The Analysis of Methods for Measuring Success in Online Social Networks – Using the Example of SoundCloud

Masters Thesis

Lindsey Alejandra Funk

Spring Term 2016

Advisor:
Andreas Lanz

Chair of Quantitative Marketing and Consumer Analytics
L5, 2 - 2. OG
68161 Mannheim
www.quantitativemarketing.org

Table of Contents

List of Figures

List of Tables

List of Abbreviations

PLC	Product Life Cycle
CD	Change-Dominance
MDA	Multiple Discriminant Analysis
OECD	Organization for Economic Cooperation and Development

Abstract

The present master's thesis deals with the analysis and evaluation of quantitative methods for measuring success in online social networks. Therefore, the online audio-distribution platform *SoundCloud* and its musicians serve as objects of study. The purpose of this work is to explore the effectiveness of the methods from different academic disciplines by empirically applying them to data collected on SoundCloud. Based on both the analysis and the results, conclusions and recommendations for marketing managers dealing with online social networks are derived. After the extensive review of academic literature, three fundamentally different methods, namely *Threshold for Takeoff*, *Structural Break Analysis*, and *Analysis of Growth Episodes*, are adjusted for the context of SoundCloud and applied empirically. The results prove that these methods show strengths and weaknesses in measuring the success of artists on SoundCloud. With some limitations, they are applicable to the context of online social network in general. Future research is needed to get a deeper understanding of the complexities of success in online social networks.

Keywords: *success, takeoff, growth, diffusion, online social networks*

1. Introduction

Today companies around the world are coping with fierce market competition. Especially, corporations in saturated markets, as it is the case in developed countries, have to shorten their products' life cycles in order to launch innovations more often. Sometimes this leads to high rates of product failures due to the company's deficiency to measure the potential of an innovation's success. Thus, it has become essential to develop different methods for measuring the success of a potential product. This facilitates the decision-making process for managers on whether or not to invest in it. Particularly important for the measurement of success is a certain point in the product's life cycle, when sales exhibit a rapid growth in the adoption curve. In academic literature this particular event is referred to as *takeoff* (Tellis, Stremersch, and Yin 2003). For managers this point is of high importance, as the identification of a *takeoff* categorizes the innovation as either successful or not (Golder and Tellis 1997; Tellis, Stremersch, and Yin 2003). This is the case not only with technological products or consumer goods, but also with digital goods.

Most academic studies focus on the diffusion of products through the market starting from the takeoff point (Golder and Tellis 1997). However, there are different operational definitions for *takeoff* as basis for success. Several approaches resulting from different streams of research, such as growth and diffusion literature, have been adopted over the past decades in the attempt to investigate both likelihood and timing of a product's success. As a result, measuring the performance based on takeoff or growth has become one of the most challenging tasks within corporations. For managers it is important to understand (1) how success can be measured, and (2) which method is most appropriate when measuring a product's success.

The present master's thesis gives answers to these questions in the context of online social networks, as their prevailing dominance in today's society should not be ignored. Therefore, different methodological frameworks identified in academic literature are evaluated in regard of their capability to measure success. Moreover, the most promising methods are applied empirically to the case of online social networks using the example of the online audio-distribution platform *SoundCloud*[1].

This work is structured as follows: section 2 gives an overview of both diffusion theory and essentials of online social networks in order to provide the conceptual and theoretical background for the following sections. Section 3 addresses the methodological approach to the literature as well as to the empirical data. In section 4 several methods from different streams of research are presented and evaluated in regard of their ability to measure success in general. Therefore, a wide range of academic publications from the fields of marketing, management, economics, finance, information systems, and biology was reviewed. In section 5 the most promising methods are applied to a provided data set of SoundCloud. In the final section 6 main findings are summarized, managerial implications for the application to the day-to-day business are presented, and limitations as well as directions for future research are discussed.

2. Conceptual & Theoretical Framework

2.1. Diffusion Theory

The diffusion of new products is of vital importance for companies and academics alike. This topic has been the focus of numerous studies across different disciplines, especially marketing, economics, and sociology (Chandrasekaran and Tellis 2007). As a result, an essential body of literature has emerged on both the theory and modeling of diffusion (Mahajan, Muller, and Bass 1990). The *takeoff*, the first rapid significant increase of a new product's sales, builds a substantial part of the diffusion of products, as it represents one of the two key

[1] See https://soundcloud.com.

turning points in the S-shaped diffusion curve (Chandrasekaran and Tellis 2007). Therefore, diffusion modeling provides the foundation for measuring the success of new products.

In general, diffusion is defined as "[…] the process in which an innovation is communicated through certain channels over time among the members of a social system" (Rogers 1995, p. 5). Including the four critical elements innovation, communication, time, and social system, the diffusion process describes the transmission of information about an innovation in the social system (Golder and Tellis 1998; Mahajan, Muller, and Bass 1990). In such a social system, personal influence is responsible for facilitating mass media effects and the speed and shape of the entire diffusion progress. Moreover, the process is expected to present a s-shaped (or *sigmoid*) pattern, although there also has been evidence for other, especially exponential, diffusion patterns. The speed of diffusion is referred to as *rate of diffusion*, which describes the velocity with which product sales emerge over an observed time (Bass 1969; Cox 1967; Gatignon and Robertson 1985).

[Insert Figure 1 about here]

Academics have approached the concept of diffusion by developing diffusion theories based on analytical models. Especially in marketing, the main focus has been to gain insights into the dynamics of new product diffusion. Early first-purchase diffusion models have been proposed by Bass (1969), Mansfield (1961), and Fourt and Woodlock (1960). They represent pioneering works on the analysis of both the penetration and saturation aspects of diffusion (Mahajan, Muller, and Bass 1990). A large part of the research follows the widely known Bass model (1969), which is based on the premise that the probability of an innovation's adoption within a social network with potential m at a given time T depends on the factors mass media and word of mouth (see figure 1). Bass refers to these factors as p and q, with p

being the coefficient of innovation or external influences, and q the coefficient of imitation or internal influences. The basic premise of the Bass model derives from a hazard function, that is " […] the probability that an adoption will occur at time t given that it has not yet occurred" (Mahajan, Muller, and Bass 1990, p. 3). Moreover, the author points out that the probability, that a consumer will adopt a product, is linear with regard to the number of previous adopters (Bass, Krishnan, and Jain 1994; Bass 1969; Mahajan, Muller, and Bass 1990; Peres, Muller, and Mahajan 2010). The model proposed by Fourt and Woodlock (1960), in contrast, states that the main driver of the diffusion process is mass-media communication (external influence) only, whereas Mansfield (1961) sees word-of-mouth communication (internal influence) as main impetus (Mahajan, Muller, and Bass 1990).

The Bass model has been revised and extended in numerous following studies, partially because of the prevailing limitations of the original model. For instance, there has been extensive research on including further model variables, such as market size and repeat purchase, into the model (e.g. Bass, Krishnan, and Jain 1994; Dodson and Muller 1978; Horsky and Simon 1983; Lilien, Rao, and Kalish 1981; Mahajan and Peterson 1978). Moreover, diffusion modeling was successfully put into international context (e.g. Gatignon, Eliashberg, and Robertson 1989; Sarvary, Parker, and Dekimpe 2000).

2.2. Online Social Networks

Today people are exposed to various external influences, which is why diffusion processes have become even more versatile and complex. Besides word-of-mouth communication, network externalities, and social signaling, new market trends, such as online social networks, are affecting consumers' everyday life. Thus, researchers have attempted to modify diffusion modeling by incorporate aforementioned influences (Peres, Muller, and Mahajan 2010).

A social network refers to a "[…] set of actors and the relationships [ties] among them" (Goldenberg et al. 2009, p. 2). Rogers (1995) states that social networks are crucial for the understanding and analysis of diffusion processes. Furthermore, although the social dimension has been recognized in the prevailing diffusion models, novel influences and trends have been of little attention in research. However, Peres, Muller, and Mahajan (2010) suggest including the dimension of social interdependencies into the traditional definition of diffusion. Besides word-of-mouth communication, the authors refer to *network externalities* (Stremersch et al. 2007) and *social signals* (Van den Bulte and Stremersch 2004) as additional influences. Network externalities emerge when "[…] the utility of a product to a consumer increases as more consumer adopt the new product." (Peres, Muller, and Mahajan 2010, p. 92), whereas social signals mean "[…] the social information that individuals infer from adoption of an innovation by others." (Peres, Muller, and Mahajan 2010, p. 92). Social interdependencies play a major role in the diffusion of innovations, as they strongly affect growth processes within a social system. Especially, people with a considerable number of ties to other people, also called *influentials*, *opinion leaders*, or *hubs* (Van Den Bulte and Wuyts 2007), are assumed to have an essential influence on the process of innovation diffusion within a social network. They tend to adopt earlier and thus affect other potential adopter's decision (Goldenberg et al. 2009). This phenomenon is also called *social contagion*, which relates to the direct or indirect effect of consumers who have adopted a product on those who have not adopted it yet. In diffusion modeling, contagion effects are incorporated as one of the main drivers of the diffusion process (Du and Kamakura 2011; Mahajan, Muller, and Bass 1990).

Since takeoff is highly affected by both consumer heterogeneity and consumer interaction, having an understanding of the underlying structure of social networks is of high importance for marketers. This have to be considered and heavily improved, when measuring success in online social networks (Ansari, Koenigsberg, and Stahl 2011; Peres, Muller, and Mahajan 2010).

3. Methodological Approach

3.1. Procedure

The methodological approach for the present work is divided into two main parts in order to give answers to the research questions of (1) how to measure success in online social networks and (2) which methods are most appropriate in doing so. Quantitative methods are applied and evaluated in the principal part of the master's thesis.

In the first part, an extensive literature review is conducted, in order to give a valuable and accurate overview of the state-of-the-art in the measurements of success. A great deal of academic papers from different disciplines is screened and evaluated. The most valuable works are then selected based on potential frameworks they provide for measuring success. They shall constitute the basis for the empirical application.

In the second part, the empirical application of the selected frameworks follows, in order to detect the most effective way(s) to identify success. The most promising methods are applied to secondary SoundCloud data that contain performance information of a sample of musicians. In so doing, the methods' analytic performance in identifying their overall success is investigated. Finally, a conclusion is given and recommendations for managers and future research are made.

3.2. Material

For the purpose of the literature review, relevant scholarly papers from classified A+ and A management and marketing journals (with focus on *Marketing Science*, *Management Science* and *Journal of Marketing Research*) were screened. Diffusion theory and methodology as well as associated concepts were the focus of the research. The takeoff phenomenon as academic topic is relatively under researched, which is why the review was extended to related disciplines. In so doing, the growth literature in economics, the epidemiology literature in computer science, physics, and biology as well as the contagion literature in sociology were extensively reviewed.

After the screening of 192 academic papers from different disciplines, 18 most relevant papers to the topic were chosen as promising basis for the empirical application. In academic research, the concept of "success" mostly modeled in terms of takeoff or growth. Thus, the frameworks found serve as basic concepts for analyzing threshold points and growth patterns.

The empirical part of the work involves the application of the frameworks to the SoundCloud data set. The Chair of Quantitative Marketing at the University of Mannheim provided the data in forefront. The data set contains detailed information on the performance of 554 artists in regard of song plays and followers collected over a time span of 123 weeks. The following information for each artist are covered:

- *Plays per day*
- *Total aggregated plays*
- *New follower(s) per day*
- *Total aggregated follower(s)*

Based on that information, the success of artists shall be identified and analyzed by means of the selected methods. The purpose is twofold: first, making conclusions from the results of the analysis and second, giving recommendations on which measurement is most efficient in identifying success in online social networks in general.

4. Measures of Success

4.1. Success as a Threshold

Various models show the diffusion of innovations through the markets indicating how customers accept new products. However, only little is known about the explicit starting point of success. The factor "success" is described and defined differently across different research streams and disciplines. In the following, two distinct approaches are presented, which measure success as a defined threshold point in the product diffusion.

4.1.1. Takeoff Models

One prevailing approach is to analyze a product's takeoff point. The following papers address this topic from different angles.

