Media judgment of entrepreneurial failure - implications for founders

Vom Promotionsausschuss der Technischen Universität Hamburg-Harburg
zur Erlangung des akademischen Grades

Doktor der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.)

genehmigte Dissertation

von

Matthias JACOBI

aus

Bad Segeberg

2018

Gutachter: Prof. Dr. Christoph Ihl, Prof. Dr. Michel Clement
Tag der mündlichen Prüfung: 02.03.2018

TECHNISCHE UNIVERSITÄT HAMBURG-HARBURG

Abstract

Media judgment of entrepreneurial failure - implications for founders

by Matthias JACOBI

Among the five most valuable firms in the world are two former startups. These two, Facebook and Google have one thing in common: they are both from the US. Other successful startups support the commonly held assumption that the US is especially entrepreneurial friendly, and the place to be for founders. However, is this actually true? Most research findings to this question are based on surveys or interviews and therefore limited in their explanatory power. In contrast to these approaches, we aim to quantify entrepreneurial friendliness in the US and compare it to Germany, which is considered less entrepreneurial friendly. We set focus on the media judgment of startups, as a surrogate of entrepreneurial friendliness, and changes of this sentiment after a startup fails. Organizational theory states that media is capable of assigning or withholding legitimacy, and therefore influences the culture of entrepreneurial activity. By employing a difference-in-difference approach, along with regression models, we find significant differences in media judgment of startups in the two countries, with permanent positive judgment in the US (even after failure), and a rather neutral judgment in Germany, which changes to the negative after failure. Our findings reveal, it is not common to report about failure in the US, whereas Germans seem to be interested in detailed reporting about failures. For both countries, we find startups with a novel business model to be judged more harshly after failure, than startups with traditional business models. In addition, startups which receive high levels of funding through investors, are also judged more harshly after failure. In a cross-comparison, we find the US to exhibit an in-group bias, meaning a favor in media judgment for their own startups over German startups. Surprisingly, we find a "now more than ever" mentality among failed founders, who have been negatively criticised by the media. This thesis enhances our understanding of entrepreneurial friendliness in the US and Germany, and adds to the social judgment of organizations theory, as it quantifies media legitimacy of startups and the influence of failure on legitimacy of startups in two different countries.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Society values risk taking, but not gambling, and what is meant by gambling is risk taking that turns out badly.

(March and Zur Shapira, 1987)

1.1 Motivation

In 2014 the largest social network Facebook bought a five year old company with 55 employees for approximately $22 billion. The acquired messaging platform WhatsApp reached 400 million users back then and created a yearly revenue of three cents per user. Facebook paid $55 for every single one of them. The social network giant was in a poor bargaining position, after unsuccessfully trying to force Facebook users into their messaging tool, called messenger. With the need to remain competitive, Facebook admitted having failed to keep abreast with the developing market trends, where due to sparse internet connectivity users prefer a messenger tool, which is much simpler than Facebook's solution. This example shows that even fairly new and small firms can compete with global players today. Companies thus invest millions of dollars into new business models, even if there is a high risk of failure. Prominent examples of such failed experiments are Google Glass, Amazon's Fire Phone or Youtube Plus. As a result, CEOs encourage their staff to take more risks. Jeff Bezos (Amazon CEO) says: "If you bet on the future, it is always an experiment. You never know, if it works. But few major successes compensate for dozens of failures." So risk taking and failure become part of a firm's culture. It is becoming a part of the company's strategy to compete, and thus moves failure away from the usual negative perception. Conferences, such as Failcon, are being held specifically to learn from and prepare for failure.

It is a process, and cultural changes need time, huge differences between societies and how they treat failure exist (Singer, Amorós, and Moska, 2014). Interestingly, those societies, which are considered resilient for failure, bring up the most promising startups. This becomes clear, when we look at the most valuable firms in the world. Among the five most valuable firms are two former US startups, with no German. Indeed, national culture influences entrepreneurship (Singer, Amorós, and Moska, 2014). Survey analyses show higher levels of fear of failure in Germany, and fewer people regard entrepreneurship as a good career choice compared to the US (Singer, Amorós, and Moska, 2014). Adding to this is a common belief that the German society does not tolerate risky activities and failure as the US does. Results from the Global Entrepreneurship Monitor show, that cultural norms of the German society favor risk avoidance and social stigma attached to failure (Reynolds et al., 2000, p. 37).

To compensate for this disadvantage, Germany and other countries run initiatives to promote entrepreneurial activity. Though some scholars argue, that most national initiatives to stimulate entrepreneurship are not based on scientifically robust findings, but on assumptions, expert assessments or beliefs (Mueller and Thomas, 2001). This situation is not satisfactory, as scholars should be able to explain the drivers of entrepreneurship, down to a regional level. Therefore, we are aiming to add to the emerging research stream of cultural implications on entrepreneurial activity with our empirical study of the judgment of startups by the media, as it is considered a surrogate for entrepreneurial friendliness. To our knowledge, no prior research has investigated the media judgment of startups, especially in the event of failure. Nor does a country comparison exist, to allow a proper classification of results. Therefore, we created a dataset comprised of failed and operating German and US startups from 1995 to 2015 and gathered the corresponding press articles about them. The press articles were investigated via sentiment analysis and rated either positively, neutrally or negatively. With secondary data about the firms and founders, we are able to gain insights into the overall sentiment towards failure, the influence it has on the entrepreneur and how it differs between the countries. The approach is fundamentally different to existing research, which is based on surveys and interviews.

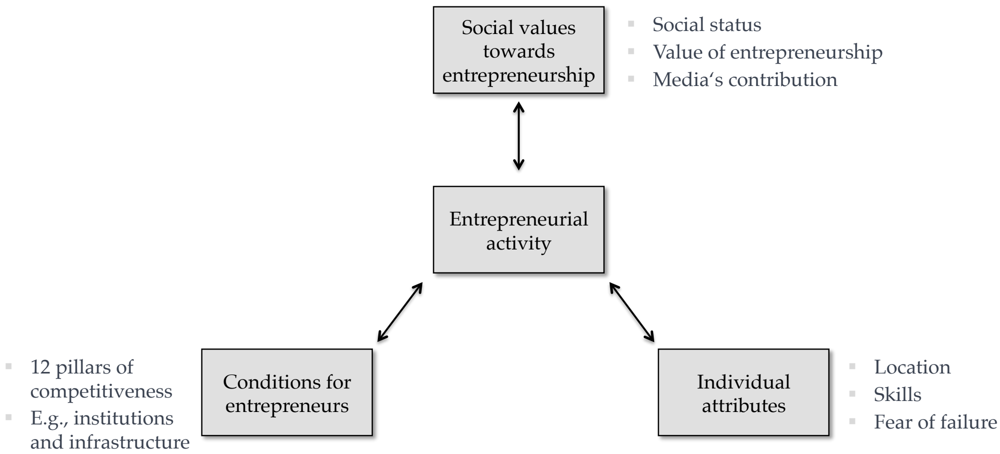
We formulate our research agenda after thoroughly studying the literature and derive thirteen hypotheses for three studies. This is being explained in more detail in the following chapter.

1.2 Research questions

Starting a new business is a risky activity, and often leads to failure. Existing research from of the US Bureau of Labor Statistics (BED) has shown, roughly 50% of all new businesses fail within the first five years, and approximately 35-45% of all new products fail (Boulding, Morgan, and Staelin, 1997). Failure is thus a common phenomenon. Some scholars have argued that failure is the prerequisite for success. It has been shown in fields such as new product development (Maidique and Zirger, 1985), internal corporate venturing (McGrath, 1995), or joint venturing (Peng and Shenkar, 2002). For future success, previous failures lead to valuable learnings for the entrepreneur or the organization as a whole. In fact, learning from failure has been extensively studied in the field of entrepreneurship research (e.g., Cope, 2011; Politis, 2005; Shepherd, 2003b). It is sometimes linked to a sensemaking process, and may eventually lead the entrepreneur to starting anew (Singh, Corner, and Pavlovich, 2015). Besides learning and sensemaking from failure, other factors need to be considered when it comes to entrepreneurial activity. Numerous studies give anecdotal evidence, that failure leads to stigmatization of entrepreneurs, mainly in Europe and some Asian countries (Jenkins et al., 2014; Lee et al., 2011). Countries like the US, on the other hand, are considered tolerant to failure and hence entrepreneurially friendly (Singer, Amorós, and Moska, 2014). Research in this field has examined entrepreneurial activity and cultural dimensions on national or regional levels (Davidsson, 1995; Davidsson and Wiklund, 1997; Shane, 1992; Shane, 1993). Similarly, Shane (1992) compared national rates of innovation to Hofstede's dimensions on individualism vs. collectivism and power distance. Most of this research focuses on early stages, such as the link between culture and new firm formation. Little empirical research focuses on cultural effects or social values and entrepreneurial failure, though a positive culture of failure is associated with higher levels of entrepreneurial activity (Shepherd and Patzelt, 2017). To our knowledge, no empirical studies exist to investigate how failed entrepreneurs are judged by society. We seek to do so, and build our research on a framework from the Global Entrepreneurship Monitor (GEM) to describe entrepreneurial activity (see Figure 1.1). According to the research of the GEM, entrepreneurial activity is driven by individual attributes of entrepreneurs, conditions for entrepreneurs (derived from the 12 pillars of competitiveness from the world's economic forum (Schwab, Sala-i-Martin, et al., 2010)), and social values towards entrepreneurship, such as the social status associated with being an entrepreneur, or how media contributes to an entrepreneurial friendly society.

As our literature search in Chapter 2 reveals, there exists no empirical research regarding the social values towards entrepreneurship dimension, which analyses the impact of media on entrepreneurial activity. Hence, our research aims to quantify entrepreneurial friendliness of society by analyzing media reports on startups, especially in the event of failure. The

FIGURE 1.1: The three components to influence entrepreneurial activity of a society (from Singer, Amorós, and Moska (2014)).



sentiment or tone portrayed in these articles will be measured and tracked over time. Starting from the first press reports of a startup, until years after its failure. We aim to identify how failure influences or impacts on media reporting. It is thus used as an indirect measure of entrepreneurial friendliness of society, since media plays a major role in shaping the public's perception and opinion building (see Chapter 3.1). In this thesis, we compare data from failed and operating startups from 1995 to 2015, to quantify the impact on failure. A cultural dimension is added by doing so for the US and Germany.

1.3 Research approach

This thesis is structured as follows. We start by studying the literature on failure and bankruptcy with emphasis on entrepreneurship (see Chapter 2). We intent to get an overview of the current state of research in this field, and to identify potential white spots. A great deal of research in this field has focused on topics such as causes of failure (Bruno and Leidecker, 1988; Everett and Watson, 1998), consequences for entrepreneurs (Politis and Gabrielsson, 2009; Singh, Corner, and Pavlovich, 2007), or cultural effects on entrepreneurship (Lee and Peterson, 2001; Wennberg, Pathak, and Autio, 2013). Most of this research is qualitative or based on surveys. To the best of our knowledge, no quantitative research has aimed to measure entrepreneurial friendliness of a society with particular focus on failure. Therefore, our intention is to close this gap. In Chapter 3, we introduce theories relevant for our work, namely, social judgment of organizations theory, as it explains how and why organizations such as startups are being judged by an external audience, stigmatization theory, which explains how stigma develops and how it impacts individuals (Jenkins et al., 2014). In addition, cultural effects are covered in two ways. First, with Hofstede's cultural dimensions theory,

TABLE 1.1: The three different analysis levels of the thesis and the particular research questions.

Study No.	Level	Research question	Model
1	Macro	Does the media in Germany judge startup failure harsher in the US?	• Difference in differences estimation • Ordinary least squares regression model with fixed effects
2	Meso	Does the media judgment vary between startups?	• Ordinary least squares regression
3	Micro	Does the media judgment influence a founder's decision to start again after failure?	• Binary probit model with selection

an approach to describe cultural influences on society's values and behavior (Lonner, Berry, and Hofstede, 1980). Second, with the in-group bias, a bias of favoring one's "in-group" over members of an "out-group". Equipped with this theory base, we develop a research agenda along three main topics, on macro, meso and micro level, respectively. Table 1.1 lists the three main research questions, the level, and the applied model.

First, we investigate media judgment of failed startups and cultural differences of this judgment between the US and Germany in Study I (see Chapter 6), by applying sentiment analysis, different matching techniques for our startups, and a difference in differences estimation paired with an ordinary least squares regression model. In Chapter 7 we investigate specific differences in media judgment, such as dependencies on a startup's novelty, funding sum, experience of the founding team, prior founding experience, and age. Our last study focuses on the founder's level. It investigates potential influences on the founding again after failure chances of a founder. Therefore, we consider media judgment, the startup's business model and its novelty, prior founding experience, and raised funding sum (Chapter 8).

Chapter 2

Entrepreneurial failure: literature review

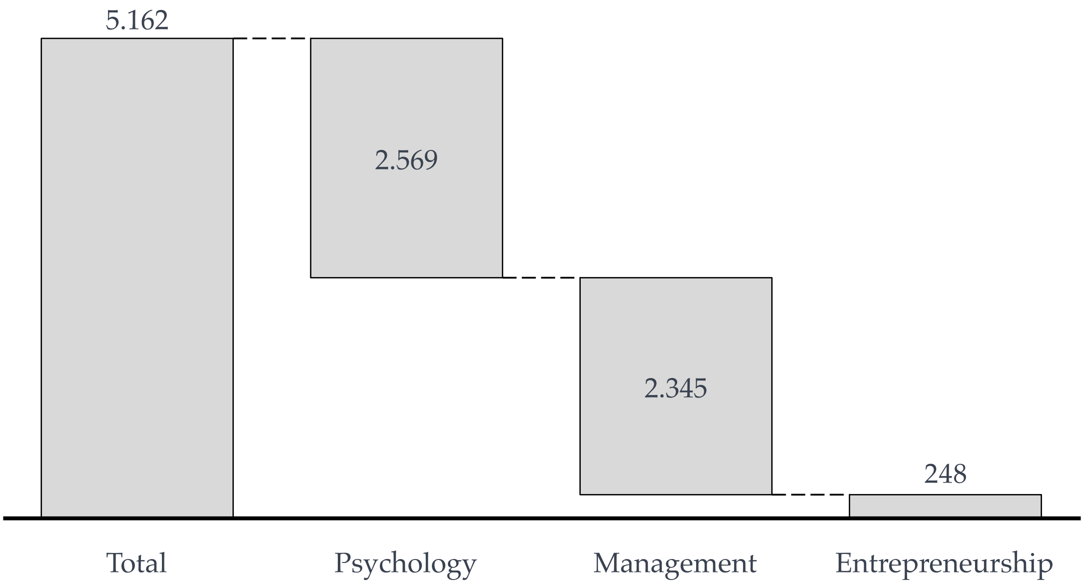
In this chapter we define and introduce the main terms used in this thesis. In between we start with a short overview of entrepreneurship and failure studies, as these are fundamental to this work. Then follows an overview of cultural aspects on entrepreneurship, with a focus on the countries in scope. The last paragraph switches to the entrepreneur's perspective to explain different types of judgment the entrepreneur experiences after he or she has failed with a startup.

2.1 Literature search and co-citation analysis

To obtain a first coarse overview of the literature on entrepreneurial failure, as well as to understand the influence of theories from the field of psychology to explain how failure impacts entrepreneurs personally, we study the literature by examining the most relevant publications on this topic. We follow an approach introduced by Short (2009) on how to do a best-practice literature review. With the help of the Web of Science platform, we searched for literature with "failure" and "bankruptcy" in title, abstract and keywords. Only top journals in the fields of entrepreneurship, management and psychology were considered. We expect findings from the field of psychology to impact recent entrepreneurship research, which is why it is included into the search. A list of these journals can be found in the Appendix E. This procedure resulted in 5.162 publications from 105 different journals, which are used for a co-citation analysis to identify the most relevant publications to our topic. Figure 2.1 breaks down the 5.162 publications by research discipline. A minority of 248 is from the field of entrepreneurship.

Two documents are co-cited if at least one other cites them (Small, 1973). It is a frequency measure and expresses the similarity of content (Gmür, 2003). In our approach, we choose the document level. Figure 2.2 shows the results in a color-coding scheme, representing the different disciplines. In yellow are publications, which include the term entrepre*,

FIGURE 2.1: Paper breakdown per category for literature review.



