2007) depicted a figure in his paper, showing the market share generated in revenue per week (see figure 1 next page). Based on the fluctuations, he described 4 different seasons during the seasons, which can also be depicted in the figure. They consist out of summer (starting from the last week of May (Memorial Day) to the first Monday of September (Labor Day)), fall, holiday season (last week of November (Thanksgiving) to mid January), and winter/spring. It can be concluded from the figure that quite a large variance in revenue creation occurs during the different seasons. This suggests that there must be one or more variables at play that cause the difference during the seasons. This research plans to expose the variable or variables affecting the difference. This will allow the distributors to figure out how much costs can be allocated to the movie without becoming unprofitable.

Figure 1: Seasonality Fluctuation



Source: Einav (2007)

Figure 2: Movie Revenue Life Cycle



Source: Einav (2007)

For the majority of the papers, the independent variables have differed considerably. This is because there are numerous factors that can influence movie revenue. The variables that are used throughout for predicting drivers of motion picture revenue success have a couple of large overarching sectors. One sector comprises out of movie quality variables like ratings (Terry et al., 2005; Basuroy et al., 2006; Dellarocas et al., 2007), and star power and award counts (Einav, 2007; Karniouchina, 2011). Another sector consists out of industry related variables, amongst others: timing (Corts, 2001; Lehmann and Weinberg, 2000; Terry et al., 2005), movie display length (Gil, 2008), and competition (Nam et al., 2015; Sullivan, 2009; Davis, 2006). The last main sector consists out of movie specific variables such as genre (Corts, 2001; Prag and Casavant, 1994), number of screens (Davis, 2006; De Vany and Walls, 1997), and budget (Hennig-Thurau et al., 2006; Eliashberg et al., 2006). These are the most frequently used variables for predicting the dependent variable, which in most cases is box office revenue.

An interesting fluctuation can be found in the revenue generated during different seasons. Krider and Weinberg (1998) found that a movie during the Christmas period, or in the summer, have been found to have a significantly higher revenue on average than during other seasons. More research into both the Christmas period and the summer season has uncovered diverse results. Where Litman (1983) found that the ideal period for movie releases is the Christmas period, Sochay (1994) established that, while both summer and Christmas launch dates have higher box office revenue, the summer period had a stronger positive relationship. He did note however, that the most successful season in terms of box office revenue could differ from year to year due to distributors trying to avoid competition. The majority of research agrees that movies released during the summer and Christmas holiday season generate more revenue on average than during winter/spring or fall. Radas and Shugan (1998) checked if the difference in revenue had anything to do with the average length that a movie is viewable in the theaters. They found no significant results. That is relatively easy to explain. Movies generally create the majority of their box office revenue within the first few weeks. Einav (2007) found that the within four weeks of the life cycle, a movie typically has generated over 80 percent of its total revenue (see figure 2, previous page). This indicates that other factors are the cause for the higher revenue. Einav also found in 2001 that “they (the distributors) may consider releasing some of their big budget movies later in the summer or in January, rather than around Memorial Day, when underlying demand has yet to peak”. What is also noteworthy from his research is that the number of movies that is released during the periods follows a different pattern. This is explained due to the smaller budget movie releases avoiding competitive seasons such as summer and the holiday period.

The last aspect that is affecting box office revenue is competition from substitute products. Sullivan (2010) notes that the motion picture industry is losing money primarily on DVD sales. Consumers are less stimulated to physically rent movies, and rather use video on demand (VOD) services via online service providers. Consumers pay a fixed monthly fee to online service providers like Netflix, Hulu, or Amazon Prime. In return they will get unlimited access to movies and television shows alike. The online service providers in their turn pay an undisclosed amount to the distributors in order to add the movie to their online library. The rise of online streaming services like Netflix can cause people to substitute theatrical movies with online streaming movies at home. This causes less revenue, and can be a good indication for the stagnant motion picture industry and rising VOD service market. Nam et al. (2015) concluded that the preference for VOD channels is similar to or greater than that for DVD. Of the online service providers out there, Netflix is the biggest in terms of consumer base and revenue generation. Netflix was founded in 2002 in the United States. The total amount of subscribers for Netflix has gone from approximately 1,932,000 total subscribers in Q1 2004 to over 41 million total subscribers in Q1 2015. With that consumer base, Netflix has become the leading subscription service provider for online movies and television shows. As of Q4 2014, 36 percent of the US households have a subscription to Netflix (Nielsen report, 2014). That makes Netflix an adequate source for measuring VOD services. Furthermore, it also shows that the online streaming market is on the rise seeing as how the subscriber amount skyrocketed from 2 million to nearly 41 million subscribers. Furthermore, Nielsen report notes that households in the US are watching more cable television over the last couple of years. A scatterplot from the data on households watching television per quarter shows that there indeed is a rise in consumers watching cable television (appendix 1). Unfortunately, there is no scientific article known to the author that has investigated if there is a causal relationship between the time consumers spend watching cable television and the increase/decrease in box office revenue during a period in the US.

The online streaming services and television usage increase are cause for a substitution effect of theatrical movies. The substitution principle contains that the conditional probability of consumers choosing product i at time t will affect the probability of choosing product j at time t. But suppose the price of product i increases, the substitution principle entails that consumers would switch from product i to product j because it would maximize the utility function of the consumer. An easy example of this principle is as follows: a consumer normally buys 10 pads of coffee and 2 thee bags during the week. But if the price of coffee increases, then the substitution principle will likely make the consumer switch to buying 5 pads of coffee and 5 thee bags. This substitution principle can also be applied to the motion picture industry. The difference in movie revenue is significantly different between the winter/spring and the holiday season of the year. Because movies are an experiential product (Cooper-Martin, 1991), it will be interesting to see if that can lead to consumers substituting it with other experiential products.

2.2 Hypotheses and Conceptual Model

This paper will investigate 3 different segments of movie revenue predictors. First, this research will investigate whether the quality of the movies differs significantly between the different seasons and whether it can explain the difference in revenue. Movie quality is split in two different hypotheses, namely ratings and star power. The second segment will determine how the substitution effect influences movie revenue. The last segment covers movie specific attributes. This is described in more detail below. This leads to four main independent variables. Lastly, the fifth hypothesis will test if the seasons have a moderator effect on the independent variables if these show to have significant variance between seasons.

As stated above, this research will split up movie quality in two different variables. The first variable that will be tested is how the rating given to movie *i* will affect the revenue of movie *i*. Movies that receive a high rating from either critics or ordinary viewers correspond with increased box office revenue according to Reinstein and Snyder (2005). Ravid (1999) found MPAA ratings to be significant variables in his regressions and that the relationship was positive. There is even more research into this relationship (e.g. Terry et al., 2005; Basuroy et al., 2006), all implying the same relationship between ratings and revenue. 'Movie reviews provide consumers with presumably professional 'objective' information and have been shown to correlate with box-office results' (Hennig-Thurau et al., 2006**).** Because the research shows that ratings lead to higher revenue, this research will expect to find that ratings in the winter/spring and fall period will be lower seeing as Einav (2007) showed how these two seasons had low revenue on average. That leads to the belief that ratings matter less in the winter/spring and fall periods. This leads to the first hypothesis being as follows:

H1: A higher rating for movies lead to higher box office revenues for theatrical movies.

The second hypothesis that is also partially explaining movie quality is star power. There is mixed results in scientific research that investigate the effect star power has on box office revenue. Albert (1998) examines the effect that stars have on a movie. He concludes that they are particularly important in getting a movie financial budget, but in no means does that entail financial success. Yang and Selvaretnam (2015) found a significant positive relationship, just as Karniouchina (2011). She found that stars have an impact on revenue, primarily due to their ability to generate buzz and drive audiences to the theaters during the opening week. Ravid (1999) also found a positive impact. He found that "stars signal high returns or at least high revenues". Seeing as there are that many papers displaying a positively correlated relationships between stars and revenue, it can be argued that star power does create bigger box office revenue. To link this to seasonality, the follow-up should be that the star power variable is most relevant during these periods. This is based on that the average revenue is higher during summer and holiday season. With that being said, the hypothesis with regards to star power will be:

H2: Bigger star power creates bigger box office revenue for movies.

The third hypothesis will test what effect substitute products have on box office revenue. Seeing as the motion picture industry has been stagnant with regards to total box office revenue over the past eight years, a likely theory is that consumers are replacing the theatrical movie for other products. There has been little empirical research whether video on demand (VOD) services and cable television have an effect on box office revenue. Related to VOD services though are DVD sales. Sullivan (2010) found that the motion picture industry is losing revenue because of a drop in DVD sales. Building on this is the research from Nam et al. (2015). They concluded that the preference for VOD channels is similar to or greater than that for DVD. Therefor, this research will analyze if VOD services have a substitution effect for theatre movies. Consumers will be stimulated less to physically rent movies. Instead, they will be more likely to make use of VOD services. Moreover, seeing as there is a fixed fee per month for VOD services, consumers are incentivized to make more use of such a service (Danaher, 2002) rather than a service that does not pertain a subscription fee[1]. Moreover, statistics from Nielsen report (2007-2014) show that US households make increased usage of cable television (appendix 1). This leads to belief that a substitution effect may be at play. Consumers are replacing theatrical movies for other products like Netflix and cable television. This paper expects to find that this relationship will be strongest during the winter, seeing as television shows are starting to air again in the beginning of January. Furthermore, we expect the relationship to be the weakest during the summer because the theatrical movies are generating a lot of revenue. This being said, it leads to the following hypothesis:

H3: Increased usage of substitute products causes a decrease in theatrical box office revenue.