According to **Golder and Tellis (1997)**, most successful products have a distinctive takeoff in the diffusion curve, as this indicates a product's transition into the mass market. This first large increase is called the *point of takeoff*. In contrast to common studies, which examine the product's diffusion, the authors offered both definitions and methods with which they are able to identify an explicit takeoff event. The focus of the study lied on the period from market introduction to takeoff. The authors first presented a conceptual definition of takeoff, after which it is described as "the point of transition from the introductory stage to the growth stage of the product life cycle" (Golder and Tellis 1997, p. 257). Moreover, an operational definition for detecting takeoff was developed. In so doing, they found that a "[…] relatively large percentage increase could occur without signaling the takeoff [and] conversely, when the base level of sales is large, the takeoff sometimes occurs with a relatively small percentage increase in sales." (Golder and Tellis 1997, p. 256)

Based on this finding, Golder and Tellis proposed a *threshold for takeoff*. This operational rule represents a plot of percentage growth in product sales relative to its base level of sales (see figure 2). The explicit definition rests upon data on consumer durables that were available to the authors. The *relative growth rate* was described as the annual growth in sales relative to the base sales. Finally, a *takeoff* was defined as the first year in which a product's growth rate in sales passes this threshold. By visual identification, it is possible to directly determine a product's takeoff.

[Insert Figure 2 about here]

Moreover, the authors modeled takeoff as a time-dependent binary event. That is, the probability of the event increases with the length of time a takeoff has not occurred. Showing good results in analyzing duration times, a *proportional hazard function* was applied to model the rate at which the event emerges. With their study, Golder and Tellis provide a pioneering work for identifying the takeoff of new products.

In another work by **Golder and Tellis (2004)**, the takeoff phenomenon was examined within the framework of product life cycles (PLCs). The researchers seek to describe specific metrics for turning points in the PLCs. Moreover, the theory of *informational cascades* was incorporated, which refers to the notion that one adopter's decision is influenced by another individuals' decision (Bikhchandani, Hirshleifer, and Welch 1992). These informational cascades can, for instance, prolong the time-to-takeoff or sharpen the takeoff itself. Consequently, hypotheses on turning points and informational cascades were tested based on data on consumer durables partly following the elaborated framework provided by Golder and Tellis (1997). Results showed that the duration of the growth stage, as an additional dimension, may be shortened or prolonged because of informational cascades.

The takeoff phenomenon was further examined in a multinational context by **Tellis, Stremersch and Yin (2003)**. In their study, the focus is on the identification of takeoff across different countries by adapting the threshold rule developed by Golder and Tellis (1997). The authors analyzed the time-to-takeoff of new products across European countries. The authors used the threshold rule for determining the year of takeoff. However, due to the complexity resulting from the large diversity of base sales across countries, market penetration was chosen as a base for percentage increase. Thus, Tellis, Stremersch, and Yin operationalized takeoff as "[…] the first year a product's growth in sales crosses the threshold" (Tellis, Stremersch, and Yin 2003, p. 197). In conclusion, the work gives good insights into the takeoff phenomenon in a multinational context. The probability of takeoff increases with previous takeoff events and, most important, takeoff in one country enhances the probability of a takeoff in another country.

Based on the findings of Tellis, Stremersch, and Yin (2003), the scholars **Van Everdingen, Fok, and Stremersch (2009)** put their focus on global spillover effects on a country's time-to-takeoff. They used penetration data of innovative technologies across 55 countries. The authors applied the methodology proposed by Tellis, Stremersch, and Yin (2003) to heuristically identify the point of a product takeoff. They specified a threshold plot for market penetration in order to detect takeoff as the first year a product's percentage increase crosses the threshold. Van Everdingen, Fok, and Stremersch found that foreign takeoffs highly affect another country's time-to-takeoff.

The takeoff event was also captured in studies on *indirect network effects*, which describe the understanding of software and hardware forming a system in mutual dependence (Katz and Shapiro 1994). In their work, **Stremersch et al. (2007)** examined the impact of indirect network effects on sales of new technological products. Therefore, they adapted the takeoff analysis along with the concept of *critical mass* (Shapiro and Varian 1998), as this theory refers to the "[…] point of no return after which nuclear fusion becomes self-sustaining" (Stremersch et al. 2007, p. 60). A product's takeoff is thus followed by a rapid self-sustained increase caused by a critical mass of adopters evolved in the course of time. The scholars modified the threshold rule in so far that takeoff was defined as the year in which "the ratio of change in the growth of sales relative to base sales reaches its maximum before the inflection point in sales" (Stremersch et al. 2007, p. 61). The study found that indirect network effects on product sales are not as pervasive as assumed initially.

Product takeoff was further explored in a study conducted by **Markovitch and Golder (2008)**. The scholars investigated stock prices as predictors for long-term survival of corporations at the point of takeoff. They considered takeoff as a critical event in any product diffusion and thus crucial for predicting a firm's long-term success. Therefore, they anticipated sales takeoff based on firms' stock prices. A stock return model was mainly applied to analyze stock prices. In order to identify the year of takeoff, again, the threshold rule developed by Golder and Tellis (1997) was used. The authors found that stock prices could actually predict sales takeoff as well as long-term firm survival. Abnormal stock returns seemed strongly positive one year before takeoff.

The scholars **Agarwal and Bayus (2002)** developed a fairly different approach to determine takeoff compared. In their study, they provided empirical evidence for the assumption that sales exhibit takeoff because of supply-side and demand-side factors. Unlike Golder and Tellis (1997), the authors used a more generalized discriminant analysis to identify sales takeoff. In so doing, it was possible to distinguish between two consecutive intervals regarding to percentage changes in annual sales. The takeoff year was then determined by partitioning the data series into three categories that reflect pre- and post-takeoff as well as the explicit takeoff point. The findings revealed that supply-side as well as demand-side factors lead to rapid takeoffs after initial periods of slow growth.

4.1.2. Turning Point Models

Turning point models represent another valuable approach to analyze success of products. Similar to the takeoff models described previously, explicit threshold are defined and used as success indicators. Once one predetermined value is reached, success is initiated. The most promising works from different disciplines are presented in the following.

In a study conducted by **Delre et al. (2010)** the social influence on innovation adoption was examined. Their analysis followed the basic idea of *percolation models*, which originate in statistical physics (see Stauffer and Aharony 1994), that there exists a network of agents showing different states. These states change and eventually result in a diffusion process. The authors developed an innovative percolation model by incorporating the dimension of social influence. By means of various computation simulations, a percolation threshold was defined based on the data that would indicate social influence in a social network. Results revealed that markets with high social influence exhibit uncertainty, as social influence can either positively or negatively affect the innovation diffusion.

Before simulations of diffusion could be conducted, the scholars **Fourt and Woodlock (1960)** described a method to predict product success based on consumer panel statistics. In so doing, they essentially contributed to the beginnings of diffusion theory and modeling. They developed a model of market penetration in which the focus is on repeat purchase behavior rather than on initial purchase only. The *first repeat ratio*, which describes "[…] the fraction of initial buyers who make a second purchase" (Fourt and Woodlock 1960, p. 32), is considered as the most important factor for a product's future success. Therefore, high values of the ratio indicate success already at an early stage. Results showed the predictive power of threshold values, even though with some limitations.

In another study by **Muller and Yogev (2006)**, success was investigated against the background of *dual-market structures*. The dual-market phenomenon describes the notion that the market for innovations "[…] is composed of early and main markets with a discontinuity in the diffusion process in between them" (Moore 1991; Muller and Yogev 2006, p. 1107). The authors examined the time it would take the main market to become the majority over the early market. The empirical analysis was based on innovation sales data and performed by means of the *Change-Dominance (CD) Time Analysis*. The CD time is described as "[…] the number of years it takes the main market adopters to outnumber the early market adopters" (Muller and Yogev 2006, p. 1108). By visual observation of annual sales data, the researchers could explicitly determine the first time when main market adopters became the majority. The authors found that the CD time could be a predictor of product takeoff, as it was highly correlated with future takeoffs in the diffusion process.

Beyond the scope of marketing research, the field of finance also provides a variety of frameworks measuring success by means of thresholds. For instance, in a study conducted by **Bai and Perron (2003)** a linear multiple structural change model was developed in order to determine so-called threshold points or *breaking dates*. These breaking dates refer to sudden shifts in time series data and may occur in the variance of the data set. In their empirical application, Bai and Perron successfully detected occurrences of such structural changes. In so doing, frequency and timing of possible threshold points were analyzed.

Another procedure that originates in the field of finance was developed by **Altman (1968)**. In his work, the author assessed a firm's performance by applying a multiple discriminant analysis. Therefore, a set of financial and economic ratios was examined as basis for predicting corporate bankruptcy. The multiple discriminant analysis (MDA) is an useful approach to "[…] classify an observation into one of several a priori groupings dependent upon […] individual characteristics" (Altman 1968, p. 591). Groups of firms were established based on different characteristics that were provided in the data, whereby a discriminant score Z classified the objects to either group. This Z-score was regarded as a cut-off point between the groups with which an object could be either in condition “bankruptcy” or “non-bankruptcy”. Altman’s approach demonstrated predictive validity for forecasting a firm’s performance.

4.2. Success As A Process

Success is often described as a process, especially as a process of growth. This notion is valid across many disciplines, especially in economics. In the following, various methods are presented in which success is approached by identifying and analyzing relevant growth processes.

4.2.1. Growth Models

In economic theory, the economic development of countries is a prevailing topic. For instance, in a work by **Easterly (2006)**, the economic growth of numerous OECD countries was investigated in regard of the phenomena *big push* and *poverty trap*, which may arise during economic development. The author claimed that a big push involves a rapid acceleration of growth, which was also referred to as takeoff. Moreover, it was suggested that “[…] the least developed countries are caught in a poverty trap, from which they need a big push involving increased aid and investment to emerge, after which they will have a takeoff into a sustained growth” (Easterly 2006, p. 292). According to Easterly, a takeoff occurs when a country is able to permanently shift its zero per capita growth to a positive one. Therefore, the scholar determined a subsequent per capita growth rate of about 1.5% per year as a result of a structural break analysis. The results showed an elusive picture of the takeoff phenomenon in economic development.

Another model stemming from economics was proposed by **Kelly (1997)**. The author especially focused on the geographical expansion of markets caused by increased specialization of labor force. This so-called *Smithian Growth* was proved to exhibit threshold behavior, since after reaching a critical density of linked markets growth acceleration rapidly arose. The scholar showed that after a slow initial economic growth the critical density of linkages was reached at a threshold time of *t(years)*=1.15. After that, markets expanded noticeably, eventually resulting in commercialization.

A more qualitative approach was developed by **Levine (1997)** in a work on the interdependencies of financial development and economic growth. In his framework, the relationship between stock markets and national growth, amongst others, was examined by measuring the *turnover ratio*. This ratio describes "[…] the total value of shares traded on a country's stock exchanges divided by stock market capitalization […]" (Levine 1997, p. 712). The author found that countries, that were active in trading, exhibited turnover ratios of 0.5. The method was found to be statistically significant in predicting economic growth.

A quite different method for growth analysis was proposed in a study by **Hausmann, Pritchett, and Rodrik (2005)** in which instances of economic growth accelerations were investigated. Based on the standard growth theory, the scholars analyzed clear shifts in growth processes. Therefore, different growth episodes and their determinants were identified and examined by means of spline regressions. *Growth accelerations* were defined as a slope in per-capita growth of 2% or higher, which had to be sustainable for a time horizon of at least 8 years. Moreover, a post-acceleration growth rate had to amount to 3.5% per year. These numbers were based on an observed OECD average.

4.2.2. Spreading Models

In epidemiological models, which originate in physics and biology, the focus lies on the spreading of information in networks. One relevant work was provided by **Bampo et al. (2008)** in the fields of computer sciences. The focus of their study was on social structures in digital networks. The authors developed a computer simulation model for predicting the spreading of viral messages within different kinds of social networks. They utilized the *epidemic threshold parameter μ* that basically measures the growth rate of infected people. If $\mu \geq 1$, the number of infectives grows exponentially. The results confirmed that the nature of social structures is of high importance for the spreading process.

In another study conducted by **Garber et al. (2004)**, word-of-mouth and imitation played a central role. The scholars proposed a model of *spatial analysis*, which is based on diffusion theory, in order to provide reliable predictions of product success. They added the spatial dimension to the temporal dimension of growth because of the correlation between geographic proximity and word-of-mouth spread in product diffusion. By measuring the "distance" between different spatial distribution functions, successful and failed product were discriminated. The discrimination was performed on the basis of a threshold of 16%: the diffusion process indicated success "[…] if 16% of the market is obtained before a specified time T" (Garber et al. 2004, p. 422). The results revealed that spatial divergence approach is an appropriate measure to predict success already at the beginning of product diffusion.

