LUND UNIVERSITY
School of Economics and Management

Did Modern Media Kill the Superstar?

A contribution to the theory of consumer behaviour in the presence of increasing information.

by

Simon Sjölander

June 2016

Master's program in Economics

Supervisor: Peter Jochumzen

Abstract

This thesis will examine the increased amount of information made available through modern media, and how this affects the diversity in consumption. The thesis will focus on how diversity in music consumption has evolved over the years. There are two contradicting theories that try to describe the impact modern media could have on consumption, the theory of the Long Tail and the Superstar effect. The Superstar effect states that people will reduce their search costs by consuming the same goods as everyone else, and that this will lead to a convergence of products available in the market. The theory of the Long Tail, on the other hand suggest that the increased information will reduce the uncertainty for products and lead to an increased variation in the market. This thesis aims to examine which of the theories that are best applied to the observed development of music consumption, and by doing so give empirical evidence for how the modern media markets are affecting consumer behaviour. The data in this study covers the historical performance of the Billboard chart The Hot 100. The chart reflects the most popular songs from the past week and has been produced since 1958. Further, the dataset was found to be a good fit in order to answer the research question. We find a strong indication in favour of the Superstar effect, and that the music industry is evolving into an industry where a few actors make up the lion-share of the market. The music industry has been defined as a frontrunner for the cultural industries. The observed results are therefore likely to be existing or evolving in the other entertainment industries as well.

Keywords: Superstar effect, the Long Tail, consumer behaviour, economics, modern media, digital media, music consumption, Billboard, word of mouth, transaction cost economics, uncertainty, bundling, unbundling.

Acknowledgements

First and foremost, I would like to thank my supervisor Peter Jochumzen for believing in this thesis and truly support this idea.

I would also like to thank my wife Amanda Sjölander for supporting me during this process.

Table of Contents

List of Figures

List of Regressions

List of Tables

1. Introduction

1.1. Background

The idea to write about modern media consumption and how the increased amount of information might be a driving factor behind the change in consumer behaviour came along last spring while reading the excellent thesis by Artursson (2015) on how online ratings of video games had an impact on sale performance. Further investigations revealed that the impact modern media have on consumption has been studied extensively on a theoretical level. Among the more interesting research on how the increasing amount of information can have an impact on consumption is the theory of the Long Tail developed by Anderson (2004). The Long Tail describes how modern media give consumers an increased opportunity to obtain relevant information for products they intend to consume. The increasing possibility to receive appropriate information decrease the uncertainty for the products, making people more diversified in their consumption. This result in an increased variation of products in the market. In contrast to the theory of Long Tail there is the Superstar effect, this theory states that the increased amount of information makes consumers reduce their search costs by consuming the same type of products as everyone else and that a few central products take up an extensive amount of the public attention on the market (Adler, 1985; Hausman & Leonard, 1997). The common ground for these theories is that they note how modern media and the introduction of internet opened up for the possibility to completely transform an industry. Innovations that impact the compositions of an industry in this way has been described as a disruption process by Christensen (2013).

According to Anderson (2009) the music industry was recently disrupted by piracy and the possibility to consume music digitally. But this is not the only occurrence of disruption for the music industry. Throughout history the industry has gone through several severe changes, for example when video killed the radio star, and has been defined as a frontrunner in how digitalization affect the cultural industries (Hesmondhalgh, 2013). The development in music consumption is therefore likely to be present or evolve in the other entertainment industries as well.

1.2. Aim and Purpose

1.2.1. Aim

This thesis aims to study if the increased flow of information has changed diversity in music consumption and furthermore if the observed behaviour can give evidence in favour of either the theory of the Long Tail or the Superstar effect. This is performed by applying a set of quantitative time-series models to a dataset covering the chart “The Hot 100”. The thesis will examine if there has been a change in the diversity of music consumption based on a decrease/increase in the number of artist and titles present on the chart throughout its history.

The music industry has been defined as a frontrunner in digital media channels, and the development of music consumption could be useful in order to predict future consumption patterns in the other entertainment industries (Hesmondhalgh, 2013). Furthermore, the observed pattern could answer important questions in how consumer behaviour has evolved. However, the main focus for this thesis will be on the development of variety in music consumption, and how the development can be explained by applying the theory of the Long Tail or the Superstar effect.

1.2.2. Purpose

The purpose of this thesis is to contribute to the growing number of empirical papers examining how digital media and the increasing amount of information could impact consumption. This will be performed by studying how music consumption has evolved over time and examine if the theories of the Long Tail or the Superstar effect can explain the observed behaviour. A set of additional theories are presented to give a better understating for the transformation of consumption. Several previous studies have used data on word of mouth and sale performance to study the change in consumer behaviour. This study aims to extend this research and fill a gap in the current research by investigate change in diversity of consumption over time, i.e. how the number of consumed artists and titles are evolving.

1.3. Outline of the Essay

This essay consists of six chapters and they are organized as followed:

1. In the first chapter there is a background presentation of how music consumption can be used to study change in consumer behaviour. Further there is an introduction to the main theories of this thesis, the aim, purpose and outline of this study. Finally, the research question is stated, together with the main result.

2. Chapter two contains a literature review covering the different theories used in this study and how they are employed in the examination and interpretation of the results from the models. There is also a review of previously conducted studies on how digital media and new technologies are changing consumer behaviour.

3. The third chapter describe the obtained data from Billboard and how it was collected. Furthermore, the calculations and alterations what have been performed on the dataset are presented in combination with the limitations in this thesis. Furthermore, there is a graphical analysis of the data. This chapter also holds a description of the methodology used in the thesis, describing the research approach.

4. In the fourth chapter there is a presentation of the empirical models used to answer the research question.

5. In the fifth chapter, the results from the models is presented together with an analysis and discussion of how the results can be interpreted to answer the research question.

6. In the final chapter the essay is summarized and the concluding remarks are stated further there are some recommendations on how this thesis could be used as a springboard towards possible future research of the consumer behaviour.

Finally, the thesis has an Appendix, mainly presenting details of the econometric test conducted in chapter 3.

1.4. Research Question & Main Result

The aim of this thesis is to answer the following research question, by using data from the Billboard chart “The Hot 100”:

Has diversity in music consumption evolved over time, and if it so can the theories of the Long Tail or the Superstar effect be applied to explain the observed development?

To be able to answer the research question a set of empirical models are tested and the results from these regressions are used to give guidance towards the answering of the question.

The results from the models gives a very clear indication for that there has been a decrease in variation of music consumption, and that the observed development is best explained by the theory of the Superstar effect. The findings suggest that the music industry has evolved into an industry there a few artist stands for a large part of the market, and that the introduction of modern media have escalated this development.

2. Theoretical Review

This chapter describes the theories that are relevant to the issues of this essay, the main theories in this thesis concerns the theory of the Long Tail and the Superstar effect. An additional set of theories related to these theories are also presented, these are: Word of Mouth, Transaction Cost of Economics and Free. The chapter also contains a review of previously conducted studies related to this study.

2.1. Consumption Variety and Sales Concentration

Webster (2011) developed a theory of how the increase in market information and user generated content allow consumers and producers to shape the modern media landscape in cooperation with each other, this leads to a duality in the formation of modern media products. According to Hennig-Thurau, et al. (2010) the interaction between consumers in modern media channels makes them harder to define as only consumers, since they can act as consumers, marketers and producers, for the same product. The interaction between consumer and producers was further discussed by Brynjolfsson, et al. (2010), they suggest that new technologies such as search engines and recommendation systems is important factors behind the change in variety of consumption and sale concentration and that the digital media channels opened up for new possibilities to express and supply demand.

This points to the question of how future consumption is affected by modern media and the ever increasing amount of information. Will consumption become more diversified or will the most popular products still dominate, or even increase their market share? There are mainly two theories that tries to answer this, the Long Tail and the Superstar effect. (Peltier & Moreau, 2012)

2.1.1. The Long Tail

Hausman (1981) found that an increased variation would have a positive effect on consumer utility. One of the first studies examining the impact internet had on variation and how this could be beneficial for consumers was conducted by Brynjolfsson, et al. (2003). They examined how online bookstores enabled an increase in the amount of products available for sale. By attaching a value to the increased diversity they found that the variety led to large upsurge in consumer welfare.

The impact that modern markets have on consumer behaviour was further developed in an article by Anderson (2004), where he lays out the foundation regarding the theory of the Long Tail. The theory suggests that a decreased cost of distribution and an increasing possibility to connect producers and consumer will make niche products stand for a larger share of the market, while hits will see their market share decrease. The theory was further evolved by Anderson (2006) where he examines the possible effect the Long Tail can have on different markets. He suggests that the Long tail will lead to an increased diversity of products and larger fragmentation of consumption. Further, he concludes that modern media enables people to in a higher degree consume the good they attach the highest value to. The decreased search costs allow for an increased opportunity to receive relevant information and this makes it easier for people to find the most suitable product for their need.

2.1.2. Superstar Effect

According to the Long Tail, modern markets allows for an increasing opportunity to produce a wider diversity of goods. However, the mere possibility to produce these products does not necessarily mean that they will be consumed (Brynjolfsson, et al., 2010). A theory by Bikhchandani, et al. (1998) suggests that people will learn from the behaviour of past agents. The popularity of a product thereby acts as a signal and influence new consumers, meaning that being present on the bestseller chart lead to an increase in sales. Further, pure volume of Word of Mouth can change consumer behaviour, because of an increased awareness for the product (Liu, 2006). There is also a strong possibility for that consumers holds a desire to buy the same basket of products and services as other individuals in their society (Brynjolfsson, et al., 2010). According to Webster (2011) modern media give individuals the opportunity to steadily consume things they enjoy, while at the same time avoid things of their disliking. Meanwhile, producers are at the same time trying to reach a profit and build a reputation while competing for the limited amount of public attention.

The theories of the Superstar effect (Rosen, 1981) and Winner-Take-All (Noe & Parker, 2005) captures this by describing how individuals tend to demand the same type of products. The Superstar effect will result in a convergence of products, this consolidate the market to a few actors, and the profit will be skewed towards them. The Superstar effect has been found both in modern and traditional media channels (Bhattacharjee, et al., 2007; Adler, 1985). According to Hausman & Leonard (1997) NFL matches there at least one superstar was playing, had significantly more viewers, compared to regular games. Suggesting that the Superstar effect is prevalent in analogue medias.

The Superstar effect can also be explained from a supply-side perspective. Economies-of-scale could allow for higher profits when producing a few number of superstars, there could also be an increasing return to advertising, which would lead to a higher concentration in the distribution of actors. Furthermore, it is possible that there is a high concentration in the supply of talent and this would lead to a skewed distribution of quality products leading to the realization of the Superstar effect. (Noe & Parker, 2005)

2.2. Word of Mouth

An area in consumer behaviour that has been studied extensively is Word of Mouth (WoM). A traditional description of WoM is defined by Arndt (1967) as "oral, person to person communication between a receiver and a communicator whom the receiver perceives as non-commercial, concerning a brand, a product or a service" (p.3). One of the earliest and most cited work concerning WoM was performed by Katz & Lazarsfeld (1955). They found that consumers attached a higher value to WoM, than regular media while purchasing food and groceries. Their article was followed by several authors studying different aspects of WoM, for example how WoM is affecting purchasing decisions and the motives behind engaging in the creation of WoM (Sundaram, et al., 1998). According to Brynjolfsson, et al. (2010) WoM is an important factor behind the Long Tail and the Superstar effect, since both of them are trying to explain how consumer is supplying and receiving information. Furthermore, volume and valence has been found to be important factors behind how WoM influence consumer behaviour (Chevalier & Mayzlin, 2006). Further, the valence by other consumers has been found to have an impact on sales, while the valence from professionals do not affect future sale performance (Artursson, 2015). According to Liu (2006) the duration and intensity of WoM might also affect consumer behaviour.

In recent years there has been an increased attention towards the interaction between individuals that occurs online, and how this electrical Word of Mouth (eWoM) affect individuals. Kietzmann & Canhoto (2013) define eWoM as "EWoM refers to any statement based on positive, neutral, or negative experiences made by potential, actual, or former consumers about a product, service, brand, or company, which is made available to a multitude of people and institutions via the Internet" (p. 147-148). Even though WoM that occurs online may be perceived as less personal, it has been shown to be more effective than regular WoM, because of its wider reach. Online interaction between individuals can be seen as a market force and eWoM has an impact on consumer's behaviour and influence which products that are consumed. (Kietzmann & Canhoto, 2013)

2.3. Transaction Cost of Economics

Transaction costs can be explained as the associated cost occurring from a transaction, traditionally from buying and selling a product. This is a contradiction compared to classic economic theory, there it is assumed that the information is the same for all market participants. In classic economic theory transactions can be made without an associated cost, other than the price of the product. (Liang & Huang, 1998)

Early research on transaction cost focused on why and when firms internalized production, instead of buying at the market (Coase, 1937). This was further developed by Williamson (1979). He laid out the foundation for factors affecting transaction costs, and how these can impact the decision for which transaction to go through with. Primarily three factors are found to affect the cost of a transaction namely, the frequency, how specialised the products are and the uncertainty. The basic idea in the theory of transaction-cost economics is that individuals choose the transaction that minimize these factors. Furthermore, Williamson showed that these factors affect all sorts of transactions, meaning that it does not only have an impact on the behaviour of companies but also affect the behaviour of regular consumers.