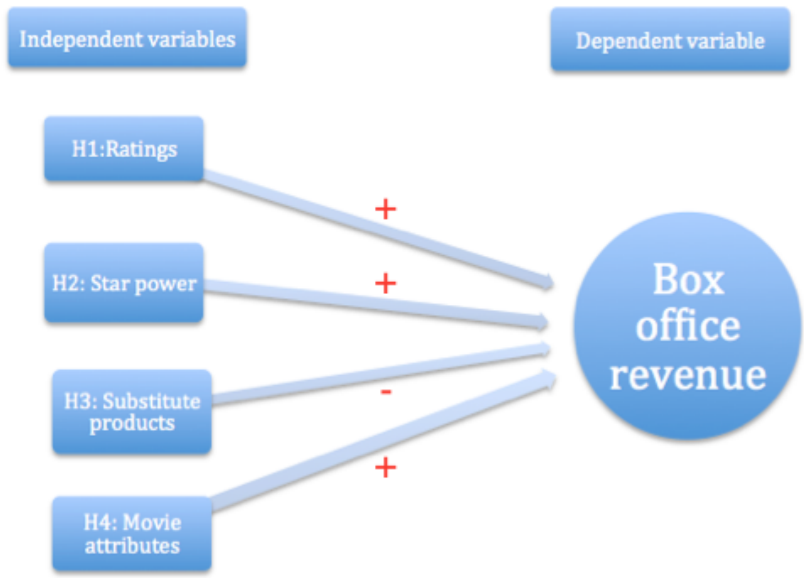
The last hypothesis will test movie specific statistics. Continuous variables like budget (Hennig-Thurau et al., 2006; Eliashberg et al., 2006) and number of screens (De Vany and Walls, 1997) have shown that there are significant positive relationships between these movie specific elements and box office revenue. For instance, the number of opening screens on the success of the movie has a positive effect according to (Hennig-Thurau et al., 2006). On the other side, it has been shown that there is stiff competition in the motion picture industry due to a saturated market. The stiff competition can lead to high entry costs and new products primarily cannibalizing the revenues of existing products rather than expanding the market (Davis, 2006). The high competition can lead to distributors allocating more costs for their movie being played in more theaters, while less revenue is simultaneously generated. Therefore, it might be possible that there is a negative relationship. However, due to more scientific papers have seen positive relationships, this paper will follow suit. Furthermore, this research expects that the relationship will be strongest for the summer and holiday season seeing as these time periods generate the most revenue on average. The last hypothesis will be as follows:

H4: Higher movie specific attributes have a positive effect on box office revenue.

---

[1] Acknowledged, there are certain memberships based on a monthly subscription fee at certain theaters, but this research assumes that the majority of visitors of theatrical movies do not own a subscription.

The hypotheses will be tested against the dependent variable box office revenue, which will be measured in a logarithmic form. The hypotheses are first measured in the baseline model, without checking the influence of seasonality. Thus the conceptual model depicted below is the model in its most basic form. The hypotheses and their correlation to the dependent variable revenue are shown in the following conceptual model, as is depicted below:



As stated at the beginning of this section, this model will be checked against seasonality seeing as that the primary goal of this paper. The fifth hypothesis is that seasonality has a significant impact on box office revenue predictability. The independent variables will be checked against seasonality if they appear to have significant fluctuation during the different seasons. This will be explained more elaborate during the methodology part. That being said, the fifth hypothesis will be as follows:

H5: Factoring for seasonality will predict the fluctuations in revenue more than the baseline model.

2.3 Variables

The following section will briefly highlight how the different hypotheses will be measured. The hypotheses ratings, substitute products, and movie aspects are measured by taking the component of multiple variables. These variables are subsequently combined to generate a variable that most effectively measures the hypothesis in question.

Rating

The variable Rating is a variable that is a combination between Critics Ratings and Amateur Ratings. Professional critics commonly provide reviews and ratings; this information signals unobservable product quality and helps consumers make good choices (Boulding and Kirmani 1993; Kirmani and Rao 2000). Although amateur consumers can obtain useful information from critics, they are sometimes at odds with critics because of some fundamental differences between the two groups in terms of experiences and preferences (Moon et al., 2010). Therefore, amateur ratings do have a slightly different view when experiencing a box office movie. However, the assumption is that both variables are highly correlated. In order to prevent multicollinearity, while at the same time take both rating systems into account, a new variable called Rating is created. Using a principal component analysis on critics ratings and amateur ratings derives this variable. Principal component analysis (PCA) forms the basis for multivariate data analysis. It combines the dominant patterns in a matrix data set that complement each other. It creates a score based on this. The closer the correlation of the variables, 'the fewer terms are needed in the expansion to achieve a certain approximation goodness' (Wold et al., 1987). The component matrix table, which contains the component loadings, can be seen in appendix 2. The component matrix shows the correlations between the variables and the component.

Star power

The hypothesis star power should be positively correlated with box office revenue according to previous research. This variable has been measured using different techniques. Consequently, different results have been found. This is due because star power is prone to subjectivity. Where one person may prefer a certain actor, another may dislike that specific actor. Karniouchina (2011) used movie buzz generated to rate star power. They measured it via IMDb searches. Yang and Selvaretnam (2015) came up with 2 different methods to measure star power; one being via Academie award count and the other via average box office revenue sales of their most memorable movies. The latter found a significant relationship and therefor this paper will see if it is consistent. The reason to use this technique is because star power fluctuates over time. By using average box office revenue, it will lead to different results. For instance, Robert Downey Jr. was not the most well known actor (known mostly for Gothica, U.S. Marshalls, and Zodiac) before he made the Iron Man trilogy and The Avenger movies. Afterwards, his star power had risen substantially. It would not be adequate to use the same level of star power over different time periods. The variable consists out of the average of the top four movies in terms of box office revenue for the actor or actress up to that point in time. If that subsequent movie generates more revenue than one of the previously used four movies, that movie will replace the least profitable movie of the four. Furthermore, seeing as the star power will be measured in average revenue, the variable will be in log form. Using the natural logarithm will account for possible outliers and will create a normally distributed variable.

Substitution products

Seeing as substitution products for movies are not widely researched, it is difficult to determine what exactly are substitutable products for theatrical movies. The research that has been done (Sullivan, 2010; Nam et al., 2015) does come up with Netflix and cable television as possible substitutes. Therefor, the variable will be a principle component of the revenue that is added per quarter to Netflix and the amount of time people watch television in hours per quarter. The variable will be generated using the same principle component analysis as with the Rating variable described above. The component matrix table for substitution can be found in appendix 3.

Movie attributes

The hypothesis movie attributes assumes that a positive relationship between movie specific attributes and revenue per movie exists. This hypothesis will be tested measuring two specific attributes for movies, which have been found to be significant in previous research. Davis (2002) found that theaters have a significant impact. Terry et al. (2005) found that an increase of theaters displaying movie *i* is significant and positively correlated to box office revenue. They argue that 'success can easily be explained by the fact that a wide release has an easier time finding an audience and is probably a product of one of the major motion picture studios with access to proper marketing channels and box office movie stars like Tom Cruise and Julia Roberts or a box office franchise like the Star Wars saga'. Therefor, this paper will use theaters as one specific aspect, while the other aspect being used is budget[2] (Hennig-Thurau et al., 2007). The two variables will likely correlate seeing as the more budget is used for creating a movie, will lead to more screens being used in order to generate revenue. Therefor, the principle component is used again to create the variable Movie attributes. The results for the component loading are found in appendix 4. The variable Movie Attributes will not make use of aspects such as genre, distributors, or directors. The reason to exclude directors and distributors is because no evidence has been found that either have a significant effect on box office revenue. The reason to eliminate genre from the variable is because it is a dichotomous character, meaning that a movie can be a romantic comedy for instance, hence it is dubious to either classify it explicitly as a comedy or as a romance.

Chapter 3 – Methodology

To test the hypotheses as outlined above, this research will test whether the independent variables ratings, star power, substitution products, and movie aspects significantly affect the dependent variable box office revenue. This will be done using a linear regression test. What sets this research apart from previous research is that it will inspect how the independent variables affect box office revenue during the different seasons. The seasons consist out of summer, fall, holiday season, and winter/spring. The seasons that will be used in this research are defined by Einav (2007). In order to test the effect that seasonality has on predicting box office revenue, there will be a linear model used in this research. The model will test the independent variables against four dummy variables, which represent seasonality. However, the independent variables will only be checked against seasonality if there is a significant difference between the means during the different seasons. Otherwise the variables are used as control variables as there is no significant difference between the groups. In order to test this difference, an analysis of variance (ANOVA) test will be performed on the independent variables to test if there is subsequent difference

[2] Because the production budget for a movie includes salary of the stars, the budget variable used will be calculated using an Ordinary Least Squares to obtain a variable that does not include the variance of salary. For more information, see data section.

between the seasons. The formula for the ANOVA test is as follows (Kerlinger, 1964):

$$\eta^2 = \frac{SS_A}{SS_T} \tag{1}$$

where SS_A is the between sum of squares for factor A, while SS_T is the total sum of squares. η^2 is simply the proportion of the total SS (or variance) associated with A (Cohen, 1973). If the ANOVA test shows that there is significant difference between the groups, then a post hoc analysis will be added to show how the groups differ from one another during the seasons. This will help towards understanding how seasonality affects the variance in box office revenue.

Afterwards, a linear regression model will be applied to measure the effects that the independent variables have on the dependent variable. The general function of a linear regression looks as follows:

$$y_{i,t} = \beta_{0,i} + \beta_1 x_{1,i,t} + \beta_2 x_{2,i,t} + \ldots + \beta_k x_{k,i,t} + \epsilon_{i,t} \tag{2}$$

Next, when inserting the suggested variables, we get the following model:

$$LogRevenue_{i,t} = \beta_{0,i} + \beta_1 Rating_{i,t} + \beta_2 LogStar_{i,t} + \beta_3 Substitution_{i,t} + \beta_4 Attributes_{i,t} + \epsilon_{i,t} \tag{3}$$

where $LogRevenue_{i,t}$ denotes the natural logarithm of total revenue for movie i at time *t*. β_0 denotes a vector of intercept parameters. The β_1 measures the effect that the rating of movies in quarter i at time t has on revenue, and β_2 measures the average star power of movies in quarter *i* at time *t*. The star power is measured in a logarithm function to account for possible outliers. β_3 signifies the substitution effect, and β_4 measures the effect that movie specific attributes has on consumers in quarter *i* during period *t*. The $\epsilon_{i,t}$ measures the error effect of quarter i at time *t*. Also note that β_3 is positive in the equation, however in the hypotheses section is already acknowledged that the expectation is that the variable will negatively impact the revenue.