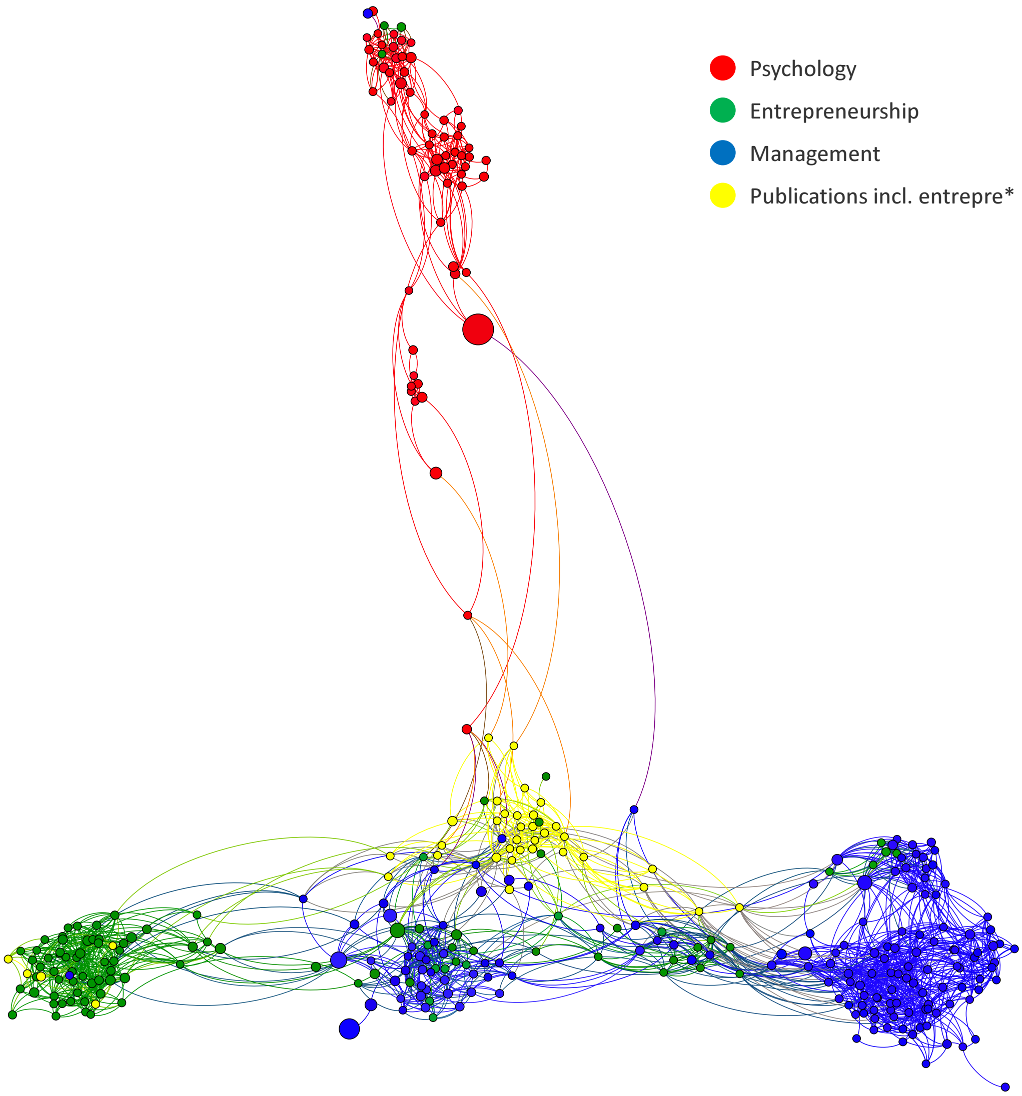
as a placeholder for all conditions related to entrepreneurship. Interestingly, these 172 publications are all from the field of entrepreneurship. No management or psychology journal of our sample published a paper including the term.

Figure 2.2 only shows papers with co-citations equivalent or more than ten in number. All disciplines are more or less separated from the others, especially the psychology papers. From this field, there are only a few connections to papers which include the term entrepre* or to the other two disciplines.

We read through those publications which are colored yellow, as they include the desired keywords mentioned above and in addition entrepre*, to get an overview of the current state of research on entrepreneurial failure. The top three topics mentioned in these papers are stigmatization of entrepreneurs, learning from failure, and the positive sides of it, as well as coping with failure on a personal level. None of these papers discussed the role of media and its influence on entrepreneurial activity, especially after failure.

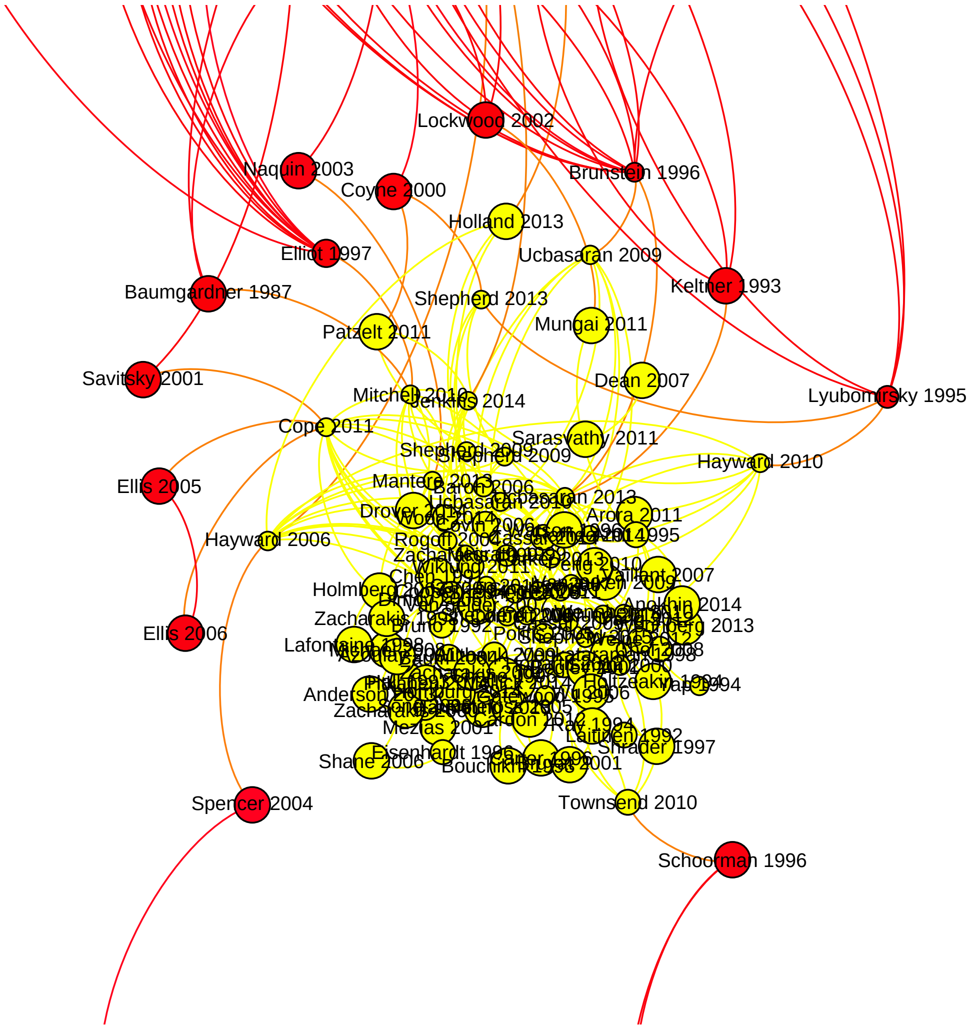
Our second target is to understand if there is any knowledge spillover from the field of psychology, to enhance our understanding of entrepreneurial failure, as it seems reasonable to assume that the entrepreneurship literature could benefit from psychological theories. Therefore, we read through all of the gatekeepers, i.e., those papers which directly connect to the yellow class. Figure 2.3 shows them, in a rearranged close-up of Figure 2.2. In total there are 13 papers from psychology directly connected to the yellow class. Again, learning from

FIGURE 2.2: The Co-citation analysis of our dataset shows little knowledge spillover from the field of psychology. The size of the dots is proportional to the overall number of co-citations within the dataset.



failure is one of the top themes discussed here, but this time related to successful and failed experiences (Ellis and Davidi, 2005; Ellis, Mendel, and Nir, 2006). Another main topic is judgment, precisely the expected judgment after failure (Savitsky, Epley, and Gilovich, 2001). Brunstein and Gollwitzer (1996) discuss the effects of failure on subsequent performance, and Lockwood, Jordan, and Kunda (2002) the motivation and impact of role models.

FIGURE 2.3: A rearranged close up of the co-citation analysis.



Throughout the 172 publications including the term entrepre*, and an additional 13 from the field of psychology, no evidence for the impact of media on entrepreneurial activity, how it reports about failure, and what influences the reporting has on the entrepreneur, can be identified. The sparse research that exists on this matter is solely based on surveys (Singer, Amorós, and Moska, 2014). We have summarized the main findings from the 26 most relevant papers out of the above mentioned 172 (see Appendix F). They are chosen, due to their role as gatekeepers to the field of psychology. In addition, the 13 psychology papers are summarized (see Appendix G). For our literature review, we use the summary as a starting point.

As we discovered this lack of empirical research, we came up with the idea to measure the media sentiment towards entrepreneurship. Therefore we choose two countries, which are seen as entrepreneurial friendly (USA), and less entrepreneurial friendly (Germany). We want to provide quantifiable facts, where there have only been surveys and descriptive approaches.

The next three sections summarize the current state of research in our field of focus; we derive it from the relevant papers identified via our co-citation analysis. We build on this summary in our following three studies.

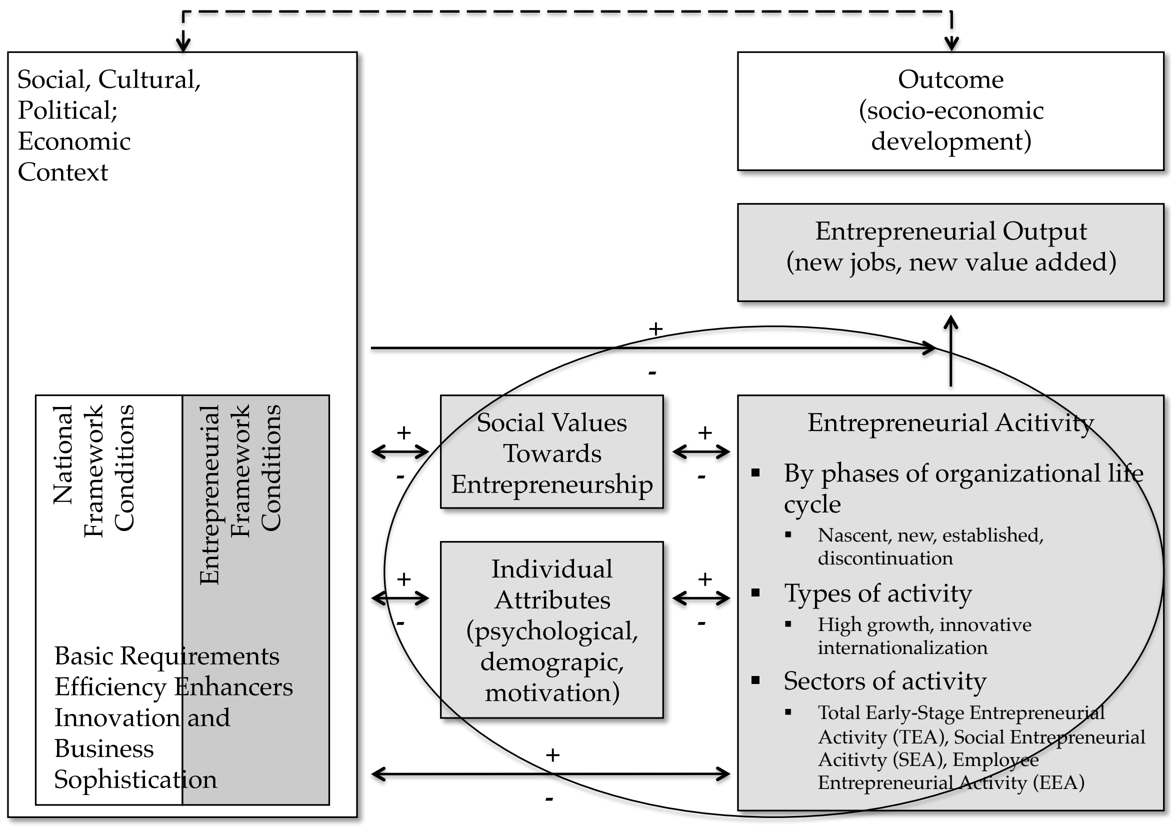
2.2 Entrepreneurship and failure

In the 1940s and 1950s business historians pioneered the study of entrepreneurship at Harvard Business School (Jones and Wadhwani, 2006). Since then, entrepreneurship has become an increasingly important phenomenon for researchers as well as policymakers and the value to society and economic development has been explored in many research projects (Acs and Audretsch, 2010; Lee et al., 2011). New inventions in the last few decades, mainly from the US, have even spurred the interest of social scientist and management scholars, as they have changed the field of entrepreneurship tremendously.

Startup firms fundamentally change business models and markets. PayPal, for instance, grew from a small cryptography company to the number one payment system online and is now larger than most of the traditional banks. Another paradigm example is Google, struggling for many years to find a reliable revenue stream until finally introducing their vastly successful AdWords program. It allowed businesses to target the advertisement to people who search for specific terms. Today Alphabet, the conglomerate Google belongs to, is the second most valuable firm in the world (Gandel, 2016). Every year startups change markets completely. As a result, there is a growing interest into the startup business. So-called venture capitalist invest into emerging firms with high growth potential, to participate in their success. Also, hubs like Silicon Valley, the epicenter of the world's largest high-tech corporations and startups is of growing interest. This evolution finally caught up to policy makers, which try to pave the way for an "entrepreneurial friendly" environment by new laws and support programs.

Social scientist and management scholars, therefore, have a vast field of new research, which is always evolving. Since we cannot tackle this enormous area all at once, we structure our approach and circumvent our focus, with the help of the Global Entrepreneurship Monitor (GEM) framework from 2014 in Figure 2.4 (Singer, Amorós, and Moska, 2014). This framework tries to picture the contribution of entrepreneurship towards national economic growth. Since its introduction in 1999, its enhancement has been continuous.

FIGURE 2.4: GEM framework for entrepreneurship. Graph from Singer, Amorós, and Moska (2014).



In this thesis, we will focus on three out of the six main clusters of the framework as indicated in the Figure 2.4. They are the "Social values towards entrepreneurship" group, "Individual Attributes", and "Entrepreneurial activity". The first cluster describes different social values that influence the entrepreneurship culture of a nation. Included here is the status of being an entrepreneur. Is it seen as a good career choice or not, and do entrepreneurs enjoy a special status? This cluster also includes how media influence entrepreneurship. Moreover, the guiding question of this thesis will be to what extent media influences entrepreneurial activity. The group "Individual Attributes" includes factors such as location, skills, fear of failure, or motivational aspects. Finally, "Entrepreneurial Activity" clusters entrepreneurship according to its characteristics, e.g., early stage startup, high growths oriented business model, or innovative technology.

The following paragraphs give an overview of research findings from the past in these areas. Thereby, we take two different perspectives. First, we take the above mentioned individual attributes view, followed by the entrepreneurial activity. Chapter 3.1 covers a tailored discussion on the social values perspective.

Individual attributes

In the research field of personal organization fit, scholars study decision criterions for career choices (Vianen, 2000). Their findings suggest some factors to be relevant for an individual's decision, such as people's values, abilities, personality, as well as factors related to the organization, such as culture (Vianen, 2000). Initial research on the question of why people decide to become entrepreneurs is predominantly in the disciplines of psychology and sociology. But with the evolution of economics and entrepreneurship research, there are growing contributions also from these disciplines (Markman and Baron, 2003; Henderson and Robertson, 1999; Welpe et al., 2012). The first attempt to derive an economic theory of entrepreneurship was started by Casson (1982). His model, however, fails to explain why people become entrepreneurs. Today research has found many prerequisites for entrepreneurial activity, which we can distinguish in two categories. First, there are benefits entrepreneurs experience in comparison to employees, and second, individual characteristics which increase or lower the probability of becoming an entrepreneur.

We start with benefits entrepreneurs experience from being self-employed. In a survey study with 2.700 US citizens, Patzelt and Shepherd (2011) found, that self-employed experience fewer negative emotions than those who are working as employees. Previous studies even link self-employment to positive emotions, such as high levels of passion (Smilor, 1997; Baum and Locke, 2004; Cardon et al., 2005; Cardon et al., 2009), experiences of happiness, flow and excitement (Komisar, 2000; Rai, 2008; Schindehutte, Morris, and Allen, 2006), or greater levels of life satisfaction in general compared to employees (Blanchflower, Oswald, and Stutzer, 2001; Bradley and Roberts, 2004; Thompson, Kopelman, and Schriesheim, 1992). Consequently, Bird and Jellinek suggest entrepreneurs enjoy their work and are willing to work longer hours (Bird and Jelinek, 1988), even though they do not appear to start entrepreneurial activities to get rich, as they expect no higher incomes as employees (Douglas and Shepherd, 2002). All these factors render entrepreneurship a desirable career choice.