2.3.1. Uncertainty

A higher degree of uncertainty is generally assumed to be associated with a larger transaction cost. There are two different sorts of uncertainty, product uncertainty, related to the characteristics of the products and process uncertainty, reflecting the uncertainty of the transaction. The process uncertainty involves search, comparison, examination, negotiation, order, payments, delivery, and post-service costs. Reducing the uncertainty for either of these factors will decrease the transaction cost and thereby increase the likelihood for that the consumers will go through with the transaction. (Liang & Huang, 1998)

2.4. Free

Throughout history there have been several industrial revolutions. A common feature for these revolutions is that factors that has been expensive and scarce becomes abundant and drastically cheaper (Kelly, 1993). According to Anderson (2009) bits, the ones and zeroes written to the hard drive that makes the software on our computers work, are becoming an abundant factor. According to Anderson Moore's law concludes that the price for processing power halves by every second year, and the price of bandwidth and storage falls even more drastically. Thereby the marginal cost for producing some extra bits is rapidly approaching zero. Anderson states that then something is approaching zero in this pace, the cost will sooner or later become negligible. As the cost of bits are approaching zero so does the cost of digital information. The term digital information is broad and includes nearly everything that can be distributed digitally. However, there is a difference in the marginal cost for abundant and scarce information. The marginal cost of producing more of the abundant information is very low, information of this type will be free, while scarce information will remain costly. (Anderson, 2009)

2.5. Previous Research

Previously conducted research has examined different aspects on how media is affecting consumer behaviour, but as far as my research goes this is one of the first studies to examine how chart performance over time might explain change in consumption behaviour.

An extensive amount of previous research has been conducted on WoM and how it can have an impact on consumer behaviour. Liu (2006) studied how weekly data of WoM affects box office performance, and suggest that volume have an impact on performance while valence do not. This study was followed up by Wenjing, et al. (2008) used daily instead of weekly data and allowed for WoM to be endogenous to performance. Their results were similar to the ones found by Liu (2006). They also suggest that the important part of WoM is the volume of posts and that WoM thereby acts as an awareness effect rather than signalling quality for a product. Furthermore, Wenjing, et al. (2008) suggest that there is an autocorrelation between sale performance and the volume of WoM. This is also in line with the findings by Salganik, et al. (2006) by creating an artificial music market, they found that the success of a song is determined by the choice made by previous listeners. They suggest that social influence is the most important factor behind the success of a song, giving strong support for the Superstar effect.

Similar results were found by Fu & Sim (2011) they examined how the amount of views for online videos had an impact on future performance. Their results suggest that videos with high view counts perform better, than videos with low amount of views. They suggest that an important reason behind their result might be the reduction in uncertainty products with high view count enjoys. The increasing amount of digital information makes individuals seek for ways to reduce the uncertainty, for which information they actually want to consume. The actions of previous agents reduce the uncertainty, and thereby acts as a signal of quality.

Webster & Ksiazek (2012), claim that both producers and consumers have a desire for top-list media products. Following the theory of free they argue for that the possibility to more easily access and consume the best content will continue to drive demand for top-list performances and lead to a realization of the Superstar effect. They also argue for that media consumption might act as a “coin-of-exchange”. Consuming the same type of products as other individuals’ in the society give an individual social value. It might be an important personal characteristic to have the capability to discuss the latest release by a top-performing artist over the coffee break. Therefore, people will consume the same type of products and the Superstar effect will be the dominating theory on the modern media market.

Unbundling is another factor that might be of importance for the explanation of the Superstar effect. Unbundling was discussed by Elberse (2010), she states that digital marketplaces opened up for the possibility to sell items separately. This was not at least observed in the music industry, consumers were no longer forced to buy the whole album for the opportunity to consume a specific song, instead they could cherry-pick the top tunes. According to Elberse this could be seen in the consumption pattern, with a steep increase of titles sold as singles, while at the same time album sales decreased.

Other studies examining bundling have however found that micropayments do not seem to be very efficient. The reason behind this seems to be that consumer will be setting up mental barrier that arises from the mental transaction costs associated with paying something at all for the usage, the actual price does not seem to matter (Szabo, 1999). Bundling products and selling them for a fixed cost i.e. setting the marginal price for usage to zero, reduce the mental transaction costs and bring down the mental barrier (Bakos & Brynjolfsson, 1999). Apart from the risk reduction consumers seem to attach a value for simplicity when choosing a pricing plan, and consumers with flat-rates subscriptions have a tendency to increase their usage(Odlyzko, 2001).

Another field of modern media that has been studied extensively is recommendation systems and how they might have an impact on consumer behaviour. An important aspect in recommendation systems might be how the receiver is linked to the sender. Vilpponen, et al. (2006) studied how relations between peers online could impact adoption behaviour. They suggest that there is no direct support for that a stronger tie between the source and receiver have an impact on adoption behaviour. Their interpretation of this result is that that peers engaging in digital communication should be seen as equals in terms of effectiveness and power of influence. How recommendation systems might impact diversity in consumption was studied further by Fleder & Hosanagar (2009). They found that recommendation systems might lead to a reduction of diversity. They suggest that an explanation behind their findings might be that common recommendation systems are based on historical sales and ratings, and this makes it difficult for these systems to value a new or unknown product. Further they argue that this will lead to a realization of the Superstar effect, since popular products become even more popular based on previous sales. According to Fleder & Hosanagar recommendation systems could make individuals explore more products, while at the same time converge all consumers to the same products. This might explain why individuals feel that their variety increased while diversity in aggregate decreased.

The effect of recommendation system is however disputed. Peltier & Moreau (2012) examined how online markets differ from traditional markets, by conducting a study on the French book market. They suggest that digital markets somewhat made consumers change their consumption from bestsellers to mid- and low sellers. A possible explanation behind this could be that it is easier for consumers to access information online. A similar study by Brynjolfsson, Hu & Simester (2011) suggest that recommendation systems leads to an increased diversity of products and that the digital markets lead to a significantly lower concentration of sales for specific products. They suggest that the main reason behind the increased diversity is the reduction in search costs, and interpret this as support for the Long Tail. This is in line with the findings of Oestreicher-Singer & Sundararajan (2010) they found that recommendation systems increase the demand for less popular products, and shrink the demand for the highest grossing goods. Lynch Jr & Ariely (2000) suggest that the main gain for consumers in the digital environment is the reduction of search costs. According to Brynjolfsson, et al. (2003) recommendation systems and search engines makes it easier for consumer to find niche products, and reduce the uncertainty for a product. This is in line with the findings by Liang & Huang (1998) they suggest that a reduction in search costs lower the process uncertainty, and that this reduce the transaction cost, making it more likely for a transaction to occur. An increased diversity of products will allow for a greater variety of consumption and this has been found in several papers to increase the level of utility, giving support for the Long Tail (Hausman, 1997; Brynjolfsson, et al., 2003; Christian & Weinstein 2006).

A few previous studies have examined chart performance. Bradlow & Fader (2001) used data from Billboard to estimate the movement and expected lifetime for the titles on the chart. They suggest that the debut position for a title was an important explanation for the expected lifetime. Further, they suggest that titles from popular artists has an extended survival on the chart. The survival for titles was further examined by Giles (2007), he studied the survival for titles reaching the number one spot. His findings suggest that there has been an increase in the survival for titles reaching number one as time has gone by. Furthermore, he suggests that stars on average has a higher survival on the chart. Another study examining the survival on the chart was performed by Bhattacharjee, et al. (2007). They studied how survival of albums was affected by the introduction of digital markets. Their findings suggest that for top performing albums, the introduction of digital markets had no effect on survival, while for the albums with a lower initial placement the introduction digital markets reduced the average survival. They interpret this as evidence for the Superstar effect.

2.6. Contribution to the Research

Benghozi & Benhamou (2010) raises the question whether the Long Tail is a myth or reality?

In this thesis I will try to answer this by examine how the consumption pattern of music has evolved and examine if the observed development is best explained by applying the theory of the Long Tail or the Superstar effect. By applying a different method compared to the ones used above I will be able to track how music consumption has changed over a long period of time. The findings can then be used both in the academia as well as for professionals to get a better understanding of how modern media is affecting consumption.

3. Method

This chapter start with a presentation of the collected data. Further, the applied method in this thesis is presented, together with a graphical analysis and an explanation of the econometric tests applied to the data.

3.1. Data Collection

The dataset used in this study contains the historical performance for the Billboard chart "The Hot 100", and contains information of music consumption over the last 57 years[1]. The data was collected from Billboards official website (Billboard, 1958 - 2015). The chart is published weekly and is compiled of 100 songs, and in total the sample contains 297 242 observations. A Python script was used to extract the records from Billboards website and transform it to manageable data.[2]

3.1.1. Data Presentation

The Hot 100 chart is compiled by Billboard, a music publication covering the music industry (Billboard, 2011). The Hot 100 ranks the most popular songs in the U.S. from the past week by using a formula which takes different kinds of music consumption into account. The chart is namely not just made up of physical sales, but weights in radio plays, number of streams and digital downloads as well. The weighting-formula assure that the chart provides an accurate description of the most popular songs from the previous week (Billboard, 2016). Billboards method for calculating The Hot 100 has gone through several changes during its nearly 60-year-old history to stay relevant and continue to be correct in its reflection of music consumption. Some of the more comprehensive changes were to allow for songs which only plays on the radio to be ranked in 1998, and to include digital sales and streaming in 2005, and 2012 respectively (Molanphy, 2013).

[1] Billboard introduced the chart in August 1958, full years are wanted for the study and therefore the first year is sorted out.
[2] The data, as well as the Python script is available upon request.

The dataset contains the following set of variables: date of publication, current position on the chart, artist, title, the position on the chart from last week, the change between this week and the last, the peak position for the title and number of weeks the title has been on the chart.

One correction has been made to the dataset concerning guest artists featuring in a song with another artist, the featuring artists were removed from the dataset. Guest performances has been a growing phenomenon in recent years, one reason behind this might be the promotion less known artists receives then performing with more famous musicians. Another reason might be that stars stays relevant by collaborating with new names. (Edwards, 2007) The original dataset confirms that guest appearances have been growing in recent years[3] and that artists might collaborate with several different artists during a short period of time. Featuring artists could thereby result in an upward bias and removing them lead to a more accurate description of the development of diversity in music consumption.

There was mainly two reason behind the selection of The Hot 100. Firstly, The Hot 100 chart covers a long period of time. This enables the study to more clearly examine how music consumption has evolved, and this allows for a better understanding of the change in consumer behaviour. An equally important reason for choosing The Hot 100 is that the music industry can be thought of as a frontrunner in the digital media environment (Papies, et al., 2011). Music consumption is thereby interesting to work with, since it point towards there the other cultural industries are heading (Hesmondhalgh, 2013).

3.1.2. Validation of Data

The Hot 100 chart is compiled by Billboard. Billboard is an established publication and I find them to be a trustworthy source of data. Further, I rely on that the information provided through their archives is correct. The data was fetched using a Python script and random sample controls were conducted throughout the dataset to confirm the validity of the fetched data. The controls showed that the fetched data contain the same information as the one stated on Billboards website.

[3] Out of a total of 1731 distinct artists with featuring performances, 1387 occurs after 2000.

3.2. Research Approach

This is a quantitative study, using linear OLS regressions on a set of dependent variables and regress them against time. The thesis aims to examine how diversity in music consumption has evolved and explain the observed pattern by applying the theory of The Long Tail or the Superstar effect. (Verbeek, 2012)

3.3. Research Design

This thesis will be studying how music consumption has evolved during the last 57 years, and examine if there has been a change in diversity of consumption, and if so which one of two theories, the Long Tail or the Superstar effect, what best captures the observed development.

Since the purpose of this thesis is not to track individual movements for specific titles but to examine the aggregated behaviour for the chart, the original panel dataset is transformed into a set of multiple time series dependent variables. The method is then straightforward to implement. The basic idea for the research design is to run a regression on all the dependent variables over time and examine how an increase in information has changed the diversity in consumption. The independent variable in all the models is time, in this thesis time act as a proxy capturing the development of produced information. Moore's law states that the price for processing power is halved by every second year, and thereby the price for producing information decrease over time, and this in turn affect the amount of information produced (Anderson, 2009). This makes time a good proxy for capturing evolvement of information.

The main focus in this thesis will be on how the dependent variables covering distinct number of artists and titles are evolving. Two supporting dependent variables are added to the thesis. These variables are included to give support for the observed effects and to build a better understanding for the transformation of diversity in consumption. In table 1, the dependent variables examined in this study are presented.

It is important to note that it is distinct observations that is counted, i.e. an artist or title is only counted for once during the specified time period. The same approach has been used in previous studies examining "The Hot 100" (Bradlow & Fader, 2001; Giles, 2007). Counting distinct observations allows for a clear interpretation regarding which one of two theories, the Long Tail or Superstar effect, that best captures the change in diversity.

Table 1 - Dependent Variables. The table gives a short description of the dependent variables controlled for in this thesis.

Variable	Description
$ArtWC_t$	Distinct number of Artist presented on the whole chart for time ***t***.
$TitleWC_t$	Distinct number of Titles presented on the whole chart for time ***t***.
$ArtTop10_t$	Distinct number of Artist reaching a peak-position of 10 or better on the chart for time ***t***.
$TitleTop10_t$	Distinct number of Titles reaching a peak-position of 10 or better on the chart for time ***t***.
$Weeks_t$	Average number of weeks a Titles stays on the chart during time ***t***.
$TitleArt_t$	Mean number of Titles per Artist during time ***t***.

3.4. Method for choosing dependent variables

Several aspects were taken into account when selecting the dependent variables examined in this thesis. Below is an explanation on what grounds the different variables were chosen and how they are used to answer the research question.