However, the model in formula (3) does not check how seasonality influences the box office revenue. In order to implement this, four dummy variables have been created to represent the four different seasons. To measure the effect of the variables during the different seasons, there are two alternative models to use. Suits (1957) explained how to implement dummy variables in linear regression model. Furthermore, he explained how an interaction effect between independent variables and dummy variables work. The first option is to create a linear model where one of the dummies is set to 0. Suppose that there is a dataset with 3 dummy variables, then the following linear model is constructed:

$$Y = \alpha_0 + \alpha_1 X + \beta_1 R_1 + \ \beta_2 R_2 \tag{4}$$

where α_0 measures the constant, $\alpha_1 X$ measures variable coefficient of variable X, and the R's measure the effect of the two dummies in regards to the third dummy, which is implemented into the base line and the $\alpha_1 X$. Another way to receive the same results is to insert the last dummy variable as well. It will subtract the coefficients of the last dummy variable from the constant, as well as all the other dummy variables. It will create the following linear model:

$$Y = \alpha_0 + \alpha_1 X + \beta_1 R_1 + \ \beta_2 R_2 + \beta_3 R_3 + \epsilon \tag{5}$$

where u is unaffected when the constant k is added to each beta value for the dummies and subtracted from the α_0. Formula (5) is easier to interpret and explain because it is possible to compare all the factors, or in this particular seasons to each other, instead of three of the dummy variables to one that is injected into the baseline. This procedure is explained more in detail in Suits (1984). The last step is to create an interaction effect between the dummy variables and the independent variables. The standard model will look as follows:

$$Y = (\alpha + d_1 R_1 + d_2 R_2)X + \ \beta_1 R_1 + \beta_2 R_2 + \epsilon \tag{6}$$

This particular formula with the interaction effect is based upon formula (4), where the last dummy variable is intertwined within the other variables as a baseline variable. For this research, and assuming that all independent variables have significant different means between groups (obtained from the ANOVA test described above), the entire linear model will be as follows:

$$LogRevenue_{i,t} = \beta_{0,t} + \sum \beta_i S_i \left(Rating_{i,t} + LogStar_{i,t} + Substitution_{i,t} + Attributes_{i,t}\right) + \sum \beta_i S_i + \epsilon_{i,t} \tag{7}$$

where $S_{i,t}$ in $\sum \beta_i S_{i,t}$ represents season i at time t, and β_i represents the corresponding coefficient of season i. As previously explained in formula (5), the final model will include all the dummy variables and its corresponding interactions (Suits, 1957). Initial analysis on the independent variables are done in the analysis section below, followed by an updated model with control variables if need be. Furthermore, the model will also be checked for interaction effects of possible moderators and/or mediators as well. This is excluding the moderating effect that seasonality has on the baseline model as that is tested regardless during this model.

Chapter 4 – Data

4.1 Dataset

The dataset used to test the model is collected solely from numbers related to the United States. This is because of the limited amount of data available across countries, as well as the scope becoming too wide. The time period for which the data has been gathered is from 2007 until 2014. The reason behind this is because that is starting point when the movie industry started to become stagnant. The dependent variable numbers is movie revenue, which have been obtained from the website boxofficemojo.com. The website describes itself as 'the leading online box-office reporting service'. Besides revenue numbers, the independent variables star power and the movie specific attributes budget as well as number of screens are collected via boxofficemojo.com as well. The data for the ratings variables is collected from the websites IMDb.com and RottenTomatoes.com for amateur and critics ratings respectively. The Netflix data is gathered from quarterly earning reports, letters to their shareholders, and financial statements. These are available on their website. The data used for Netflix is from the domestic balance sheets only as this is a study in regards to United States numbers. The cable television statistics are gathered from independent website Nielsen.com. They include only US household statistics with regards to cable television hours. They acquire their data through panels and other databases.

To recap, the dependent variable box office revenue, as well as the variables star power, critics ratings, Netflix revenue, television usage, and number of theaters are all based on United States statistical numbers solely. The amateur ratings variable unfortunately takes into account foreign voters as well, but because the voting frequency per movie is considerably large, it is argued that United States voters do not differ significantly in opinion from other foreign voters. The variable production budget does not include the costs of advertising, nor distribution. It is purely associated with the cost to produce a movie. Therefor, the budget is also allowed and does not need to be transformed or excluded from the dataset.

There are two constraints to the data. The first is in regards to the theory of Einav. Where Einav (2007) defined the 4 major seasons at the start and end of major American federal holidays, which are for winter starting halfway through January, the summer season starts roughly around Memorial Day, fall starts after Labor Day, and the holiday season starts around Thanksgiving and lasts till mid January. The dataset differs slightly from these seasons. In the dataset, winter starts from the first weekend after New Year, summer starts the first Friday of May and lasts till Labor Day, and the holiday season starts the first Friday of November. This is because boxofficemojo, the provider of movie revenue data, defined seasons these particular starting dates. It differs slightly from the theory by Einav however. It is not a major issue, but should be noted nonetheless that it differs slightly. The other constraint to the data is that the substitution variable is measured per quarter. So the substitution data for movie i at time t is the same as for movie j. This implies that it will be far less accurate to explain the variance. However, it is included nonetheless to try and see if it affects box office revenue.

As mentioned before, the time period is from 2007 until 2014. That is 8 years with 4 different seasons. Because it is impossible to inspect every movie created during this time period, a sample of the population has been taken to represent a particular season. To adequately account for the revenue made during the different seasons, the sample size per season consists out of the 20 biggest revenue-creating movies. The only exception to this is in regards to the season Holiday. Initially, 20 movies during the holiday season were measured as well. The problem arose in skewness of the data for this particular season. Whereas the other seasons would measure around 70 percent of the total profit with the top twenty movies, using the top twenty most profitable movies during the holiday season would measure around 81 percent. However, the latter ten movies added around 20 percent of total revenue. This is due to the fact that the skewness in this season is very high. There is no particular explanation for this skewness. Nonetheless, if this paper were to use twenty movies during the holiday period as well, the average would drop considerably and therefor would not represent the holiday season adequately in terms of revenue generation. Therefor, using more than ten movies results in less proficient representation of the seasonal fluctuation in revenue averages. Overall, the sample size accounts for 68 percent of the total revenue generated during all the seasons.

4.2 Variable Computation

Before the descriptive statistics of the variables can be executed, the individual variables must first be computed. As is stated in the variables section above, the ratings, substitution, and movie attributes variables must be extracted. Ratings will be a variable that will be the component of the critics and amateur ratings. Appendix 5 displays the bivariate correlation of both variables. There is a high correlation between both variables in the data sample as was expected, namely .753. There are 554 critic ratings and 553 amateur ratings. This is due to 6 movies being 3D remakes, which have no rating and one movie that had no amateur rating. Because of the high correlation, a principal component analysis has been done to create the variable Rating. The same has been done in order to create the variable substitute products. The correlation between both variables is not as high as the critics and amateur ratings (.439, appendix 6), but is still substantially correlated. Therefor it is valid to use the component between the two variables in order to explain box office revenue. As for movie specific attributes, the two variables production budget and number of screens are used. A high budget means that the movie can employ high-profile stars, but high-profile stars generally also attract financing, which in turn enables a higher production budget. Seeing as how actors with higher salary will produce more revenue (Wallace et al., 1993), multicollinearity might become a problem. In order to counter the threat of multicollinearity, an Ordinary Least Squares method will be used to obtain a budget variable that is not influenced by the effect of star power. This procedure of eliminating the threat of multicollinearity by reducing the variance of budget explained by star power has been done in research before (Ter Braak et al., 2013; Batra et al., 2000). The following model is used to obtain the new variable:

$$Budget = \beta_0 + \beta_1 Star + \epsilon \tag{8}$$

where the error term of the model explains the variance in the dependent variable Budget that is not explained by the independent variable Star. These residuals make up the new variable Starless Budget. Starless Budget is the budget for movie i that is not influenced by the star power of movie i. The use of the residuals makes sure that there is no correlation between star power and budget. This new budget variable is subsequently checked for correlation with theaters in appendix 7. The correlation of .586 is relatively high. Because of that, a new variable based on the principal component method has been created. The variable is named movie attributes.

4.3 Descriptive Statistics

Seeing as the time frame is 8 years, with 70 movies analyzed per year, the movie sample size is 560. Appendix 8 shows the descriptive statistics for the dependent variable Log Revenue. Furthermore, an ANOVA test has been done in order to see if the means between the different seasons differ significantly from another (Appendix 9). Seeing as seasons have a significant effect on box office revenue (sig. <.001), a post hoc analysis has been added to see how the seasons differ. The homogeneity test of variance is significant, therefor a Dunnett C post hoc test has been used for further analysis. The results of the Dunnett C test are displayed in the table on the next page. It shows that all seasons differ significantly (measured at .05 significance level) in revenue means from another. The statistics for the independent variables are shown in appendix 10. As can be seen from the table, there is some missing data. The reason for missing data for the attributes variable is because there was no data made public on the budget for some particular movies. The sample size of 407 is still regarded as sufficient. The rule of thumb is that the sample size should be multiplied by 5 for each parameter (Janssens et al., 2008). The reason that the sample size of star power is 495 is because animated movies are not measured for star power. Furthermore, what can be seen from the descriptive statistics of the independent variables is that the variables are normally distributed (skewness and kurtosis levels within the value of -2 and 2 are accepted). Furthermore, a bivariate correlation has been performed (appendix 11) to test multicollinearity. As can be depicted from the output, the highest collinearity is between rating and log star power (.225). Janssens et al. (2008) say that a correlation of .60 can be used to define multicollinearity. Seeing as the independent variables do not even come close to that level, no multicollinearity exists. Another alternative to check for multicollinearity is to check the VIF statistics. The independent variables are tested against each other, and none show a VIF level exceeding 2 (appendix 12), indicating that there is no multicollinearity (Kahn and Mentzer, 1998). Therefore, it is safe to say that multicollinearity will not be a problem.