Besides these benefits, there are also individual characteristics, such as a positive attitude towards risk and independence. In previous research, Douglas and Shepherd

(2002) found evidence, that individuals with a higher tolerance for risk, and a positive attitude of decision-making autonomy share a stronger intention to become entrepreneurs.

Entrepreneurial activity

Economic activity is always moving. As new opportunities emerge, entrepreneurs seek to exploit them. One of the major trends associated with this constant change is that these activities move from big business to rather small companies. A prominent example here is the employment share of the Fortune 500 (Americas 500 largest firms). It has dropped from 20% in the 1970s to 8.5% in 1996 (Carlsson et al., 1992; Carlsson, 1999). In the following, we summarize findings concerning the influence of entrepreneurship on economic growth.

Many definitions of entrepreneurship exist. One of the most used is from Venkataraman (1997):

> ***Entrepreneurship****. The scholarly examination of how, by whom, and with what effects opportunities to create future goods and services are discovered, evaluated, and exploited.*

As of today, a much-debated question amongst researchers is, whether entrepreneurial activity contributes significantly to a country's economic growth? A measure of entrepreneurial activity is turbulence (entry-exit turnover). Reynolds (1999) has found some evidence for higher growth rates due to turbulence in the United States. Contradicting this finding is a study by Audretsch and Fritsch (2002), which found the opposite to be true for Germany. Shane (2009) argues that, typical startups are not innovative, create few jobs, and, hence, generate little wealth.

Up to 92% of startups fail within the first three years. This number is derived from a study by Marmer et al. (2011), who investigated 3.200 high growth, mostly tech startups, and concluded with this high rate of failure. Nevertheless, some entrepreneurs decide to start again, the so-called serial entrepreneurs who will be fore-mentioned in the next paragraph.

TABLE 2.1: Types of entrepreneurship (see Westhead et al. (2005)).

No.	Types of entrepreneurship	Explanation
1	Novice	Individuals with no prior minority or majority business ownership experience either as a business founder, an inheritor or a purchaser of an independent business, but who currently own a minority or majority equity stake in an independent business that is either new, purchased or inherited.
2	Serial	Individuals who have sold/closed a business in which they had a minority or majority ownership stake, and they currently have a minority or majority ownership stake in a single independent business that is either new, purchased or inherited.
3	Portfolio	Individuals who currently have minority or majority ownership stakes in two or more independent businesses that are either new, purchased and/or inherited.

Serial entrepreneurship

There are three categories of entrepreneurs. First, there are novice entrepreneurs, who start an enterprise for the first time. Second, there are entrepreneurs, who restart after having sold or closed a business. Third are portfolio entrepreneurs who aim for minority or majority ownership stakes in two or more independent companies. Westhead et al. (2005) suggest a definition for the three types, which the Table 2.1 summarizes.

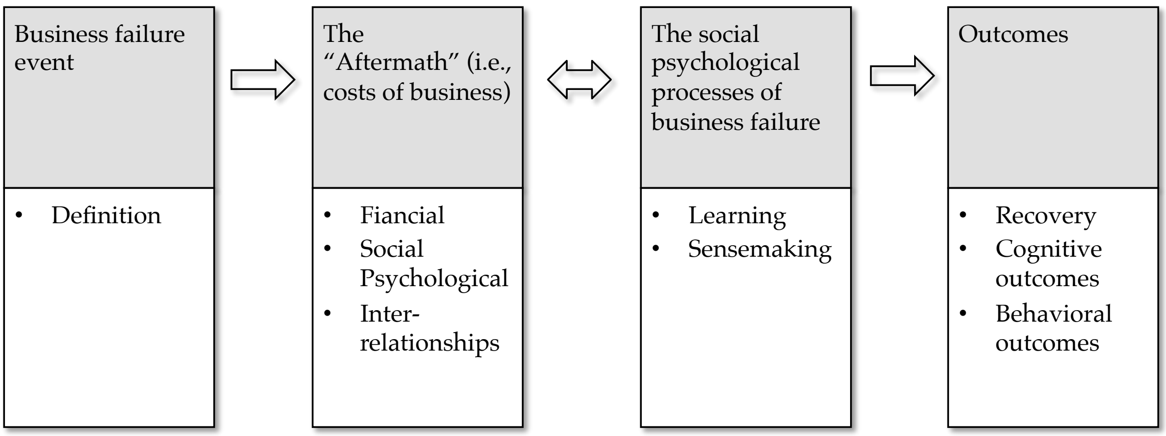
Our focus is serial entrepreneurs. Following the pioneering work of MacMillan (1986), pointing out the necessity of serial entrepreneurship to understand entrepreneurship, many more studies have been conducted in this field of research (e.g., Westhead et al., 2005; Zhang, 2011; Zhang, 2011). First of all, it appears to be fairly common for entrepreneurs to start anew. Serial entrepreneurs make up for a non-negligible proportion, between 18 to 25% of all entrepreneurs start again (Westhead et al., 2005; Westhead and Wright, 1998; Wagner, 2002). They have different characteristics than novice or portfolio entrepreneurs (Westhead et al., 2005). For instance, serial entrepreneurs exit significantly more businesses than portfolio entrepreneurs. A potential reason could be their motivation to start a business, which has the highest potential to be successful (Stokes and Blackburn, 2002). This behavior could also be interpreted by missing entrepreneurial skills to grow a business. Some cultures consider exits as failures, although a venture termination can have multiple reasons (see Chapter 2.3). Up to now, we know little about the individual's motivation and reason to start anew. Consequently, no theoretical framework exists to explain serial entrepreneurship. Researchers have mainly focused on individual topics, e.g., performance of serial entrepreneurs (Parker, 2013). Eggers and Song (2015) studied a common phenomenon among failed entrepreneurs. These tend to blame the external environment for their businesses failure, and particular rare their leadership, decision-making or planning-style. As a consequence, these entrepreneurs change industries for subsequent ventures. However, some studies on industry changes argue that these are costly, since valuable knowledge as the key to success, becomes obsolete in the new industry (Agarwal et al., 2004; Chatterji, 2009).

In the next chapter, we deep dive into entrepreneurial failure, as our core research focus.

2.3 Entrepreneurial failure

Over half of the ventures founded by entrepreneurs fail within five years after they started (Cooper, Carolyn Y. Woo, and William C. Dunkelberg, 1988). The Oxford dictionary defines failure as "lack of success". However, a sole simplification of entrepreneurial failure to this would certainly be too easy. The reasons for the high probability of entrepreneurs to fail are of different nature, as well as the consequences for the entrepreneur. For instance, some experience valuable learnings, while for others it may be an emotional or traumatic experience (McGrath, 1999; Cope, 2011; Shepherd, 2003a). For an overview on the current state of research on entrepreneurial failure, we follow an organizing schema for research on entrepreneurial business failure by Ucbasaran et al. (2012) (see Figure 2.5). They divide the research of entrepreneurial failure into four categories (A-D).

FIGURE 2.5: Business failure model by Shepherd (see Shepherd (2003b)).



A. Business failure definition

Before we deep dive into the subject, we have to answer questions such as: when is an entrepreneur failed? How do we define failure? In the following, we thus give a definition of failure regarding entrepreneurship. The literature provides different definitions for a venture's failure, each one being adapted to the specific research needs. The most important ones are:

1. **Discontinuity of ownership** (Singh, Corner, and Pavlovich, 2007; Watson, John and Jim E. Everett, 1996). In this case, the entrepreneur exits his or her business. This includes a business closure for instance due to retirement or health reasons of the entrepreneur. In addition, the entrepreneur might sell the business and therefore discontinues his ownership. Although this definition also applies to earlier studies (e.g., Star and Massel (1981)), discontinuity can not be put at the same level as failure. As Wennberg et al. (2010) stated, it can also be the result of success.

2. **Bankruptcy** (Shepherd, Haynie 2011). This definition refers to a poor economic performance. It is measurable, and therefore a popular definition of failure (Zacharakis, G. Dale Meyer, and Julio DeCastro, 1999). Not included are businesses that do not provide a reasonable income for the entrepreneur who, as a consequence, closes down the business.

3. **Discontinuity of ownership due to insolvency or performance below the threshold**. This definition combines 1 and 2. Due to poor performance, compared to the entrepreneur's expectations or goals the business goes bankrupt. The entrepreneur's expectations involve a subjective assessment of possible investment strategies.

From these different definitions, we can see that there is no one single definition. A precise definition though, has important implications for the outcome of the research project. We, therefore, choose definition 3 as it is most broadly applicable, as well as measurable.

B. The "Aftermath"

When a business fails, the entrepreneur experiences various different negative emotions, such as pain, remorse, shame, blame, anger, guilt, humiliation and fear (Cardon and McGrath, 1999; Cope, 2011; Harris and Sutton, 1986; Shepherd, 2003b; Singh, Corner, and Pavlovich, 2007). This is very understandable, as they tend to actively connect personally to their own business and some even build an emotional relationship towards it. A business failure, therefore, provokes a feeling of personal loss within the entrepreneur and creates grief

(Shepherd, 2003b). Shepherd defines grief as "[...] an umbrella term characterizing a number of negative emotions generated from losses associated with failure" (Shepherd, 2003b). To recover from grief, Shepherd describes the “loss orientation” and the “restoration orientation” modes and suggests a combination of both to recover quickly. As a result of this recovery, the entrepreneur may react positively to the failure of his or her business, in a sense that some even experience epiphanies after realizing a sudden deep insight into why and how they contributed to their firm’s failure, through, e.g., ego-based thinking and behavior (Shepherd, 2003b).

Besides these negative emotions, there is also a financial aftermath. Extensive research has been done on the influence of bankruptcy laws on entrepreneurship (Lee et al., 2011; Lee, Peng, and Barney, 2007; Armour and Cumming, 2008). It turns out that risk-averse entrepreneurs benefit from higher exemption levels due to these laws, by providing partial wealth insurance (Fan and White, 2003). Lee et al. (2011) found that entrepreneurial-friendly laws even lead to higher rates of venture funding. As the entrepreneur invests personally into his or her venture, there is also a financial loss related to a failure. Interestingly, entrepreneurs with high opportunity costs (i.e., those with multiple alternatives to the current venture) appear to be less patient with waiting for success, even if this means to fail faster (Arora and Nandkumar, 2011). The result of this behavior are increased chances of significant financial losses or gains.

With considerable economic costs, come great social costs involved into a venture failure. With social costs we refer to all negative effects on personal or professional relationships of the entrepreneur. Indeed, the venture failure can be connected with a loss of one’s social arena or network (Harris and Sutton, 1986). A lot of times these two are related and a result of stigma, which is of increasing interest to scholars and policy makers. The stigma related to entrepreneurial failure can be defined as "some attribute or characteristic that conveys a social identity that is devalued in a particular context" (Crocker et al., 1998). Shepherd (2003b) and Cope (2011) describe it as a "potentially painful and traumatic experience for entrepreneurs". But why does stigma exist, and why do entrepreneurs get stigmatized? One reason is that failure is contradicting commonly accepted norms and values. A deviation from them is considered as a thread for the overall society; the entrepreneur, therefore gets sanctioned. Stigmatization is seen as a collective punishment by society (Jones, Hesterly, and Borgatti, 1997, p. 931). As research on this topic focuses mainly on the socio-economic level of stigmatization, little is known about the actual effects on the individual entrepreneur. Singh, Corner, and Pavlovich (2015) are pioneers in this field, as they studied the influences of stigma on failed entrepreneurs qualitatively. They derive a process flow from their findings and divide the venture failure into three episodes, the anticipating, the meeting, and the transforming failure episode. One of their conclusions is, that stigmatization should be seen as a dynamic process rather than a static label. It is, therefore, developing over time. Accordingly, entrepreneurs may first experience stigmatization when realizing that failure cannot be ruled out. These conclusions contradict other researcher's findings, stating that stigma only occurs after the venture has failed (Cardon, Stevens, and Potter, 2011; Lee, Peng, and Barney, 2007). Other empirical studies point to the socio-cultural component of stigma and, for instance, highlight that the associated shame of entrepreneurial failure is stronger in Asia than in Anglo countries (Begley and Tan, 2001). For a broader introduction to the concept of stigma, we refer to Chapter 3.3.

C. The social psychological processes

Learning from failure and sense-making are the two most studied psychological processes in management and entrepreneurship literature (Cope, 2011; Shepherd, 2003a; Coelho and McClure, 2005). Prior research has shown, that failure can point out to the entrepreneur that something went wrong clearly, and it may start a process of reflection (Sitkin, 1992). At the end of this process, an individual understands potential causes of the failure and changes his or her mental models (Minniti and Bygrave, 2001; Politis, 2005; Ucbasaran, Westhead, and Wright, 2009). As failure can be a traumatic experience for entrepreneurs, learning does not necessarily occur immediately. Entrepreneurs need a recovery time to reflect on the causes of failure Cope (2011), before they can start with a reflection of failure causes. Still, learning from failure is considered difficult and sometimes does not occur, as it requires a critical self-view, agreement, and acknowledgment of failure causes (Cannon and Edmondson, 2001). Similar to this Cassar and Craig (2009) found evidence for a hindsight bias, which is capable of preventing entrepreneurs from starting self-reflection. It builds on the assumption that what individuals believe they experienced and what truly happened, may differ.

Several lines of evidence suggest that learning is part of a sense-making process, following scanning and interpreting (Gioia and Chittipeddi, 1991; Thomas, Clark, and Gioia, 1993; Weick, 1979). With scanning, scholars refer to an information gathering process, while interpreting requires an understanding of the information gathered. (Thomas, Clark, and Gioia, 1993). At the end of this process, individuals assign meaning to past experiences (Gioia and Chittipeddi, 1991). The concept of sense-making was first introduced by Dervin (1983), in information science. Weick (1988) applied it to organizational theory, and, e.g., Filion (1991), Shepherd (2009) and Shepherd (2003a) applied it to entrepreneurial failure. Shepherd (2003a) suggests confronting the loss associated with failure for accelerated recovery. As many different interpretations for a business failure exist, and entrepreneurs seek quick recovery, a sense-making approach focuses on plausibility over accuracy (Ucbasaran et al., 2012). Over time, impression and insights gained compensates for the missing accuracy, which resembles a learning process (Daft and Weick, 1984). Hence, through more and more plausible explanations for a venture failure, sense-making may help the entrepreneur to learn and move on.

D. Outcomes

The last step in the organizing schema on entrepreneurial failure by Ucbasaran et al. (2012) is the outcome. It splits into three types: recovery, cognitive, and behavioral outcomes.

Recovery from grief after failure is achieved, when the entrepreneurs thoughts concerning the event no longer generate negative feelings (Shepherd, 2003a). Previous research by Cope (2011) suggests a time in which entrepreneurs let go psychologically, followed by sense-making via reflection period, and finally a move-on period with ongoing reflection. The last stage enables entrepreneurs to start anew (see Chapter 2.2).

Cognitive constructs such as optimism or confidence have also been studied with failure (Feather, 1968; Ucbasaran et al., 2010; Hayward et al., 2010). Scholars studying optimism, mainly focused on over-optimism, i.e., the tendency to believe that one experiences negative events less likely than others, but positive events more likely. Entrepreneurs tend to be over-optimistic compared to non-entrepreneurs. This cognitive bias is considered a prerequisite for entrepreneurship (Ucbasaran et al., 2010). Previous studies on confidence concluded, entrepreneurs equipped with more confidence (or overconfidence) are more likely to start again (Hayward et al., 2010). These results indicate why it is a common phenomenon in entrepreneurship. With the various negative effects such as commitment to risky, new projects or "throwing good money after bad", overconfidence is seen to be part of it (La Hayward and Hambrick, 1997; Staw and Ross, 1989). Besides these severe effects, entrepreneurship requires some degree of resilience in entrepreneurs. Since many startups fail (see Chapter 2.2), it can be beneficial to display some power of endurance (Hayward et al., 2010).

Behavioral outcomes imply potential changes in an entrepreneur's behavior after failure. As we have seen above, some entrepreneurs experience valuable learning and develop plans to start anew (Schutjens and Stam, 2006). Unfortunately, research to explain the impact of learning from failure on subsequent founding performance is scarce. Ucbasaran, Westhead, and Wright (2009) found entrepreneurs with previous failure experience to identify more business opportunities than novice entrepreneurs, while surprisingly, not being more innovative.