3.4.1. Number of distinct observations for Artist and Titles

The most central variables in this thesis are the ones that examine the number of distinct observations for artist and titles. A central question regarding how to handle these variables was whether to break up the sample in different time periods or not. Allowing for breaks means that an artist or title will be counted as a distinct observation for every time-period were it is present. Not breaking up the sample would mean that the observation is counted for then it occurs and thereby the distinct observation is only counted for once in the whole sample. In this thesis there is a break for every year. Presenting the variables with a break result in a more accurate description for the development of The Hot 100.

Star musicians might persist over a long period of time, without breaks they will be counted for once and this might result in a bias then estimating the development in diversity. A large fraction of the artists on the chart could have had their first appearance on the chart in another time period. The persistency of popular artist could thereby result in a faulty estimation and lead to incorrect conclusions in the analysis for how diversity in music consumption has evolved. Setting breaks for every year result in a more accurate reflection of the development, since the number of observations will not jump if there in a given year is a high number of new observations, followed by a period where a low amount of new artists makes it to the chart.

Displaying the number of distinct observation for every year results in an accurate description for how diversity in music consumption is evolving. The Hot 100 is announced for every week and thereby it is impossible to give a presentation for distinct observations with a break for every week. A possible alteration would be a break for every year but present the number of new distinct observation by week throughout the year. This would however be problematic since it would return a skewed distribution, with a high number distinct observation in the first week after the break. Instead of years one could use months, half a year and so forth, but this only makes the results more difficult to interpret, and does not change the results.

Another aspect to take into consideration for the variables covering the number of observations of titles and artists is the restriction based on peak-performance i.e. the highest rank a specific artist or title has reached. A set of alternatives might be interesting to examine, but some form of restriction had to be implemented and two variables for artist and titles respectively were finally chosen to be examined. The first variable covers all the distinct artist and titles that were ranked on the chart for a given year i.e. reaching Top 100. The second variable is examining the number of observations for artists and titles reaching a rank of ten or better.

Examining the development for artists and titles presented on the whole chart for a given year is central, since it displays how consumption at large is evolving. If more artists and/or titles are circulating as time goes by it would be an indication towards the Long Tail, while a decline in number of observations would indicate that music consumption is evolving into a market that is best described by the Superstar effect. The variables containing artist and titles with a peak-performance of ten or better were chosen since it could be relevant to examine the development for the the top ten-percentile of the chart and study if there is a difference compared to how the chart as a whole is evolving. According to Brynjolfsson, et al. (2010) there could be a difference in concentration of products throughout the distribution of sales. This can be linked to the performance for artists and titles on the chart. There could be a Superstar effect prevalent in the top ten-percentile of the distribution of the chart. At the same time the chart as a whole could display a Long Tail effect. Examine the top ten-percentile instead of the the artist or titles reaching number one was because of the larger sample size and the higher variance in the data.

3.4.2. Supporting variables

Two additional variables were included for the final models. These supporting variables were added to give support for the results and findings from the models examining the distinct artists and titles, and to give a better understanding of the transformation in consumer behaviour. The break for the supporting variables are set to one year, the same as for the variables covering the distinct number of artists and titles. Applying the same break period to all variables is a natural choice since it makes it easy to compare the development for different variables in the sample.

The first supporting variables is the average amount of weeks a title stays on the chart for a given year. The inclusion of this variable enables a better understanding for how the performance of titles have evolved, and how the survival on the chart is developing at large. A growing average number of weeks would act as a signal of higher persistency on the chart, and lower circulation, which would be in line with the theory of the Superstar effect.

The second supporting variable is examining the mean number of titles per artist for a given year. This variable is included since it enables a better understanding for how consumption is changing. A higher amount of titles per artist could be seen as evidence for the Superstar effect, with an increased public attention put towards a few artists (Webster & Ksiazek, 2012).

3.4.3. Calculations

The calculations for the dependent variables are performed in the original panel dataset. The distinct number of observations is counted for each year, resulting in a time series dataset holding 57 observations. The supporting variables is also calculated for each year and holds 57 observations. Following is a more detailed description for how the dependent variables were calculated.

The dependent variable covering the number of artists on The Hot 100 for a given year is calculated by sorting the dataset by date of publication. As was discussed in section 3.4.1. there is a break for every year in the sample, so the distinct number of observations is calculated for each year. The number of observations for distinct artist is summed up by counting the number of new distinct observation for every week during the year, this of natural reasons return a skewed distribution for the first week, but since the relevant number is the sum of distinct observations during the year this is not a problem. The resulting variable holds 57 observations, covering the development for the number of artists. The process for calculating the variable covering the number of artists reaching top ten is the same as for the chart as whole but with an added restriction for that the artist should have reached a peak-position of ten or better in the given year.

The procedure for titles is a little more complicated compared to the calculation for artists, since the same name for a title could be used by several artists[4]. Instead of just sorting the data by its date of publication an additional criterion is added to the sorting order, there the title is linked to its specific artist to make sure that all titles receive a unique identity. The sorting thereby ensures that the distinct titles are linked to its distinct artist and this ensures that all the distinct titles for a given year is included. This is further controlled for by a duplicate test in Stata. In other aspects the procedure for counting distinct titles is the same as for counting distinct artists. The procedure for calculating the titles reaching a position of ten or better in the given year, is the same as for the titles on the whole chart, but with the added restriction for that the titles should have reached a peak position on the chart of ten or better.

Average number of weeks a title stays on the chart is computed in three steps, first the distinct titles are sorted by year. This is followed by finding the maximum amount of weeks each title has been on the chart, and from this the average is calculated by summing up all the maximum values and divide it by the number of titles for that year.

The mean number of titles per artists for a given year is calculated in the following way. First the distinct artists for a given year is sorted out. Thereafter, each of the titles a distinct artist has on the chart for that year is counted, from this the mean number of titles per artist is calculated by summing up the number of titles per artist and divide it by the number of artists.

3.5. Limitation

The dataset cover the historical performance of the Billboard chart The Hot 100 and holds a comprehensive amount of information. This is of course good news for the purpose of this study, since it makes it easy to find interesting topics to study and the dataset allow for many different forms of investigation. However, it also means that all the information in the data cannot be studied in a single thesis, and some limitations has to be implemented. The chosen method for the thesis is to run a regression for the number of artist and titles over time and add a number of supporting variables to control for the observed behaviour. This method is chosen since it in an accurate way describes the development of music consumption at large, and can answer the question of whether the Long Tail or the Superstar effect is the prevalent one.

[4] The same title used by different artists is seen several times throughout the sample, even in the same weeks.

It is important to note that the data consist of rankings and not actual numbers of sold copies, played streams or radio plays. This means that there are some limitations in drawing conclusions on absolute performance for a specific title, since the ranking is based on relative performance. Further, because of the lack of absolute figures for performance, the study cannot say anything about the aggregated levels of music consumption. The position on the chart is based on a weighting scheme calculated by Billboard, and I have not been able to access the exact details for how the calculations are performed when determining the position on the chart for a specific title. Neither do I have information on what the dominating factor behind a specific titles current rank is, i.e. radio plays, sold records or number of streams. This is problematic in several ways, since the weighting might be an important factor in explaining why the list have changed during the years. Furthermore, it could be useful to have knowledge about the weighting scheme since it would allow for investigation into how specific factors of consumption could explain a change in consumer behaviour.

3.5.1. Alternative method

This study is as far as my research goes one of the first to study the development of music consumption over time and link it to consumer behaviour. Applying the chosen research design is a conscious choice and fulfils its purpose of answering the research question. However, for further investigation into consumer behaviour it could be fruitful to apply an alternative model. The survival rate for the titles ranked as number one was examined by Giles (2007), applying the same method on different rankings could be a useful way to further examine the Superstar effect. However, this method will have a hard time explaining the prevalence of the Long Tail, since a mere decrease in the survival rate cannot be directly linked to the Long Tail. The dependent variable examining the average number of weeks a title stays on the chart in this thesis can be seen as simplified approach to the one used by Giles.

Furthermore, it could be interesting to apply a Bayesian Lifetime Model as proposed by Bradlow & Fader (2001), they use a Markov Chain Monte Carlo simulation to attach values to different factors and use these to assign a worth to each title. They further examine how different factors affect the titles behaviour on the chart during the year of 1993. Extending their research to cover more time periods and study how the factors affecting the movement have changed over time could be a useful way to understand why we observe the Superstar effect.

3.6. Graphical Analysis

Below is a graphical presentation for the different regressions. The figures display the observed values in combination with the fitted linear trend lines. For detailed results, see to chapter 5. Plotting the observed values is a commonly used approach, since it is a convenient way to get a basic idea of the results and can give an indication toward what type of econometric tests one should perform on the dataset. (Enders, 2015)

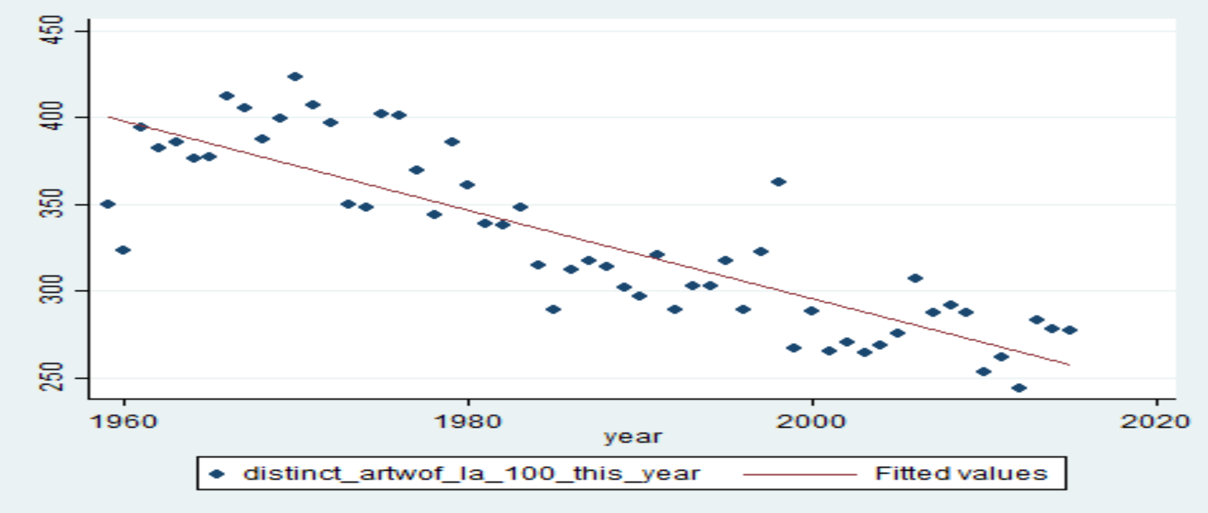


Figure 1 - Number of distinct artists, per year. The figure displays the distinct number of artist represented on the chart per year, and there is a negative trend in number of observations.

Figure 1 shows a clear trend for that there has been a decline in number of artist presented on the chart, suggesting that there has been a reduced diversity in music consumption.

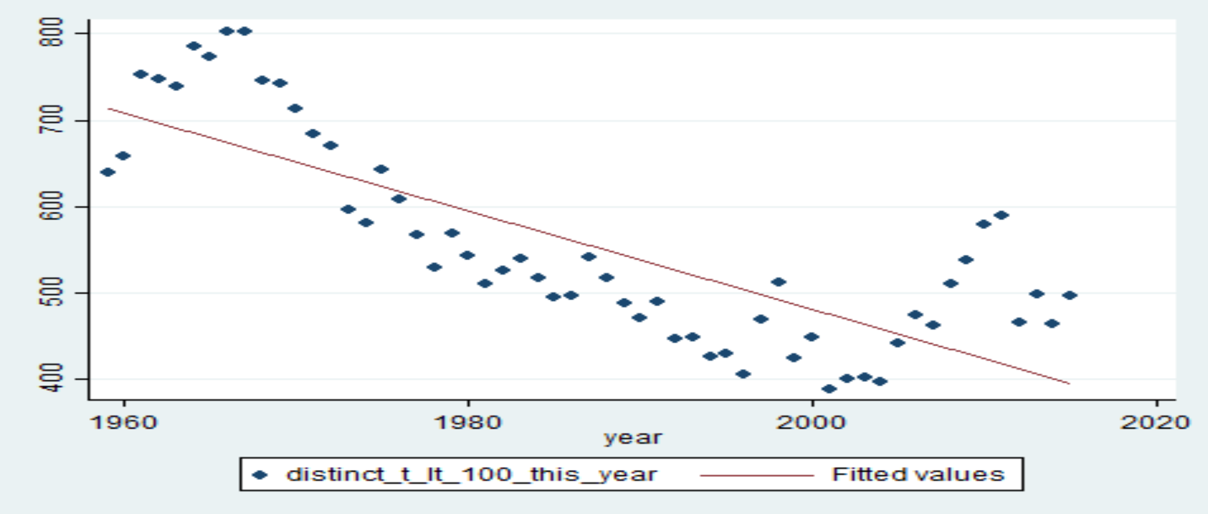


Figure 2 - Number of distinct titles per year. The figure displays the number of distinct titles reaching the chart per year, there is a negative trend in the data.

As can be seen in figure 2, the pattern for the observed values of distinct titles follows the same trend as for the artists. There is a negative trend in the amount of titles what makes it to the chart per year. However, there might be a break in the data in 2005, which needs to be further examined.