Table 1: Post Hoc Results for Revenue Means

Multiple Comparisons

Dependent Variable: Log Revenue
Dunnett C

(I) Quarter	(J) Quarter	Mean Difference (I-J)	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Winter/Spring	Summer	-.40412*	.07075	-.5878	-.2204
	Fall	.61369*	.06601	.4423	.7851
	Holiday	-.69759*	.06536	-.8685	-.5267
Summer	Winter/Spring	.40412*	.07075	.2204	.5878
	Fall	1.01782*	.08031	.8093	1.2263
	Holiday	-.29346*	.07978	-.5016	-.0853
Fall	Winter/Spring	-.61369*	.06601	-.7851	-.4423
	Summer	-1.01782*	.08031	-1.2263	-.8093
	Holiday	-1.31128*	.07561	-1.5086	-1.1139
Holiday	Winter/Spring	.69759*	.06536	.5267	.8685
	Summer	.29346*	.07978	.0853	.5016
	Fall	1.31128*	.07561	1.1139	1.5086

*. The mean difference is significant at the 0.05 level.

Chapter 5 – Analysis

An initial linear regression has been performed to see how the variables fit without checking for seasonality. The results can be found in appendix 13. The R^2 shows that the variance in box office revenue is for 61,4 % explained by the four independent variables. The results also show that all four variables have a significant effect on box office revenue as can be seen from table 2. The variables rating, star power, and attributes are all significant (<.001), while the substitution variable is significant at .017. Furthermore, the model has been checked for non-linear relationship and whether there is additive relationship between the dependent and independent variables. None had a significant change in R^2 according to the model summary. Therefore, no interaction or squared variable has to be added to the model.

Table 2: Baseline Regression Model Coefficients

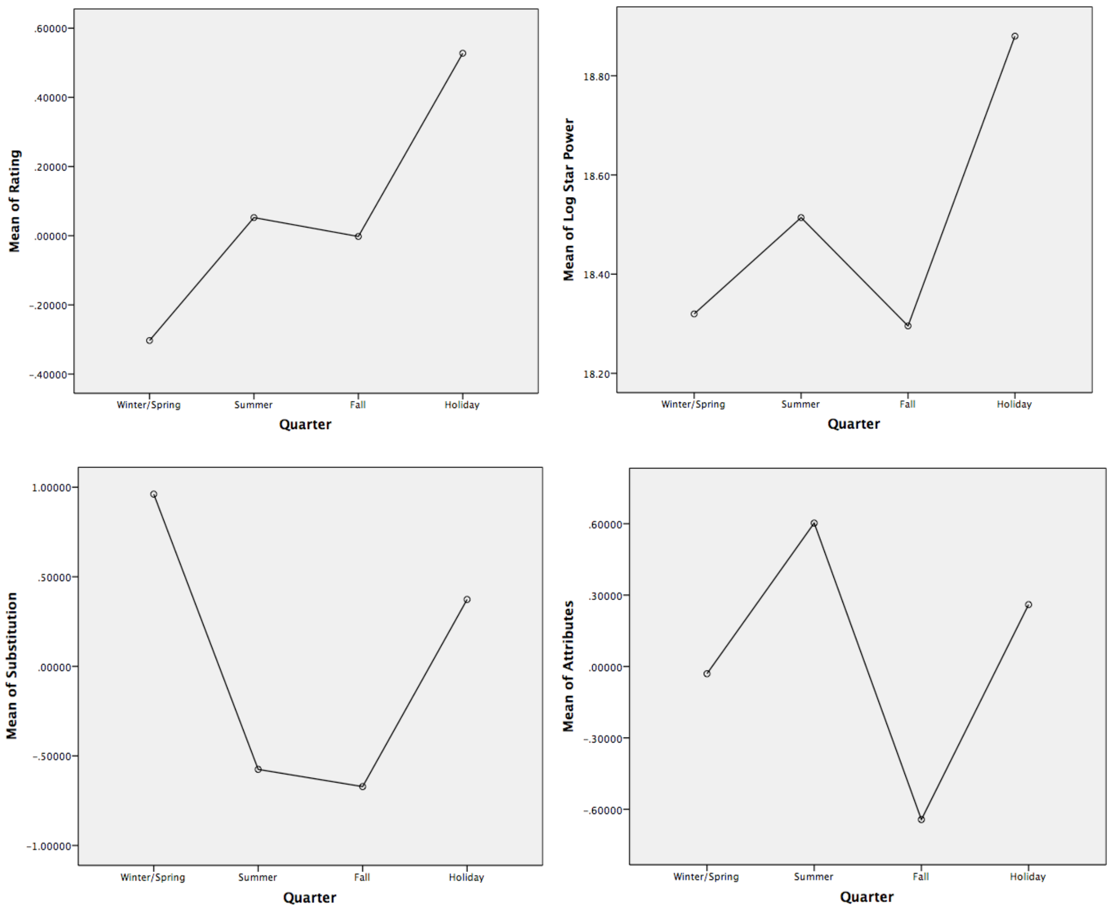
Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	14.686	.693		21.178	.000
	Rating	.192	.026	.237	7.453	.000
	Log Star Power	.191	.038	.165	5.095	.000
	Substitution	.056	.023	.075	2.403	.017
	Attributes	.496	.024	.658	20.603	.000

a. Dependent Variable: Log Revenue

In order to check which of the independent variables should be set as control variable an ANOVA test is used. The ANOVA test will check if there is significant variance between the four different seasons. If the variance is not significant, the variable will function as control variable in the eventual model. The results for the ANOVA test can be found in appendix 14. The ANOVA test shows that all the variables are significant (<.001). That indicates that the variance between the means of the seasons differ significantly. The homogeneity test shows that only the variable rating should assume equal variances (appendix 14). Therefore, for the post hoc analysis of rating, Tukey will be used. For the remainder of the variables, Dunnett C will be used. The results for Rating can be found in appendix 15. It shows that only the ratings for summer and fall are the only statistically insignificant means. The post hoc results (appendix 16) for star power shows that the holiday season is significantly higher than all other variables. Other than that, the test also shows that winter/spring is insignificant from summer and fall, while summer has significantly higher star power than fall. For substitution, the Dunnett C post hoc test (appendix 17) shows that the only insignificant seasons are summer and fall. Other than that, the summer season has the highest substitution usage among consumers. The last post hoc test shows whether the means of the holidays for attributes are significant among each other (appendix 18). The means for the independent variables can be seen in figure 3. What can be seen from the graphs is that rating and star power follow a comparable pattern, which is to be expected as both measure movie quality. Furthermore, substitute product levels are highest during the winter as was expected, and lowest during the fall.

Figure 3: Means of the Independent Variables



Now that independent variables have been found to have significant differences among the groups, the linear model can be made for seasonality. Because all the variables are significant, they will be used to see how seasonality affects box office revenue. The dummy variables, that represent the different seasons in the linear model, will be multiplied with the independent variables. This will create 4 variables for each independent variable to represent that variable during a specific season. The linear model will be the same as the one from formula (7). The expectation for the model is that it will explain the variance in box office revenue more than the basic model. The results for the baseline model are found in table 2. The R^2 in the final model should be higher than the .614 from the basic model in order for the fifth hypothesis to be accepted. The reason is because there is significant difference in seasonality for all the variables.

Chapter 6 – Results

The results for the linear model with seasonality are depicted in appendix 19. As expected, the inclusion of seasonality on the basic model creates a more accurate prediction of the linear model. The R^2 of the linear model is .713. Because there are so many more variables inserted, the adjusted R^2 will be used to adequately predict the variance explained by the independent variables. In this case, the adjusted R^2 predicts 69,9 percent of the variable. This is significantly higher (<.001) than the 61,4 percent that predicts the basic model as can be seen in table 3.

Table 3: R Squared Change Test

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.784[a]	.614	.610	.47749	.614	158.897	4	399	.000
2	.844[b]	.713	.699	.41992	.099	8.794	15	384	.000

a. Predictors: (Constant), Attributes, Substitution, Rating, Log Star Power

b. Predictors: (Constant), Attributes, Substitution, Rating, Log Star Power, Holiday, Attributes Winter, Rating Winter, Substitution Summer, Rating Summer, Attributes Fall, Substitution Fall, Star Power Winter, Star Power Fall, Rating Fall, Attributes Summer, Substitution Winter, Fall, Winter/Spring, Star Power Summer

Following this result, hypothesis 5 is confirmed. Factoring the data for the different seasons does improve the linear model. The seasonal linear regression coefficients can be found on the next page. One note from the results is that the dummy variable winter has been excluded from the model. This is because it is correlated too much with the constant in the model. Other than that, all variables are approved in the model. From the results, the respective seasonal regression models look as follows:

$$LogRevenue_{i,winter} = 15.627 + .072Ratings_i + .142Star_i - .009Substitution + .322Attributes_i$$

$$LogRevenue_{i,summer} = 13.840 + .254Ratings_i + .242Star_i - .064Substitution + .392Attributes_i$$

$$LogRevenue_{i,fall} = 17.025 + .237Ratings_i + .050Star_i - .066Substitution + .614Attributes_i$$

$$LogRevenue_{i,holiday} = 12.990 + .120Ratings_i + .302Star_i + .002Substitution + .239Attributes_i$$

In the equations, the dummy variables are included within the constant. Furthermore, the independent variables have been plotted against the dependent variable in a scatterplot. These can be found in appendix 20. They can be observed to see how the independent variables influence the revenue, while being controlled for per season. Table 4 shows the coefficients and their significance levels for the model.