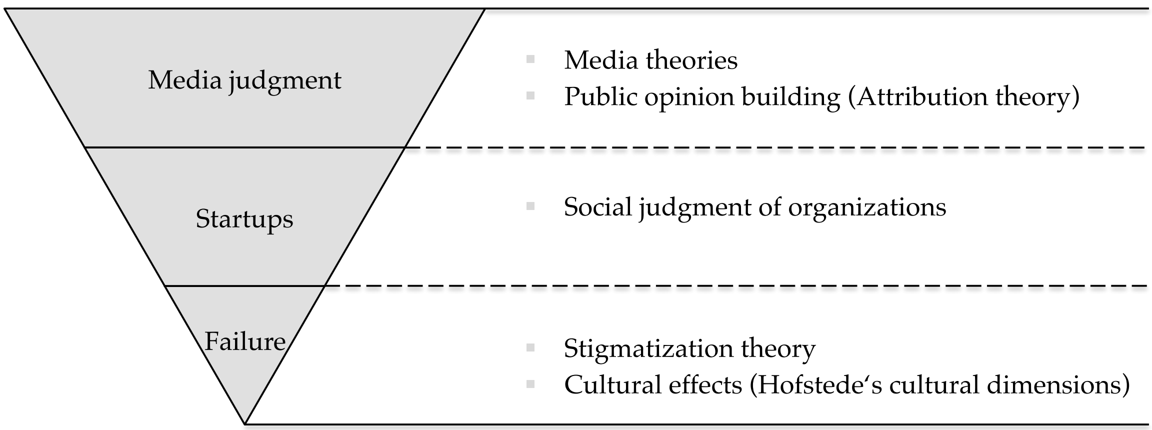
The following chapter integrates the role of media and explains how it judges entrepreneurs. Then it introduces the main theories on which this thesis builds on the following three studies.

Chapter 3

Theoretical foundation

The previous chapter gave an overview of the current state of research in the field of entrepreneurial failure. In this chapter we introduce how media judges organizations and actors, as well as the three main theories, on which this thesis is build on. First to mention is the social judgment of organizations theory, a sub-field of organizational theory. Followed by an overview on stigmatization, where we make use of findings from psychology. The last chapter is dedicated to cultural effects, especially in work-related topics. Figure 3.1 explains the logic of this approach.

FIGURE 3.1: Level and associated theory for our research approach.



3.1 Media judgment of failed organizations and entrepreneurs

This chapter introduces the term of *mass media* and relevant theories explaining the influence on public opinion building. We then provide an overview of the so-called social judgment and attribution theory to build the link to stigmatization.

Media and public opinion building

We focus our research on print and online press articles, released by major news corporations, newspapers or specialized journals. The influence of messages and opinions portrayed to the readers and, hence, public opinion has been subject to numerous research projects in fields such as election (McCombs and Shaw, 1972), television news (Iyengar, 1990), and image of companies (Carroll and McCombs, 2003). Many contending and sometimes overlapping theories have been developed to explain the findings. Here, we focus on the three most prominent theories, namely, the agenda-setting theory, the reinforcement theory, and the bullet theory. As of today, the field of mass communication lacks a comprehensive theoretical framework, while existing theories leave many unresolved issues.

Agenda-setting theory

To study the influence of mass media, McCombs and Shaw (1972) investigated its agenda-setting capacity in the US presidential campaign in a pioneering quantitative study. Since these initial steps, their theory has been refined. Today, two levels of agenda-setting theory exist.

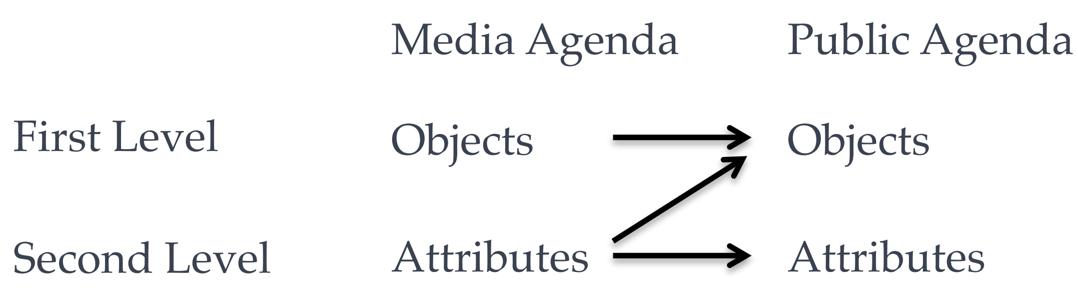
First level agenda-setting builds on the above-mentioned study by McCombs and Shaw (1972). Through this empirical study of voter's behavior, they discovered a relationship between media reporting on issues the public's salience of these topics. In other words, news items which are frequently and prominently covered by the media, will be regarded as more important by voters. The claim of media telling the public what to think about is called first level agenda setting. This finding was contradicting the long-held belief of weak media effects and led to further research in this field. Therefore, McCombs and Shaw (1972) formally developed the associated theory of agenda-setting. In a different agenda-setting study (which was lacking labels as such) Funkhouser (1973) discovered, that the visibility of topics for the public is strongly influenced by the amount of media attention given to it (Funkhouser, 1973). Comparing the covered topics to objective conditions, he showed even when issues already improved, the media coverage could still be negative or vice versa (Funkhouser, 1973).

Second level agenda-setting has developed more recently. Its main statement is that media not only tell the public what to think about (as in first level agenda-setting), media also influences how to think about the covered topics. This is the fundamental principle of second level agenda-setting, which focuses not simply on the themes being presented by the media, but also on how authors describe them and the attributes employed to do so. Therefore, second level agenda-setting is also called attribute agenda-setting. A lot of research is done on this topic (McCombs et al., 1997; Golan and Wanta, 2001; Kiousis, Bantimaroudis, and Ban, 1999). Ghanem (1997, p. 4) suggests to summarize it with two major hypotheses about attribute salience:

1. The way the media covers an issue or object (the attributes emphasized in the news) affects the way the public thinks about that object.
2. The way the media covers an issue or object (the attributes emphasized in the news) affects the salience of that object on the public agenda.

Ghanem illustrates first and second level agenda-setting, with Figure 3.2. In first level agenda-setting the dependent variable is media agenda, and considered in terms of objects. Hence, it does not link to any judgment, but merely the issue or topic itself. The same holds for the public agenda. It is different from second level agenda-setting, as media agenda (dependent variable) is considered in terms of attributes or perspectives (Ghanem, 1997). For both cases, the public agenda is the independent variable. Figure 3.2 illustrates these effects with two horizontal arrows. The diagonal arrow points out to the above mentioned second hypothesis from Ghanem on second level agenda-setting. Attributes used in news coverage influence the salience of issues and topics on the public agenda.

FIGURE 3.2: The two levels of agenda-setting. Graph from Ghanem (1997).



In an independently performed study, Noelle-Neumann and Mathes (1987), suggested to examine media content on three levels: agenda-setting, focusing, and evaluation. Here, agenda-setting refers to observing issues or topics reported by the media. Focusing relates to emphasized topics, and evaluation to opinions or views being put into the reporting. These three steps are similar to first and second level agenda-setting. Ghanem (1997) argues it is feasible to substitute agenda-setting with first level agenda-setting, as well as focusing and evaluation with second level agenda-setting.

Apart from the initial study by McCombs and Shaw (1972), several other scholars study agenda-setting effects in the US. Iyengar (1990) investigated television news coverage of main topics between 1974 and 1980 and compared it with data from the American public referring to the mentioned issues. They found significant effects of news coverage and confirmed agenda-setting effects. In a further project Iyengar and Kinder (2010) investigate how television news influences America's conceptions for the political reality. Participants in the study were exposed to different news information, each one receiving altered reports. Their results show, changed focus of news considerably influenced the degree of problem importance evaluation by participants (Iyengar and Kinder, 2010). Therefore, media does influence viewers understanding of political reality and even shapes it. It is reasonable to assume, that the same holds for the startup business. In the same vein, Carroll and McCombs (2003) argue, that even public's perception is focused by journalist's selection of news to cover.

A variety of studies on this topic has been performed in Germany. In a 1986 study on the agenda-setting function of television news, Brosius and Kepplinger (1990) performed a content analysis of main television news shows, paired with weekly surveys investigating the problem awareness of 16 prominent issues of that time. As a result the scholars concluded, problem awareness to generally correspond to media coverage. In addition, some opinion trends caused by problem awareness of the public have been counterbalanced by the media. In other words, first level agenda-setting also means the media is able to reduce problem awareness of a topic by refocussing the media coverage to others. This can be investigated deeper with our data in the sense that less media coverage of failure may pull away the attention from it, and, thus, may be seen as less likely or crucial by society.

Brosius and Kepplinger (1992) explore effects of media agenda and partisanship of German voters. Their results from surveys and television news data show how media coverage shapes the problem awareness of certain issues and voter's party preferences.

Agenda-setting theory describes how media reporting and public opinion are connected. The expressed sentiment is a surrogate for entrepreneurial friendliness. Strongly positive media coverage can thus lead to a positive public opinion on the startup business or vice-versa. Hence, the media has a non-negligible effect on the founding culture and, potentially, on entrepreneurial activity. Our research questions from Chapter 1 aim to help shed some light on this field. We continue with the next theory on media and public opinion building, the reinforcement theory.

Bullet theory

The phenomenon of mass media rose significantly in the early decades of the 20th century. Nothing similar has ever existed before. Therefore, theories emerged in the attempt to explain consequences of mass media. The bullet theory is one of them. According to Hindle and Klyver (2007), it derives its genesis on two primary assumptions:

1. Media is a very powerful social, political, and cultural institution.
2. Media can directly influence the behavior, and thinking of audiences.

These two assumptions reflect the strength associated with mass media in the bullet theory model. Based on this idea, the public therefore, perceives information more or less uniformly thus making it highly susceptible to messages or opinions that are expressed (Hindle and Klyver, 2007). Thereby, members of society are considered isolated from each other due to an impersonal and ruthlessly industry focused community.

Reinforcement theory

Reinforcement theory was originally developed by Klapper (1960). The theory dwells on three assumptions:

1. Audiences are current recipients of media content and question it.
2. The media's influence on changing values or beliefs of audiences is very limited.
3. The media can only reinforce opinions and ideas of audiences.

Klapper justifies the considerably low power of mass media with regards to far more powerful agencies, such as family, friends, school or political institutions. As these are real life factors, they are more relevant for the members of society. We will compare the above-mentioned hypotheses to our measured results.

Public opinion building and attribution theory

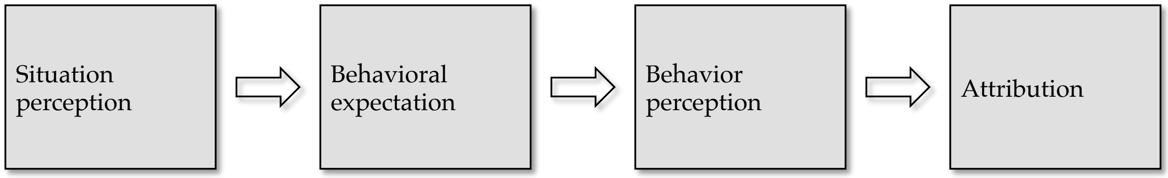
Contrary to some popular beliefs, judgment by journalists is often subliminal, as they seek to report objectively and fact-based and will refrain from using too positive or too negative statements (Balahur et al., 2013). Nevertheless, they are capable of expressing opinions, by citing other people, omitting or highlighting certain facts, or underlying statements. Different theories exist to explain the motivation for this type of judgment and its consequences. We will focus on the so-called attribution theory, which receives the most attention in this field (Kelley and Michela, 1980).

Attribution theory was introduced by the pioneering work of Heider et al. (1958) and incorporates ideas from Jones and Davis (1965), as well as Kelley (1973) and Kelley and Michela (1980). The associated field of study refers to the perception or inference of cause (Kelley and Michela, 1980). In other words, the attribution theory provides us with ideas of how we come to conclusions about the causes of actions. A fundamental belief of this theory is that everybody is a nonprofessional psychologist, who tries to explain the behavior of others (Jonas and Hewstone, 2007, p. 72). There is no single attribution theory but much different. These share the same idea of people interpreting the behavior of a person regarding its causes, and the interpretation of the observed behavior is relevant for the resulting reaction to it (Kelley and Michela, 1980). This process is called causal attribution and can be defined as (Jonas and Hewstone, 2007):

> ***Causal attribution.*** *The process through which the observer comes to conclusions about the causes of a person‘s behavior.*

In most of the attribution research, two main characters exist, the observer, who explains the behavior towards another person or object, and the actor (Jonas and Hewstone, 2007). In our case, the observer is the media consumer, who reads an article about a startup, and makes up his mind about the described behavior of an entrepreneur, which the media consumer may know or not. Hence an attribution is when the observer assigns causality to a person, object or situation. According to Gilbert, an attribution follows a four-step process, as Figure 3.3 displays. To execute a proper attribution, the observer has to go through all these steps. The first is called situation perception. Here, the observer tries to understand the situation an actor is in. For example, an entrepreneur facing economic issues. Next, the observer considers prior knowledge and beliefs, such as a poor economy. These beliefs raise expectations towards the actor's behavior. In the third step, an observer processes the actor's behavior and categorizes it. For example, the entrepreneur may have put the startup on the line by making wrong decisions.

FIGURE 3.3: Event sequence for an attribution made by an observer. Graph from Gilbert and Malone (1995).



Eventually, the observer determines whether or not the actor's behavior contradicts his or her expectations. If it does contradict, the observer draws dispositional inferences about the actor, e.g., the entrepreneur not being capable of finding a good strategy for his startup. If it does not contradict, the observer will not draw inferences and just attribute the entrepreneur's issues to the overall economy.

The above-mentioned example illustrates how people make use of prior knowledge and beliefs when drawing conclusions of other people's behavior. Kelley and Michela (1980) argue, that the final attribution rarely occurs without the influence of pre-existing suppositions. As demonstrated in Chapter 3.1, consumption of media can impact this knowledge and the beliefs of observers. This effect is not negligible, as observers do not always follow the four steps in Figure 3.3, but only make use of expectations about how incidents proceed and take them as a reference point for their attribution (Jonas and Hewstone, 2007). In this case, learned attribution styles is used by the observer (Jonas and Hewstone, 2007). These styles could be the result of a dominant media coverage orientation, such as an imbalance of press articles with the predominantly negative sentiment. Schachter and Singer (1962) add to this belief, as they show that observers are prone to manipulation. Since manipulation may lead to false inferences, observers sometimes are unable to review their point of view, even if they receive additional information. This phenomenon was shown by Gilbert and Malone (1995).

In the more than fifty years passed since the first steps of attribution theory, scholars have come up with many interesting results. For example, in his 1979 study, Zuckerman concludes that outcomes, which are unexpected by observers, are more likely to be attributed to luck than skill or ability (Zuckerman, 1979). This seems especially suitable for uncertainty avoidance oriented societies, where people refrain from taking risk and envy success (see Chapter 3.4). In general, though, success is more likely attributed to personal skills and failure to external factors (Miller and Ross, 1975; Zuckerman, 1979). Going into more detail, this conclusion does not hold true for normative expectations. Deaux (1976) shows in his study, that for many tasks, the success of men and failure of women is more likely to be expected, and attributed to personal ability. In addition failure of men and success of women for the same tasks is more likely to be attributed to luck.

Judgment of failure

In the previous paragraph, we have seen how audiences build opinions and how the judicial process works. In the following, we will answer the question of how media reporting in general affects entrepreneurship, and how negative media judgment in particular, directly and indirectly, affects entrepreneurs.

Besides the above-mentioned findings, media communication, and hence the portrayed sentiment is expected to influence entrepreneurial activity, as it impacts culture and social behavior (Macnamara, 2003; McDonald, 2004). In fact, prior research has shown, institutional norms, inclusive of formal rules and informal cultural values are key determinants of a countries citizens to engage in entrepreneurial activity and societal wealth generation (Acs et al., 2005). By doing so, institutional norms shape the frame of rules and expectations the entrepreneur has to follow and accept. According to Aldrich and Fiol (1994), these norms or standards can be seen as a dictation of what is legitimate to do and what is not. As the entrepreneur's interest is to ensure economic growth and continuity of his venture, he faces pressure to act in accordance with these normative expectations (Dimaggio and Powell, 1983). A derivation from normative expectations is not tolerated by societies and leads to negative judgment of entrepreneurs (Devers et al., 2009), for instance by the media. Hence, negative judgment occurs when entrepreneurs do not act as they are expected. Negative judgment is therefore not the result of failure and can also be observed for successful entrepreneurs, especially so, when the success is unexpected, as described in the previous paragraph.