4.3. Evaluation of Methods

In the previous section, different methodological frameworks from a wide range of disciplines were presented in detail. The distinct approaches provide a fruitful basis for answering the research questions of this work. Due to the explicit problem formulations of each study, not all frameworks are transferrable equally well to the actual case of SoundCloud. Therefore, the following evaluation focuses on selected parts of the frameworks only in order to identify the most promising approaches for the purpose of this work.

Takeoff Models

In the study proposed by Golder and Tellis (1997), the authors developed a threshold plot for identifying new product takeoffs based on sales data. This plot depicts percentage growth in product sales relative to its base level of sales with which the explicit point of takeoff is visually identifiable. The application of the approach is easy and generalizable, which makes it transferrable to the SoundCloud case. The available SoundCloud data include information on followers and plays rates that can be regarded as "substitutes" for sales data. In so doing, a *threshold plot* for takeoff can likewise be constructed for followers and plays data in order to identify an artist's point of success once the threshold is exceeded. Consequently, the threshold rule may serve as a useful and efficient method for measuring success of artists on SoundCloud.

Another framework presented by Golder and Tellis (2004) deals with product takeoffs embedded in the context of product life cycles and informational cascades. The key events and stages in a product life cycle were determined. By applying the framework to followers and plays data from SoundCloud, the long-term success of an artist could be measured. However, the theory of informational cascades will not be applied to the actual context, as it is related to more special economic issues.

In the approach worked out by Tellis, Stremersch, and Yin (2003), the operational threshold rule was applied to measure product takeoff across different countries. Instead of sales data, the authors applied market penetration data as a baseline for growth rate in the threshold plot. This variable, however, is problematic in regard of the SoundCloud context, as measuring explicit market size and penetration levels is a quite challenging task in online social networks. Especially, in regard of a musician's success, which is highly influenced by variable consumer taste, the overall "market" size is split into different minor markets changing from day to day. Vague assumptions have to be made in order to determine the market penetration of SoundCloud. Thus, the model proposed by Tellis, Stremersch, and Yin is not transferrable adequately to the SoundCloud data.

This is also the case with the approach suggested by Van Everdingen, Fok, and Stremersch (2009). Similar to the work discussed before, the authors investigated global spill-over effects occurring in international product introduction and takeoff. Therefore, they used penetration data instead of sales data for applying the threshold plot. As already mentioned, the market size is not explicitly quantifiable on SoundCloud. Therefore, the developed framework is not applicable to in the actual context.

In the work conducted by Stremersch et al. (2007), the takeoff phenomenon was investigated against the background of indirect network effects of software and hardware sales. The concept of indirect network effects represents an interesting approach to identify and measure success of interdependent products and services. This is also the case for social online networks such as SoundCloud. For instance, the performance of an artist may be dependent from the overall success of SoundCloud, or vice versa. However, the data set available for the present work does not include information on the performance of the platform itself; rather is the focus on the success of artists independently of the performance of other users or of SoundCloud.

The scholars Markovitch and Golder (2008) investigated stock prices as predictors for long-term survival of corporations at the point of takeoff. In so doing, they were able to measure the duration of the post-takeoff survival of corporations. In order to identify the year of sales takeoff, the threshold rule developed by Golder and Tellis (1997) was applied. Although this method is a fruitful approach for further research, it is not applicable to the actual SoundCloud context. Another parameter than stock price observations that is more related to online social networks could be determined in order to measure the long-term potential of artists.

The authors Agarwal and Bayus (2002) proposed a different approach to measure sales takeoff based on a generalized discriminant analysis. By investigating pre- and post-takeoff intervals in the data, the transition point from one interval to another was identified as a takeoff point. This statistical method may be an efficient framework for measuring the success of musicians on SoundCloud when dividing the data into various intervals.

Turning Point Models

In the study conducted by Delre et al. (2010) the authors did not measure success as such; rather did they develop a framework that measures the process of diffusion, which becomes unleashed once a defined percolation threshold is reached. By combining agent-based models and percolation models, social influence was included in the simulation. However, although the results of the study were highly promising, the framework is hardly transferrable to the SoundCloud context. Network typology and different social influences must be known in order to conduct the analysis.

The scholars Fourt and Woodlock (1960) developed a model of penetration with focus on repetitive purchase behavior to predict new product success. A high first repeat ratio value indicated product success at an early stage. The proposed method enables the quick discrimination of successful products from failed ones. However, the available user data are anonymous and it is not possible to track, if a user plays a song more than once. In general, it is a highly challenging task to track users on online social networks and therefore not feasible within the framework of this work.

In the study conducted by Bai and Perron (2003), sudden shifts in data were analyzed and evaluated. By applying structural break tests based on dynamic programming, the scholars could determine the explicit turning point indicating a shift in data. The method is a versatile tool for identifying sudden structural breaks, especially because of its adaptability to various data types. Although its applicability is complex and requires several different steps, this framework may be a promising method to identify sudden shifts as starting points of success.

Against the background of product diffusion, the scholars Muller and Yogev (2006) investigated dual-market structures and the time it would take the main market adopters to outnumber early market adopters. Therefore, by analyzing product diffusion processes, an average market penetration of 16% was assumed as threshold for the transition from one to another market. This methodological framework is based on classic diffusion theory and is easy applicable. However, as discussed before, market penetration data is not available for the present work.

In a study originating in the fields of finance, the scholar Altman (1968) applied a multiple discriminant analysis in order to predict potential corporate bankruptcies. In so doing, the resulting Z-scores were used to classify firms as either bankrupt or not. This method is hardly transferrable to other frameworks; however, a similar approach with a Z-score as threshold could be developed in order to classify artists as either successful or not.

Growth Models

The scholar Easterly (2006) investigated economic success of countries. He defined economic takeoff as a sequence of periods of positive per capita growth that takes place after a sequence of almost zero growth. A subsequent growth rate of 1.5% per years was determined as an indicator for economic success. By looking for a continuous sequence of growth in followers or plays data, this approach is regarded as a promising method.

The model of geographical expansion of markets, as proposed by the author Kelly (1997), deals with the critical density of linked markets after which rapid growth accelerates. The idea of reaching a critical density in a defined time may be useful for measuring success in an online social network. However, as the available data do not shed light on the structure and related linkages of the SoundCloud network, this method is not applicable to the actual context.

The qualitative approach suggested by the scholar Levine (1997) focused on the interdependencies of financial development and economic growth. The author examined the *turn-over ratios* as a liquidity measure, which indicates that a country has been active in trading. Applied to the SoundCloud context, it may be possible to calculate, for instance, the ratio of listeners turned into followers relative to the total follower base. However, this is not possible within the scope of this work, as there is no tracking of network users.

The scholars Hausmann, Pritchett, and Rodrik (2005) conducted growth analyses by investigating episodes of growth accelerations in economic development. An increase in per-capita growth of 2% and higher sustained for a time horizon of 8 years and a post-acceleration growth rate of 3.5% were defined as indicators for success. The presented method serves as a fruitful basis that is applicable to the available SoundCloud data.

Spreading Models

Bampo et al. (2008) proposed a model for predicting the spread of viral messages in digital networks. By running different computer simulations in online social networks, the spreading rate was analyzed. One crucial parameter for measuring the epidemic-like spread was the epidemic threshold parameter, indicating the number of users who have received the message. The reach of the viral message affects this epidemic threshold. However, as the actual SoundCloud case is based on user activity rather than on user-to-user communication, the proposed method is not applicable to the present context. However, it provides a promising basis for future research.

In the diffusion model developed by Garber et al. (2004), the geographical dimension of growth was added to product diffusion. By observing the spatial patterns of a product's spread, successful products were identified. A product was considered successful, if it was adopted by at least 16% of the potential adopters. Even though this approach represents a simple tool for identifying success on SoundCloud, it is not applicable due to the lack of market penetration data. Moreover, the assumed uniform distribution of customers in an online social network is problematic, as it may distort actual results.

5. Empirical Application to the *SoundCloud* Context

5.1. Facts on SoundCloud

The online social network *SoundCloud* is a leading music-sharing platform launched in 2008. In the following years, it has evolved to the market leader among aggregators of audio-only content (Allington, Dueck, and Jordanous 2015). It allows users, mostly artists, to "(…) collaborate, promote, and distribute their music [and] to upload their own works" (Bogdanov et al. 2011, p. 250). This innovative combination of social online networking and online audio-sharing has made SoundCloud to the world's leading social sound platform (SoundCloud.com 2016). Figures 3 and 4 show the SoundCloud landing page and an exemplary user desktop.

[Insert Figure 3 about here]

[Insert Figure 4 about here]

On the simple, accessible, and feature-rich platform both artist and listener can be engaged in different activities such as following a user, reposting tracks, "favoring" tracks, commenting on tracks, and collaborating as well as interacting with other users in comments (Allington, Dueck, and Jordanous 2015; Bhadane et al. 2014). In so doing, artists can easily spread their music and quickly get analytics and feedback from other musicians or users. The overall performance of an artist can then be evaluated in regard of played songs or total followers.

Within the scope of this work, data on total song plays and total numbers of followers of 554 artists are utilized to apply the methodological frameworks discussed in the previous sections. The data were collected over a time span of 123 weeks. Based on the results of the empirical application, the most reliable method(s) for identifying and measuring an artist's success on SoundCloud shall be determined.

5.2. Empirical Application of Methods on Empirical Data

In the previous sections, different methods were evaluated regarding their applicability to the SoundCloud data in order to identify takeoffs or growth indicating an artist's overall success. By reference to the results, the most conclusive methods shall be found.

5.2.1. Design

The data set includes information on *plays per day*, *total aggregated plays*, *new follower(s) per day*, *total aggregated follower(s)* of in total 554 musicians who were tracked over 123 weeks. *Plays* refers to the number of times a song (of a particular artist) was played by other users, whereas *followers* refers to users who followed an artist's profile and his or her music. In order to clear the data from random large fluctuation, the data is analyzed on a monthly instead of a weekly basis.

The presented figures from the data range from 0 to 206,759 (plays, aggregated) and 12,297 (followers, aggregated) after 123 weeks. Due to the large variance in the data, the 90^{th} percentile of each variable constitutes at a low level: 1,479 for *total plays* and 213 for *total followers*. Therefore, in order to build a viable foundation for the application, the 90^{th} percentile is also chosen to discriminate the best 10% of all musicians. The variable *total aggregated* *followers* is taken as the basis for the selection of top artists, as these numbers reveal the actual success. By contrast, the number of *total plays* may show random events and fluctuations in the course of time.

The best 10% involves a sample of 56 musicians who shall be examined in regard of their overall success in plays and followers rates. The selected group along with the data builds the foundation for the application of the quantitative methods. Analysis and results are shown in the next sections.

5.2.2. Analysis

In this section, the different frameworks are applied to the sampled group of artists in an exemplary way[2]. The methods serve as patterns that are mostly modified in a way that suits to the actual SoundCloud context. For this work, *RStudio* is used for conducting the quantitative analyses. RStudio is an integrated development environment for the statistical programming language R.

I) Threshold for Takeoff (Golder and Tellis, 1997)

The scholars Golder and Tellis (1997) call the first large increase in the diffusion of product sales the point of takeoff. In order to identify this particular point, they heuristically developed a threshold for takeoff. It represents a threshold plot of percentage increase in product sales relative to its base level of sales. The growth rate of product sales per year is set in relation to the base sales. Thus, the first time a product's percentage increase in sales crosses the "threshold" is referred to as a takeoff.

Following this framework, a threshold plot is developed and applied to the sampled best 10% of artists. Therefore, the percentage increase in both absolute plays and absolute

[2] In an exemplary way means that the total sample of 56 artists is relatively small for applying quantitative models and that the applied frameworks are not executed in detail as shown in the academic papers.

followers is plotted relative to the respective base[3]. The numbers are based on the analysis of diffusions of plays and followers for each artist. Figure 5 and 6 respectively show the different courses of percentage increase relative to the base. It becomes evident that most of the musicians exhibit a large increase at the beginning still at a relatively low base.

[Insert Figure 5 about here]

[Insert Figure 6 about here]

However, according to Golder and Tellis (1997), this does not explicitly imply a takeoff, as the base level must be large for a takeoff to occur. The best 5% of the artists, who mostly show high percentage increase still at a large base, are more interesting. In turn, especially in the case of followers (figure 6), no artist marked as green (90th percentile) demonstrates a relevant percentage increase at a large base. This finding arises from the fact that the best 5% of musicians only show more than 500 followers in total.