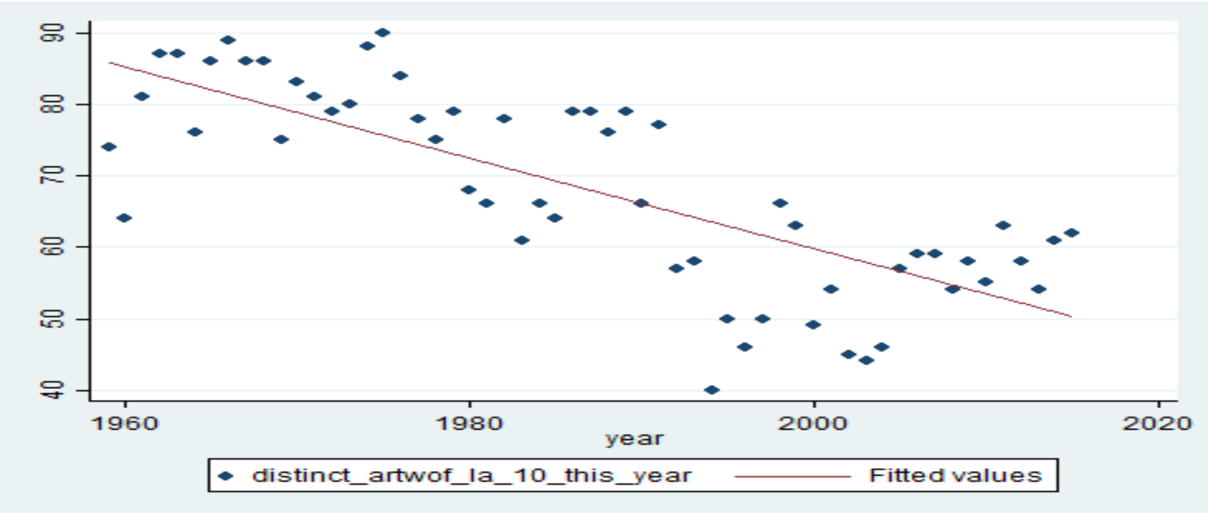


Figure 3 - Number of distinct artist reaching Top 10, per year. The figure displays the number of distinct artist reaching a placement on the chart of ten or better, there is a negative trend in the data.

The same trend as for the artist on the whole can be seen in the plot of artist reaching a ranking of ten or better, this is displayed in figure 3.

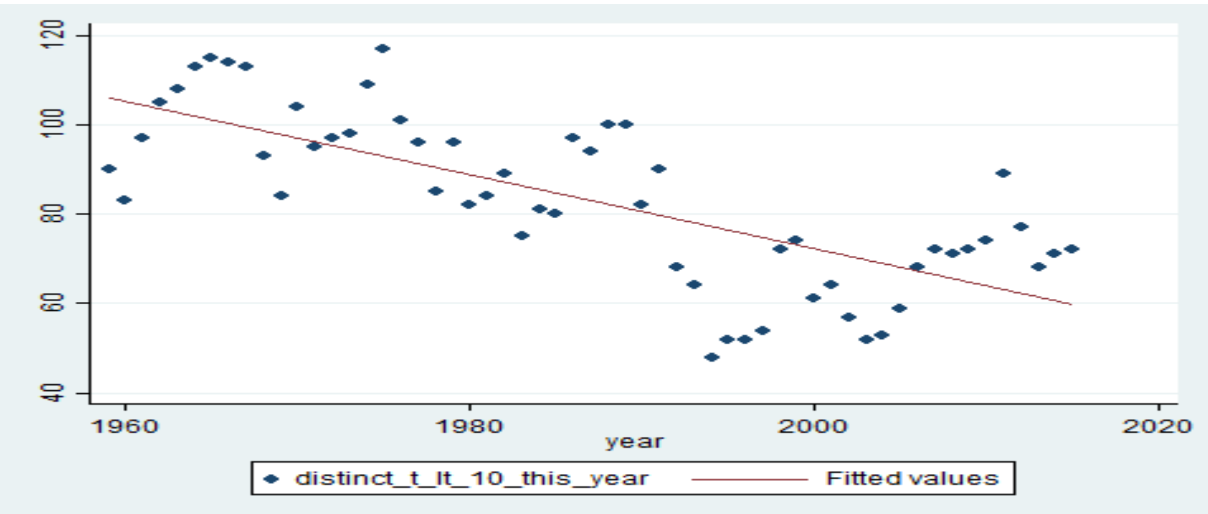


Figure 4 - Number of distinct titles reaching Top 10, per year. The figure displays the number of distinct titles reaching a placement on the chart of ten or better, there is a negative trend in the data.

Figure 4 shows the number of titles reaching a rank of ten or better and it follows the same pattern as in the previously displayed figures, variation of consumption goes down.

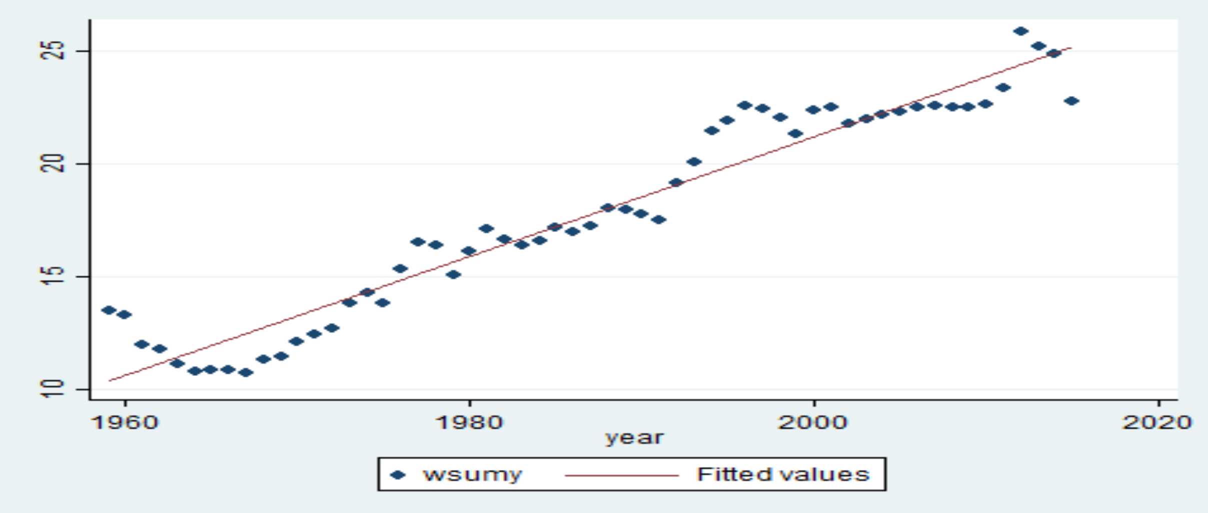


Figure 5 - Mean number of weeks title stay on chart, per year. The figure shows the plotted observed values for the average number of weeks a title stays on the chart, there is an increase in the average number of weeks a title stays on the chart.

Figure 5 shows a clear trend for that the average number of weeks a title stays on the chart is increasing over time. This indicate that there is a higher survival rate on the chart, and a lower circulation of titles.

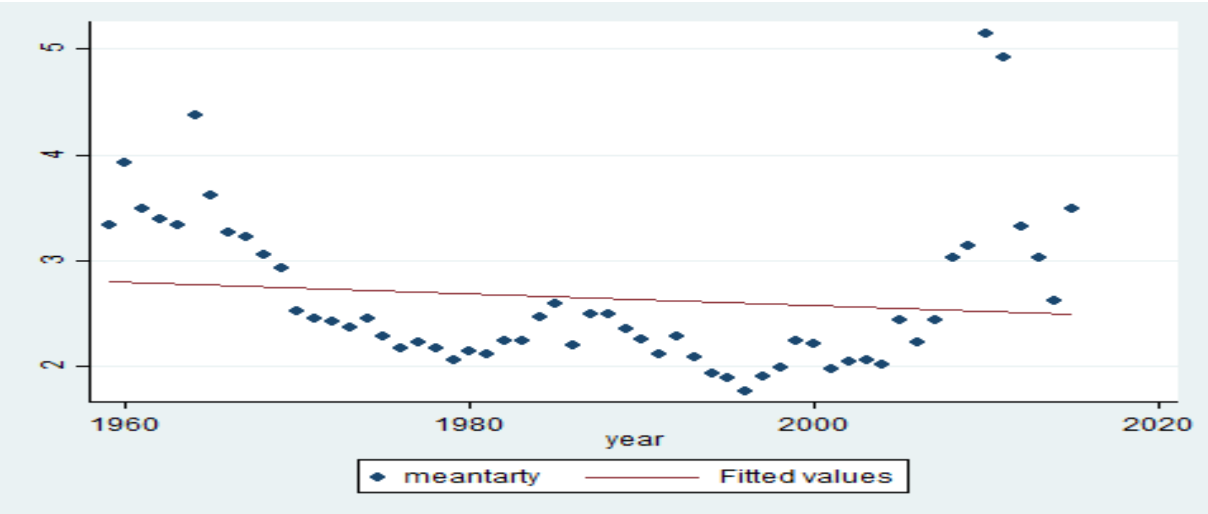


Figure 6 - Number of Titles per Artist, per year. The figure displays the mean number of titles each distinct artist has on the chart for a given year.

Figure 6 shows the number of titles each distinct artist has on the chart for a given year. The plot suggests that there has been a shift in consumption, from an initial decrease in the mean number of titles each artist holds on the chart towards the reversed effect. This could suggest that the model might be better explained by a quadratic function, this will need to be further examined.

3.7. Econometric Tests

The following set of econometric tests is applied to the data to confirm its validity: test for autocorrelation, heteroscedasticity, for non-linear models, trend stationarity, and for structural breaks. Test for structural breaks is however only applied then it is found to be relevant in a combination of a graphical examination and evidence from theory. In this chapter I will go through the applied tests and the possible corrections, for full test statistics see to the appendix.

3.7.1. Autocorrelation

Autocorrelation arise when the error terms correlates with their own past values and can be a severe problem, since it can make the OLS estimation inefficient and result in a faulty estimation of the standard errors. Furthermore autocorrelation is commonly found in time series data. (Verbeek, 2012)

Since the models used in this thesis are based on several times series regressions it was essential to thoroughly test the possible occurrence of autocorrelation. The first test was to look at the Durbin-Watson statistic for the different regressions, for all regressions the Durbin-Watson statistic displayed values much lower than 2, indicating a positive autocorrelation (Verbeek, 2012). Further tests were then conducted to determine what type of autocorrelation that could be prevalent in the regressions. The performed test was the Ljung-Box test and the Breusch-Godfrey Lagrange Multiplier test. An examination of the Ljung-Box test statics returns a similar result for all the regressions, displaying a high spike for the partial autocorrelation function at the first lag and a decaying trend for the autocorrelation. This gives strong support for the inclusion of a first order autoregressive model, AR(1), in the models (Enders, 2015). The Breusch-Godfrey LM test was then performed to test for first-order autocorrelation, adding support for the findings in the Ljung-Box test. The test results showed that one could reject the null hypothesis of no autocorrelation (Verbeek, 2012).

There is strong support for that there is an autocorrelation in the model of order one, therefore all the models were adjusted for an AR(1) process. By adjusting for the AR(1) process, the Durbin-Watson static comes closer to 2, the Ljung-Box shows no sign of autocorrelation and one cannot reject the null then performing the Breusch-Godfrey LM tests. The models therefore seems to be correctly specified when adjusted for an AR(1) process. (Enders, 2015)

3.7.2. Heteroscedasticity

Heteroscedasticity is present when the variance of the residuals is not constant over the sample. Heteroscedasticity might result in problems similar to the ones for autocorrelation. In time-series data the most common form of heteroscedasticity is that the variance of residuals may differ across time (Wooldridge, 2003). The test for heteroscedasticity was performed after the test and correction for autocorrelation as suggested by Wooldridge (2003).

To control for heteroscedasticity, Whites-test for heteroscedasticity was performed. In two regressions, the null for homoscedasticity at a 5% level had to be rejected, suggesting a possible heteroscedasticity in the data (Wooldridge, 2003). Another test was then carried out for these regression, namely for Autoregressive Conditional Heteroscedasticity (ARCH). For both regressions the null hypothesis of no ARCH was rejected, the standard errors were adjusted by applying the Newey-West HAC standard errors as suggested by Wooldridge (2003).

3.7.3. Non-Linear models

By performing a graphical examination, it stands clear that one regression, mean number of titles per artist, might be better fitted with a non-linear model. Following the approach by Enders (2015) and include higher order polynomials and examine their significance exhibit that the variable is better specified with a quadratic model. Allowing for a higher degree of polynomials result in a better specification for the model, and the specified coefficient is significant.

3.7.4. Trend stationarity

The graphical examination shows that there are strong trends presented in the regressions, and therefore it is not relevant to test for stationary. However, by detrending the data, one can control for trend stationarity, meaning that once the trend is removed the process is stationary. To determine if the observations are trend stationary or just random walks with drift, the Augmented Dickey Fuller test is carried out to control for unit-root, the null-hypothesis being that there is a unit-root in the data. (Greene, 2012)

The test results show that we can easily reject the presence of unit-root for all the detrended variables, giving support for that the variables are in fact trend stationary. (Greene, 2012)

3.7.5. Structural Breaks

The plotted data displays a possible structural break at 2005 for the variable, number of titles presented on the chart as a whole. However, only examine the graphical representation of the data is not enough then controlling for structural breaks, it must be some plausible explanation and reason behind the break (Verbeek, 2012). At this point there is quite strong support for that some underlying factors are causing the break in the data. According to Hesmondhalgh (2013) the music industry faced a disruption in the way music was consumed around the millennia, not at least by the introduction of the iTunes store in 2003, allowing people to buy singles as they pleased. Billboard reacted to this change in February 2005 by including digital music sales in its weighting (Molanphy, 2013). Thereby it becomes relevant to look for a structural break at this point. The test for a possible structural break is carried out by performing the Chow Breakpoint Test (Verbeek, 2012).