The consequences of adverse judgment and the associated stigmatization can be fatal. Hardly any punishment of people is as painful as expulsion from one's social network (Baumeister and Leary, 1995). This goes back to prehistoric times as banishment from one's group meant certain death. The associated fear is still tangible in modern times and people, as social animals, therefore pay keen attention to other's opinions. As described in Chapter 3.1, these opinion's are influenced by the media. Cardon, Stevens, and Potter (2011) found that negative media coverage and the transmitted stigmatization can result in criticism of unsuccessful entrepreneurs. One of the most prominent examples for such stigmatization of recent times is the founder of the internet platform Megaupload Kim Schmitz. Following his arrest in New Zealand, hundreds of articles accusing Schmitz of criminal acts are available in print form.

The results of stigmatization can be very different. It effects the entrepreneurship-culture, as criticism of entrepreneurs can deter subsequent venture start (Politis and Gabrielsson, 2009). On a personal level, the negative judgment may enhance negative feelings associated with a venture failure. As a consequence, entrepreneurs may experience emotions such as denial, anger, personal pain, sadness, dismay, worry, anxiety, annoyance, frustration and even depression (Dillon, 1998) and this is only the personal emotions of the entrepreneur. Adding to these personal emotions of the entrepreneur is the external view from family members, friends and society (e.g., neighbors, business partners, former colleagues, etc.) via indirect or direct judgment of the entrepreneur, and resulting stigma (Jenkins et al., 2014; Simmons, Wiklund, and Levie, 2014; Landier, 2005).

The consequences of judgment may deter subsequent venture creation, and therefore has already captured the attention of policy makers. The European Commission launched a “Second Chance” policy which attempts to reduce the negative effects, and entrepreneurial-friendly bankruptcy laws are put in place (Singh, Corner, and Pavlovich, 2015). Nevertheless, the judgment of entrepreneurs after a business failure can be extremely harmful. Entrepreneurs may experience judgment such as blame (Moulya and Sankaranb, 2000), discrimination (Singh, Corner, and Pavlovich, 2007; Cope, 2011; Shepherd and Haynie, 2011), malicious joy or devaluation (Wiesenfeld, Wurthmann, and Hambrick, 2008). But besides these harsh judgments, positive signs such as support or praise for the entrepreneur exist (Singh, Corner, and Pavlovich, 2015).

3.2 Social judgment of organizations

The field of organizational theory aims to answer questions such as why organizations come into existence, remain in business, or how they function (Jones et al., 2010; Shafritz, Ott, and Jang, 2015; Jensen, 1983). From numerous publications in this field of study, factors for venture growth, such as sufficient capital, advanced technology, qualified personnel, or networks have been identified (Zimmerman and Zeitz, 2002). This thesis concentrates on a unique part of the organizational theory, the social judgment of organizations. To the "hard facts" of venture growth, it introduces so-called "conditions". They include legitimacy, reputation, and status. These "conditions" are only partly under the control of an entrepreneur, as they depend on an evaluator's judgment. From an evaluator's perspective, legitimacy, reputation, and status are different dimensions to form and express social evaluation or judgment of an organization. We introduce these concepts, as we rely on them for interpreting our findings.

Before we delve into these concepts to explain their relevance for entrepreneurs and ventures, we define what is meant by them. As researchers have studied these concepts with various intentions, many different definitions exist for legitimacy, reputation, and status (for an overview see Bitektine (2011)). This conundrum has motivated researchers to strive for a more consistent framework in this field (Bitektine, 2011), and raised criticism of existing approaches (Hudson, 2008). Here, we summarize those definitions, which are most suitable for our focus (see Table 3.1).

With the definitions in Table 3.1 one can understand legitimacy as the acceptance of an organization, a justification for its existence, or a social judgment about it after an evaluation process. Legitimacy is, therefore, essential to an organization's success. Evaluators can grant or withhold legitimacy to an organization, and are, therefore, equipped with a mechanism of social control (Bitektine, 2011). As different evaluating audiences exist, they are not all equally relevant for the organization. Research has focused on fields such as media legitimacy, where journalists evaluate and present their view to the general public (Brown and Deegan, 1998; Deephouse and Carter, 2005; Pollock and Rindova, 2003; Wry, Deephouse, and McNamara, 2006). The sentiment or tone which media portrays is used as an indicator for legitimacy, with the targeted audience (Deephouse, 1996). Hence, the media has a monitoring and reporting function of illegitimacy, e.g., a deviation from social norms (Hybels, 1994). This becomes even more apparent as studies show that investors react quickly to newly released information on firm's performance (Klassen and McLaughlin, 1996; Konar and Cohen, 1997; Muoghalu, Robison, and Glascock, 1990). Another field of research is legitimacy with regulators (Deephouse, 1996; Rao, 2004; Baum and Oliver, 1991), which examines legitimacy with executives of rules an organization has to obey. Typically, the pronounced sanctions to organizations are used as a measure of their legitimacy (Deephouse, 1996).

Aldrich and Fiol (1994) further distinguish two types of legitimacy and refer to cognitive and sociopolitical legitimacy. Cognitive legitimacy displays the level of public knowledge about a new activity, whereas sociopolitical legitimacy reflects the evaluation of a venture to be appropriate, by, e.g., the general public, or key stakeholders. Hudson (2008) links this type of legitimacy to stigma. She argues that illegitimacy is the result of negative social evaluation of

TABLE 3.1: Definitions for legitimacy, reputation, and status, according to current research.

Concept	Definition	Reference
Legitimacy	Justification of organization's "right to exist".	(Maurer, 1971)
	"A social judgment of appropriateness, acceptance, and/or desirability".	(Zimmerman and Zeitz, 2002)
	"Acceptance of the organization by its environment".	(Kostova and Roth, 2002)
Reputation	"Expectation of some behavior or behaviors based on past demonstrations of those same behaviors".	(Podolny, 1993)
	"A set of attributes inferred from the firm's past actions and ascribed to the firm".	(Weigelt and Camerer, 1988)
	"Perceived ability of the firm to create value for stakeholders".	(Rindova, Pollock, and La Hayward, 2006)
Status	"Prestige accorded firms because of the hierarchical positions they occupy in a social structure".	(Jensen and Roy, 2008)
	"Prominence of an actor's relative position within a population of actors".	(Wejnert, 2002)
	"Rank-ordered relationship among people associated with prestige and deference behavior".	(Huberman, Loch, and Önçüler, 2004)

organizations in form of stigma (see also Chapter 3.3).

Reputation is the result of past behavior and builds over time. Large cooperations aim for a high reputation, and hence positive social evaluation, and actively seek to manage it, especially after scandals such as the Volkswagen emission crisis (Tihanyi, Graffin, and George, 2014). Startups, on the other hand, are to a lesser degree able to actively manage their reputation, and one could argue that a positive social evaluation, thus becomes even more important for them.

Status is seen as a relative professional position or social standing. George et al. (2016) found that high status results in different benefits. It may attract more resources, more capable collaborators, and recognition from the general public. On the other hand, it is not for free, as high status also raises expectations, and hence may result in more visibility and scrutiny. Moreover, as actors may give reason to question their status, a decline could be quick and severe.

In summary, legitimacy, reputation, and status are of high importance for a startup's success. All three factors are influenced by the media and the portrayed judgment. Organizations, such as startups, may benefit in these dimensions through the media, or may experience negative social evaluation, and thus become illegitimate, lose reputation or status. Hudson (2008) shows that negative social evaluations of organizations also affect the main actors, such as founders, that are associated with the organization. Here we seek to add to this emerging research stream by quantifying social evaluation. A key concept linked to social judgment of organizations theory, is stigmatization theory, which we introduce in the following chapter.

3.3 Stigmatization

Stigma research primarily evolved in the field of psychology (Jones, 1984; Crocker and Major, 1989; Kurzban and Leary, 2001). More recent studies of stigma research focus on entrepreneurship (Jenkins et al., 2014; Landier, 2005; Simmons and Wiklund, 2011). A broad definition is given by Crocker et al. (1998):

> ***Stigmatization***. *The outcome of a process whereby social audiences form collective judgments about the consequences of bearing a particular marking and*

> *whereby persons who bear that marking are socialized to incorporate the judgments of the wider society into their conception of self.*

Stigmatization is, thus, a process, which develops over time and results in a marking of individuals or groups. Prior research has investigated stigmatization in fields such as ethical minorities (Brigham, 1974; Hartsough and Fontana, 1970; Samuels, 1973), women (Heilbrun, 1976; Broverman et al., 1994), unattractive persons (Berscheid and Walster, 1974; Dion, 1972), or homosexuality (Levitt and Klassen, Jr, Albert D, 1976; Herek, 1984).

People who experience stigmatization are disadvantaged compared to non-stigmatized people, as they suffer from, e.g., negative social, economic, and psychological consequences (Treiman and Hartmann, 1981; Braddock and McPartland, 1987; Dion, 1972). Sociologists study theses consequences, and developed the so-called self-concept (Cooley, 1956). The self-concept consists of one's awareness of other people's evaluation and the adoption of their views. Following this theory, stigmatized individuals absorb negative evaluations by others into the self-concept. Mead (1934) found that two types of evaluators who judge and influence the self-concept exist. First, individuals with whom the evaluated person interacts (e.g. peers, colleagues), and second, there is the sociocultural environment as a whole (Mead, 1934). As a result, stigmatized people should have a lower self-esteem (Cooley, 1956).

Surprisingly, empirical evidence shows, that stigma does not lead to a reduction in self-esteem expressed in form of the lower self-acceptance, worthiness, or self-respect (Rosenberg, 1965; Rosenberg, 1979; Wylie, 1979). Different possible explanations for this inconsistency exist. Stigmatization may have little effect on self-esteem, as it develops fairly early in life. This interpretation is questionable, as some scholars show, that self-esteem varies with age, and other incidences for example feedback on performance, social context or educational transitions (Gergen, 1971; Rosenberg, 1979; Harter, 1986). Another explanation argues, stigma does not result in lower self-esteem when it comes from sources, which are not influential on the individual. As described above, only certain groups affect the individual in case of stigmatization. Besides these groups, researcher additionally found evidence for negative depictions of stigmatized groups in the media, and suggest a bias for explanation (Taylor, Lee, and Stern, 1995; Niven, 2001; McCarthy, McPhail, and Smith, 1996; Watts et al., 1999). We will discuss this in more detail in Study I. A last potential explanation for the lack of empirical evidence of lowered self-esteem in stigmatized groups might also be, that evaluators can suppress their negative opinions about them. Negative and positive feelings can also exist at the same time (Gergen and Jones, 1963). When these evaluators interact with members of a stigmatized group, they may not articulate their negative feelings or opinions. Again, there is empirical research, which contradicts these ideas, arguing that people are not fully capable of suppressing prejudice in their behavior (Farina et al., 1976).

Besides explanations of why stigma has little effect on self-esteem, stigmatized individuals develop self-protection mechanisms to maintain their self-esteem (Snyder, Higgins, and Stucky, 1983; Taylor and Brown, 1988; Tesser and Campbell, 1980). Berglas and Jones (1978) suggests some individuals, who may experience stigmatization in case of failure, anticipate this failure by handicapping themselves in advance, which equips them with an excuse. In addition, Bradley (1978) found, that these individuals also tend to attribute success and failure to subjective causes in favor of themselves.

Since the aforementioned examples show the little effect of stigmatization on self-esteem, there does exist a connection to the overall feelings or global feelings of people. Global feelings and self-worth tend to be connected, as for instance, one's physical abilities or math skills, are considered by individuals when they assess themselves. From a research point of view, these concepts have to be distinguished (Marsh, 1986; Rosenberg, 1979). For instance, one may consider his or her appearance as negatively but still, experiences high global self-esteem. Beyond the aforementioned examples, also entire groups suffer devaluation in form of stigma. Typically, the stigmatizing audience is the broader society, or it is part of a countries culture (Crocker and Major, 1989).

These empirical findings do not infer that stigmatized individuals are immune to criticism. In fact, as Crocker and Major (1989) argue, prejudice and discrimination may be harmful in other psychological ways (e.g. self-confidence, performance expectancies, achievement motivation, and susceptibility to certain forms of mental and physical illness). Also, there is evidence that certain types of stigma do indeed lower self-esteem. As Jones (1984) found, stigma varies substantially in severity, concealability, and social disruptiveness. Consequently, reactions of stigmatized individuals vary depending on the dimension of stigma. Hence, stigma processing is not the same for everyone, and there is evidence for stigma causing lower self-esteem indeed. Verba and Scholozman (1979) show that becoming unemployed may cause an erosion of self-esteem. Similar findings show that going on welfare may also lower self-esteem (Briar, 1966). Now, the question arises, what distinguishes stigma which lowers self-esteem, from stigma which does not? One explanation focuses on the stigmatized individual. While some people do experience stigma since the beginning of their lives (e.g., physically handicapped or religious minorities), others do not experience it from the beginning, but later due to, e.g., disabilities caused by accidents or diseases. Scholars argue the psychological consequences of stigma for these two groups differ (Janoff-Bulman and Frieze, 1983; Jones, 1984). As members of a stigmatized group have developed protection strategies to save them from negative effects onto their self-esteem, individuals, who experience stigma suddenly in life, are ill-prepared for coping with stigma. Hence, it may be a long learning process for them to devitalize negative feelings associated with the experienced stigma. Similarly, Jones (1984) suggests, that these sudden stigmas may be less endurable to the stigmatized person than little by little stigma. By transferring these ideas to our research topic, it is reasonable to assume entrepreneurs who experience a huge negative change in media judgment after their business fails, to suffer more from failure than those, where the change is smaller. We will have a deeper look into this relationship in Study II.

In summary, all the above studies highlight the diversity of stigma. For our research topic, we can derive three assumptions. First, as business failure is new for many of the entrepreneurs in our dataset, we would expect them to experience difficulties in dealing with a potential arising stigma. Hence, we expect those entrepreneurs, who experience a more negative media sentiment change, i.e., before and after failure, to be less likely to start again. Second, as stigma does not necessarily lower self-esteem, entrepreneurs who experience setbacks, such as failure, may have self-protective strategies to prevent them from negative psychological effects, and thus a potential restart is independent of the overall media sentiment of the entrepreneur's startup. Third, as the culture of a society reflects the judgment, and therefore a potential stigmatization, we expect the judgment to vary between countries. Such differences should be particularly pronounced for a comparison of Germany and the USA. The problem of how to address cultural differences and in what dimensions they differ will be part of the next chapter.

3.4 Cultural differences

In 1980, one of the most influencial studies on the before culture of the last century: "Cultural Consequences: International Differences in Work-Related Values" was published (Lonner, Berry, and Hofstede, 1980). The study relies on data from surveys of employees in 40 countries. Hofstede and his team found that across societies cultural differences can be explained and quantified using four dimensions or cultural values (Hofstede, Hofstede, and Minkov, 1991; Hofstede, Hofstede, and Minkov, 2010). In the following years, G. J. Hofstede complemented these and came up with six dimensions in total:

1. **Power distance:** The extent to which power differs within the society, institutions, or organizations are accepted by those members with less power.

2. **Individualism vs. collectivism:** Individualism pertains to societies in which the ties between individuals are loose: everyone is expected to look after oneself and one's immediate family. Collectivism, as its opposite, pertains to societies in which people from birth onward are integrated into vigorous and cohesive in-groups, which, throughout people's lifetimes, continue to protect them in exchange for loyalty.

3. **Femininity vs. masculinity:** A society is masculine when it displays a preference for achievement, heroism, assertiveness, and material rewards for success. Femininity relates to preferences for cooperation, modesty, caring for weaker members of society, and quality of life.

4. **Uncertainty avoidance:** The extent to which the members of a culture feel threatened by ambiguous or unknown situations.

5. **Long term vs. short term orientation:** Related to the choice of focus for people's efforts: the future or the present and the past.

6. **Indulgence vs. restraint:** Indulgence oriented societies allows for the relatively free gratification of basic human desires related to having fun and enjoying life. Restraint oriented societies control gratification of needs and regulate it via strict social norms.