Based on the descriptive, a threshold plot is developed heuristically, in order to visually identify relevant takeoffs in the diffusion of plays and followers. The separate plots for plays and followers are drawn along the curves of percentage increase, mainly for the top 5%, as they are regarded as the most successful artists in the sample. Figures 7 and 8 demonstrate the explicit course of the developed threshold plots to detect takeoff. Following the definition of Golder and Tellis (1997), takeoff is defined as the first time in which the plays' (or followers', respectively) percentage increase passes the threshold plot.

[Insert Figure 7 about here]

[Insert Figure 8 about here]

---

[3] Meaning percentage increase of plays relative to base plays and, accordingly, percentage increase of followers relative to base followers based on the definition by Golder and Tellis (1997).

The figures 9 to 11 exemplify the threshold plot applied to different artists. In figure 9, a musician reaches a takeoff after crossing the required threshold, whereas in figure 10, this is not the case. Moreover, by reference to the evident takeoffs, it is possible to visually detect the exact month of takeoff. In so doing, an artist's starting point of success can be identified. For example, as exhibited in figure 11, the artist shows a takeoff in month 9 based on the identified takeoff in figure 9.

[Insert Figure 9 about here]

[Insert Figure 10 about here]

[Insert Figure 11 about here]

II) Structural Break Analysis (Bai and Perron, 2003) and Discriminant Analysis (Agarwal and Bayus, 2002)

The scholars Bai and Perron (2003) developed a structural change model to determine so-called threshold points, or breaking points, in a particular data set. These points indicate explicit structural shifts in time series data that can be utilized to identify successful and failing events. Thus, the occurrence of an abrupt structural change can be regarded as the starting point for success (or failure). Both the length and timing of the emerging breaking points reveal more detailed information on the type of success an artist shows.

The scholars Agarwal and Bayus (2002) investigated takeoff of product sales by means of a generalized discriminant analysis. The statistical method allows differentiating two consecutive intervals by exploring the annual percentage change in sales. In order to determine the takeoff year, the time series data are divided into different intervals. These intervals are then categorized into three groups reflecting point of takeoff, pre-takeoff, and post-takeoff periods based on mean values.

Both structural break analysis and discriminant analysis represent similar approaches for measuring explicit transition points that indicate rapid growth. By partitioning time series data into different stages, or intervals, the point of data shift is determined. Therefore, both methods are fused together to one potential framework for identifying success. The model provided by Bai and Perron (2003) is based on dynamic programming and requires several complex steps. Hence, this framework is simplified for the purpose of the present work.

The focus is on the structural changes in the diffusion of both plays and followers of the best 10% of artists. A structural shift in the data is defined as a rapid, large increase in the diffusion curve to a significantly higher level (also called *plateau*) of plays and followers as compared to the time period before. A breaking point reveals a takeoff and indicates the beginning of a new period. By finding the break that is followed by a rapid increase, the period before is labeled *pre-takeoff interval*, whereas the period afterwards is called *post-takeoff interval*. Based on the examination of diffusion curves of plays and followers as well as simulations with varying time variables, the following definitions for the analysis are developed.

A structural break is assumed,

(1) when the mean of period 2 ($mean_2$) minus its standard deviation (SD_2) is at least 14 standard deviations above the mean of period 1 ($mean_1$), and

a structural break is confirmed,

(2) when the plateau in period 2 is maintained for at least 5 months.

Both definitions must be fulfilled for an artist to be considered successful in regard of their performance in plays and followers. The figures 12 and 13 show the defined framework applied to actual SoundCloud cases.

[Insert Figure 12 about here]

[Insert Figure 13 about here]

The developed framework becomes evident in figure 12. The diffusion curve of plays shows a relevant break in month 9 revealing a subsequent, rapid increase. Hence, the first required definition is fulfilled, as *$mean_2$* minus its *SD_2* is at least 14 standard deviations above *$mean_1$*. Moreover, the second interval is sustained for at least 5 months. In regard of total plays, 4 artists show relevant breaking points in their diffusion curves confirming the developed framework and thus an overall successful performance[4]. Figure 13 demonstrates the importance of analyzing the mean of period 2 along with its standard deviation as reference point. If *$mean_2$* would only be the reference point, the analysis would indicate a relevant break without including the large variance in period 2. A break at this point, however, is not the case. Figure 14 illustrates an example for which the definitions do not apply. A relevant breaking point is found only for a standard deviation of 3 or less, but this does not represent a significant increase for success[5].

[Insert Figure 14 about here]

Applied to the diffusion curves of followers, only 2 artists show relevant breaks. However, by adjusting the definitions (1) and (2) through simulations, the number of total artists who exhibited relevant breaks did only slightly increase to 3 in total. This finding additionally approves the initially developed framework.

III) Analysis of Growth Episodes (Easterly, 2006; Hausmann, Pritchett, and Rodrik, 2005)

Measuring success is akin to analyzing growth processes. The scholars Hausmann, Pritchett, and Rodrik (2005) investigated instances of economic growth accelerations of OECD coun-

---

[4] Tables with on overview of all means, standard deviations, and months of break are available in Appendix B.

[5] This conclusion is based on the large variance found in the data; therefore, for an artist to be successful, he or she has to exhibit a relevant increase of at least 14 standard deviations.

tries by identifying growth episodes and their determinants. According to the authors, growth acceleration occurs when an increase in growth of 2% or higher is maintained for at least 8 years. Moreover, after the acceleration stage, the growth rate must be sustainable for at least 3.5% per year.

Economic takeoff was the focus also in a study conducted by Easterly (2006). The takeoff of a country, referred to as "big push", occurs when a country is able to boost its zero per capita growth to a positive one. Therefore, the author defined a consecutive growth rate of 1.5% per year as an indicator for takeoff based on a series of structural break analyses.

The defined growth rates, provided in the two papers, apply for economic growth episodes of OECD countries. For the purpose of this work, they are adjusted to the actual context on the basis of the SoundCloud data. The best 10% of artists are examined regarding their diffusions of plays and followers. As a result, those artists are considered successful

(1) who present a first period of little growth,

(2) followed by a major leap in plays and followers of at least 25% and 20%, respectively.

Moreover, this takeoff is referred to as *growth acceleration*,

(3) when it is sustained for at least 4 months.

Little growth is defined as growth rates under 10%, whereas growth rates of 20% for followers and 25% for plays or larger indicate a takeoff[6]. Moreover, a period of at least 4 months[7] confirms acceleration in growth and thus the success of the artist.

[Insert Figure 15 about here]

[Insert Figure 16 about here]

[6] The respective growth rates follow the definitions of the analysis in I): 10% growth was presented by poorly performing artists; 20% for followers and 25% for plays were regarded as minimum percentage increase for an artist to reach a takeoff. By means of simulations in RStudio these numbers were confirmed.

[7] Based on the analysis in II), 5 months were taken as reference for a sustainable period. However, no artist exhibited constant growth over 5 periods. By means of simulations, 4 months were detected as an optimal period.

Figures 15 and 16 show musicians who fulfill the formulated definitions. In total, 6 artists exhibited relevant growth episodes in their diffusions of plays, whereby only 2 of them demonstrated success in their performance of followers. However, the results show that the definitions following the frameworks by Hausmann, Pritchett, and Rodrik (2005) and Easterly (2006) are applicable to the SoundCloud context to measure successful episodes.

5.2.3. Results

The empirical application of the developed frameworks to the data demonstrates 3 fairly different approaches to how to measure the success of artists on SoundCloud. All 3 methods prove to be efficient in detecting takeoffs and growth processes in the diffusion of plays and followers[8]. Nevertheless, some limitations in the implementation of each framework have appeared.

The developed threshold plot in I) is an easy-to-use and valid method to quickly identify takeoff. It was easily applicable to the SoundCloud data and it proved the success of the most successful artists. The threshold plot for takeoff can be regarded as a universal tool for measures of success also in other contexts of online social networks. However, as the plot is developed on a heuristically basis, it is highly exposed to arbitrariness. The determination of takeoff may be not as accurate as with merely quantitative procedures. The framework still represents a viable and effective way to identify success. As only method it has detected the best performing artists from the initially selected 10%.

In II), a method mainly based on a combination of a classical structural break analysis and a generalized discriminant analysis was derived. It required several more steps in order to identify relevant breaking points in the diffusion of plays and followers. By calculating the means and standard deviations of different time periods, significant positive shifts in the data

[8] See Appendix B for an overview of results of the empirical applications in I), II), and III).

are denoted as success. The identification of these shifts rests upon quantifiable measures through which the results become empirically significant. This approach is applicable to a variety of data, which makes it universally usable in other contexts. However, the main limitation is that relevant breaking points can only be determined in retrospective, as the analysis highly depends on the length of the period of interest. As a consequence, the analysis only works by using historical data. Moreover, the method showed only decent conclusiveness in detecting and measuring the success of the best performing artists from the 56 in total.

The third measurement presented in III) was derived from approaches mainly used in the fields of economics. The analysis of growth episodes is a simple way of identifying successful artists by looking for significant growth stages in the diffusion of plays and followers. The straightforwardness of this method, however, results in some limitations. First, as described for method II), historical data is needed to identify periods as successful growth episodes, since a defined period must maintain a certain growth rate. Second, this analysis shows only the relative growth rate and thus the relative success of an artist. Thus, as indicated in the results, some artists were considered successful also with only a small number of base plays or followers.

6. Discussion

6.1. Summary of Findings

The main purpose of this master's thesis was to investigate, analyze, and evaluate different methods across various academic disciplines with which it becomes feasible to measure success in online social networks. The online audio-sharing platform SoundCloud and its users served as objects for this investigation, since the success of musicians plays a central role on this social network.

After providing a theoretical and conceptual foundation about diffusion theory and modeling as well as social networks in general, the main part of the work followed consisting of two major parts. First, an extensive literature review of numerous academic papers across different disciplines was performed, including marketing, management, physics, and biology. In so doing, the most promising and fruitful methodological frameworks for measuring success were selected. It has become evident that the term "success" is defined in various ways, mainly as takeoff or growth. Therefore, these two concepts were regarded as definitions of success for the purpose of the work. Second, based on the preceding review of literature and after an extensive evaluation, 5 methods built the foundation for the following analysis. Modified and partly combined for the SoundCloud context, the methodological frameworks were empirically applied. In so doing, their efficiency and reliability in terms of measuring the success of artists were tested.

The analysis revealed that the final methods applied to the data set showed strong conclusiveness in measuring success. The *Threshold Plot for Takeoff* was heuristically developed on the basis of the available data. It was possible to visually detect an artist's starting point of success, also called takeoff, once the percentage increase in plays or followers reached a defined minimum of the plot. The method's strength especially was the fact that it is easy-to-use and transferrable to different contexts. However, one limitation was found to be the possible arbitrariness in developing the threshold plot and the resulting dependence of the method's efficiency from the data. The second examined methodological framework was the *Structural Break Analysis*. The method was modified extensively for the application to the data and it showed decent results in measuring the success of artists by detecting relevant breaks in the data structure. It exhibited a major weakness in being highly dependent from historical data, though. Moreover, its application is more complex as that of the other methods due to the numerous statistical simulations it requires. The third and last method was the *Analysis of*

Growth Episodes, which was composed of two major frameworks from the fields of economics. Based on previous descriptive analyses of the data, explicit definitions were determined as requirements for success. This method was a simple tool for detecting successful growth episodes. However, two major limitations involve its dependence on historical data as well as its capability of measuring only the relative instead of the absolute success. This partly led to elusive results.

In conclusion, both the literature review and the empirical application of the methodological frameworks shed light on different dimensions of the term success in online social networks in general, and on SoundCloud in particular. In the empirical part of this work, all 3 methods showed strengths and weaknesses. However, the *Threshold Plot for Takeoff*, following the pioneering work by Golder and Tellis (1997), appears to be the most robust and valid method for the analysis of success. The threshold plot is a tool that is capable of grasping all the different dimensions of success, be it takeoff or percentage increase. With this method it was possible to detect the success of the highest performing artists out of the selected 10%. Moreover, in contrast to the other frameworks, the threshold plot enables to identify success already at its initiation stage. This makes this methodological framework, despite of possible arbitrariness, a powerful tool to measure overall success in online social networks.