Table 4: Coefficients for Seasonal Variables

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	15.627	1.151		13.581	.000
	Summer	-1.787	1.646	-1.069	-1.086	.278
	Fall	1.398	1.575	.819	.887	.375
	Holiday	-2.637	2.668	-1.193	-.988	.324
	Rating Winter	.072	.051	.041	1.410	.159
	Rating Summer	.254	.045	.162	5.637	.000
	Rating Fall	.237	.040	.172	5.992	.000
	Rating Holiday	.120	.066	.059	1.825	.069
	Star Power Winter	.142	.063	1.545	2.261	.024
	Star Power Summer	.242	.063	2.684	3.805	.000
	Star Power Fall	.050	.058	.538	.857	.392
	Star Power Holiday	.302	.127	2.572	2.368	.018
	Substitution Winter	-.009	.038	-.009	-.240	.810
	Substitution Summer	-.064	.069	-.033	-.926	.355
	Substitution Fall	-.066	.068	-.036	-.974	.331
	Substitution Holiday	.002	.121	.001	.019	.985
	Attributes Winter	.322	.051	.177	6.365	.000
	Attributes Summer	.392	.039	.337	10.075	.000
	Attributes Fall	.614	.054	.389	11.301	.000
	Attributes Holiday	.239	.065	.109	3.702	.000

a. Dependent Variable: Log Revenue

There are a couple of interesting results in the table. The first noteworthy result is that the dummy variables are insignificant. This implies that the differences in revenue between the seasons are not significant on their own. The fluctuation between the seasons becomes significant because of the variance in the independent variables. Another result, which is remarkable when comparing to the baseline linear model from above, is that all the substitution variables have become insignificant. This can possibly be explained due to the data having identical numerical values for different movies in the same quarter. Unfortunately this is a flaw in the dataset.

When looking at the rating coefficients for the different seasons, also a couple of statistics are noticeable. Ratings are insignificant during the winter and holiday seasons. These are different from the expectations of the hypothesis section. This paper reasoned that ratings were particularly important during the summer and holiday seasons because of the higher revenue averages. A reason for these particular seasons having insignificant ratings can be found in the means for ratings (see figure 3, page 23). The ratings for movies during the winter periods are considerably low (-.303). This can indicate that the movies during the winter period are of such low quality, that it does not matter to the audience how good the movie is. Opposite of this is that the ratings during the holiday period are relatively good on average (.528). This can indicate that because the movie quality is so high, it is less relevant as a deciding factor to determine why consumers decide to go to a particular movie during the period.

Another peculiar statistic in the results is the fact that star power is insignificant during the fall. The star power level during the fall is very low. As can be seen from the scatter plot (appendix 20), the mean lies relatively lower than the other variables. The coefficients for summer and holiday seasons are relatively high at .242 and .302 respectively. It makes sense that the coefficients are this large. The average star power during both seasons is already quite high. That entails that the higher the star power of the movie during these seasons; the more it will lead to increased revenue. This is insightful for studios as it displays that star power is a relatively important factor in order to differentiate from competition during these two seasons. What is also interesting is that the star power during the winter period is also significant at .024. The mean for star power during the winter is the second lowest at 18.32

approximately (appendix 14). This implies that there is an opportunity for studios to distinguish themselves. Seeing as the competition in star power is low during the winter, studios that will differentiate their movies by using bigger stars will create more revenue.

Chapter 7 – Conclusion

7.1 Main Findings

The main finding from this research is that seasonality significantly impacts box office revenue patterns. Where the basic model depicts that all independent variables are significantly influencing box office revenue, the seasonality model shows different results.

In order to gain a competitive advantage, studios must develop a strategy based on their release date. The findings of this research conclude that the summer and holiday period are the most competitive seasons in regards to box office revenue creation. Einav (2001) came to the conclusion that distributors may want to consider releasing some of their big budget movies later in the summer or during the winter dump months. This is because, according to him, the underlying demand has yet to peak. This research finds similar results. There are two strategies to be used during the less competitive seasons. If studios are planning on a fall release for their movie, they should invest more of their resources in the production budget and amount of screens, rather than movie stars. However, when the studios are contemplating a winter release, they should try and differentiate from the other movies in this season. The way to differentiate is through attracting a movie star. The reason behind that is because star power has a significantly positive effect, while ratings seems not to significantly impact the box office revenue during the winter.

During the more competitive seasons however, studios must employ different strategies. During the summer period, the best strategy is to differentiate with both a good story plot, as well as the use of a movie star. The coefficient for rating is .254, while the movie star coefficients is .242. Both are relatively high, but if you have to pick one or the other, studios should decide to go with solid story line (rating). For the holiday season, studios should primarily focus on stars. The coefficient is .302, while the rating variable does not significantly predict box office revenue.

Another finding is that the substitution variable is highly insignificant. Apparently, the increased usage of cable television and Netflix revenue has no significant effect on the stagnant motion picture industry. This implies consumers are actually increasing their overall leisure time spend watching television, Netflix, and watching theatrical movies.

The last variable that has been investigated in this research is movie specific attributes. This variable is the strongest indicator for box office revenue of a particular movie. The higher the investment that has been made by the studio for a movie, the more revenue that particular movie will earn. All the seasons are significantly influenced by this variable. The most noteworthy finding is that fall is influenced the most. This is remarkable because, as can be seen from appendix 11, it has the lowest average revenue of all the seasons. This implies that another dump period is occurring during the fall, just as in the winter period right after new years. If studios were to invest more during the fall on production budget and number of screens, it would generate even more revenue for them.

7.2 Managerial Implications

The purpose of this research was to find opportunities within the motion picture industry for movies to become successful. The parties that benefit the most from this particular research are the studios and producers. This research creates an overview how the variables rating, star power, and attributes affect box office revenue, while factoring for the different seasons. The biggest opportunity for studios to create a big box office hit lies within the fall period. Producing a movie, not with particularly big stars but with a good plot, will presumably lead to a fruitful financial scenario based on the statistics. Other opportunities lie within producing a movie during the winter season. Just like with the fall season, the competition with regards to big box office revenue creating movie is limited. Whenever a star during the winter period is used, revenues have increased significantly. This is an opportunity for studios. Lastly, there lies an opportunity for studios within investment itself. The research has shown that the more studios invest, the higher the revenue margins will become. This also implies that studios are reluctant to invest over the past couple of years. Seeing as economics for a major share revolves around confidence, studios need to become less risk averse in their investment strategies. It will be more lucrative for them.

7.3 Future Research

For future research there are a couple of significant improvements that can be made to the research. Primarily, it will be interesting to see how demographic variables like age, sex, and ethnicity will influence the model. It will probably allow the model to predict the variance in box office revenue more accurately. Following on this, studios will be able to target their audience with more precision. Another opportunity is to improve upon the movie specific attributes. Only two variables have been used in this paper. Previous research has, for instance, shown genre and sequels to have significant effect on revenue as well. Including more control variables will likely improve the model created in this research. The last improvement provided in this paper that future research can make, revolves around the substitution variable. Such a variable is difficult to measure the way it has been in this research. By introducing either an aggregate of the different seasons in order to analyze the effect substitution has, or finding a way to directly link substitution to specific movies, a major improvement will be made upon this research.

Chapter 8 - Reference List

1. Albert, S. (1998). Movie stars and the distribution of financially successful films in the motion picture industry. *Journal of Cultural Economics*, *22*(4), 249-270.

2. Basuroy, S., Desai, K. K., & Talukdar, D. (2006). An empirical investigation of signaling in the motion picture industry. *Journal of Marketing Research*, *43*(2), 287-295.

3. Batra, R., Ramaswamy, V., Alden, D. L., Steenkamp, J. B. E., & Ramachander, S. (2014). Effects of brand local and non-local origin on consumer attitudes in developing countries. *Journal of Consumer Psychology*, *9*, 83-95.

4. Boulding, W., & Kirmani, A. (1993). A consumer-side experimental examination of signaling theory: do consumers perceive warranties as signals of quality?. *Journal of Consumer Research*, 111-123.

5. Ter Braak, A., Dekimpe, M. G., & Geyskens, I. (2013). Retailer private-label margins: The role of supplier and quality-tier differentiation. *Journal of Marketing*, *77*(4), 86-103.

6. Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and psychological measurement*.

7. Cooper-Martin, E. (1991). Consumers and movies: Some findings on experiential products. *Advances in Consumer Research*, *18*(1), 372-378.

8. Corts, K. S. (2001). "The Strategic Effects of Vertical Market Structure: Common Agency and Divisionalization in the U.S. Motion Picture Industry," Journal of Economics and Management Strategy, 10(4), 509-528.

9. Danaher, P. J. (2002). Optimal pricing of new subscription services: Analysis of a market experiment. *Marketing Science*, *21*(2), 119-138.

10. Davis, P. (2006). Measuring the business of stealing, cannibalization and market expansion effects of entry in the US motion picture exhibition market. *The Journal of Industrial Economics*, *54*(3), 293-321.

11. De Vany, A. S., & Walls, W. D. (1997). The market for motion pictures: rank, revenue, and survival. *Economic inquiry*, *35*(4), 783-797.

12. Dellarocas, C., Zhang, X. M., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive marketing*, *21*(4), 23-45.

13. Einav, L. (2003). Not all rivals look alike: Estimating an equilibrium model of the release date timing game. *manuscript, Stanford University*.

14. Einav, L. (2001). Seasonality and competition in time: An empirical analysis of release date decisions in the US motion picture industry. *Woiking Papei, Harvard University*.

15. Einav, L. (2007). Seasonality in the US motion picture industry. *RAND Journal of Economics*, 127-145.

16. Eliashberg, J., Elberse, A., & Leenders, M. A. (2006). The motion picture industry: Critical issues in practice, current research, and new research directions. *Marketing Science*, *25*(6), 638-661.

17. Gil, A. (2008). Breaking the Studios: Antitrust and the Motion Picture Industry. *NYUJL & Liberty*, *3*, 83.

18. Hennig-Thurau, T., Henning, V., Sattler, H., Eggers, F., & Houston, M. B. (2007). The last picture show? Timing and order of movie distribution channels. *Journal of Marketing*, *71*(4), 63-83.