The result from the Chow Breakpoint Test shows that we can easily reject the null and thereby giving support for a possible break in the data at 2005. To correct for this in the model a dummy is added to the regression for 2005 and onwards to capture the impact of digital sales.

3.7.6. Normality

To control for a normal distribution of the residuals the Jarque-Bera test is applied, and for most of the regressions the residuals are found to be normally distributed (Verbeek, 2012). Due to the small sample size combined with a graphical examination, I find the results for the residuals found not to be normally distributed to be trustworthy. (Yazici & Yolacan, 2007)

4. Empirical Model

This chapter describe the models in this thesis, together with a presentation of the null hypotheses, and how the regression could be used to answer the research question.

4.1. Regressions and Variables

The effect of time on diversity in music consumption were estimated by using the following regressions on the dependent variables described in chapter 3. In all the regressions, the independent variable $Year_t$ states the year t. Furthermore, all the regressions are adjusted for autocorrelation since this was found for all models in in the econometrical tests carried out in chapter 3, the models are corrected for by the inclusion of an AR(1) process.

4.1.1. Effect of Years on Number of Artists on the Whole Chart

Regression 1 – Artists Whole Chart

$$ArtWC_t = \beta_0 + \beta_1 Year_t + \beta_2 ArtWC_{t-1}$$

The dependent variable in this regression is $ArtWC_t$, it states the number of distinct artist represented on the whole chart at year t.

H0: There is no change in the number of distinct Artists on the whole chart as time prevails.

An increase in the number of observations would add support for the theory of the Long Tail, since an increase in distinct observations could be interpreted as a development towards a more diversified music consumption. While a decrease in the number of distinct observations, could be seen as evidence for the Superstar effect.

4.1.2. Effect of Year on Number of Titles on the Whole Chart

Regression 2 - Titles Whole Chart

$$TitleWC_t = \beta_0 + \beta_1 Year_t + \beta_2 TitleWC_{t-1} + \beta_3 Dummy * Year_t$$

The dependent variable, $TitleWC_t$, states the number of distinct titles represented on the whole chart at year t. The Dummy variable is included since a structural brake was found in 2005. The dummy variable captures the possible effect from the year 2005 and onwards.

H0: There is no change in the number of distinct Titles on the whole chart as time prevails and there is no effect from the dummy.

Generally, an increase in number of observations would add support for the Long Tail. However, a decrease in the number of artist in regression 1, combined with an increase in number of titles could also be seen as support for the Superstar effect. Since this could indicate that a lower amount of artists stands for a higher fractions of the songs presented on the chart. A mere decrease in number of observations would give clear support for the Superstar effect.

4.1.3. Effect of Year on Number of Artist Reaching Top 10

Regression 3 - Artist Top 10

$$ArtTop10_t = \beta_0 + \beta_1 Year_t + \beta_2 ArtTop10_{t-1}$$

$ArtTop10_t$, is the dependent variable in this regression, it states the number of distinct artists reaching top ten on the chart at year t.

H0: There is no change in the number of distinct Top 10 Artists on the chart as time prevails.

This regression examines the artists reaching a rank of ten or better and follows the same idea as for the artist presented on the whole chart. However, it is interesting to examine if the same pattern is observed for the top ten-percentile of the chart as for the chart as a whole. An increase in the number of observations would be an indication towards the Long Tail, while a decrease would be an indication for that the observation is best explained by the Superstar effect.

4.1.4. Effect of Year on Number of Titles Reaching Top 10

Regression 4 - Title Top 10

$$TitleTop10_t = \beta_0 + \beta_1 Year_t + \beta_2 TitleTop10_{t-1}$$

The dependent variable in regression 4 is $TitleTop10_t$, it states the number of distinct titles reaching top ten on the chart at year t.

H0: There is no change in the number of Top 10 Titles on the chart as time prevails.

Examining the development for the number titles reaching a position of ten or better follows the same logic as for regression 3. An increase in number of titles reaching top ten would generally be seen as an indication for the Long Tail. However, just as in regression 2, an increase in titles reaching top ten in combination with a decrease in the number of artist reaching top ten, could be seen as evidence for the Superstar effect, with only a higher circulation of titles. A decrease in observations would add support for the Superstar effect.

4.1.5. Effect of Year on Average Amount of Weeks a Title Stays on the Chart

Regression 5 - Average Weeks

$$Weeks_t = \beta_0 + \beta_1 Year_t + \beta_2 Weeks_{t-1}$$

The dependent variable in regression 5, $Weeks_t$, states the average number of weeks a title stays on the chart at year t.

H0: There is no change in the number of weeks a title stays on the chart as time prevails.

This regression is examining the development for titles survival on the chart. Not being able to reject H0 would indicate that there is no prevalence of the Long Tail or Superstar effect on the chart. An increase in the average number of weeks a title stays on the charts would indicate that the observed behaviour is best explained by the Superstar effect. Since a title in average would hold its position on the chart for a longer period of time and thereby reduce the possibility for new songs to enter the chart. A decrease in the number of weeks a title stays on the chart could on the other hand indicate that there is a higher circulation on the chart. Hence, there is a higher degree of diversity in music consumption, giving support for the Long Tail.

4.1.6. Effect of Year on Mean Number of Titles per Artist

Regression 6 - Titles per Artist

$$TitlesArt_t = \beta_0 + \beta_1 Year_t + \beta_2 Year_t^2 + \beta_3 TitlesArt_{t-1}$$

The dependent variable in this regression is $TitlesArt_t$, it states the mean number of distinct titles per distinct artist at a given year.

H0: The mean number of Titles per Artist on the chart is best explained by a linear function.

This model is included give further support for the findings in the regressions examining distinct observations, and furthermore to give an indication for how the chart might be evolving. Not being able to reject H0 would suggest that the number of titles per artist could be explained as a linear function.

5. Empirical Results & Analysis

This chapter start with a presentation of the results from all the regressions, this is followed by an analysis and discussion of how the results could be interpreted, and how they are related to the theories and previously conducted studies, presented in chapter 2.

5.1. Regression Results

Following is a presentation of the regression results.

Table 2 – Result from Regression 1. The table displays the effect of years on the number of distinct artists present on the whole chart.

Variable	*Coefficient*	*Std. Error*	*t-Statistic*	*Prob.*
Intercept	405.7558	10.71212	37.87819	0.0000
Year	-2.689318	0.299484	-8.979825	0.0000
AR(1)	0.415532	0.118086	3.518902	0.0009
R-squared	0.792510			
Adjusted R-squared	0.784680			

In table 2 the results from regression 1 is presented, we can easily reject the null hypothesis, the result show that there has been a decrease in the number of artist presented on the chart of 2,68 per year and the result is significant on the 1% level.

Table 3 – Result from Regression 2. The table shows the effect of year on the number of distinct titles present on the whole chart. Note the effect of the dummy, displaying the impact the opportunity to consume singles has on the number of titles.

Variable	*Coefficient*	*Std. Error*	*t-Statistic*	*Prob.*
Intercept	763.6075	47.86876	15.95210	0.0000
Year	-7.878089	1.520711	-5.180530	0.0000
Dummy	111.0584	52.74908	2.105409	0.0401
AR(1)	0.701259	0.091801	7.638930	0.0000
R-squared	0.792510			
Adjusted R-squared	0.784680			

Table 3 present the results from regression 2. We can reject the null hypothesis, the result show that there has been a decrease in the number of titles presented on the chart by 7,88 titles per year, the result is significant at the 1% level. The structural break in 2005 suggest that there is a strong increase in the number of titles present on the chart. However, the standard errors are quite high due to the small sample size, nevertheless the dummy variable is found to be significant on the 5% level.

Table 4 – Result from Regression 3. The table displays the effect of year on the distinct number of artist reaching a position of ten or better.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Intercept	87.14362	4.701945	18.53353	0.0000
Year	-0.657093	0.138777	-4.734872	0.0000
AR(1)	0.550114	0.114725	4.795057	0.0000
R-squared	0.716256			
Adjusted R-squared	0.705549			

Table 4 present the results from regression 3, the null hypothesis is rejected. The result show that there is a strong downward trend in the data, with a decrease of 0.66 artists reaching top ten per year, the result is significant at the 1% level.

Table 5 - Result from Regression 4. The table displays the effect of year on the distinct number of titles reaching a position of ten or better.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Intercept	109.2206	9.115718	11.98157	0.0000
Year	-0.876629	0.259917	-3.372734	0.0014
AR(1)	0.710357	0.095371	7.448367	0.0000
R-squared	0.777661			
Adjusted R-squared	0.769271			

In table 5 the results from regression 4 is presented, the null hypothesis is as in the other regressions easily rejected. The result show that there is a decrease of 0.88 in the number of titles reaching top ten per year, the result is significant on the 1% level.

Table 6 – Result from Regression 5. The table displays the effect of year on the average amount of weeks a title stays on the chart.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Intercept	9.754691	0.823868	11.84012	0.0000
Year	0.273636	0.022874	11.96292	0.0000
AR(1)	0.733962	0.087438	8.394096	0.0000
R-squared	0.974795			
Adjusted R-squared	0.973844			

In table 6 the result from regression 5 is presented, as can be seen in the table we reject the null hypothesis. The result show that there is an increase in the average number of weeks a title stays on the chart by 0.27 weeks per year and the result is found to be significant on the 1% level.

Table 7 –Result from Regression 6. Displays the effect of year on the mean number of titles each artist has on the list. Note that the model is explained by a quadratic function.

Variable	***Coefficient***	***Std. Error***	***t-Statistic***	***Prob.***
Intercept	4.195707	0.292179	14.36005	0.0000
Year	-0.146026	0.029251	-4.992214	0.0000
Year^2	0.002440	0.000597	4.084574	0.0002
AR(1)	0.550575	0.096323	5.715905	0.0000
R-squared	0.706767			
Adjusted R-squared	0.689849			
F-statistic	41.77770			

Table 7 present the result from regression 6, and the null hypothesis is rejected. There is strong support for that the function is best described by a quadratic function. In the first part of sample the mean number of titles per artist decrease. However, in the more recent years the effect has been reversed. This suggests that a lower number of artists produce a larger fraction of titles reaching the chart. All the variables are individually significant on the 1% level, and so is the F-statistic.

5.2. Analysis & Discussion

The empirical models were estimated to give guidance in the answering of the research question:

Has diversity in music consumption evolved over time, and if it so can the theories of the Long Tail or the Superstar effect be applied to explain the observed development?

As a starting point, it could be useful with a reminder of why music consumption was chosen to be examined in the first place. The music industry has been defined as a frontrunner in the digitalization of cultural industries and the observed development for music consumption is likely to be present in the other entertainment industries as well (Hesmondhalgh, 2013). This of natural reasons makes the music industry interesting to study.

The results from the models clearly suggests that the increased amount of information has led to a change in the variety of music consumption. There has been a strong decline in the number of artist and titles reaching the chart, and the same pattern can be observed for the top ten-percentile, a declining number of artists and titles is making it to the top-position of the chart. The development in music consumption thereby clearly follows the theory of the Superstar effect as described by Rosen (1981) and the Winner-Take-All effect (Noe & Parker, 2005), with a converge of consumption and an increased public attention put towards a few actors. The tendency for consumption to follow the theory of Superstar effect has been prevalent in nearly the whole sample period as can be seen in figure 1-4. The introduction of modern media has not changed this, rather on the contrary. The results from regression 6 seems to suggest that the Superstar effect is even stronger in the modern era with an increase in the mean number of titles per artist.

There is a structural break in the observation for the number of titles reaching the chart from 2005 and onwards, the structural break is explained by the introduction of digital markets. The result suggests that the possibility to consume the preferred singles rather than the whole albums increase the number of titles consumed. This could be interpreted as evidence for the Long Tail effect, but this would be a severe misjudgement. According to Webster & Ksiazek (2012) the modern media market is fragmented, there a few artists stands for a large share of the media attention, and can reach out with their products more easily. As can be seen in regression 1 and figure 1 there is no structural break for the number of artist presented on the chart. The number of artist presented on the chart follows a declining trend throughout the whole sample. This is a central finding in this thesis, since it gives very strong support for the Superstar effect. Even though there is an increase in the number of titles following the introduction of modern media markets, the number of artist keeps declining, suggesting that fewer artist stands for a larger share of the total consumption. This result is supported by Elberse (2010), she suggests that the possibility of unbundling, i.e. to buy singles rather than albums, led to that the Superstar effect prevailed, and this is also what we found.

The results from regression 6 shows that initially in the sample period there was a decline in the mean number of titles per artist, but in more recent years the effect has been the reversed, with an increasing mean amount of titles per artist. This gives further support for the Superstar effect and be understood by the explanation laid out by Webster & Ksiazek (2012), they argue that the possibility to consume top-list products for a low or even free cost will lead to a realization of the Superstar effect. The findings from regression 6 combined with the result from the structural break in regression 2, suggesting that people want to consume titles produced by a declining number of artists, Superstars.