6.2. Managerial Implications

The findings obtained in the previous chapters have significant implications for companies and their managers around the world. More and more managers have to cope with the prevailing dominance of online social networks. The stake of consumers who are present on online social networks and get informed about brands and products is getting larger. Therefore, it is highly important for managers to get to know their “digital consumers” by learning their habits, needs, and buying behavior. The different online social platforms enable firms to collect relevant data of (potential) customers. In so doing, online marketing strategies can be developed that address the preferable markets on such platforms directly.

Measuring the success of these strategies or their effects on product sales is vitally important for companies. In recent years, plenty of approaches have been published that should help managers to cope with the flood of data being available at once. Converting these data to diagnostic figures, in order to make recommendations, have become one major obstacle for managers. Although there are relevant contributions to this topic, there still exists a substantial lack of research.

The present work and its analysis of different methodological frameworks draw implications for overcoming this major hurdle. The partially quite abstract methods were derived and applied to the actual context of an online social network. The music-sharing platform SoundCloud is especially qualified for the analysis, as its users, mostly musicians, can easily be assessed based on their rates of plays and followers. In so doing, the success of an artist can be examined. This measurement of success, as seen in the empirical part of this work, can be transferred and applied to other similar contexts in the Internet environment. For instance, it is possible to measure the success of marketing campaigns or products assessed by consumers on product forums. This enables managers to react more quickly and accurately to the consumers' behavior and needs. It is of high importance to collect relevant data for the measurements of success; however, this is not always possible due to security and privacy issues. Therefore, it is inevitable for operators of online social networks and companies to provide a safe and trustworthy online environment for their users, as this is one major precondition for online social networks to be used.

6.3. Limitations and Future Research

This master's thesis has presented valuable insights into the measurements of success in online social networks. However, the work also has encountered some limitations. First, the empirical application of the methodological frameworks to online social networks is exemplified by the audio-sharing platform SoundCloud. This has the main advantage that the concept of success is easy to grasp by means of the artists' performance. Thus, SoundCloud represents a special case of online social network. On other platforms, such as Facebook or Twitter, a user or a post is not considered "successful" as such. This may lead to possible biases in the perception of what post or which user is finally objectively successful. As a consequence, the application of the developed frameworks to other social networks may be not easy to grasp. Second, the underlying network structure of SoundCloud is not known and thus not included in the analysis of the methods. However, as described in the theory as well as investigated in relevant academic papers, this can have significant influence on both the application of methods and the interpretation of success on SoundCloud. Third, the nature of the available SoundCloud data may have affected the final results of the analysis due to high fluctuations as well as the partly poor artist performance. Fourth, the concept of success in general is a widely ranged topic that is defined and handled in multifold ways across disciplines. For the purpose of this master's thesis, success was defined as takeoff and growth; however, in other contexts this may vary significantly. It is crucial to clarify what kind of success will be analyzed. Thus, the results of this work are eligible only for the present research context.

The present master's thesis sheds light on an issue often neglected in academic research. The definition and measurement of success is of high importance for companies and managers alike. Therefore, there is still scope for future research in this field. The limitations pointed out before serve as starting points for further studies. Especially, the enlargement of the data set and the application of methods to other online social networks are important improvements. Moreover, it may be interesting to incorporate data on the network structures and network penetration to the empirical analysis.

Bibliography

Agarwal, Rajshree and Barry L. Bayus (2002), “The Market Evolution and Sales Takeoff of Product Innovations,” *Management Science*, 48 (8), 1024–41.

Allington, Daniel, Byron Dueck, and Anna Jordanous (2015), “Networks of value in electronic music: SoundCloud, London, and the importance of place,” *Cultural Trends*, 24 (3), 211–22.

Altman, Edward I. (1968), “Financial Rations, Discriminant Analysis and the Prediction of Corporate Bankruptcy,” *Journal of Finance*, 23 (4), 589–609.

Ansari, Asim, Oded Koenigsberg, and Florian Stahl (2011), “Modeling Multiple Relationships in Social Networks,” *Journal of Marketing Research*, XLVIII (August), 713–28.

Bai, Jushan and Pierre Perron (2003), “Computation and Analysis of Multiple Structural Change Models,” *Journal of Applied Econometrics*, 18 (1), 1–22.

Bampo, Mauro, Michael T. Ewing, Dineli R. Mather, David Stewart, and Mark Wallace (2008), “The Effects of the Social Structure of Digital Networks on Viral Marketing Performance,” *Information Systems Research*, 19 (3), 273–90.

Bass, Frank M. (1969), “A New Product Growth for Model Consumer Durables,” *Management Science*, 15, 215–27.

———, Trichy V. Krishnan, and Dipak C. Jain (1994), “Why the Bass Model Fits without Decision Variables,” *Marketing Science*, 13 (3), 203–23.

Bhadane, C., H. A. Mody, D. U. Shah, and P. R. Sheth (2014), “Use of Elastic Search for Intelligent Algorithms to Ease the Healthcare Industry,” *International Journal of Soft Computing and Engineering*, 3 (6), 222–25.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992), “A Theory of Fads, Fashion, and Cultural Change as Information Cascades,” *Journal of Political Economy*, 100 (5), 992–1026.

Bogdanov, Dmitry, Martin Haro, Ferdinand Fuhrmann, Anna Xambo, Emilia Gomez, and Perfecto Herrera (2011), “A content-based system for music recommendation and visualization of user preferences working on semantic notions,” *Proceedings - International Workshop on Content-Based Multimedia Indexing*, 249–52.

Van den Bulte, Christophe and Stefan Stremersch (2004), “Social Contagion and Income Heterogeneity in New Product Diffusion: A Meta-Analytic Test,” *Marketing Science*, 23 (4), 530–44.

Van Den Bulte, Christophe and Stefan Wuyts (2007), “Social Networks and Marketing,” *Internetworking Research And Experience*.

Chandrasekaran, Deepa and Gerard J. Tellis (2007), “A Critical Review of Marketing Research on Diffusion of New Products,” in *Review of Marketing Research*, N. K. Malhotra, ed., Emerald Group Publishing Limited, 39–80.

Cox, William E . Jr. (1967), “Product Life Cycles as Marketing Models,” *Journal of Business*, 40 (4), 375–84.

Delre, Sebastiano. A., Wander Jager, Tammo H. A. Bijmolt, and Marco A. Janssen (2010), “Will It Spread or Not? The Effects of Social Influences and Network Topology on Innovation Diffusion,” *Journal of Product Innovation Management*, 27 (1), 267–82.

Dodson, Joe A. and Eitan Muller (1978), “Models of New Product Diffusion Through Advertising and Word-of-Mouth,” *Management Science*, 24 (15), 1568–78.

Du, Rex Y. and Wagner A. Kamakura (2011), “Measuring Contagion in the Diffusion of Consumer Packaged Goods,” *Journal of Marketing Research*, 48 (1), 28–47.

Easterly, William (2006), “Reliving the 1950s: The Big Push, Poverty Traps, and Takeoffs in Economic Development,” *Journal of Economic Growth*, 11 (4), 289–318.

Van Everdingen, Yvonne, Dennis Fok, and Stefan Stremersch (2009), “Modeling Global

Spillover of New Product Takeoff," *Journal of Marketing Research*, 46 (5), 637–52.

Fourt, Louis A and Joseph W Woodlock (1960), "Early Predictions of Market Success for New Grocery Products," *Journal of Marketing*, 25 (2), 31–38.

Garber, Tal, Jacob Goldenberg, Barak Libai, and Eitan Muller (2004), "From Density to Destiny: Using Spatial Dimension of Sales Data for Early Prediction of New Product Success," *Marketing Science*, 23 (3), 419–28.

Gatignon, Hubert, Jehoshua Eliashberg, and Thomas S. Robertson (1989), "Modeling Multinational Diffusion Patterns: An Efficient Methodology," *Marketing Science*, 8 (3), 231–47.

——— and Thomas S. Robertson (1985), "A Propositional Inventory for New Diffusion Research," *Journal of Consumer Research*, 11 (March), 849–67.

Goldenberg, Jacob, Sangman Han, Donald R. Lehmann, and Jae W. Hong (2009), "The Role of Hubs in the Adoption Process," *Journal of Marketing*, 73 (2), 1–13.

Golder, Peter N. and Gerard J. Tellis (1997), "Will It Ever Fly? Modeling the Takeoff of Really New Consumer Durables," *Marketing Science*, 16 (3), 256–70.

——— and ——— (1998), "Beyond Diffusion : An Affordability Model of the Growth of New Consumer Durables," *Journal of Forecasting*, 280 (17), 259–80.

——— and ——— (2004), "Growing, Growing, Gone: Cascades, Diffusion, and Turning Points in the Product Life Cycle," *Marketing Science*, 23 (2), 207–18.

Hausmann, Ricardo, Lant Prtichett, and Dani Rodrik (2005), "Growth Accelerations," *Journal of Economic Growth*, 10 (4), 303–29.

Horsky, Dan and Leonard S. Simon (1983), "Advertising and the Diffusion of New Products," *Marketing Science*, 2 (1), 1–17.

Katz, Michael L and Carl Shapiro (1994), "Systems Competition and Network Effects," *Journal of Economic Perspectives*, 8 (2), 93–115.

Kelly, Morgan (1997), “The Dynamics of Smithian Growth,” *Quarterly Journal of Economics*, 112 (3), 939–64.

Levine, Ross (1997), “Financial Development and Economic Growth: Views and Agenda,” *Journal of Economic Literature*, 35 (2), 688–726.

Lilien, Gary L., Amber G. Rao, and Shlomo Kalish (1981), “Bayesian Estimation and Control of Detailing Effort in a Repeat Purchase Diffusion Environment,” *Management Science*, 27 (5), 493–506.

Mahajan, Vijay, Eitan Muller, and Frank M. Bass (1990), “New Product Diffusion Models in Marketing: A Review and Directions for Research,” *The Journal of Marketing*, 54 (1), 1–26.

——— and Robert A. Peterson (1978), “Innovation Diffusion in a Dynamic Potential Adopter Population,” *Management Science*, 24 (15), 1589–97.

Mansfield, Edwin (1961), “Technical Change and the Rate of Imitation,” *Econometrica*, 29 (4), 741–66.

Markovitch, D. G. and P. N. Golder (2008), “Using Stock Prices to Predict Market Events: Evidence on Sales Takeoff and Long-Term Firm Survival,” *Marketing Science*, 27 (4), 717–29.

Moore, Geoffrey A. (1991), *Crossing the Chasm: Marketing and Selling High-Tech Products to Mainstream Customers*, *Harper Business*, New York.

Muller, Eitan and Guy Yogev (2006), “When does the majority become a majority? Empirical analysis of the time at which main market adopters purchase the bulk of our sales,” *Technological Forecasting and Social Change*, 73 (9), 1107–20.

Peres, Renana, Eitan Muller, and Vijay Mahajan (2010), “Innovation Diffusion and New Product Growth Models: A Critical Review and Research Directions,” *International Journal of Research in Marketing*, 27, 91–106.

Rogers, Everett M. (1995), *Diffusion of Innovations*, *The Free Press*, New York: The Free Press.

Sarvary, M, M Parker, and Marnik G. Dekimpe (2000), “Global Diffusion of Technological Innovations: A Coupled-Hazard Approach,” *Journal of Marketing Research*, 37 (1), 47–59.

Shapiro, Carl and Hal R. Varian (1998), *Information Rules: A Strategic Guide to the Network Economy*, *Harvard Business School Press*, Boston.

Stauffer, Dietrich and Amnon Aharony (1994), “Introduction to Percolation Theory,” *Computer*, 1 (4), 192.

Stremersch, Stefan, Gerard J. Tellis, Philip H. Franses, and Jeroen L. G. Binken (2007), “Indirect Network Effects in New Product Growth,” *Journal of Marketing*, 71 (3), 52–74.

Tellis, Gerard J., Stefan Stremersch, and Eden Yin (2003), “The International Takeoff of New Products: The Role of Economics, Culture, and Country Innovativeness,” *Marketing Science*, 22 (2), 188–208.

Appendix

Appendix A: Figures

Figure 1: Bass Diffusion Model

Illustration of the analytical structure of the Bass model (Mahajan, Muller, and Bass 1990).