19. Hennig-Thurau, T., Houston, M. B., & Sridhar, S. (2006). Can good marketing carry a bad product? Evidence from the motion picture industry. *Marketing Letters*, *17*(3), 205-219.

20. Janssens, W., Wijnen, K., De Pelsmacker, P., & Van Kenhove, P. (2008). *Marketing research with SPSS*. Pearson.

21. Kahn, K. B., & Mentzer, J. T. (1998). Marketing's integration with other departments. *Journal of Business Research*, *42*(1), 53-62.

22. Karniouchina, E. V. (2011). Impact of star and movie buzz on motion picture distribution and box office revenue. *International Journal of Research in Marketing*, *28*(1), 62-74.

23. Kerlinger, F. N., & Lee, H. B. (1964). *Foundations of behavioral research: Educational and psychological inquiry*. New York: Holt, Rinehart and Winston.

24. King, T. (2007). Does film criticism affect box office earnings? Evidence from movies released in the US in 2003. *Journal of Cultural Economics*, *31*(3), 171-186.

25. Kirmani, A., & Rao, A. R. (2000). No pain, no gain: A critical review of the literature on signaling unobservable product quality. *Journal of marketing*, *64*(2), 66-79.

26. Krider, R. E., & Weinberg, C. B. (1998). Competitive dynamics and the introduction of new products: The motion picture timing game. *Journal of Marketing Research*, 1-15.

27. Lehmann, D. R., & Weinberg, C. B. (2000). Sales through sequential distribution channels: An application to movies and videos. *Journal of Marketing*, *64*(3), 18-33.

28. Litman, B. R. (1983). Predicting success of theatrical movies: An empirical study. *The Journal of Popular Culture*, *16*(4), 159-175.

29. Moon, S., Bergey, P. K., & Iacobucci, D. (2010). Dynamic effects among movie ratings, movie revenues, and viewer satisfaction. *Journal of Marketing*, *74*(1), 108-121.

30. Nam, S. H., Chang, B. H., & Park, J. Y. (2015). The Potential Effect of VOD on the Sequential Process of Theatrical Movies. *International Journal of Arts Management*, *17*(2), 19.

31. Prag, J., & Casavant, J. (1994). An empirical study of the determinants of revenues and marketing expenditures in the motion picture industry. *Journal of Cultural Economics*, *18*(3), 217-235.

32. Radas, S., & Shugan, S. M. (1998). Seasonal marketing and timing new product introductions. *Journal of Marketing Research*, 296-315.

33. Ravid, S. A. (1999). Information, Blockbusters, and Stars: A Study of the Film Industry*. *The Journal of Business*, *72*(4), 463-492.

34. Reinstein, D. A., & Snyder, C. M. (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.

35. Sochay, S. (1994). Predicting the performance of motion pictures. *Journal of Media Economics*, *7*(4), 1-20.

36. Suits, D. B. (1957). Use of dummy variables in regression equations. *Journal of the American Statistical Association*, *52*(280), 548-551.

37. Suits, D. B. (1984). Dummy variables: Mechanics v. interpretation. *The Review of Economics and Statistics*, 177-180.

38. Sullivan, R. (2009). Rental Epidemic of the Twenty-First Century: A Look at How Netflix and Redbox are Damaging the Health of the Hollywood Film Industry and How to Stop It, The. *Loy. LA Ent. L. Rev.*, *30*, 327.

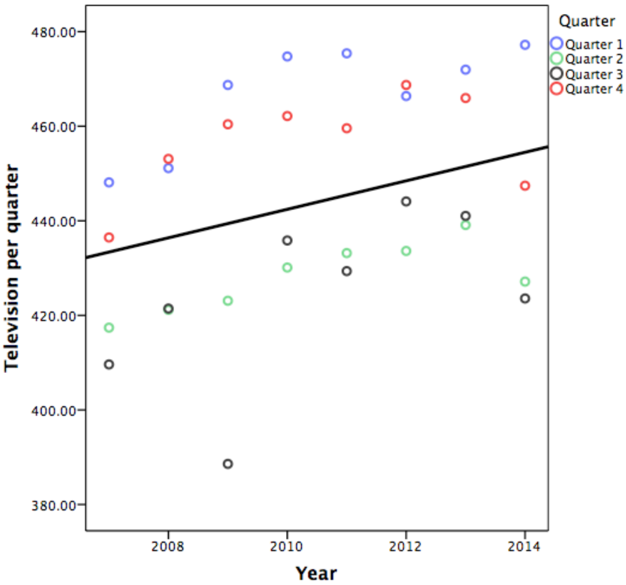
39. Selvaretnam, G., & Yang, J. Y. (2015). Factors Affecting the Financial Success of Motion Pictures: What is the Role of Star Power?.

40. Wallace, W. T., Seigerman, A., & Holbrook, M. B. (1993). The role of actors and actresses in the success of films: How much is a movie star worth?. *Journal of cultural economics*, *17*(1), 1-27.

41. Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, *2*(1), 37-52.

Appendices

Appendix 1: Television Views in Hours per Quarter (Nielsen report)



Appendix 2 – Principal Component Matrix Rating

Component Matrix[a]

	Component
	1
Critics Rating	.936
Amateur Rating	.936

Extraction Method: Principal Component Analysis.

a. 1 components extracted.

Appendix 3 – Principal Component Matrix Substitution

Component Matrix[a]

	Component
	1
Television	.848
Netflix Revenue Added	.848

Extraction Method: Principal Component Analysis.

a. 1 components extracted.

Appendix 4 – Principal Component Matrix Attributes

Component Matrix[a]

	Component
	1
Theaters	.916
Budget	.916

Extraction Method: Principal Component Analysis.

a. 1 components extracted.

Appendix 5 – Ratings Correlation

Correlations

		Critics Rating	Amateur Rating
Critics Rating	Pearson Correlation	1	.752**
	Sig. (2-tailed)		.000
	N	555	553
Amateur Rating	Pearson Correlation	.752**	1
	Sig. (2-tailed)	.000	
	N	553	554

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix 6 – Substitution Correlation

Correlations

		Television	Revenue Added
Television	Pearson Correlation	1	.439**
	Sig. (2-tailed)		.000
	N	560	560
Revenue Added	Pearson Correlation	.439**	1
	Sig. (2-tailed)	.000	
	N	560	560

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix 7 – Attributes Correlation

Correlations

		Theaters	Starless Budget
Theaters	Pearson Correlation	1	.586**
	Sig. (2-tailed)		.000
	N	559	407
Starless Budget	Pearson Correlation	.586**	1
	Sig. (2-tailed)	.000	
	N	407	407

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix 8 – Descriptive Statistics Dependent Variable Log Revenue

Descriptive Statistics

	N	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Log Revenue	560	18.1884	.78374	-.015	.103	-.059	.206
Valid N (listwise)	560						

Appendix 9 – ANOVA Test Dependent Variable Log Revenue

Descriptives

Log Revenue

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Winter/Spring	160	18.1486	.48272	.03816	18.0732	18.2240	17.04	19.83
Summer	160	18.5527	.75355	.05957	18.4350	18.6704	17.32	20.25
Fall	160	17.5349	.68127	.05386	17.4285	17.6413	15.73	19.43
Holiday	80	18.8462	.47465	.05307	18.7405	18.9518	18.07	20.44
Total	560	18.1884	.78374	.03312	18.1233	18.2534	15.73	20.44

Test of Homogeneity of Variances

Log Revenue

Levene Statistic	df1	df2	Sig.
18.290	3	556	.000

ANOVA

Log Revenue

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	124.434	3	41.478	105.337	.000
Within Groups	218.932	556	.394		
Total	343.365	559			

Appendix 10 – Descriptive Statistics Independent Variables

Descriptive Statistics

	N	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Rating	552	.0048295	.99651243	-.074	.104	-.519	.208
Log Star Power	495	18.4427	.67039	-.786	.110	1.500	.219
Substitution	560	-.0282228	1.02524036	.799	.103	1.058	.206
Attributes	407	.0289389	1.01293654	.553	.121	.307	.241
Valid N (listwise)	404						

Appendix 11 – Multicollinearity Test Independent Variables

Correlations

		Rating	Log Star Power	Substitution	Attributes
Rating	Pearson Correlation	1	.225**	-.015	.097
	Sig. (2-tailed)		.000	.723	.051
	N	552	495	552	407
Log Star Power	Pearson Correlation	.225**	1	.072	.224**
	Sig. (2-tailed)	.000		.108	.000
	N	495	495	495	404
Substitution	Pearson Correlation	-.015	.072	1	-.008
	Sig. (2-tailed)	.723	.108		.868
	N	552	495	560	407
Attributes	Pearson Correlation	.097	.224**	-.008	1
	Sig. (2-tailed)	.051	.000	.868	
	N	407	404	407	407

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix 12 – VIF statistics – Multicollinearity Check

Coefficients[a]

Model		Collinearity Statistics Tolerance	VIF
1	Log Star Power	.948	1.055
	Substitution	.997	1.003
	Attributes	.950	1.053

a. Dependent Variable: Rating

Coefficients[a]

Model		Collinearity Statistics Tolerance	VIF
1	Rating	.990	1.010
	Substitution	.999	1.001
	Attributes	.991	1.009

a. Dependent Variable: Log Star Power

Coefficients[a]

Model		Collinearity Statistics Tolerance	VIF
1	Rating	.962	1.040
	Log Star Power	.922	1.085
	Attributes	.947	1.056

a. Dependent Variable: Substitution

Coefficients[a]