An important part in the explanation behind the findings of why modern media market has increased the market power for a few artists and resulted in a convergence of music consumption, could be in how WoM spreads through modern media channels. According to Liu (2006) and Wenjing, et al. (2008) the important part of WoM is volume, and WoM thereby mainly acts as an awareness effect. The increasing amount of information makes it hard to reach out through the buzz, and the Superstars might create a sufficient amount of WoM to be recognized. The results from regression 6 support this, the increasing mean number of titles per artist suggests that popular artist have an easier time reaching out with their titles. This could also be explained by how the increasing amount of information on the digital media markets allow producers and consumers to shape the market in duality (Webster, 2011). People will demand more titles from the artist they prefer and this will also be supplied by the producers, leading to a convergence on the market.

The results from regression 5 shows that there has been an increase in the average number of weeks a title stays on the chart. This could indicate that that there is a lower circulation on the chart, since the titles in average stays on the chart for a longer period of time and thereby also shut out new entrants. This result is supported by the findings of Giles (2007) and Bhattacharjee, et al. (2007), they found that there has been an increase in the survival for the titles and albums on the top position of the chart, and interpret this as evidence for the Superstar effect.

The results from regression 5 could also be explained by the findings of Wenjing, et al. (2008). They showed that there is an autocorrelation between sales and the volume of WoM. This suggests that a high number of sales, which could be perceived as a placement on the chart, will lead to an increased volume of WoM, and volume is a main driver behind how WoM affect future sales. This suggest that being ranked on the chart is becoming increasingly important, since it acts as an awareness effect towards other consumers, and this lead to a longer persistency on the chart since in aggregate more people will consume the product.

Similar findings were found by Fu & Sim (2011) and Salganik, et al. (2006) their findings suggest that the success for a product is determined by the choice made by previous consumers. According to Fu & Sim (2011) an import explanation behind the observed behaviour is the reduction in uncertainty products that has been previously consumed enjoy. This can be part of the explanation behind the increased lifetime for the titles on the chart. The title has already been consumed and has thereby been "approved" by previous consumers, and this reduce the uncertainty for the title. The reduction in uncertainty surge in the number of people consuming the title and this increase the the title's probability of staying on the chart, further strengthening the awareness effect, and result in a process were success breeds success. This can also be seen in that duration is an important factor behind WoM (Liu, 2006). The increased persistency for the titles can be understood by their own increased duration, since the longer duration will increase the WoM, and this result in an increased sale performance. The reasoning above could be part of the explanation for why consumers according to Fleder & Hosanagar (2009) feel that their variation has increased, while variety in aggregate have decreased. The increase in volume and duration of WoM leads to an increased awareness for the titles, and thereby a larger fraction of people will consume the same title. This could be part of the explanation for why people feel that their individual consumption has increased., while in aggregate it has decreased. The results from regression 5 give further support for that the observed development is best explained by the theory of the Superstar effect. The public attention will be focused on a few artists, which stands for a large fraction of the produced media (Webster & Ksiazek, 2012).

The reduction of uncertainty could also be explained by the reduction in search costs (Liang & Huang, 1998). According to Lynch Jr & Ariely (2000) the reduction in search costs is the main utility improvement for consumers in the digital environment. Several studies have examined how recommendation systems have reduced search costs and how this affect diversity in consumption, the studies have however returned mixed results. According to Fleder & Hosanagar (2009) recommendation systems lead to a reduced diversity, while Oestreicher-Singer & Sundararajan (2010) and Brynjolfsson, et al. (2011) suggest that recommendation systems increase variation of consumption. The results from this thesis clearly support the findings from Fleder & Hosanagar, suggesting that there has been a reduction in variety. According to Fleder & Hosanagar (2009) the recommendation systems have a hard time assigning values to new artist, and this could be an important part of the explanation behind the results from regression 2 and 6. The increasing amount of consumed titles combined with the decreasing amount of artists, could suggest that there is a bias in the recommendation systems, resulting in a decreased variation.

The convergence in music consumption could also be understood by the theory of Free. According to Anderson (2009) there is a difference in cost for abundant and scarce information, and this can be part of why recommendation systems lead to a consolidation of consumption.

The information for top-list artists or titles could be seen as abundant information and can therefore be consumed at a low cost, while the possibility to consume a specialised basket could involve scarce information, and thereby become more expansive. Hence, artists that has previously been ranked could have an easier time reaching out, and this could be a part of the explanation behind the findings from regression 6. Another possible explanation behind the result form regression 6 could be that there is a skewed distribution of talent as suggested by Noe & Parker (2005). The reduction in search cost makes it easier for consumers to find the “best” artists, and this would explain the convergence of music consumption, and the rise of the Superstar effect.

While the results from the empirical models give strong support for the Superstar effect, it in a high regard also discards the theory of the Long Tail, and the possibility to discard the contradicting theory was one of the main reasons for choosing the applied research design. Anderson (2004, 2006) suggest that the introduction of modern media platforms would result in a decreased concentration pattern for the most popular product and that there would be a larger diversity of consumption. The results from the empirical models strictly rejects this statement, there is no tendency for an increased variety in consumption in the examined models. The only result that could be thought of as to be in line with the Long Tail is the observed break for the number of titles presented on the whole chart, but since the decline in artist persist this cannot be interpreted as a result giving support for the Long Tail.

6. Conclusion

The main focus in this thesis has been to study how an increasing amount of information has affected diversity in music consumption. Prior research on how modern media affect consumer behaviour has mainly focused on how WoM impact sales performance. This study instead examined if there has been a change in diversity of consumption over time, and how the observed development can be described by applying the theory of the Long Tail or the Superstar effect. The music industry was chosen since it has been defined as a frontrunner for technological disruption within the entrainment industries, and the development is thereby likely to be seen in other cultural industries as well (Hesmondhalgh, 2013). The results from the models return a very clearly statement, there has been a decreased diversity in music consumption. Furthermore, there has been an increased in the persistency for the titles reaching the chart indicating a lower circulation. The observed behaviour is best explained by applying the theory of the Superstar effect, and the introduction of internet has made the Superstar effect even stronger. Anderson (2009) states the following regarding the the theory of the Long Tail:

“Hit-driven economics is a creation of an age without enough room to carry everything for everybody. Not enough shelf space for all the CDs, DVDs, and games produced. Not enough screens to show all the available movies. Not enough channels to broadcast all the TV programs, not enough radio waves to play all the music created, and not enough hours in the day to squeeze everything out through either of those sets of slots. This is the world of scarcity. Now, with online distribution and retail, we are entering a world of abundance. And the differences are profound.” (Anderson, 2004)

This can, in contrast to the intention by Anderson, be part of the explanation behind the Superstar effect. There is an ever increasing amount of information on the market, but an individual is only able to take in so much information during a day. Being played on the radio or being recommended through Netflix’s recommendation system might be more important today than ever before, since it is the only way to reach out through the buzz.

The findings in this essay have several practical implications. Managers and professionals within the entertainment industry can do well in acknowledging that there is a strong convergence on the music market. The same effect is likely to be present or evolve in the other entrainment industries as well. Previous conducted studies have found that being present on top-lists increase sales, and this thesis suggest that there has been an increased lifetime for the titles on the chart as time has gone by. This suggest that there is an increasing importance for making it to the chart. Therefore, it is central to reach out in the modern media market and create volume of WoM for the product. This will be even more important when introducing a new artist, the results shows that there has been an increase in the number of titles produced per artist. This could be interpreted as what previous popularity gives an edge, making it even harder for new artists to break through. This could also be an indication for that it is important to be present on all possible platforms and not limit the distribution to specific providers or platforms, as has been done recently by a set of artists, since this will reduce the possibility to reach out (Rys, 2016). This is equally important for the platforms, the number of artist consumed has been declining, and in order to stay relevant it is important to have these Superstars on one’s platform.

Further, the findings from this thesis holds information for how recommendation systems could alter their algorithms to become more relevant. Firstly, there is an increased persistency on the chart, and the titles should thereby be recommended for a longer period of time. Secondly, there is a decrease in the number of artist that are consumed. This could be implemented in the recommendation systems as that there is a few Superstar which people actually tend to consume. In later years there has been an increase in titles on the chart, while the decline of artists has continued, this suggest that people consume an artist that they have some previous relationship to, and this could be implemented in the recommendation systems as well.

The applied method in this thesis had several benefits, it was successful in answering the main question, determining if diversity in music consumption has evolved. Furthermore, the applied research design was successful in determining which of the two mainly discussed theories one should apply to the observed behaviour and at the same time reject the other theory. The applied method was thereby successful in its attempt to lay out a foundation for future research to stand on.

However, the applied method is somewhat simplified and future research could do well in further study the Superstar effect by implementing different sorts of survival models on the data to examine how the persistency for titles and artist on the chart might have changed during the years. Future research could also try to find explanations for why the observed development is best explained by the Superstar effect and give stronger support for possible reasons behind the development. Several topics and possible reasons behind the observed development was disused in section 5.2 and future research could do well in trying to find empirical evidence for the discussed issues. An example of this would be to link chart performance to the volume of WoM generated on internet by using traffic from Google analytics. It would also be interesting to extend the research conducted by Bradlow & Fader (2001) and investigate if there has been a change in chart performance over time, and if so how it could be linked to the decreasing diversity of consumption.

Further it could be interesting to examine if there has been a change in variety of consumption on fields outside of the digital markets. For example, one could examine how the increased amount of information affect the tourist industry. Top-list attractions in a city could be perceived as the Superstars in the tourist sector. Being ranked at webpages and apps that recommends restaurants, sightings and so forth could be increasingly important. Since being highly ranked could attract new consumers, and this could lead to that the ones that are recommend creates even more WoM, leading to a realization of the Superstar effect. According to Fu & Sim (2011), the increasing amount of information flowing through internet will make people learn from the behaviour of previous agents and thereby they could end up consuming the same type of attractions as visitors before them has done, resulting in the prevalence of the Superstar effect. This is just a suggesting and the observed behaviour could possibly be linked to many different types of industries, with a reduction in variation and a realization of the Superstar effect.

Another area which would be interesting to examine would be to link the Superstar effect to chain stores, with the same type of chains spreading throughout the world. An explanation for this could be that the possibility to consume the same type of coffee at all the Starbucks around the world reduce the uncertainty. Another explanation could be that a few companies receives an increasing amount of public attention, giving rise to the Superstar effect. This is something for future research to give evidence to.

7. Reference list

Adler, M., 1985. Stardom and talent. *The American economic review,* 75(1), pp. 208-212.

Anderson, C., 2004. *The Long Tail | Wired.* [Online]
Available at: http://www.wired.com/2004/10/tail/
[Accessed 12 May 2016].

Anderson, C., 2006. *The long tail: how endless choice is creating unlimited demand.* London: Random House Business.

Anderson, C., 2009. *Free : radikalt pris - ny ekonomisk modell.* Stockholm: Modernista.

Arndt, J., 1967. *Word of mouth advertising: a review of the literature.* New York: Advertising Research Foundation.

Artursson, E., 2015. *Online Ratings – who decides what games you buy?,* Lund: s.n.

Bakos, Y. & Brynjolfsson, E., 1999. Bundling information goods: Pricing, profits, and efficiency. *Management science,* 45(12), pp. 1613-1630.

Benghozi, P.-J. & Benhamou, F., 2010. The long tail: Myth or reality?. *International Journal of Arts Management,* 2010(March), pp. 43-53.

Bhattacharjee, S. et al., 2007. The Effect of Digital Sharing Technologies on Music Markets: A Survival Analysis of Albums on Ranking Charts. *Management Science,* 53(9), pp. 1359-1374.

Bikhchandani, S., Hirshleifer, D. & Welch, I., 1998. Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades. *Journal of Economic Perspectiv,* 12(3), pp. 151-170.

Billboard, 1958 - 2015. *Billboard The Hot 100.* [Online]
Available at: http://www.billboard.com/archive/charts/2015/hot-100
[Accessed 15 February 2016].

Billboard, 2011. *Billboard - About us.* [Online]
Available at: http://www.billboard.com/articles/events/467859/about-us
[Accessed 16 April 2016].

Billboard, 2016. *Billboard Hot 100 Chart.* [Online]
Available at: http://www.billboard.com/charts/hot-100
[Accessed 18 April 2016].

Bradlow, E. T. & Fader, P. S., 2001. A Bayesian Lifetime Model for the "Hot 100" Billboard Songs. *Journal of the American Statistical Association,* 96(454), pp. 368-381.

Brynjolfsson, E., Hu, Y. & Simester, D., 2011. Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science,* 57(8), pp. 1373-1386.

Brynjolfsson, E., Hu, Y. & Smith, M. D., 2003. Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers. *Management Science,* 49(11), pp. 1580-1596.

Brynjolfsson, E., Hu, Y. & Smith, M. D., 2010. Long Tails vs. Superstars: The Effect of Information Technology on Product Variety and Sales Concentration Patterns. *Information Systems Research,* 21(4), pp. 736-747.

Chevalier, J. A. & Mayzlin, D., 2006. The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of marketing research,* 43(3), pp. 345-354.

Christensen, C. M., 2013. *The innovator's dilemma : when new technologies cause great firms to fail.* 2000 ed. Boston(Mass.): Harvard Business Review.

Christian, B. & Weinstein, D. E., 2006. Globalization and the Gains from Variety. *The Quarterly Journal of Economics,* May, 121(2), pp. 541-585.

Coase, R. H., 1937. The nature of the firm. *economica,* 4(16), pp. 386-405.

Edwards, M., 2007. *Voices find gilt by association | THE AUSTRALIAN.* [Online]
Available at: http://www.theaustralian.com.au/arts/voices-find-gilt-by-association/story-e6frg8n6-1111114930938
[Accessed 18 April 2016].