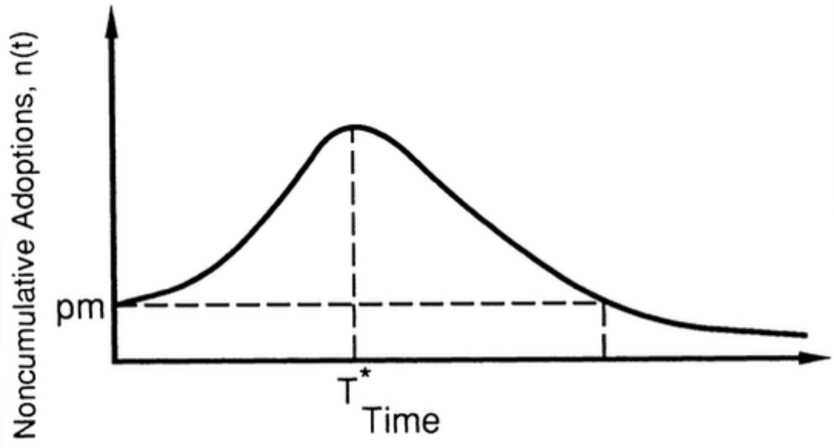


Figure 2: Threshold Plot for Takeoff

Determined threshold plot by Golder and Tellis (1997) based on sales data.

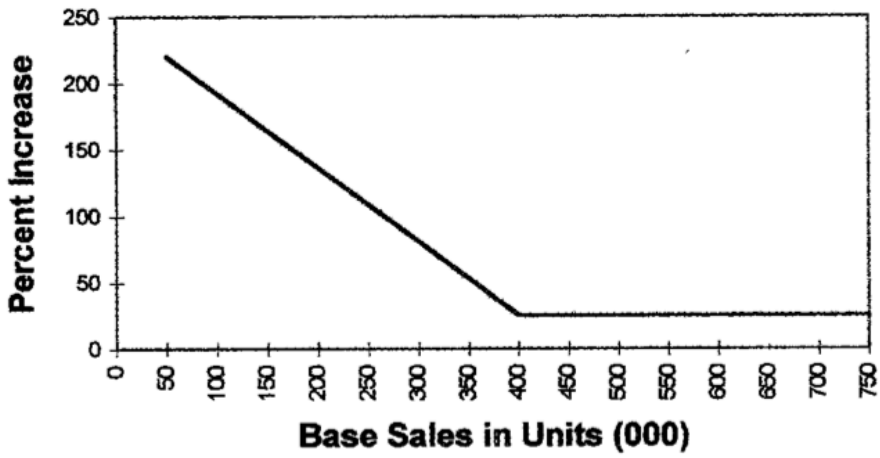


Figure 3: SoundCloud Landing Page

Landing Page of SoundCloud. Users need to have an account (SoundCloud.com).

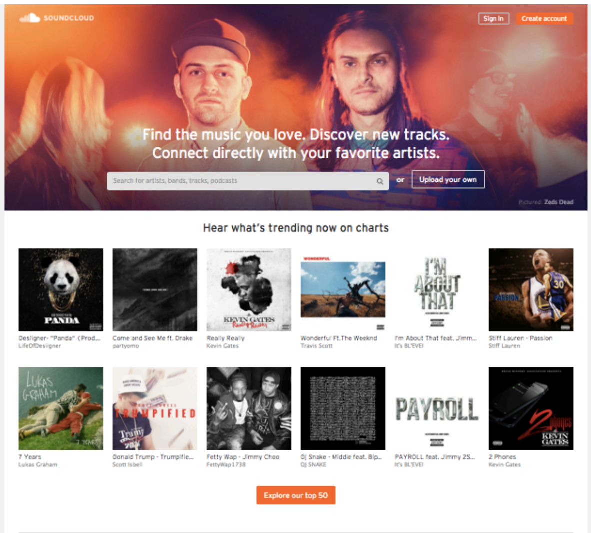


Figure 4: SoundCloud User Desktop

An exemplary desktop of a user on SoundCloud (SoundCloud.com).

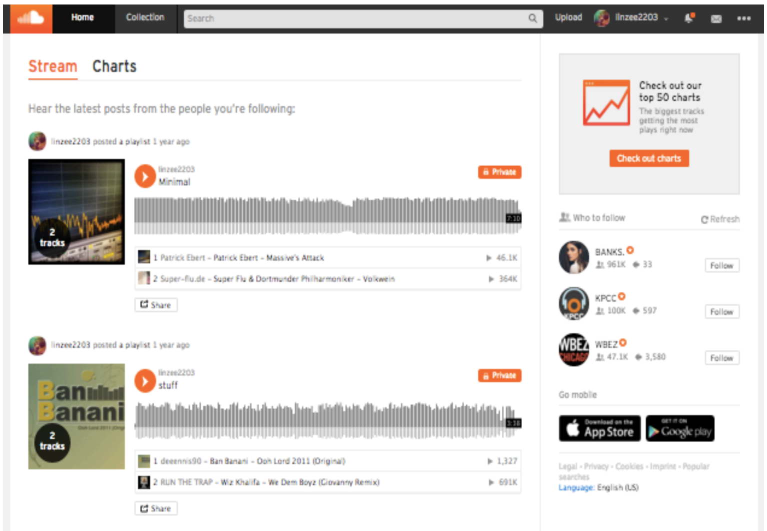


Figure 5: Overview Percentage Increase of Plays for Best 10%

Overview of the percentage increase in relation to base plays of total best 10% artists classified into 90th, 95th, and 99th percentiles (author's own work).

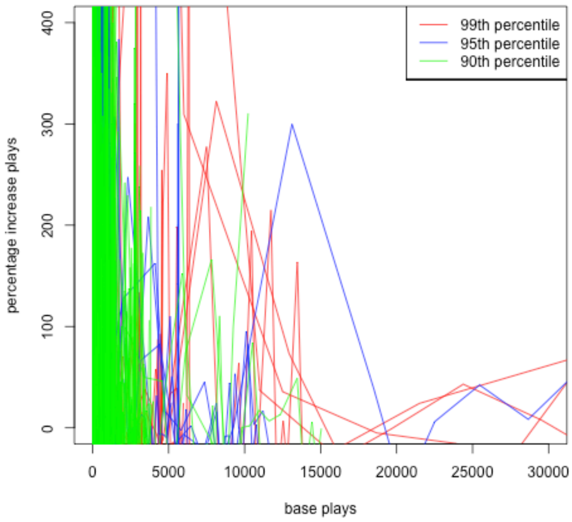


Figure 6: Overview Percentage Increase of Followers for Best 10%

Overview of the percentage increase in relation to base followers of total best 10% artists classified into 90th, 95th, and 99th percentiles.

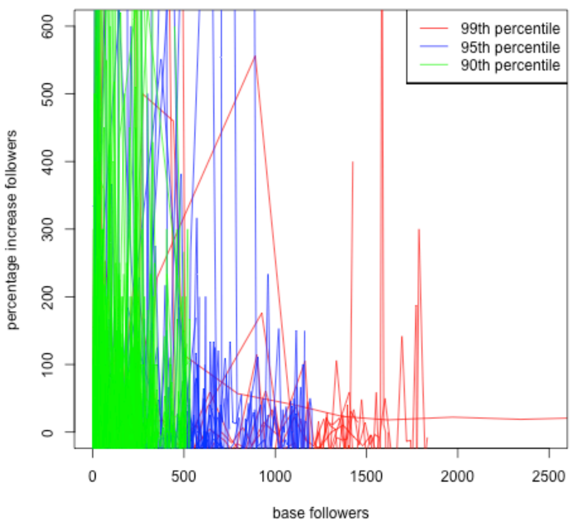


Figure 7: Developed Threshold Plot for Takeoff of Plays

Scale of base plays and percentage increase based on available data; set elbow point at 10,000 plays and 25% percentage increase as required minimum for takeoff.

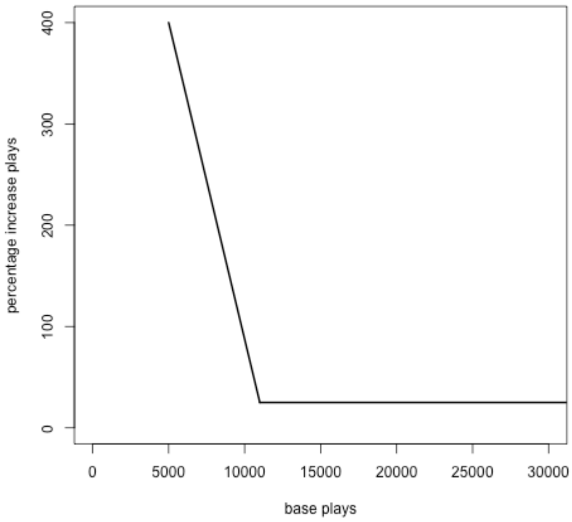


Figure 8: Developed Threshold Plot for Takeoff of Followers

Scale of base followers and percentage increase based on available data; set elbow point at 500 followers and 20% percentage increase as required minimum for takeoff.

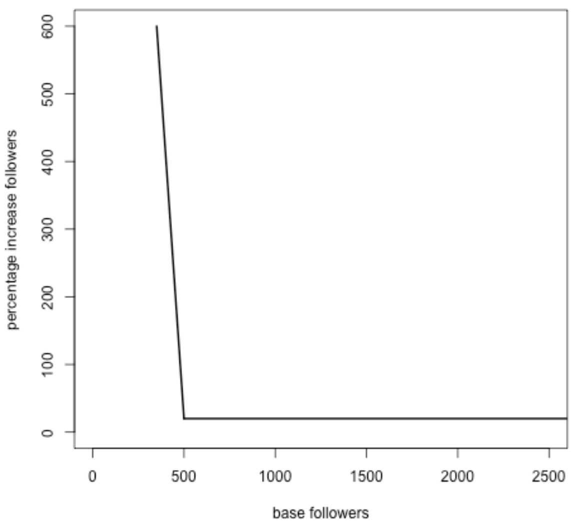


Figure 9: Application of Threshold Plot

Percentage increase in relation to base plays; the graph crosses the defined threshold at 5,000 plays with a percentage increase of 240 % indicating a takeoff at this point.

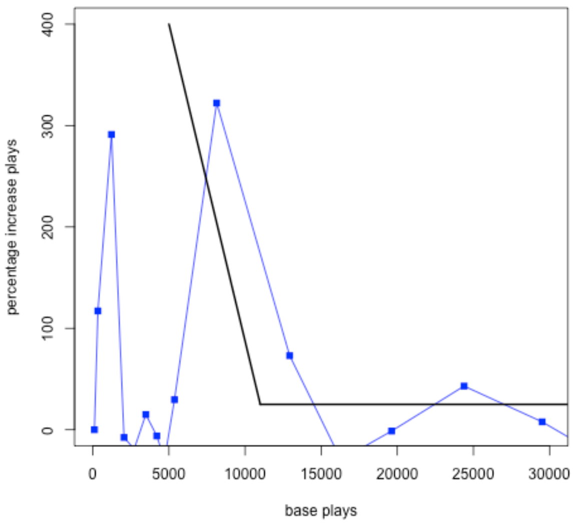


Figure 10: No Exceedance of Threshold Plot

There is no takeoff in the diffusion of plays, as the required threshold is not exceeded.

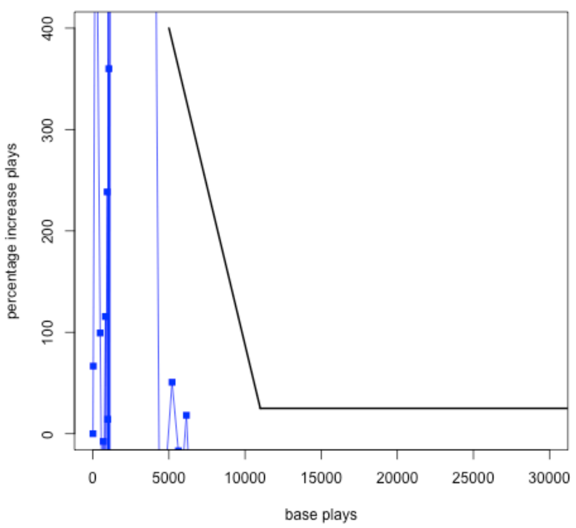


Figure 11: Exemplary Diffusion of Plays

Based on figure 9, takeoff occurs as of reaching 5,500 plays, thus in month 9 of the observed period.