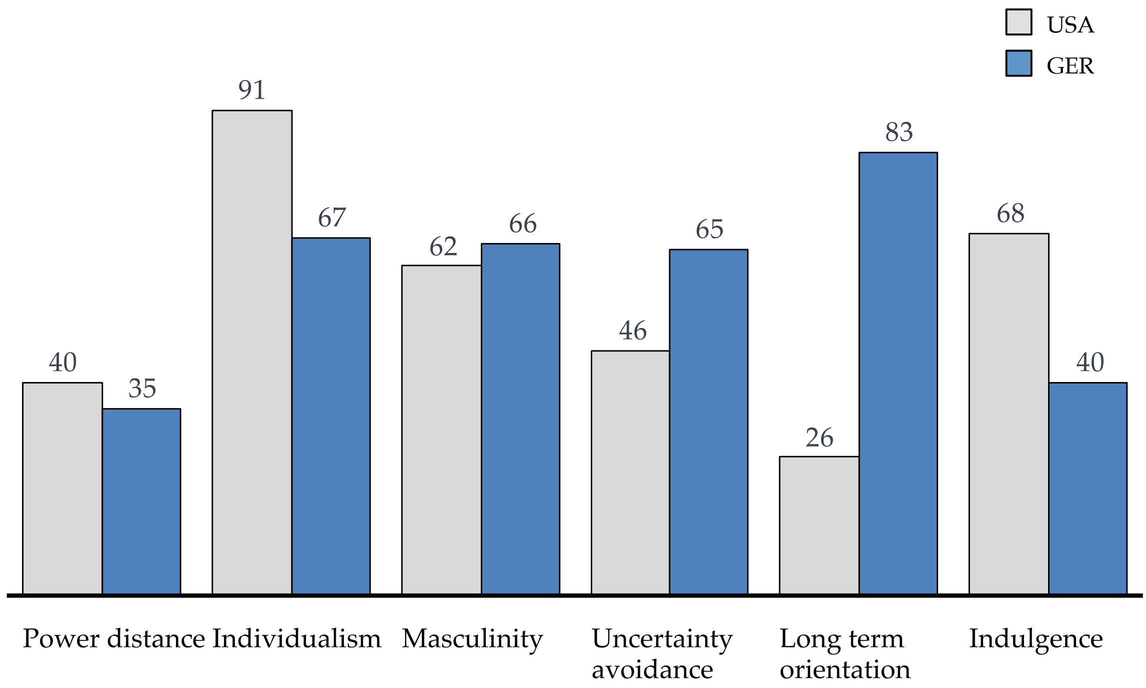
The uneven distribution of cultural values explains a culture's attitude towards entrepreneurship and how it deals with failure. The two most promising dimensions to measure entrepreneurial friendliness are individualism vs. collectivism, and uncertainty avoidance. The latter describes how a society deals with ambiguity, and should not be mistaken for risk avoidance (Hofstede, Hofstede, and Minkov, 2010). The level of uncertainty avoidance indicates how the members of a culture feel in unstructured situations, e.g., novel, unknown, surprising, or different from the usual (Hofstede, Hofstede, and Minkov, 2010). As entrepreneurship can be defined as "[...] the scholarly examination of how, by whom, and with what effects opportunities to create future goods and services are discovered, evaluated, and exploited" (Venkataraman, 1997), it is strongly related to uncertainty and ambiguity. There are two characteristics of this social value. Countries can either have a low or high uncertainty avoidance. In low uncertainty avoidance cultures, uncertainty is well excepted and more easily dealt with. For example, a teacher might rather admit not to know, instead of pretending to have all the answers (Hofstede, Hofstede, and Minkov, 2010). Consequently, these cultures treat deviation from conventional rules as not threatening, and display a greater tolerance, for instance to mistakes or failure. These cultures are also more willing to take risks, and success is well recognized (Lonner, Berry, and Hofstede, 1980). High uncertainty avoidance cultures try to avoid the above-mentioned situations, as they have behavioral codes, laws, and rules in place that are more strict. A deviation from this tight corset is seen as a potential danger, causing members of societies with high levels of uncertainty avoidance, to respond with intolerance as their personal aim is a high level of security. Lonner, Berry, and Hofstede (1980) also found that these societies experience a greater fear of failure. We, therefore, expect to replicate these findings in this research project.

The second dimension of interest is individualism vs. collectivism. Hofstede measures each dimension on a scale from 0 to 100. A high value in this dimension stands for a high level of individualism. Such societies value extra efforts, and it is a common belief that those members of a society should be rewarded for it via financial remuneration or recognition for their achievements (Lonner, Berry, and Hofstede, 1980). In contrast, collectivistic societies do not support this type of recognition (Lonner, Berry, and Hofstede, 1980). Collectivistic societies also have a tendency to attribute a high importance to getting ahead in industry, whereas individualistic societies value personal contacts, i.e., knowing the relevant people, as more important in order to lead them further to achieve goals (Lonner, Berry, and Hofstede, 1980). Shane (1993) found, that highly innovative societies display intense individualism and accept uncertainty. He explains these findings with the dominating values of society and suggests investing primarily in social values, rather than research and development.

It is possible to compare countries and their values of uncertainty avoidance on Hofstede's website (https://geert-hofstede.com), based on the data used in his book (Hofstede, Hofstede, and Minkov, 2010). The U.S. scores 46, while Germany reaches a staggering 65 (see Figure 3.4). This means the U.S. has a significant lower uncertainty avoidance compared to Germany and is, therefore, more open for new ideas and innovative products. As the U.S. scores below the average of 50, it can be seen as a country with low uncertainty avoidance. Germany is at the upper end in a global comparison. The value of 65 indicates a preference for uncertainty avoidance (Hofstede, Hofstede, and Minkov, 2010). Looking at the values for individualism vs. collectivism, the two countries also vary, but this time the values are switched. The US scores 91, which is one of the highest scores, and Hofstede concludes: "The [US] society is loosely-knit in which the expectation is that people look after themselves and their immediate families only and should not rely (too much) on authorities for support" (Lonner, Berry, and Hofstede, 1980). Germany also scores fairly high (67) in comparison to the overall average of 50, and Hofstede concludes that "there is a strong belief in the ideal of self-actualization" (Lonner, Berry, and Hofstede, 1980). This attitude builds a solid foundation for entrepreneurial activity and friendliness in both countries.

We can therefore reason, that both, the US and German society welcome entrepreneurial activity, though the German has a smaller tendency to enter uncertain situations and consider security as an asset.

FIGURE 3.4: Hofstede's six dimensions compared between the US and Germany.



Societies with high levels of uncertainty avoidance display a so-called in-group bias (Billig and Tajfel, 1973). As these in-groups are also relevant in Hofstede's individualism vs. collectivism dimension, we briefly introduce the concept of in-group bias in the next paragraph.

In the 1970s a series of research projects investigated a phenomenon which nowadays is called in-group bias (Billig and Tajfel, 1973; Tajfel et al., 1971). These studies aimed to answer the question why people tend to favor their own group over others. In-group bias is thus a more positive evaluation of the in-group compared to the out-group (Brewer, 1979). An in-group can be understood as a social group, to which an individual psychologically relates to, and considers oneself as being a member in (Tajfel and Turner, 1979). The psychological relation can be due to, e.g., age, race, or religion. Today the in-group is a highly confirmed finding within the field of psychology. The bias expresses itself in ways such as a preference for resources allocation to members of one's group, a positive differentiation from the out-group, or higher levels of cooperation with in-group members. Thus, researchers tend to relate this type of behavior to ethnocentrism (Brewer, 1979; Mullen, Brown, and Smith, 1992), which describes the prejudice of individuals towards foreign groups.

From research on in-group bias emerged the field of social identity theory (SIT) (Tajfel and Turner, 1979; Abrams and Hogg, 1990). An individual's identity is part of the self-concept discussed in Chapter 3.3. According to social identity theory, individuals aim for positive social identities to retain or improve their self-esteem (Tajfel and Turner, 1979). The best way to achieve this is to compare one's group to others, and to positively differentiate it from them. Hence, by differentiation, the in-group seeks to achieve or maintain a superior outcome in some dimensions (Tajfel and Turner, 1979). Aberson, Healy, and Romero (2000), as well as, Hogg (2000) suggest a more complicated relationship between in-group bias and effects on self-esteem. They argue that individuals tend to identify themselves with groups primarily due to self-uncertainty. Individuals aim to reduce uncertainty, as it is considered a reluctant psychological state (Hogg, 2000; Hogg, 2007). Hogg argues, that social groups equip the individual with guidelines and norms and therefore help them to steer through uncertainty. This uncertainty reduction technique can be related to Hofstede's cultural dimension uncertainty avoidance (see above).

Despite an ongoing debate on requirements for an in-group bias to occur, there has been little doubt that the effect truly exists. It is proven to be robust and stable in a number of studies (Bettencourt et al., 2001; Mummendey and Otten, 1998; Mullen, Brown, and Smith, 1992), and is accepted as a relevant theoretical foundation for explaining intergroup relationships (Hewstone, Rubin, and Willis, 2002).

As the above-mentioned examples show, Hofstede's uncertainty avoidance dimension is linked to the in-group bias, and the social identity theory. Hogg distinguishes two different types of uncertainty (Hogg, 2007; Hogg, 2000), one immediate, and one in larger social context.

Uncertainty avoidance in a larger social context refers to Hofstede's uncertainty avoidance. Fischer and Derham (2016) found that greater uncertainty avoidance can be associated with more in-group bias. With the scores displayed in Figure 3.4, we expect Germany to display higher levels of in-group bias than the US. We will investigate this further in Study I.

Findings regarding Hofstede's individualism-collectivism dimension and in-group bias have, thus far, been inconsistent. Yamagishi, Jin, and Miller (1998) found that collectivistic societies exhibit a tendency for higher levels of in-group bias, as reluctant behavior is frowned upon by the in-group. On the contrary, Hogg (2000) suggests, that a need for individuals to identify with various possible in-groups in individualistic societies exists.

Chapter 4

Data collection

We build our dataset in two main steps. First, we select startups from the CrunchBase database. Second, we download the associated press articles. The next chapter introduces the selection of startups, and the selection of articles in 4.2. After that, Chapter 4.2, and 4.3 describe the preprocessing, and cleaning of press articles. Finally, we give an overview of the resulting dataset, and descriptive statistics in Chapter 4.4.

4.1 Selection of startups

One of the much-used databases for information on startups is CrunchBase (Liang and Yuan, 2016; Marra et al., 2015). The CrunchBase database was founded in 2007 by Mike Arrington in San Francisco. It provides data concerning investors, incubators, and startups. The company claims to comprise data about roughly 50.000 startups, founders, funds, fundings, and events. CrunchBase started as a crowdsourced database to track high-tech startups covered on TechCrunch, which is one of the most popular and regarded blogs concerning technological innovation.

The database is constantly growing and updated by active contributors. To this day, CrunchBase claims to have about 2 million users per month, accessing the database.

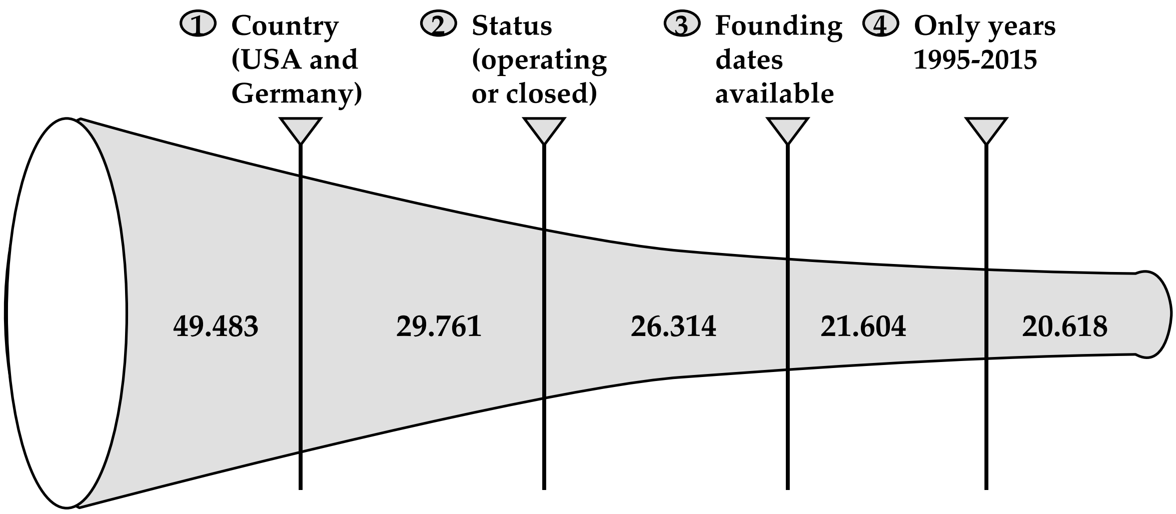
A member of the CrunchBase community can access and download excerpts of the collected data in XLSX-format. This excerpt includes information such as:

- Name and URL of startups
- Category and market information
- Total funding sums, number, and dates of funding rounds

- Operating status (acquired, closed, operating)
- Country and city of origin
- Names of investors

The starting point is a CrunchBase excerpt from December 2014 with data on 49.438 startups in total. With filter criteria as country, i.e., the USA and Germany, the operating status closed (failed) and operating, the availability of founding dates, and the timeline 1995-2015, we can narrow the list down for our purposes (see Figure 4.1). This results in a total number of 1.186 closed startups, 1.150 from the US and 36 from Germany. Closure dates are not provided by the CrunchBase excerpt, but available in part on the CrunchBase website. The closing dates are, therefore, looked up manually online via search engines. Due to the low number of closed German startups, the dataset is enriched with listed offline startups from Deutsche-Startups.de on November 3, 2015 (the day we started downloading press articles). This list includes 176 closed German startups, plus additional information such as founding year, founders, CEO, the name of the investor, head office and business model. Again, closing dates are looked up manually online.

FIGURE 4.1: The applied filter logic for the CrunchBase dataset.

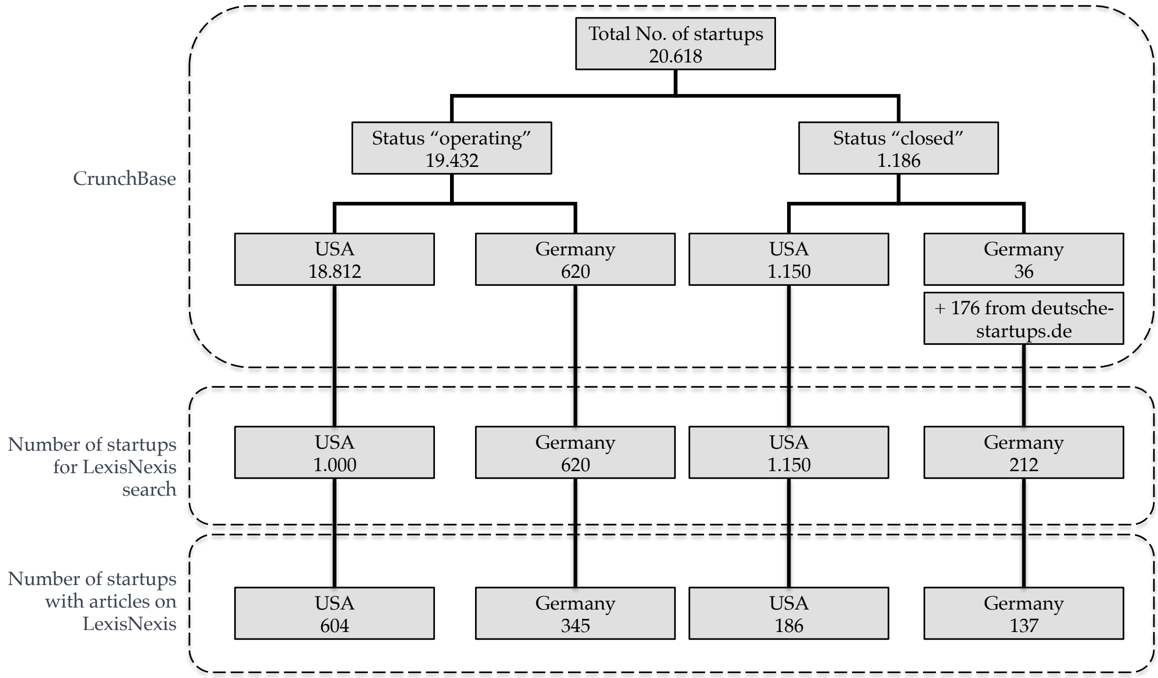


The resulting set of data includes for 1326 closed startups (1150 from the USA, 212 from Germany).

To compare failed to operating startups, the same logic as above is applied, in order to gather a dataset of operating startups. The CrunchBase excerpt includes 19.432 operating startups from the respective countries, 18.812 from the US, and 620 from Germany. In a typical research design of a so-called treatment group (i.e., all closed startups) and the corresponding control group (i.e., all operating startups), the control group is about two to three times the size of the treatment group. This ratio leads to a much better precision in estimates and tests in comparison to one to one assignment (Woodward, 2013, p.226) .

A test with 100 startups and the LexisNexis database, the chosen source for press articles (see Chapter 4.2), shows, articles exist for roughly 50% of the operating, and 25% of the failed startups. Hence, out of the 18.812 US operating startups, 1.000 startups are randomly selected (to reach roughly 500 with articles), while all 620 German startups are included (to reach approximately 300). An overview of the resulting number of startups is given in Figure 4.2.

FIGURE 4.2: Number of startups per operating status and country included in the dataset.