Elberse, A., 2010. Bye-bye bundles: The unbundling of music in digital channels. *Journal of Marketing,* 74(3), pp. 107-123.

Enders, W., 2015. *Applied Econometric Time Series.* 4. ed. Hoboken(N.J.): Wiley.

Fleder, D. & Hosanagar, K., 2009. Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management science,* 55(5), pp. 697-712.

Fu, W. W. & Sim, C. C., 2011. Aggregate bandwagon effect on online videos' viewership: Value uncertainty, popularity cues, and heuristics. *Journal of the American Society for Information Science and Technology ,* 62(12), pp. 2382-2395.

Giles, D. E., 2007. Survival of the hippest: life at the top of the hot 100. *Applied Economics,* 39(15), pp. 1877-1887.

Greene, W. H., 2012. *Econometric analysis.* 7. ed. ed. Boston: Pearson.

Hausman, J. A., 1981. Exact Consumer's Surplus and Deadweight Loss. *The American Economic Review,* 71(4), pp. 662-676.

Hausman, J. A., 1997. Valuing the Effect of Regulation on New Services in Telecommunications. *Brookings Papers on Economic Activity. Microeconomics,* 1997(1997), pp. 1-54.

Hausman, J. A. & Leonard, G. K., 1997. Superstars in the National Basketball Association: Economic value and policy. *Journal of Labor Economics,* 15(4), pp. 586-624.

Hennig-Thurau, T. et al., 2010. The Impact of New Media on Customer Relationships. *Journal of Service Research,* 13(3), pp. 311-330.

Hesmondhalgh, D., 2013. *The cultural industries.* 3rd ed. London: SAGE.

Katz, E. & Lazarsfeld, P. F., 1955. *Personal Influence, The part played by people in the flow of mass communications.* Glencoe(IL): Free Press.

Kelly, K., 1993. George gilder: When bandwidth is free. *Wired,* September/October.

Kietzmann, J. & Canhoto, A., 2013. Bittersweet! Understanding and Managing Electronic Word of Mouth. *Journal of Public Affairs,* 13(2), pp. 146-159.

Liang, T.-P. & Huang, J.-S., 1998. An empirical study on consumer acceptance of products in electronic markets: a transaction cost model. *Decision support systems,* 24(1), pp. 29-43.

Liu, Y., 2006. Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. *Journal of marketing,* 70(3), pp. 74-89.

Lynch Jr, J. G. & Ariely, D., 2000. Wine online: Search costs affect competition on price, quality, and distribution. *Marketing Science,* 19(1).

Molanphy, C., 2013. *How The Hot 100 Became America's Hit Barometer : The Record : NPR.* [Online]
Available at: http://www.npr.org/sections/therecord/2013/08/16/207879695/how-the-hot-100-became-americas-hit-barometer
[Accessed 18 April 2016].

Noe, T. & Parker, G., 2005. Winner take all: Competition, strategy, and the structure of returns in the internet economy. *Journal of Economics & Management Strategy,* 14(1), pp. 141-164.

Odlyzko, A., 2001. Internet pricing and the history of communications. *Computer networks,* 36, Issue 5, pp. 493-517.

Oestreicher-Singer, G. & Sundararajan, A., 2010. *Recommendation networks and the long tail of electronic commerce.* [Online]
Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1324064
[Accessed 13 May 2016].

Papies, D., Eggers, F. & Wlömert, N., 2011. Music for free? How free ad-funded downloads affect consumer choice. *Journal of the Academy of Marketing Science,* 39(5), pp. 777-794.

Peltier, S. & Moreau, F., 2012. Internet and the ‘Long Tail versus superstar effect’ debate: evidence from the French book market. *Applied Economics Letters,* 19(8), pp. 711-715.

Rosen, S., 1981. The economics of superstars. *The American economic review,* 71(5), pp. 845-858.

Rys, D., 2016. *How Taylor Swift's Apple Marketing Muscle Helped Drake Get Out of the Gate | Billboard.* [Online]
Available at: http://www.billboard.com/articles/business/7378147/apple-music-taylor-swift-marketing-drake-views
[Accessed 23 May 2016].

Salganik, M. J., Dodds, P. S. & Watts, D. J., 2006. Experimental study of inequality and unpredictability in an artificial cultural market. *science,* 311(5762), pp. 854-856.

Sundaram, D. S., Mitra, K. & Webster, C., 1998. Word-of-Mouth Communications: A Motivational Analysis. *Advances in consumer research,* 25(1), pp. 527-531.

Szabo, N., 1999. Micropayments and mental transaction costs. *2nd Berlin Internet Economics Workshop.*

Webster, J. G., 2011. The duality of media: A structurational theory of public attention. *Communication Theory,* 21(1), pp. 43-66.

Webster, J. G. & Ksiazek, T. B., 2012. The dynamics of audience fragmentation: Public attention in an age of digital media. *Journal of communication,* 62(1), pp. 39-56.

Wenjing, D., Gu, B. & Whinston, A. B., 2008. Do online reviews matter?—An empirical investigation of panel data. *Decision support systems,* 45(4), pp. 1007-1016.

Verbeek, M., 2012. *A guide to modern econometrics.* 4. ed. ed. Hoboken(NJ): Wiley.

Williamson, O. E., 1979. Transaction-cost economics: the governance of contractual relations. *The journal of law & economics,* October, 22(2), pp. 233-261.

Vilpponen, A., Winter, S. & Sundqvist, S., 2006. Electronic word-of-mouth in online environments: Exploring referral networks structure and adoption behaviour. *Journal of Interactive Advertising,* 6(2), pp. 8-77.

Wooldridge, J. M., 2003. *Introductory econometrics: a modern approach.* 2. ed. ed. Mason(Ohio): Thomson South-Western.

Yazici, B. & Yolacan, S., 2007. A comparison of various tests of normality. *Journal of Statistical Computation and Simulation,* 77(2), pp. 175-183.

8. Appendix

8.1. Heteroscedasticity

Table 8 - Test Statistics, White Test

Variable	P-Value F-Statistics White	P-Value Chi^2 White
ArtWC	0.0489	0.0491
TitleWC	0.0203	0.0229
ArtTop10	0.9212	0.9171
TitleTop10	0.7363	0.7250
Weeks	0.0873	0.0853
TitlesArt	0.0528	0.0542

Table 9 - Test Statistics, ARCH Test

Variable	P Value F-Test	P-Value F-Test	RESID ^2(-1)
ArtWC	0.0001	0.0004	0.0000
TitleWC	0.0000	0.0000	0.0000

8.2. Autocorrelation

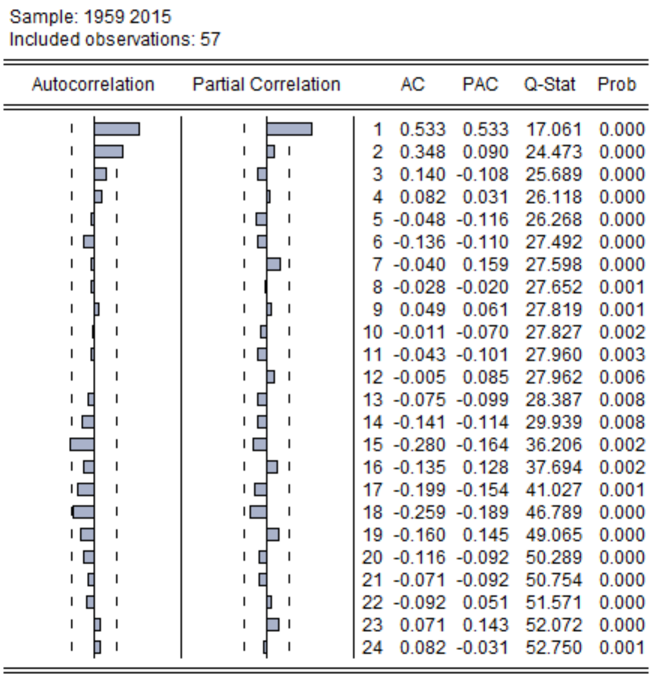
Table 10- Test Statistics, Durbin Watson

Variable	Before Configuration	After Configuration for AR1
ArtWC	1.099588	2.054812
TitleWC	0.719667	2.053616
ArtTop10	0.872280	2.149066
TitleTop10	0.557945	1.996775
Weeks	0.443710	1.508937
TitlesArt	0.869961	1.690096

Table 11 - Test Statistics, Serial Correlation Lagrange Multiplier Test

Variable	***Prob. F***	***Prob. Chi^2***	***Resid -1 Prob***
ArtWC	0.0065	0.0072	0.0027
TitleWC	0.0000	0.0001	0.0001
ArtTop10	0.0001	0.0002	0.0007
TitleTop10	0.0000	0.0000	0.0000
Weeks	0.0000	0.0000	0.0000
TitlesArt	0.0000	0.0001	0.0000

8.2.3. Q-Stat



Sample: 1959 2015
Included observations: 57

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.533	0.533	17.061	0.000
		2	0.348	0.090	24.473	0.000
		3	0.140	-0.108	25.689	0.000
		4	0.082	0.031	26.118	0.000
		5	-0.048	-0.116	26.268	0.000
		6	-0.136	-0.110	27.492	0.000
		7	-0.040	0.159	27.598	0.000
		8	-0.028	-0.020	27.652	0.001
		9	0.049	0.061	27.819	0.001
		10	-0.011	-0.070	27.827	0.002
		11	-0.043	-0.101	27.960	0.003
		12	-0.005	0.085	27.962	0.006
		13	-0.075	-0.099	28.387	0.008
		14	-0.141	-0.114	29.939	0.008
		15	-0.280	-0.164	36.206	0.002
		16	-0.135	0.128	37.694	0.002
		17	-0.199	-0.154	41.027	0.001
		18	-0.259	-0.189	46.789	0.000
		19	-0.160	0.145	49.065	0.000
		20	-0.116	-0.092	50.289	0.000
		21	-0.071	-0.092	50.754	0.000
		22	-0.092	0.051	51.571	0.000
		23	0.071	0.143	52.072	0.000
		24	0.082	-0.031	52.750	0.001

Figure 7 - Artist Top 10 Before

Sample: 1959 2015
Included observations: 56
Q-statistic probabilities adjusted for 1 ARMA term

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.119	-0.119	0.8405	
		2	0.170	0.158	2.5861	0.108
		3	-0.066	-0.031	2.8536	0.240
		4	0.090	0.056	3.3606	0.339
		5	-0.067	-0.039	3.6456	0.456
		6	-0.166	-0.210	5.4371	0.365
		7	0.086	0.080	5.9286	0.431
		8	-0.065	-0.002	6.2167	0.515
		9	0.141	0.111	7.5829	0.475
		10	-0.052	0.015	7.7767	0.557
		11	-0.031	-0.128	7.8464	0.644
		12	0.078	0.072	8.2914	0.687
		13	-0.041	-0.007	8.4174	0.752
		14	0.001	-0.030	8.4175	0.815
		15	-0.275	-0.228	14.396	0.421
		16	0.137	0.052	15.914	0.388
		17	-0.058	0.049	16.193	0.440
		18	-0.185	-0.254	19.118	0.322
		19	-0.011	-0.007	19.130	0.384
		20	-0.033	-0.008	19.226	0.442
		21	0.005	-0.120	19.228	0.507
		22	-0.235	-0.170	24.495	0.270
		23	0.183	0.154	27.801	0.182
		24	-0.005	0.095	27.803	0.223

Figure 8 - Artist Top 10 After

Sample: 1959 2015
Included observations: 57

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.412	0.412	10.185	0.001
		2	0.134	-0.043	11.282	0.004
		3	0.136	0.116	12.430	0.006
		4	0.171	0.094	14.293	0.006
		5	0.075	-0.043	14.656	0.012
		6	-0.089	-0.141	15.178	0.019
		7	-0.026	0.056	15.224	0.033
		8	0.018	-0.003	15.247	0.055
		9	0.012	0.018	15.257	0.084
		10	-0.146	-0.160	16.791	0.079
		11	-0.094	0.037	17.442	0.095
		12	-0.183	-0.223	19.932	0.068
		13	-0.178	-0.013	22.366	0.050
		14	-0.138	-0.021	23.861	0.048
		15	-0.141	-0.038	25.448	0.044
		16	-0.188	-0.145	28.350	0.029
		17	-0.167	0.001	30.692	0.022
		18	-0.040	0.021	30.833	0.030
		19	-0.112	-0.115	31.939	0.032
		20	-0.187	-0.127	35.129	0.019
		21	-0.101	0.072	36.080	0.021
		22	-0.031	-0.104	36.172	0.029
		23	-0.081	-0.081	36.827	0.034
		24	-0.143	-0.095	38.904	0.028

Figure 9 - Artist Whole Chart Before

Sample: 1959 2015
Included observations: 56
Q-statistic probabilities adjusted for 1 ARMA term

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.087	-0.087	0.4502	
		2	-0.047	-0.055	0.5854	0.444
		3	0.021	0.012	0.6129	0.736
		4	0.133	0.135	1.7148	0.634
		5	0.041	0.069	1.8197	0.769
		6	-0.190	-0.173	4.1582	0.527
		7	0.048	0.013	4.3093	0.635
		8	0.053	0.027	4.4987	0.721
		9	0.077	0.089	4.9108	0.767
		10	-0.146	-0.093	6.4195	0.697
		11	0.101	0.097	7.1533	0.711
		12	-0.109	-0.164	8.0250	0.711
		13	-0.080	-0.106	8.5048	0.745
		14	-0.092	-0.097	9.1529	0.761
		15	-0.063	-0.063	9.4663	0.800
		16	-0.034	-0.080	9.5575	0.847
		17	-0.095	-0.040	10.308	0.850
		18	0.105	0.083	11.257	0.843
		19	-0.062	-0.045	11.591	0.868
		20	-0.100	-0.141	12.498	0.863
		21	-0.026	-0.007	12.561	0.895
		22	0.039	-0.017	12.708	0.919
		23	-0.026	-0.015	12.772	0.939
		24	-0.175	-0.139	15.867	0.861