Model		Collinearity Statistics Tolerance	VIF
1	Rating	.962	1.039
	Log Star Power	.961	1.040
	Substitution	.996	1.004

a. Dependent Variable: Attributes

Appendix 13 – Regression Model without Seasonality

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.784[a]	.614	.610	.47749

a. Predictors: (Constant), Attributes, Substitution, Rating, Log Star Power

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	144.914	4	36.228	158.897	.000[b]
	Residual	90.972	399	.228		
	Total	235.885	403			

a. Dependent Variable: Log Revenue

b. Predictors: (Constant), Attributes, Substitution, Rating, Log Star Power

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	14.686	.693		21.178	.000
	Rating	.192	.026	.237	7.453	.000
	Log Star Power	.191	.038	.165	5.095	.000
	Substitution	.056	.023	.075	2.403	.017
	Attributes	.496	.024	.658	20.603	.000

a. Dependent Variable: Log Revenue

Appendix 14 – ANOVA Test of Independent Variables

Descriptives

						95% Confidence Interval for Mean			
		N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
Rating	Winter/Spring	157	-.3030180	.90245628	.07202385	-.4452858	-.1607502	-3.15217	1.84387
	Summer	160	.0523933	.96465023	.07626230	-.0982244	.2030111	-2.80331	2.24378
	Fall	155	-.0022421	1.03054122	.08277506	-.1657632	.1612790	-2.01898	2.14553
	Holiday	80	.5275534	.95132878	.10636179	.3158455	.7392612	-1.85036	1.99470
	Total	552	.0048295	.99651243	.04241439	-.0784842	.0881431	-3.15217	2.24378
Log Star Power	Winter/Spring	146	18.3198	.62232	.05150	18.2180	18.4216	15.56	19.65
	Summer	138	18.5140	.66982	.05702	18.4013	18.6268	16.84	19.72
	Fall	144	18.2955	.70443	.05870	18.1795	18.4116	15.33	19.85
	Holiday	67	18.8798	.47038	.05747	18.7651	18.9946	17.73	19.61
	Total	495	18.4427	.67039	.03013	18.3835	18.5019	15.33	19.85
Substitution	Winter/Spring	160	.9612219	1.02806192	.08127543	.8007032	1.1217405	-.51827	3.04287
	Summer	160	-.5752778	.57284373	.04528727	-.6647200	-.4858356	-1.32943	.61531
	Fall	160	-.6713955	.64114410	.05068689	-.7715019	-.5712891	-1.84737	.13059
	Holiday	80	.3733435	.48234185	.05392746	.2660036	.4806834	-.64982	.95929
	Total	560	-.0282228	1.02524036	.04332431	-.1133211	.0568756	-1.84737	3.04287
Attributes	Winter/Spring	117	-.0304354	.78524661	.07259607	-.1742211	.1133503	-1.41127	2.52993
	Summer	120	.6026844	1.09808710	.10024118	.4041968	.8011719	-1.41889	3.18063
	Fall	112	-.6434481	.74175653	.07008940	-.7823349	-.5045613	-2.74165	1.32640
	Holiday	58	.2600543	.90645852	.11902383	.0217132	.4983954	-1.40810	2.49453
	Total	407	.0289389	1.01293654	.05020940	-.0697639	.1276418	-2.74165	3.18063

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
Rating	1.693	3	548	.167
Log Star Power	2.950	3	491	.032
Substitution	22.580	3	556	.000
Attributes	13.468	3	403	.000

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Rating	Between Groups	37.108	3	12.369	13.289	.000
	Within Groups	510.056	548	.931		
	Total	547.163	551			
Log Star Power	Between Groups	18.828	3	6.276	15.166	.000
	Within Groups	203.185	491	.414		
	Total	222.014	494			
Substitution	Between Groups	283.611	3	94.537	172.924	.000
	Within Groups	303.964	556	.547		
	Total	587.575	559			
Attributes	Between Groups	93.648	3	31.216	38.957	.000
	Within Groups	322.924	403	.801		
	Total	416.572	406			

Appendix 15 – Post hoc Analysis for Rating

Multiple Comparisons

Dependent Variable: Rating
Tukey HSD

(I) Quarter	(J) Quarter	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Winter/Spring	Summer	-.3554113*	.10837729	.006	-.6346921	-.0761305
	Fall	-.3007759*	.10923960	.031	-.5822788	-.0192730
	Holiday	-.8305713*	.13252500	.000	-1.1720791	-.4890635
Summer	Winter/Spring	.35541132*	.10837729	.006	.0761305	.6346921
	Fall	.05463544	.10872963	.958	-.2255533	.3348242
	Holiday	-.4751600*	.13210495	.002	-.8155854	-.1347347
Fall	Winter/Spring	.30077587*	.10923960	.031	.0192730	.5822788
	Summer	-.05463544	.10872963	.958	-.3348242	.2255533
	Holiday	-.5297955*	.13281329	.000	-.8720462	-.1875448
Holiday	Winter/Spring	.83057134*	.13252500	.000	.4890635	1.1720791
	Summer	.47516003*	.13210495	.002	.1347347	.8155854
	Fall	.52979547*	.13281329	.000	.1875448	.8720462

*. The mean difference is significant at the 0.05 level.

Appendix 16 – Post Hoc Test for Log Star Power

Multiple Comparisons

Dependent Variable: Log Star Power
Dunnett C

(I) Quarter	(J) Quarter	Mean Difference (I-J)	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
Winter/Spring	Summer	-.19422	.07684	-.3940	.0056
	Fall	.02424	.07809	-.1787	.2272
	Holiday	-.56005*	.07717	-.7622	-.3579
Summer	Winter/Spring	.19422	.07684	-.0056	.3940
	Fall	.21847*	.08184	.0057	.4313
	Holiday	-.36582*	.08095	-.5778	-.1539
Fall	Winter/Spring	-.02424	.07809	-.2272	.1787
	Summer	-.21847*	.08184	-.4313	-.0057
	Holiday	-.58429*	.08215	-.7993	-.3693
Holiday	Winter/Spring	.56005*	.07717	.3579	.7622
	Summer	.36582*	.08095	.1539	.5778
	Fall	.58429*	.08215	.3693	.7993

*. The mean difference is significant at the 0.05 level.

Appendix 17 – Post Hoc test for Substitution

Multiple Comparisons

Dependent Variable: Substitution
Dunnett C

(I) Quarter	(J) Quarter	Mean Difference (I-J)	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
Winter/Spring	Summer	1.5364997*	.09304103	1.2949272	1.7780722
	Fall	1.6326174*	.09578547	1.3839192	1.8813156
	Holiday	.58787836*	.09753905	.3337876	.8419692
Summer	Winter/Spring	-1.536500*	.09304103	-1.7780722	-1.2949272
	Fall	.09611769	.06797130	-.0803636	.2725989
	Holiday	-.9486213*	.07042093	-1.1326258	-.7646169
Fall	Winter/Spring	-1.632617*	.09578547	-1.8813156	-1.3839192
	Summer	-.09611769	.06797130	-.2725989	.0803636
	Holiday	-1.044739*	.07400900	-1.2380031	-.8514749
Holiday	Winter/Spring	-.5878784*	.09753905	-.8419692	-.3337876
	Summer	.94862134*	.07042093	.7646169	1.1326258
	Fall	1.0447390*	.07400900	.8514749	1.2380031

*. The mean difference is significant at the 0.05 level.

Appendix 18 – Post Hoc Test for Attributes

Multiple Comparisons

Dependent Variable: Attributes
Dunnett C

(I) Quarter	(J) Quarter	Mean Difference (I-J)	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
Winter/Spring	Summer	-.6331197*	.12376786	-.9556633	-.3105762
	Fall	.61301269*	.10090944	.3498921	.8761332
	Holiday	-.29048968	.13941615	-.6579460	.0769666
Summer	Winter/Spring	.63311974*	.12376786	.3105762	.9556633
	Fall	1.2461324*	.12231442	.9273096	1.5649553
	Holiday	.34263006	.15561159	-.0665595	.7518196
Fall	Winter/Spring	-.6130127*	.10090944	-.8761332	-.3498921
	Summer	-1.246132*	.12231442	-1.5649553	-.9273096
	Holiday	-.9035024*	.13812747	-1.2676982	-.5393065
Holiday	Winter/Spring	.29048968	.13941615	-.0769666	.6579460
	Summer	-.34263006	.15561159	-.7518196	.0665595
	Fall	.90350238*	.13812747	.5393065	1.2676982

*. The mean difference is significant at the 0.05 level.

Appendix 19 – Regression Model including Seasonality

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.844[a]	.713	.699	.41992

a. Predictors: (Constant), Attributes Holiday, Rating Fall, Attributes Winter, Rating Summer, Substitution Winter, Rating Holiday, Rating Winter, Attributes Fall, Substitution Summer, Substitution Holiday, Attributes Summer, Substitution Fall, Summer, Star Power Holiday, Star Power Fall, Star Power Winter, Star Power Summer, Fall, Holiday