4.2 Selection and preprocessing of articles

Building on the selection of startups from the previous chapter, the next step is to obtain the corresponding media articles. There are mainly two providers for media articles of the desired topic and timeline; one is 1999 founded Dow Jones & Company owned Factiva, the other 1977 founded LexisNexis Group belonging to RELX Group. LexisNexus has a slightly higher number of sources available (36.000), compared to Factiva (32.000), and is therefore the source of choice. The database allows to select specifically major newspapers (including online news portals) for the US and Germany respectively. A total number of 160.668 articles about our startups in scope are internet sourced.

The preprocessing of the media articles can be divided into three main steps, which transform the data from the original MS-Word document format into a panel data structure in CSV-format.

The files include not merely the media article itself, but also additional information, such as date of publication, the name of the media source and logo, title and length of the article in words, as well as pictures and tables (see Figure 4.3).

FIGURE 4.3: Extract from a typical downloaded article with source, title, author, category and length.

Dokument 4 von 50

Hamburger Abendblatt
www.abendblatt.de

abendblatt.de - Hamburger Abendblatt Online

Montag 30. August 2010 11:49 AM GMT

**Exzellenzserie - Teil 7;
Web-TV unter der Lupe**

AUTOR: Marc Hasse

RUBRIK: RATGEBER

LÄNGE: 1126 Wörter

HIGHLIGHT: Medienwissenschaftlerin Johanna Leuschen erforscht Internetfernsehen

Auf ihre Homepage hat Johanna Leuschen zwei Fotos gestellt. Das eine zeigt sie mit schwarzer Baskenmütze und senffarbenem Kurzmantel, Motiv Modeprospekt. Darunter das andere Foto: ein halbdunkler Gang, sie - Sportjacke, Jeans, Turnschuhe - lässig an einer Mauer lehnend, auf der Graffiti prangt. So lassen sich gerne Musiker ablichten, die möglichst cool aussehen wollen. Musikerin ist Leuschen zwar nicht, erfährt der Besucher, aber sie arbeitet als Moderatorin, die Musiker interviewt, als Producerin und Redakteurin. Außerdem betreibt sie ihre eigene Internetsendung. Und als sei das nicht genug, arbeitet sie auch noch an ihrem neuesten Projekt, das alles miteinander vereint: einer Doktorarbeit über Web-TV - Internetfernsehen.

The first step is to remove all images and tables in the articles from the downloaded MS-Word documents, since they add no information to the subsequent sentiment analysis. This is achieved by applying a VBA-macro onto the MS-Word documents. The macro opens each document, scans it for pictures and tables and deletes them. After it has scanned a document, it saves it and opens the next document in the folder. The macro code can be found in Appendix A.

The second step is to transfer the data into an easily processable panel data structure. For this purpose, a second macro is applied (see Appendix B). The macro reads document by document, including all auxiliary information, into an XLSX-file. Table 4.1 shows the resulting information.

TABLE 4.1: Overview of the panel data structure.

Parameter	Explanation
Article Index Number	Running index number for all articles
Document name	Name of the Microsoft Word document
File name	Name and number of the article within the Microsoft Word document
Company name	Name of the startup
Date	Date of publication of the article
Source	Name of the publisher
Title	Title of the article
Article	The full article itself

In a third step, duplicate articles are removed. LexisNexis allows selecting a variety of different media sources. As a result, some articles can be found more than once in the dataset. They have been uploaded multiple times due to e.g., typing or spelling errors. For duplicate articles, title, length and the article itself are the same, only the published date or time differs. In order to not double count these articles, they have to be removed from the dataset, with a simple script, written here in the Python programming language (see Appendix C for the code). The duplicate articles have the same startup name, media source, title, and length. The script searches for duplicate entries of these parameters and removes them from the dataset. The resulting dataset may include the same articles twice, but only from the LexisNexis search for two different startups, which is the desired outcome.

TABLE 4.2: Dataset overview after cleaning.

	Country	No. of Startups	No. of articles	No. of sentences	Sentences / Startup
Treatment group	USA	186	11.781	14.873	80
	GER	137	11.141	13.132	96
Control group	USA	604	99.406	124.813	207
	GER	345	38.340	52.705	153
Total		1272	160.668	205.523	162

To gain insight on the research hypotheses from section 6.2, the full article that mentions the startup's name can not be the unit of analysis. Many articles report on different startups at the same time. Most of them are not fully dedicated to one startup. An extract from an article about the startup "Halotechnics" reads:

> "...I am delighted to see these first projects getting funded, which will enable California to meet the goals of its low-carbon energy future. **Halotechnics** and UCLA will receive $1.5 million each to advance thermal energy storage technologies that will help cut costs and improve the efficiency of thermal energy storage, leading to increased capacity and operational benefits for concentrated solar power plants. Also related to solar technologies..."

The article deals with the benefits for California, the UCLA, solar technology, and about the startup Halotechnics. As we are only interested in the media judgment about the startup, only sentences which mention the startup's name should be used to derive the sentiment from. To extract these sentences we use a suite of libraries for the Python programming language called Natural Language Toolkit, or NLTK (Bird, Klein, and Loper, 2009). NLTK combines a variety of instruments for natural language processing (NLP). One of the tools in NLTK can extract individual sentences from an article; this process is called tokenization. More generally, tokenization is the process of breaking up a stream of text into words, phrases, symbols, or other meaningful elements called tokens (Perkins, 2014, p. 7).

The last step is to apply a Python script to tokenize the articles. The script searches articles for the startup in scope and extracts the sentences that name the startup explicitly, resulting in a panel dataset on sentence level. The Python script can be found in Appendix D.

Table 4.2 summarizes the output of the script. In total there are 1.272 startups, with 160.668 articles and 205.523 sentences remaining after three steps cleaning procedure, which gets introduced in the next chapter. On average there are 162 sentences per startup.

TABLE 4.3: The three applied cleaning steps.

Step No.	Explanation
1	Remove all sentences with less than three and more than fifty words
2	Clean sentences from unnecessary terms
3	Manual cleaning of sentences

4.3 Cleaning of media articles

As described in Chapter 4.2, the sentiment analysis is on the sentence level. 205.523 sentences about startups remained after a three step cleaning procedure, which will be explained in the following chapter (see Table 4.3 for an overview).

In the interest of reliable research results, it has to be ensured, only to use sentences, which deal with the startup in scope and are long enough to convey a message. A sentence can convey a message with as little as a subject, a predicate, and an object. Hence, the minimum length of a sentence has to be three words, sentences with fewer words are removed. In addition, enumerations are removed, as we set the upper bar to a maximum of 50 words. The NLTK tokenization function creates sentence tokens of any length, i.e., up from one single word. And indeed, there are single words or two-word sentences in the dataset. This is explained with the following example.

The list of startups includes the German "Arago" (underlined). An article from "Mitteldeutsche Zeitung" published on July 30, 2014, shows the results of a crossword puzzle:

Waagerecht: 1. Sitar, 6. Paar, 7. Steig, 8. Sohn, 9. Art, 11. Partie, 14. Orla, 15. Kost, 17. Staude, 19. Ene, 20. Turf, 22. Oleum, 23. Ebro, 24. <u>Arago</u>.

The NLTK tokenizer recognizes the startup between two dots and extracts the sentence, even though it is just part of a list and not a sentence. There is a separation of the short sentences from the data.

As a second step, we remove roughly fifty additional terms in the sentences, such as "http://", "Dateline:", and "Language", as they add no further information for the sentiment analysis.

In a final, step we clean the remaining sentences manually. Here, we mostly deal with a single problem. As some startup names are not unique, sentences possibly refer to other objects, but the startup itself. One way of dealing with this issue is to spot those startups, which has a relatively large number of sentences compared to others. It might be, for example, that the name of the startup has multiple meanings, such that the extracted sentences can relate not merely to the startup. One example is the German startup "The Chicken", while the name *chicken* is frequently in use in many areas. As most of the resulting sentences in the dataset do not relate to this startup, we remove "The Chicken" from the dataset.

4.4 Data structure and overview

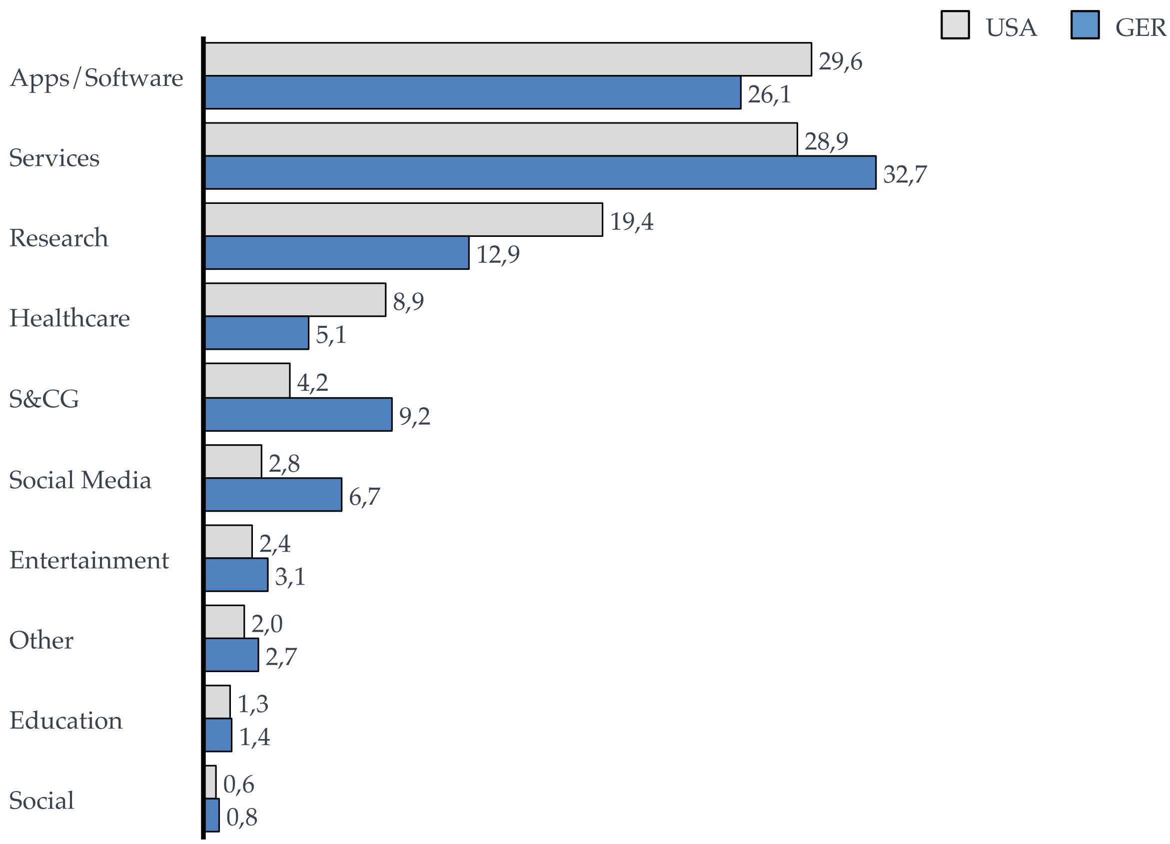
The collected dataset includes more than 200.000 sentences, and more than 30 independent, and control variables. Before one can start to analyze the dataset a number of robustness checks need to be done. Is the industry distribution of US and German startups comparable? Are the startups spread evenly across the considered time period? Are the number of articles comparable over time? To address these questions, this chapter presents some descriptive statistics, starting with an industry overview per country.

Industry distribution

Figure 4.4 shows a distribution of industries of all US and German startups in the dataset. Most startups offer a service, followed by the group of Apps/Software, and the research group. 77,9% of all US startups fall within these three categories.

In order to ensure comparability between the US and German data, Figure 4.4 also displays the distribution for Germany. One can see that the top three categories are the same, with a sum of 81,7% for Germany. Merely the order of the first two, Services and Apps/Software is different. In the US, the Healthcare group is ranked fourth, whereas in Germany this spot is occupied by Shopping and Consumer Goods (S&CG). The social media and entertainment groups for both countries are in the middle of the distribution, and the last three are other, education and social. Consequently, one can conclude that the allocation of the startups across industries is similar, which is a prerequisite for comparability.

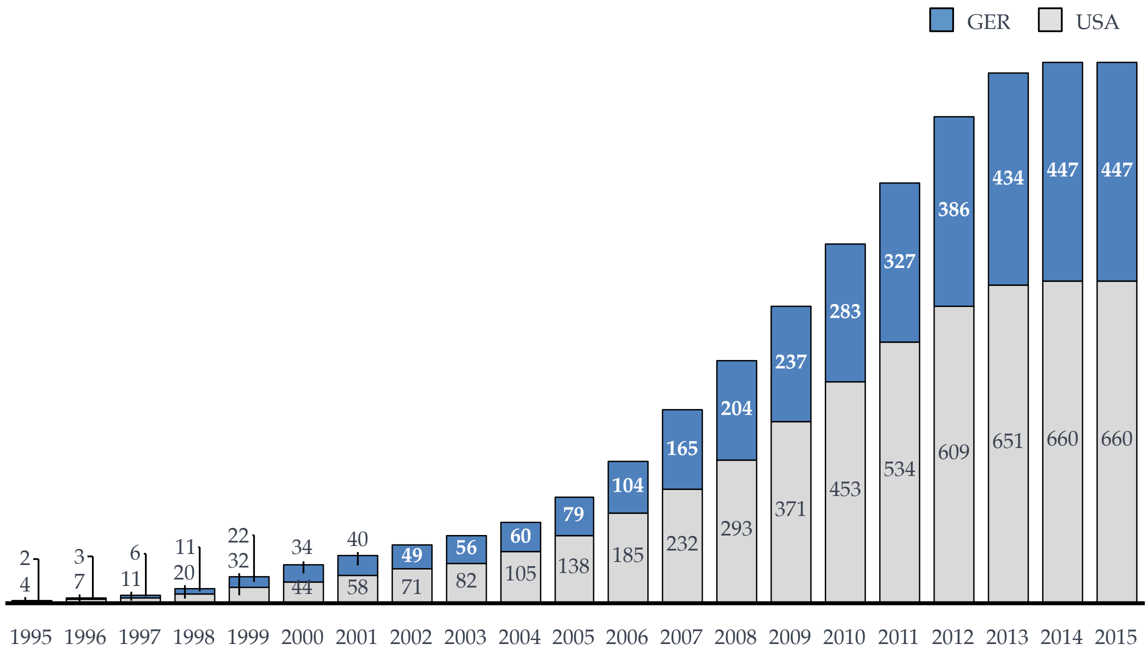
FIGURE 4.4: US and German industry distribution of startups in our dataset in percent.



Time period

A further, but less important check for robustness of our data is the number of startups in business per year and country. Figure 4.5 shows a bar-plot of the number of startups in business per country as a function of the years in focus. Starting from 1995 the number of startups in business constantly increases and flattens out after 2013. The majority of startups in our dataset thus came into existence in the last decade, which is expectable as CrunchBase was founded in 2007. Throughout the entire time frame, we have similar growth rates for both countries, which is desired to compare for temporary effects later.

FIGURE 4.5: Number of US and German startups in business per year.



Publications and sources

As described in Chapter 4.2, all articles stem from US or German news sources. Table 4.4 summarizes the number of articles, sentences about our startups and news sources associated with them. One can see that the US has about ten times the number of news sources compared to Germany, but only approximately three times as many articles. Also, as the numbers show, US media report more about German startups than vice versa, even if we take into consideration the number of startups from each country in our dataset (US: 790, GER: 482).

TABLE 4.4: Number of articles and news sources. The columns are labeled as such: country of media source - country of origin of the startup.

	USA-USA	GER-GER	USA-GER	GER-USA	Total
No. of articles	51.057	18.473	9.907	1.782	81.219
No. of sentences	120.052	56.837	25.762	2.872	205.523
No. of sources	2.620	259	569	169	3.617

Figure 4.6 displays the average number of sentences published per startup, divided by operating status. In addition, the sentences for failed startups are split into before and after failure. One can see, that reporting about operating startups far exceeds the reporting about failed startups (183 compared to 47). Moreover, most reporting about failed startups occurs before the startup fails (81.8%).

FIGURE 4.6: Average number of sentences (US press) per startup and operating status. Sentences about failed startups are in addition split into before and after the failure occurred.