Figure 10 - Artist Whole Chart After

Sample: 1959 2015
Included observations: 57

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.676	0.676	27.474	0.000
		2	0.394	-0.118	36.951	0.000
		3	0.231	0.024	40.278	0.000
		4	0.173	0.064	42.167	0.000
		5	0.015	-0.226	42.181	0.000
		6	-0.108	-0.047	42.944	0.000
		7	-0.174	-0.055	44.989	0.000
		8	-0.241	-0.152	48.971	0.000
		9	-0.325	-0.121	56.390	0.000
		10	-0.355	-0.064	65.420	0.000
		11	-0.274	0.056	70.900	0.000
		12	-0.225	-0.084	74.697	0.000
		13	-0.232	-0.103	78.793	0.000
		14	-0.169	0.061	81.030	0.000
		15	-0.056	0.004	81.280	0.000
		16	0.088	0.128	81.912	0.000
		17	0.159	0.024	84.039	0.000
		18	0.199	-0.009	87.441	0.000
		19	0.152	-0.125	89.486	0.000
		20	0.097	-0.070	90.337	0.000
		21	0.034	-0.060	90.443	0.000
		22	0.029	0.004	90.522	0.000
		23	-0.001	-0.054	90.522	0.000
		24	-0.016	0.051	90.548	0.000

Figure 11 - Average Number of Weeks Before

Sample: 1959 2015
Included observations: 56
Q-statistic probabilities adjusted for 1 ARMA term

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	0.140	0.140	1.1648	
		2	-0.052	-0.073	1.3259	0.250
		3	-0.123	-0.108	2.2522	0.324
		4	0.172	0.210	4.1070	0.250
		5	-0.007	-0.084	4.1102	0.391
		6	-0.065	-0.052	4.3826	0.496
		7	0.058	0.138	4.6090	0.595
		8	0.099	0.012	5.2684	0.627
		9	-0.137	-0.174	6.5639	0.584
		10	-0.216	-0.118	9.8559	0.362
		11	0.031	0.075	9.9240	0.447
		12	0.106	0.010	10.753	0.464
		13	-0.136	-0.165	12.159	0.433
		14	-0.146	-0.019	13.817	0.387
		15	-0.092	-0.112	14.482	0.414
		16	0.189	0.172	17.382	0.297
		17	0.024	0.055	17.430	0.358
		18	0.097	0.101	18.228	0.375
		19	0.028	-0.006	18.299	0.436
		20	0.123	0.070	19.668	0.415
		21	-0.221	-0.214	24.220	0.233
		22	-0.122	-0.077	25.653	0.220
		23	-0.021	-0.064	25.698	0.265
		24	0.160	0.037	28.299	0.205

Figure 12 - Average Number of Weeks After

Sample: 1959 2015
Included observations: 57

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.550	0.550	18.158	0.000
		2	0.161	-0.203	19.735	0.000
		3	-0.085	-0.119	20.188	0.000
		4	-0.190	-0.076	22.470	0.000
		5	-0.120	0.062	23.408	0.000
		6	-0.129	-0.150	24.500	0.000
		7	-0.122	-0.039	25.503	0.001
		8	-0.109	-0.047	26.312	0.001
		9	-0.124	-0.087	27.388	0.001
		10	-0.067	0.005	27.709	0.002
		11	-0.048	-0.062	27.874	0.003
		12	-0.105	-0.150	28.692	0.004
		13	-0.151	-0.107	30.432	0.004
		14	-0.201	-0.137	33.606	0.002
		15	-0.187	-0.115	36.393	0.002
		16	-0.117	-0.088	37.526	0.002
		17	-0.062	-0.106	37.845	0.003
		18	0.023	-0.052	37.888	0.004
		19	0.029	-0.141	37.963	0.006
		20	0.080	-0.008	38.542	0.008
		21	0.107	-0.091	39.610	0.008
		22	0.108	-0.059	40.731	0.009
		23	0.173	0.048	43.693	0.006
		24	0.126	-0.105	45.311	0.005

Figure 13 - Titles per Artist Before

Sample: 1959 2015
Included observations: 56
Q-statistic probabilities adjusted for 1 ARMA term

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	0.142	0.142	1.1951	
		2	-0.089	-0.111	1.6669	0.197
		3	-0.153	-0.128	3.1068	0.212
		4	-0.208	-0.184	5.8000	0.122
		5	0.099	0.135	6.4198	0.170
		6	-0.080	-0.182	6.8399	0.233
		7	-0.047	-0.040	6.9853	0.322
		8	0.000	-0.027	6.9853	0.430
		9	-0.091	-0.089	7.5534	0.478
		10	0.021	-0.041	7.5851	0.576
		11	0.030	0.023	7.6506	0.663
		12	-0.043	-0.097	7.7888	0.732
		13	-0.045	-0.083	7.9393	0.790
		14	-0.120	-0.114	9.0508	0.769
		15	-0.094	-0.130	9.7569	0.780
		16	-0.035	-0.121	9.8588	0.829
		17	-0.056	-0.138	10.118	0.860
		18	0.069	-0.041	10.520	0.880
		19	-0.036	-0.189	10.631	0.909
		20	0.031	-0.052	10.716	0.933
		21	0.050	-0.107	10.948	0.948
		22	-0.029	-0.146	11.027	0.962
		23	0.147	0.012	13.150	0.929
		24	0.002	-0.122	13.150	0.949

Figure 14 - Titles per Artist After

Sample: 1959 2015
Included observations: 57

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.699	0.699	29.346	0.000
		2	0.468	-0.041	42.726	0.000
		3	0.237	-0.146	46.229	0.000
		4	0.032	-0.135	46.293	0.000
		5	-0.121	-0.088	47.235	0.000
		6	-0.247	-0.123	51.251	0.000
		7	-0.145	0.279	52.673	0.000
		8	-0.089	-0.037	53.220	0.000
		9	0.052	0.151	53.406	0.000
		10	0.043	-0.241	53.540	0.000
		11	0.012	-0.079	53.550	0.000
		12	0.048	0.109	53.725	0.000
		13	-0.070	-0.157	54.095	0.000
		14	-0.170	-0.117	56.350	0.000
		15	-0.295	-0.089	63.337	0.000
		16	-0.271	0.020	69.359	0.000
		17	-0.298	-0.155	76.837	0.000
		18	-0.293	-0.027	84.242	0.000
		19	-0.227	-0.076	88.803	0.000
		20	-0.150	0.061	90.848	0.000
		21	-0.038	-0.075	90.984	0.000
		22	0.053	0.195	91.255	0.000
		23	0.156	0.062	93.666	0.000
		24	0.190	-0.011	97.345	0.000

Figure 15 - Titles Reaching Top 10 Before

Sample: 1959 2015
Included observations: 56
Q-statistic probabilities adjusted for 1 ARMA term

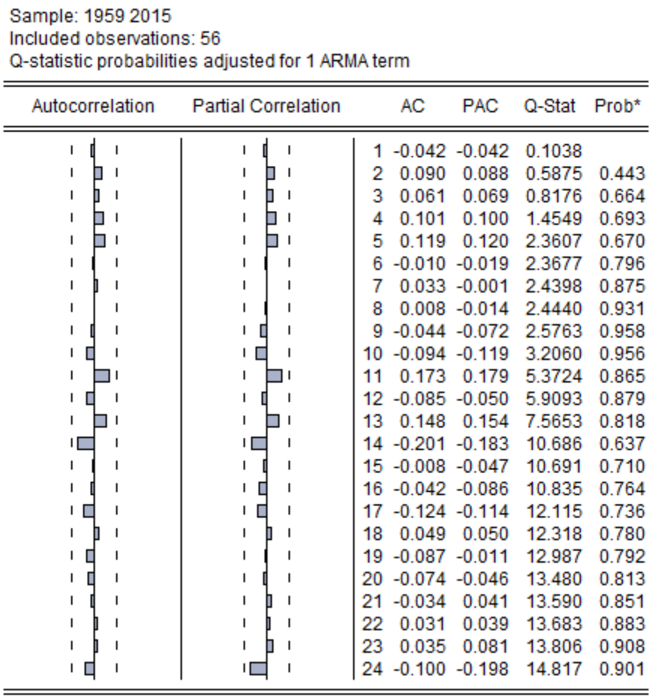
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.017	-0.017	0.0162	
		2	0.120	0.120	0.8825	0.348
		3	0.018	0.022	0.9024	0.637
		4	-0.067	-0.082	1.1837	0.757
		5	0.002	-0.005	1.1840	0.881
		6	-0.352	-0.342	9.2500	0.099
		7	0.064	0.065	9.5228	0.146
		8	-0.124	-0.057	10.567	0.159
		9	0.185	0.219	12.940	0.114
		10	0.017	-0.014	12.960	0.164
		11	-0.033	-0.067	13.040	0.221
		12	0.209	0.073	16.278	0.131
		13	-0.047	0.016	16.443	0.172
		14	-0.001	-0.095	16.444	0.226
		15	-0.227	-0.128	20.543	0.114
		16	0.059	0.074	20.824	0.143
		17	-0.125	-0.108	22.120	0.139
		18	-0.122	-0.060	23.385	0.137
		19	-0.097	-0.170	24.203	0.148
		20	-0.051	0.003	24.437	0.180
		21	-0.025	-0.214	24.495	0.221
		22	-0.043	0.028	24.668	0.262
		23	0.137	0.092	26.513	0.230
		24	0.112	0.186	27.787	0.224

Figure 16 - Titles Reaching Top 10 After

Sample: 1959 2015
Included observations: 57

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.583	0.583	20.436	0.000
		2	0.326	-0.021	26.948	0.000
		3	0.168	-0.022	28.701	0.000
		4	0.076	-0.012	29.071	0.000
		5	-0.067	-0.150	29.359	0.000
		6	-0.148	-0.066	30.804	0.000
		7	-0.210	-0.095	33.779	0.000
		8	-0.254	-0.100	38.214	0.000
		9	-0.204	0.032	41.131	0.000
		10	-0.175	-0.056	43.328	0.000
		11	-0.029	0.152	43.387	0.000
		12	-0.047	-0.140	43.555	0.000
		13	-0.008	0.023	43.560	0.000
		14	-0.090	-0.188	44.192	0.000
		15	-0.092	-0.042	44.876	0.000
		16	-0.142	-0.114	46.529	0.000
		17	-0.119	-0.003	47.723	0.000
		18	-0.067	0.046	48.115	0.000
		19	-0.067	-0.052	48.517	0.000
		20	-0.054	-0.028	48.780	0.000
		21	0.013	0.067	48.796	0.001
		22	0.055	-0.096	49.088	0.001
		23	0.026	-0.038	49.155	0.001
		24	-0.029	-0.196	49.241	0.002

Figure 17 - Number of Titles Whole Chart Before



Sample: 1959 2015
Included observations: 56
Q-statistic probabilities adjusted for 1 ARMA term

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.042	-0.042	0.1038	
		2	0.090	0.088	0.5875	0.443
		3	0.061	0.069	0.8176	0.664
		4	0.101	0.100	1.4549	0.693
		5	0.119	0.120	2.3607	0.670
		6	-0.010	-0.019	2.3677	0.796
		7	0.033	-0.001	2.4398	0.875
		8	0.008	-0.014	2.4440	0.931
		9	-0.044	-0.072	2.5763	0.958
		10	-0.094	-0.119	3.2060	0.956
		11	0.173	0.179	5.3724	0.865
		12	-0.085	-0.050	5.9093	0.879
		13	0.148	0.154	7.5653	0.818
		14	-0.201	-0.183	10.686	0.637
		15	-0.008	-0.047	10.691	0.710
		16	-0.042	-0.086	10.835	0.764
		17	-0.124	-0.114	12.115	0.736
		18	0.049	0.050	12.318	0.780
		19	-0.087	-0.011	12.987	0.792
		20	-0.074	-0.046	13.480	0.813
		21	-0.034	0.041	13.590	0.851
		22	0.031	0.039	13.683	0.883
		23	0.035	0.081	13.806	0.908
		24	-0.100	-0.198	14.817	0.901

Figure 18 - Number of Titles Whole Chart After

8.3. Trend Stationarity

Table 12 - Test Statistics, Augmented Dickey-Fuller

Variable	***ADF Unit Root***
ArtWC	0.0000
TitleWC	0.0000
ArtTop10	0.0000
TitleTop10	0.0000
Weeks	0.0000
TitlesArt	0.0013

8.4. Structural Break

Table 13 - Test Statistics, Chow Breakpoint Test

F-Statistics	Year 2004 P Value
TitleWC	0.0079
F-stat	4.405830
Log Likelihood	13.13525

8.5. Normality

Table 14 - Test Statistics, Jarque-Bera

Variable	Jarque-Bera	P-Value
ArtWC	0.2272	0.8925
TitleWC	1.3860	0.5001
ArtTop10	3.30	0.1919
TitleTop10	0.9476	0.6226
Weeks	6.617287	0.0365
TitlesArt	232.68	0.0000