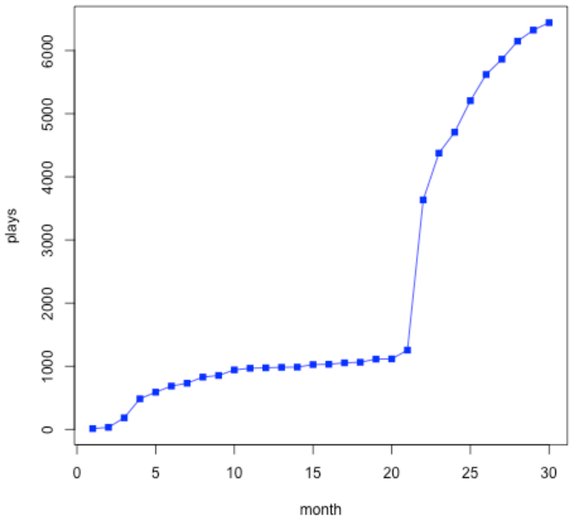


Figure 12: Application of Structural Break Analysis

Application of the definitions (1) and (2) to a actual case; indication of a relevant break in month 9 represented by the vertical line, which also marks the cutting line between period 1 and period 2 (plateau).

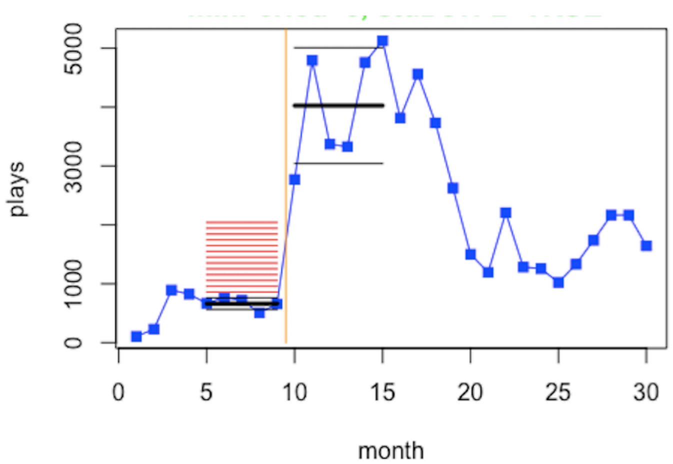


Figure 13: No Indication of Structural Break

Definition (1) is not fulfilled, as the value of mean2 minus SD2 is below the required 14 standard deviations of mean1.

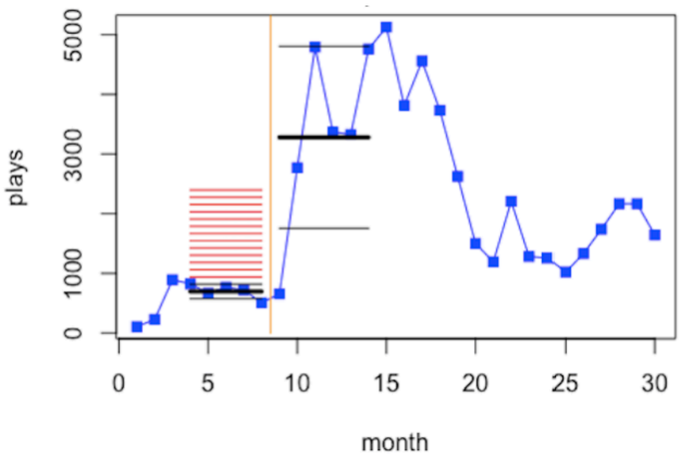


Figure 14: No Indication of Structural Break

Definition (1) and (2) are not fulfilled in this case; the increase of mean from period 1 to period 2 is not significant.

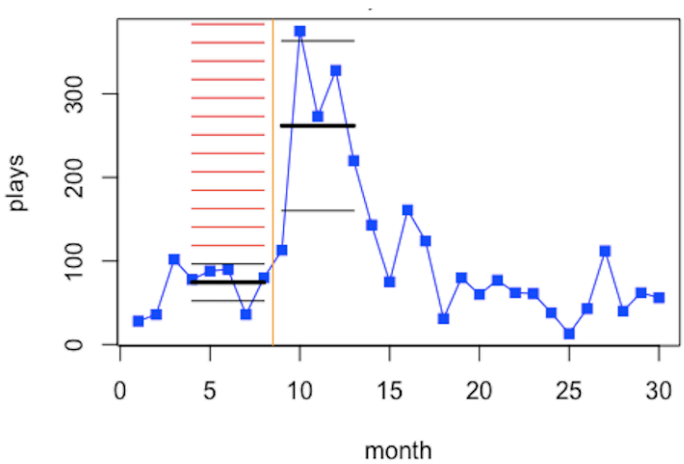


Figure 15: Analysis of Growth Episodes

Definitions (1) to (3) are fulfilled in this case; the vertical lines mark the relevant episode from month 13 until month 17 showing growth rates for plays of $t_{13\text{-}14}$=900%, $t_{14\text{-}15}$=150%, $t_{15\text{-}16}$=8%, t:4, and $t_{16\text{-}17}$=382.6%.

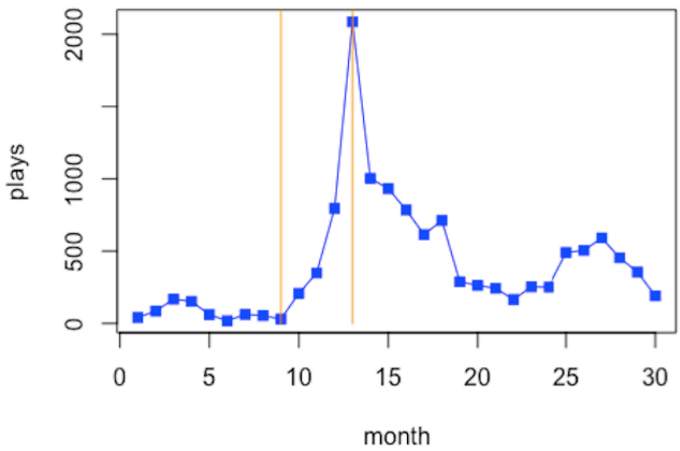
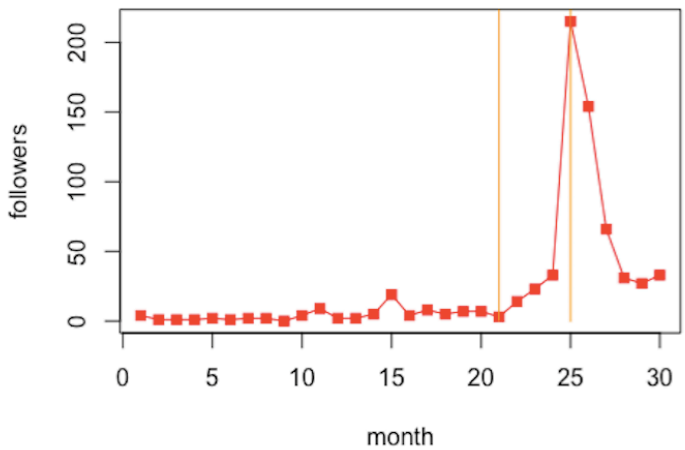


Figure 16: Analysis of Growth Episodes

Definitions (1) to (3) are fulfilled in this case; the vertical lines mark the relevant episode from month 21 until month 25 showing growth rates for followers of $t_{13\text{-}14}$=366.7%, $t_{14\text{-}15}$=64.3%, $t_{15\text{-}16}$=43.5%, , and $t_{16\text{-}17}$=551.5%.



Appendix B: Tables

1. Results of Structural Break Analysis

Performance Plays

Table 1: Break Analysis for Plays

Artist 37959294: Break month = 9	
$mean_1$	663.2
SD_1	98.462
$mean_2$	3804
SD_2	918.476
Actual gap	22.57

Table 2: Break Analysis for Plays

Artist 38473401: Break month = 6	
$mean_1$	23.8
SD_1	16.574
$mean_2$	506.2
SD_2	159.329
Actual gap	19.493

Table 3: Break Analysis for Plays

Artist 38407964: Break month = 4	
$mean_1$	5
SD_1	4.546
$mean_2$	98.6
SD_2	21.138
Actual gap	15.94

Table 4: Break Analysis for Plays

Artist 38124375: Break month = 7	
$mean_1$	5.4
SD_1	6.95
$mean_2$	256.4
SD_2	125.051
Actual gap	17.98

Performance Followers

Table 5: Break Analysis for Followers

Artist 38410207: Break month = 9	
$mean_1$	6
SD_1	3.082
$mean_2$	137.2
SD_2	85.786
Actual gap	14.734

Table 6: Break Analysis for Followers

Artist 38124375: Break month = 7	
$mean_1$	5.4
SD_1	6.95
$mean_2$	256.4
SD_2	125.051
Actual gap	17.98

Table 7: Break Analysis for Followers

Artist 38501163: Break month = 3	
$mean_1$	0.666
SD_1	1.155
$mean_2$	46.2
SD_2	26.809
Actual gap	16.216

2. Results from Empirical Application

Table 8: Overview of Results from Empirical Application

x = success		*Plays*			*Followers*		
No.	**Artist ID (Top 10%)**	**I)**	**II)**	**III)**	**I)**	**II)**	**III)**
1.	38393304	x			x		
2.	38027041				x		
3.	38135744				x		
4.	37842894	x			x		
5.	37959294	x	x		x		
6.	38473401	x	x		x		
7.	38410207	x			x	x	
8.	38407964		x		x		
9.	38116582				x		
10.	38234913				x		
11.	38381746				x		
12.	38501163				x	x	
13.	38091855				x		
14.	38283040				x		
15.	38197736				x		
16.	38124375	x	x		x	x	
17.	38396943	x			x		
18.	37981193						x
19.	38362352	x		x	x		

20.	38291393				x		
21.	38239638	x					
22.	38029406						
23.	38282863				x		
24.	38078138						
25.	38049253						
26.	37931748						
27.	38227436						
28.	38111279						
29.	38208282						
30.	38381896						
31.	38000777						
32.	38478315						
33.	38468420			x			
34.	37866462						
35.	38373631						
36.	38475322						
37.	38280345						
38.	38495996	x					
39.	38215491						
40.	37963825						
41.	38205763						
42.	38163713						
43.	38258548						

44.	38449129	x		x			
45.	37868261						
46.	38336919						
47.	37925952	x					
48.	38055574						
49.	38440951						x
50.	38312601						
51.	38501574						
52.	38349859						
53.	38187647						
54.	37945695			x			
55.	38114756						
56.	37954567						

Appendix C: Literature Review Table

Author/s (Year), *Journal*	Theoretical Background	Method/Analysis	Rationale as Success Measure
Agarwal and Bayus (2002), *Management Science*	• Identification of the transition point between two consecutive intervals • Intervals representing sales dates of annual percentage change	• Application of a generalized discriminant analysis • Definition of the transition point between two consecutive intervals of a time series based on mean values	• Intervals representing pre- and post-takeoff • Transition point is determined as take-off point when positive percentage change is significant (based on mean values)
Altman (1968), *Journal of Finance*	• Based on discriminant analysis, variables can be discriminated to particular categories • Categories are labeled as success-ful or failed	• Application of a discriminant analysis • Simulating different linear combinations to find the discriminant criterion Z • Discriminant criterion serves as threshold value to classify variables to particular categories	• Discriminant criterion Z is derived for classification • Discrimination to the categories suc-cessful or failed

Author/s (Year), ***Journal***	**Theoretical Background**	**Method/Analysis**	**Rationale as Success Measure**
Bai and Perron (2003), *Journal of Applied Econometrics*	• A sudden shift in time series data is referred to as structural break • Depending on how data breaks, this is regarded as turning point	• Computation of different simulations for the estimation of potential breaking points in time series data • Structural breaks in a data pattern are regarded as turning points	• A structural break indicating a turning point in the time series • If turning point from moderate to rapid, growth can be interpreted as success
Bampo et al. (2008), *Information Systems Research*	• Spread of viral messages in a digital network as basis for analysis • Spread is defined by network connectedness, activation and forwarding parameters	• Calculation of the epidemic threshold parameter μ measuring the growth rate of infectives in a network • "Forwarding a message" is equivalent to growth rate of infectives	• "Forwarding a message" as action within a digital network • "Following an artist" as equivalent action • Calculation of epidemic threshold parameter for examining growth • Success, if epidemic threshold parameter high enough