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	168.174	19	8.851	50.197	.000[b]
	Residual	67.711	384	.176		
	Total	235.885	403			

a. Dependent Variable: Log Revenue

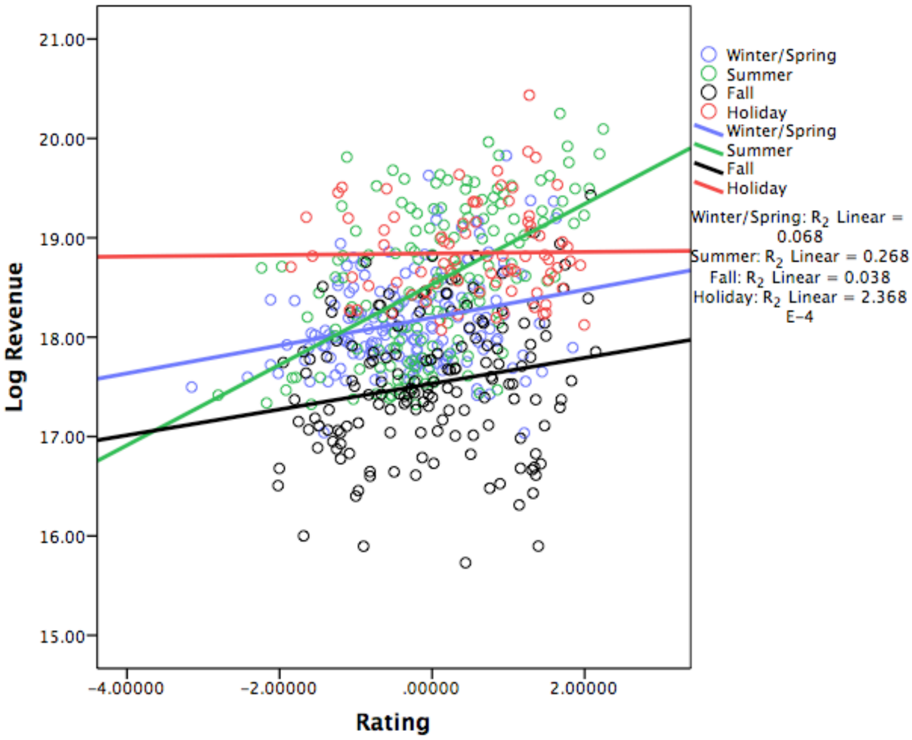
b. Predictors: (Constant), Attributes Holiday, Rating Fall, Attributes Winter, Rating Summer, Substitution Winter, Rating Holiday, Rating Winter, Attributes Fall, Substitution Summer, Substitution Holiday, Attributes Summer, Substitution Fall, Summer, Star Power Holiday, Star Power Fall, Star Power Winter, Star Power Summer, Fall, Holiday

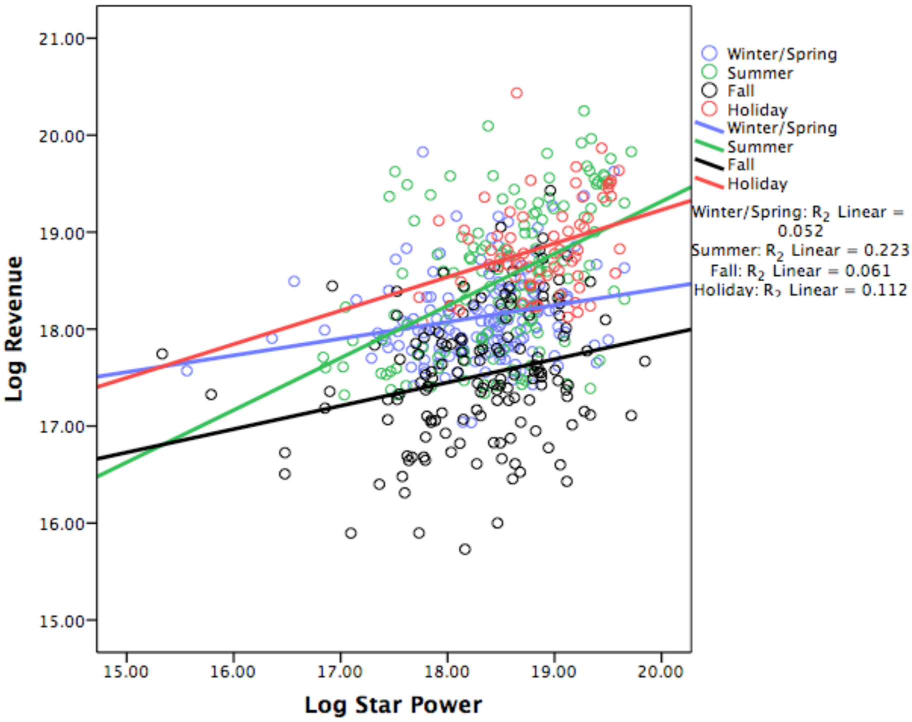
Coefficients[a]

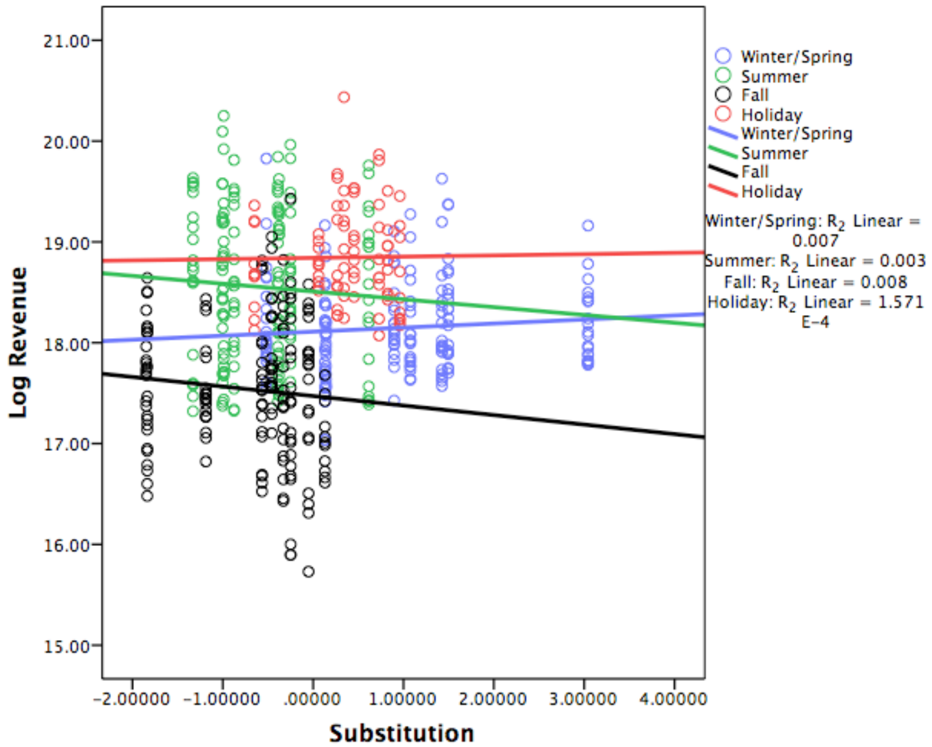
Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	15.627	1.151		13.581	.000
	Summer	-1.787	1.646	-1.069	-1.086	.278
	Fall	1.398	1.575	.819	.887	.375
	Holiday	-2.637	2.668	-1.193	-.988	.324
	Rating Winter	.072	.051	.041	1.410	.159
	Rating Summer	.254	.045	.162	5.637	.000
	Rating Fall	.237	.040	.172	5.992	.000
	Rating Holiday	.120	.066	.059	1.825	.069
	Star Power Winter	.142	.063	1.545	2.261	.024
	Star Power Summer	.242	.063	2.684	3.805	.000
	Star Power Fall	.050	.058	.538	.857	.392
	Star Power Holiday	.302	.127	2.572	2.368	.018
	Substitution Winter	-.009	.038	-.009	-.240	.810
	Substitution Summer	-.064	.069	-.033	-.926	.355
	Substitution Fall	-.066	.068	-.036	-.974	.331
	Substitution Holiday	.002	.121	.001	.019	.985
	Attributes Winter	.322	.051	.177	6.365	.000
	Attributes Summer	.392	.039	.337	10.075	.000
	Attributes Fall	.614	.054	.389	11.301	.000
	Attributes Holiday	.239	.065	.109	3.702	.000

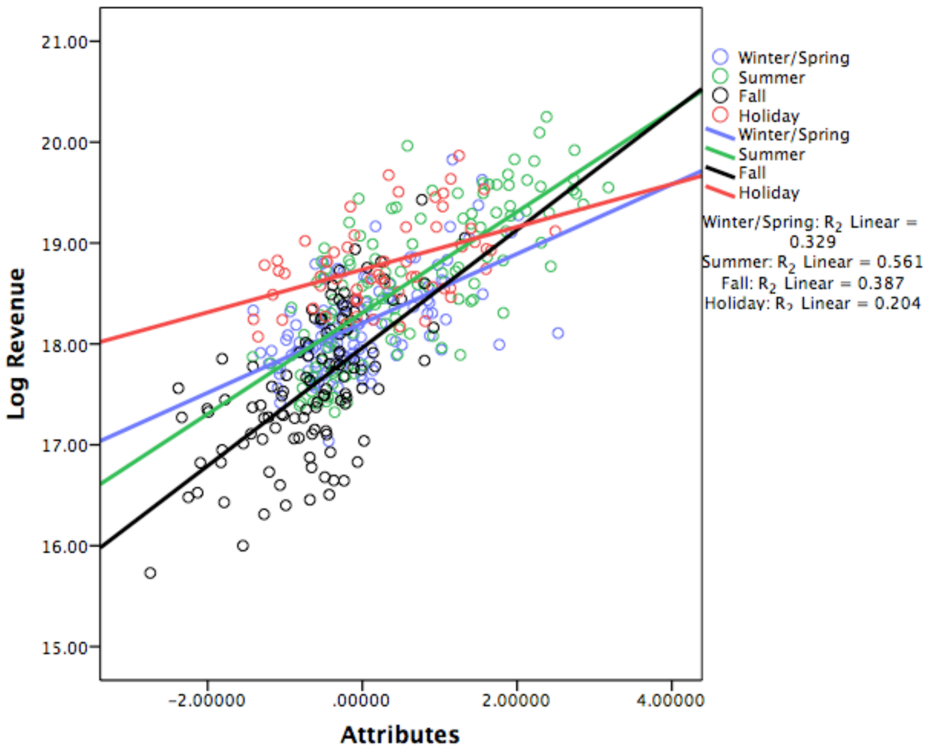
a. Dependent Variable: Log Revenue

Appendix 20 – Scatterplots for the Independent Variables






Log Revenue
Substitution
Winter/Spring
Summer
Fall
Holiday
Winter/Spring: R2 Linear = 0.007
Summer: R2 Linear = 0.003
Fall: R2 Linear = 0.008
Holiday: R2 Linear = 1.571 E-4


Log Revenue
Attributes
Winter/Spring
Summer
Fall
Holiday
Winter/Spring: R2 Linear = 0.329
Summer: R2 Linear = 0.561
Fall: R2 Linear = 0.387
Holiday: R2 Linear = 0.204