Author/s (Year), *Journal*	**Theoretical Background**	**Method/Analysis**	**Rationale as Success Measure**
Delre et al. (2010), *Journal of Product Innovation Management*	• In a percolation model there exists a network of agents • Percolation threshold *r* defines value for which a percolation (or spreading) is observed	• Network of agents; agents are activated or not • Percolation threshold must be exceeded for a agent to be activated	• Definition of a percolation or spreading threshold that must be exceeded • Success, if this threshold is reached
Easterly (2006), *Journal of Economic Growth*	• Analysis of economic growth patterns of OECD countries • A sequence of zero growth is followed by a continuous sequence of positive growth	• Application of a structural break analysis • Based on the analysis, determination of a consistent definition for successful growth	• Determining a formal definition of successful growth
Fourt and Woodlock (1960), *Journal of Marketing*	• First repeat ratio as fraction of initial buyers making a second product purchase • Successful products attract almost half early buyers	• Incorporation of the repeat ratio into the diffusion model • Repeat ratios of more than 0.5 or more indicating successful products	• Analysis of penetration of digital network • Looking for repeated actions and calculating repeat ratio that must be reached for success

Author/s (Year), *Journal*	**Theoretical Background**	**Method/Analysis**	**Rationale as Success Measure**
Garber et al. (2004), *Marketing Science*	• Focus on spatial dimension of sales data • Execution of small-world simulations for determining cluster growth patterns of products	• Application of small-world simulations and divergence measures • Examination of spatial patterns of successful and failed products • Determining "cluster growth" of products	• Investigating the spread patterns of users within a digital network • Success, if emergence of cluster growth
Golder and Tellis (1997), *Marketing Science*	• Takeoff is regarded as event when growth rate of sales crosses a defined threshold	• Definition of a threshold plot based on given data • Takeoff occurs when the percentage increase in sales relative to its base sales exceeds defined threshold	• Determination of a threshold plot based on available data from the digital network • Exploration of diffusion data and the relative percentage increase
Golder and Tellis (2004), *Marketing Science*	• See Golder and Tellis (1997)	• Measuring the time between takeoff and slowdown • Slowdown occurs when in two consecutive years after takeoff sales are lower than at peak	• See Golder and Tells (1997) • Measuring time periods between takeoff and slowdown • The longer the time to slowdown, the more success

Author/s (Year), *Journal*	Theoretical Background	Method/Analysis	Rationale as Success Measure
Hausmann, Pritchett, and Rodrik (2005), *Journal of Economic Growth*	• Finding relevant initiation of economic growth acceleration • Growth acceleration is referred to a rapid growth stage • OECD average as basis for analysis	• Determination of starting point of growth acceleration • Measuring growth rate a country would need for economic success • Defining a threshold for growth that must be exceeded to be successful	• Defining a threshold for growth • Defining a growth rate • Success, if determined threshold and rate are reached
Kelly (1997), *Quarterly Journal of Economics*	• Smithian growth shows critical behavior • Smithian growth leads to a sudden takeoff when a critical density of markets is reached	• For an economy with sites and linkage probabilities, there is a critical value • If critical value is reached, the economy of local markets fuse together to an economy-wide market • Growth accelerations occur	• Linkages in the digital network are examined • The higher the amount of linkages, the higher the probability that a critical density is reached • Determination of critical value is required

Author/s (Year), ***Journal***	**Theoretical Background**	**Method/Analysis**	**Rationale as Success Measure**
Levine (1997), *Journal of Economic Literature*	• Measures of trading activity for a country • Active markets exhibit turnover ratios as trading activity relative to market size	• Measuring the turnover ratios of countries and their markets • Defined turnover ratios based on historical data of trading	• Determination of turnover ratios based on available data • Turnover ratio is dependent on interrelationship between two variables • Success, if turnover ratio is reached
Markovitch and Golder (2008), *Marketing Science*	• Long-term survival of a firm as basis for long-term success	• Application of an hazard survival model for observation of stock returns • Observation allows for a reliable prediction of a firm's future success/failure • Determination of post-takeoff firm survival measured in years based on stock return observations	• Application of hazard survival model • Estimation of long-term potential on a digital network based on the analysis of available data

Author/s (Year), *Journal*	Theoretical Background	Method/Analysis	Rationale as Success Measure
Muller and Yogev (2005) *Technological Forecasting and Social Change*	• Determination of transition point from the early market to the main market • Bass diffusion model as basis for the analysis	• Measuring the changing dominance time from early market to main market • Determination of changing dominance time based on market penetration data	• Analysis of the market penetration data of a digital network • Determination of the time the early market needs to shift to a main market
Stremersch et al. (2007), *Journal of Marketing*	• Definition of threshold for takeoff (see Golder and Tellis, 1997) • Including theory of informational cascades, after which two complementary goods affect each other's sales	• Definition of threshold plot for takeoff (see Golder and Tellis, 1997) • Incorporation of cascading effects of complementary goods	• Determination of takeoff time by means of threshold plot • Inclusion of effects from network performance and user interactions to overall success
Tellis, Stremersch, and Yin (2003), *Marketing Science*	• Determination of international takeoff of products • Defining threshold for takeoff (see Golder and Tellis, 1997)	• Definition of threshold plot for takeoff (see Golder and Tellis, 1997) • Market penetration rather than sales as basis for takeoff definition	• Determination of takeoff time by means of threshold plot • Measuring market penetration of the digital network as basis for success

Author/s (Year), *Journal*	Theoretical Background	Method/Analysis	Rationale as Success Measure
Van Everdingen, Fok, and Stremersch (2009), *Journal of Marketing Research*	• Modeling international spillover effects for the determination of product takeoff • Defining threshold for takeoff (see Golder and Tellis, 1997)	• See Tellis, Stremersch, and Yin (2003)	• See Tellis, Stremersch, and Yin (2003)

Appendix D: RStudio Script

Listing 1: *Testing plays and followers for takeoff based on threshold plot*

```
#load functions and data
source("functions.R")

#testing growth rate for every top artist
for(a in 1:nrow(topArtists))
{
  row <- topArtists[a,]
  artistID = row$creator_id

  loadDataForArtist(artistID)

  #threshold plot for plays
  elbowX = 11000
  elbowY = 25
  topX = 5000
  topY = 400

  #plot growth rate based on plays
  plot(playsPerMonthAgg,growthRatePlays*100,
       main = bquote("Percentage increase of base plays for " ~ .(artistID) ~
""),
       xlab = "base plays", ylab = "percentage increase plays",
       log="",type="l", col="blue",
       ylim = c(0,400), xlim = c(0,30000))
  #slope
  segments(topX, topY, elbowX,elbowY, lwd = 2)
  #horizontal
  segments(elbowX, elbowY, 50000, elbowY, lwd = 2)


  #threshold plot for followers
  elbowX = 500
  elbowY = 20
  topX = 350
  topY = 600


  plot(followerPerMonthAgg,growthRateFollower*100,
       main = bquote("Percentage increase of base follows for " ~ .(artistID) ~
""),
       xlab = "base followers", ylab = "percentage increase followers",
       #pch=22, lty=1,
       log="",type="l", col="red",
       ylim = c(0,600), xlim = c(0,2500))
  #slope
  segments(topX, topY, elbowX,elbowY, lwd = 2)
  #horizontal
  segments(elbowX, elbowY, 50000, elbowY, lwd = 2)
}
```

```
#plot growth rate based on follows
  plot(followerPerMonthAgg,growthRateFollower*100,
       main = bquote("Percentage increase of base follows for " ~ .(artistID) ~
""),
       xlab = "base follows", ylab = "percentage increase follower",
       log="", type="l", col="red",
       ylim = c(0,600), xlim = c(0,2500))
  #slope
  segments(topX, topY, elbowX,elbowY, lwd = 2)
  #horizontal
  segments(elbowX, elbowY, 50000, elbowY, lwd = 2)
}
```

Listing 2: *Simulation for Structural Break Analysis*

```
#load functions and data
source("functions.R")

#variables to test for structural break
takeStdDevoOfPeriod2IntoAccount <- TRUE
maxLengthFirstPeriod            <- 5
monthsToTest                    <- 3:20
periodLengthsToTest             <- 4:5
nbrStdDevGapToTest              <- 10:14

#testing for structurel break for every top artist
for(a in 1:nrow(topArtists))
{
  row <- topArtists[a,]
  artistID = row$creator_id

  loadDataForArtist(artistID)

  for(per in periodLengthsToTest)
  {
    for(std in nbrStdDevGapToTest)
    {
      for(mth in monthsToTest)
      {
        #has the artist a structural break in plays for the given parameters
        hasBreakPlays    <-    hasStrBreak(mth,    playsPerMonth,    std,    per,
takeStdDevoOfPeriod2IntoAccount, maxLengthFirstPeriod)
        #has  the  artist  a  structural  break  in  followers  for  the  given
parameters
        hasBreakFollowers  <-  hasStrBreak(mth,  followerPerMonth,  std,  per,
takeStdDevoOfPeriod2IntoAccount, maxLengthFirstPeriod)
      }
    }
  }
}
```

Listing 3: *Simulation of Growth Episodes for Analysis*

```
#load functions and data
source("functions.R")

#variables to test growth rate
threshold     <- 20
period        <- 4

#testing growth rate for every top artist
for(a in 1:nrow(topArtists))
{
  row <- topArtists[a,]
  artistID = row$creator_id
  loadDataForArtist(artistID)
  #check growth rate for plays
  growthPlays <- checkGrowthRate(playsPerMonth, threshold, period)
  #check growth rate for followers
  growthFollowers <- checkGrowthRate(followerPerMonth, threshold, period)
}
```

Listing 4: *Functions for Structural Break Simulation and Growth Episodes Analysis*

```
#  Returns  whether  or  not  [TRUE/FALSE]  growth  episodes  for  the  provided
parameters were found
# Parameters:
# - data                 the data (e.g. plays or follower per month)
# - minGrowthInPercent  the minimum growth in percent from month to month
# - overPeriod           the period length of the growth episode (e.g. how many
months in a row has the growt be > minGrowthInPercent)
checkGrowthRate <- function(data, minGrowthInPercent, overPeriod)
{ source
  overallSuccessFound = FALSE
  for(m in 1:(length(data)-overPeriod))
  {
    successInThisPeriod = TRUE

    #check for break in current period
    for(p in m:(m+overPeriod-1))
    {
      playsThisMonth = data[p]
      playsNextMonth = data[(p+1)]

      if(playsThisMonth == 0)#avoid dividing by 0
        growthPercent = 0
      else
        growthPercent  =  (playsNextMonth  -  playsThisMonth)  /  playsThisMonth  *
100

      if(growthPercent >= minGrowthInPercent)
      {
        #all good, growth rate is sufficient
      }
      else
      {
        #growth rate not big enough
        successInThisPeriod = FALSE
        break
      }
    }

    if(successInThisPeriod)
    {
      overallSuccessFound = TRUE
      cat("successInThisPeriod ",m,"-",(m+overPeriod),"")
    }
  }
  return(overallSuccessFound)
}
```

```
# Returns whether or not [TRUE/FALSE] a structural break exists for the provided parameters
# Parameters:
# - month            check for a structural break at this month.
#                    This month divides the first period from the second.
#                    period 1: [2 - month]
#                    period 2: [month+1 - month+1+minPeriodLength]
# - data             the data (plays or follower per month)
# - nbrStdDev        the number of standard deviations of the first period
# - minPeriodLength  The length that the second period has to have in oder to count as a break  (Period2)
# - stdDevPeriod2     Take standard deviation of periot 2 into account when calculating the gap [TRUE / FALSE]
# - maxLengthPeriod1 How long period 1 can be at max
hasStrBreak  <-  function(month,  data,  nbrStdDevGap,  minPeriodLength, stdDevPeriod2, maxLengthPeriod1)
{
  if(length(data) < month+minPeriodLength)
  {
    print("data length to short")
    return(FALSE)
  }

  p1Start = month - maxLengthPeriod1 + 1
  if(p1Start <= 0)
    p1Start = 1

  #divide data in periods
  period1 = data[p1Start:month]
  period2 = data[(month+1):(month+minPeriodLength)]
  #mean and standard deviation for period 1
  mean1  = mean(period1)
  sigma1 = sqrt(var(period1))
  #mean and standard deviation for period 2
  mean2  = mean(period2)
  sigma2 = sqrt(var(period2))
  #check for invalid values
  if(mean1 == 0)
    return (FALSE)
  if(sigma1 == 0)
    return (FALSE)

  #check condition 1: gap in number of standard deviations
  if(stdDevPeriod2) #take sigma2 into account
    actualGap <- (mean2 - sigma2 - mean1) / sigma1;
  else #without sigma2
    actualGap <- (mean2 - mean1) / sigma1;

  if(actualGap > nbrStdDevGap) #gap between means is enough
    return(TRUE)
  else #gap between means is NOT enough
    return(FALSE)
}
```