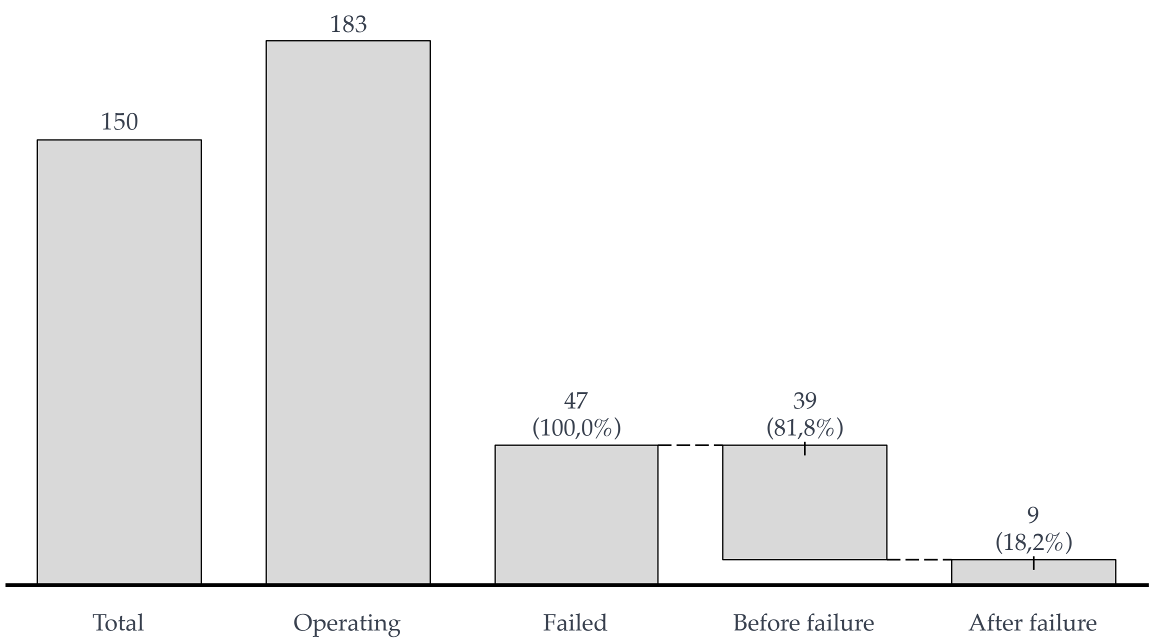
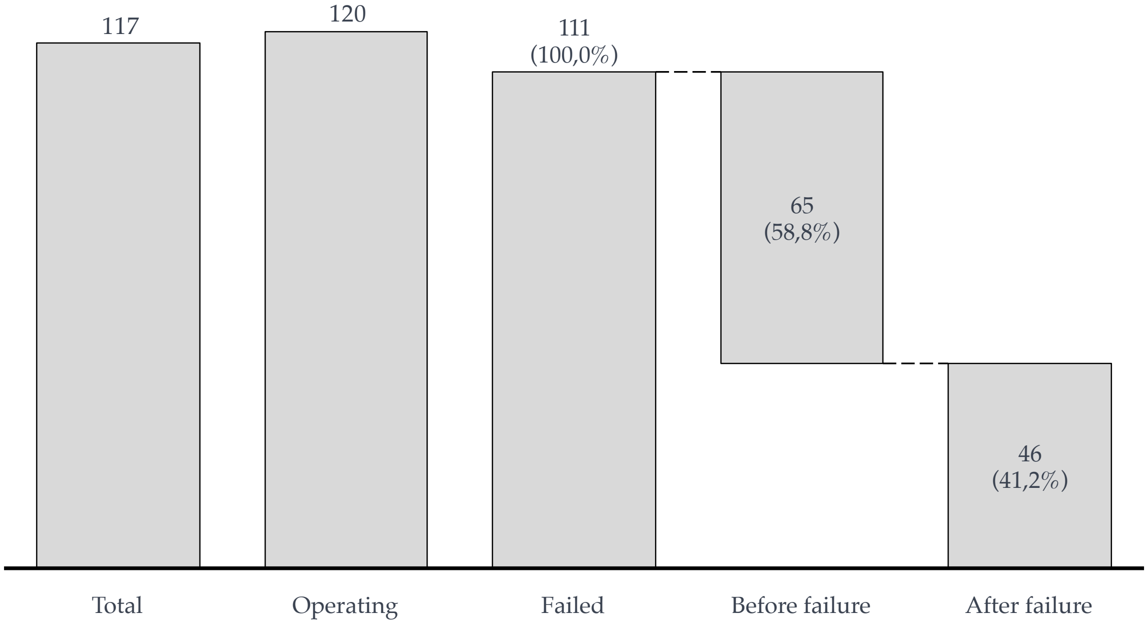


Figure 4.7 shows the same results for Germany. One can see, that German media publishes similar amounts of sentences about operating and failed startups (120 compared to 111). Interestingly, 41.2% of all reporting about failed startups occurs after failure, which is more than twice as much in percentage as in the US (41.2% compared to 18.2%), and more than five times in real numbers (46 compared to 9). These results indicate that reporting about failure is less common in the US, whereas in Germany extensive reporting takes place after failure.

FIGURE 4.7: Average number of sentences (German press) per startup and operating status. Sentences about failed startups are in addition split into before and after the failure occurred.



Chapter 5

Methodology and empirical framework

5.1 Sentiment analysis

Starting in the early 2000s, researchers began to systematically analyze the opinions expressed in documents (Pang, Lee, and Vaithyanathan, 2002), paragraphs (O'Hare et al., 2009), sentences (Seki et al., 2007), phrases (Wilson, Wiebe, and Hoffmann, 2005) and words (Hatzivassiloglou and McKeown, 1997). This is possible due to the availability of large datasets, and reasonably enough processing power to handle the data. The associated category of analysis is called sentiment analysis or opinion mining. These terms were first introduced by (Yi et al., 2003) and (Dave, Lawrence, and Pennock, 2003). A definition of sentiment analysis is given by (Liu, 2012):

> ***Sentiment analysis.*** *The field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes.*

Up to this day, sentiment analysis is applied to many different fields in business and for research. It can be automatized and is therefore seen by companies as a solid alternative to surveys, focus groups or consulting, to receive a public opinion about products (Indurkhya and Damerau, 2010). In research, sentiment analysis is used, e.g., in finance (Taylor, Schroeder, and Meyer, 2014), for stock market prediction (Van de Kauter, Marjan, Breesch, and Hoste, 2015), or product reviews (Tripathy, Agrawal, and Rath, 2016).

In general, there are three types of sentiment analysis: unsupervised learning, supervised learning and hybrid methods. The unsupervised learning methods include the so-called lexicon-based methods and fixed syntactic pattern methods. Supervised learnings methods include basically any existing learning methods, e.g., naive Bayes classification, or support vector machines (Liu, 2012, p. 31). Hybrid methods make use of both other types (Liu, 2012, p. 74). Unsupervised and supervised methods are applied in this thesis, and briefly explained in the following chapters, including the achieved accuracy results.

Sentiment analysis based on supervised learning requires a so-called training and testing dataset. As the names indicate, these datasets are used to train the algorithm and to test its accuracy. Table 5.1 lists the number of train- and test-sentences and compares it to the total number of sentences in the dataset. For both languages we have 1.2 and 1.9% of the data labeled manually, to achieve the desired accuracy level of 70% and more (Andreevskaia and Bergler, 2008). This is because humans tend to disagree in about 20% of all cases, hence, an accuracy level of 70% is nearly as good as the human judgment (Ogneva, 2010).

TABLE 5.1: Number of sentences for training and testing dataset, split by language and in relation to the total number of sentences in the dataset.

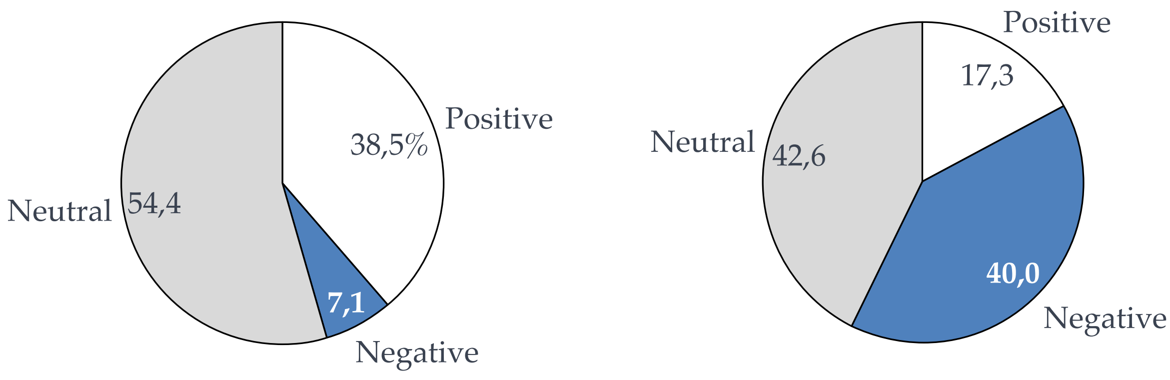
Language	Train and test sentences	Total no. of sentences	Ratio (%)
English	1705	137.945	1.2
German	1308	67.578	1.9
Total	3013	205.523	1.5

The results of manual labeling are displayed in Figure 5.1. In both countries, most of the sentences are neutral, with 54.4% in the US and 42.6% in Germany. Astonishing is, 38.5% of positive sentences in the US compare to only 17.3% in Germany, therefore the US press uses more than twice as much positive sentences compared to the German press. This becomes even more severe in the negative class. Only 7.1% of the US sentences fall into this category, compared to 40.0% in Germany. In other words, German sentences are more than five times more likely to be negative. Less present classes in the training dataset may result in a lower performance of the algorithm in that specific class (see Chapter 5.5).

5.2 Preprocessing

The sentiment analysis is performed on a sentence level. To achieve a high level of accuracy, different preprocessing techniques or tools are applied to the sentences. All of them target to reduce the overall complexity of the classification problem, as they lower the number of different words, and remove unnecessary information from the text. Table 5.2 lists the tools applied to our data and explains them.

FIGURE 5.1: US and GER training and testing data split by class.



(a) USA

(a) GER

5.3 Sentiment analysis performance evaluation

The machine learning algorithm calculates a sentiment value for each of the more than 200.000 sentences in our dataset. Such an evaluation is not 100% accurate, as it is the result of a prior training process. To make a statement on the accuracy, we must evaluate the algorithm's performance, and the classifiers. Up to this day, an ongoing debate about the correct evaluation among researchers exists (Sokolova, Japkowicz, and Szpakowicz, 2006). Hence, we chose the most suitable for this research project. All these different measures are deducted from the so-called confusion matrix. This matrix displays for a classification problem, the number of correct and incorrect assigned objects. Table 5.3 introduces the basic confusion matrix for our three-class problem.

There are only two choices possible, either the classifier correctly predicts the class or not. A correct prediction results in a TP_j value with $j \in \{p,i,n\}$, whereas TP stands for "true positive". An incorrect classification results in an error $E_{j,k}$ with $j,k \in \{p,i,n\}$. With these two different types of values, we can evaluate a classifiers performance and derive all the necessary measures.

We can generally calculate two different measure types to compare the performance of an algorithm. One is, to not focus on one specific class, but to take all into account. This measure is called accuracy. It can be defined for our three-class problem as

$$Accuracy = \frac{\sum\limits_{j} TP_j}{\sum\limits_{j} TP_j + \sum\limits_{j,k} E_{jk}}. \tag{5.1}$$

TABLE 5.2: Overview of applied preprocessing steps.

Tool	Explanation	Example
Tokenization	The process of splitting a string into a list of pieces or tokens. Tokens are pieces of a whole (e.g., a word in a sentence or a sentence in a paragraph).	“He likes chocolate.” - “He”; “likes”; “chocolate”; “.”
Stemming	A method to remove affixes from a word, leaving only the stem.	“random” or “randomly” or “randomization” - “random”
Lemmatization	Lemmatization normalizes words to their basic form by considering context and the POS of the given word, then applies rules specific to grammatical variants.	“houses” - “house”
Stop words removal	Stop words are common words that in general do not contribute to the meaning of sentences and can hence be removed.	“be” or “but” or “and”
Rare words removal	Removes rare words to reduce dimensionality and hence complexity.	"Adidas" or "me@email.com"

TABLE 5.3: Confusion matrix for three dimensions.

		Predicted class		
		Positive	Indifferent	Negative
True class	Positive	TP_p	E_{pi}	E_{pn}
	Indifferent	E_{ip}	TP_i	E_{in}
	Negative	E_{np}	E_{ni}	TP_n

Accuracy is the sum of all correctly assigned objects divided by the total number of assignments. The measure does not distinguish between the number of correct assignments per class. This is a weakness, especially in cases where one class dominates in terms of occurrences over others. If the classifier turns out to be accurate for the dominant class, but fails to predict other classes well, we would still receive a high accuracy level.

In natural language processing, algorithms often get analyzed by measures that focus on one specific class. This class is usually of special interest, and the most common measures are called precision, recall and F-measure (Sokolova, Japkowicz, and Szpakowicz, 2006). Precision can be defined as

$$Precision_k = \frac{TP_k}{TP_k + \sum_j E_{jk}}, \tag{5.2}$$

whereas the sum $\sum_j E_{jk}$ is also called the "false positive" results. Precision is thus a measure for the classifier's accuracy for one specific class. Hence, we can use it to validate our accuracy results, by calculating the precision for each class. The last new measure is called recall. It is defined as

$$Recall_j = \frac{TP_j}{TP_j + \sum_k E_{jk}}. \tag{5.3}$$

Here, the sum $\sum_k E_{jk}$ is also referred to as "false negative" results. Recall is also called sensitivity, or true positive rate, as it divides the correct assigned results by the total number of occurrences of that class. The remaining F-measure or F-score is a combination of Equation 5.2 and 5.3

$$F - Score_j = \alpha \times \frac{Precision_j \times Recall_j}{Precision_j + Recall_j}, \tag{5.4}$$

with $\alpha = \frac{\beta^2+1}{\beta}$, and β as a factor to balance the influence of precision and recall (Sokolova, Japkowicz, and Szpakowicz, 2006). For $\beta = 1$ the F-Score is balanced evenly, for $\beta < 1$ recall is more represented, and precision otherwise. As we can see from Equation 5.4, the F-Score will be between 0 and 1, where larger values correspond to a better classification quality (Özgür, Özgür, and Güngör, 2005). There are two ways to calculate the F-score. The first is called micro-averaged, the second macro-averaged F-Score (Özgür, Özgür, and Güngör, 2005).

Macro-averaged F-Score. Here, we calculate each category first and then average over the number of categories. The resulting formula can be written as

$$F - Score - Macro = \frac{\sum_j^n F - Score_j}{n}, \tag{5.5}$$

with n as the number of categories. As the macro-averaged F-Score treats every class equally, the resulting score strongly depends on the classifier's performance on the minority classes. A typical problem that goes along with minority classes, is that the classification performance is weaker, since there are less examples in the training data. Hence, the macro-averaged F-Score is a good measure to evaluate the overall performance of the classification task. Another option to calculate the F-Score is the micro-averaged approach.

Micro-averaged F-Score. In this case, the first step is to calculate the global value of precision and recall as

$$Precision_{Micro} = \frac{\sum_{k}^{n} TP_k}{\sum_{k}^{n}(TP_k + \sum_{j} E_{jk})}, \tag{5.6}$$

$$Recall_{Micro} = \frac{\sum_{j}^{n} TP_j}{\sum_{j}^{n}(TP_j + \sum_{k} E_{jk})}. \tag{5.7}$$

With 5.6 and 5.7 we can calculate the micro-averaged F-Score in the same way as in 5.5

$$F - Score - Micro = \alpha \times \frac{Precision_{Micro} \times Recall_{Micro}}{Precision_{Micro} + Recall_{Micro}}. \tag{5.8}$$

As we build the sum for all individual values in formula 5.8, this measure tends to be more influenced by the dominating categories. Our introduction to different performance measures shows, no single measure describes the classifier's performance completely. Therefore, we provide different measures in the results chapter to get a better understanding of the true performance of the classifiers. With regards to the F-Score, we choose the Macro-averaged F-Score, since it treats every class equally and, therefore, reflects the performance of our classifier best.

5.4 Unsupervised learning sentiment analysis

A lexicon-based approach classifies text or sentences based on a dictionary, consisting of words with attributed polarity and strength. The polarity determines whether a word is either positive, or negative, whereas the strength determines the degree of polarity. Each word in these dictionaries has a numeric value, e.g., continuously between -1 and +1, where -1 denotes the most negative words and +1 the most positive. These values are used to classify a whole text or phrase, e.g., via mean value. Lexicon-based dictionaries can be created manually or automatically. The words in a manually created lexicon are being hand-ranked by more than one annotator independently. When the annotators come to different results, they discuss those cases to achieve agreement. In the early days of lexicon-based sentiment analysis, researches focused on adjectives or adjective phrases as the primary source of the sentiment, verbs and nouns were added in a second step (Hatzivassiloglou and McKeown, 1997; Hu and Liu, 2004).

Another option to classify text is to use fixed syntactic patterns that are likely to be used by the author to express his or her opinion. This method was introduced by (Turney, 2002). He applies part-of-speech tagging (see also Chapter 5.2) to extract syntactic patterns from a phrase. The sentiment orientation of a phrase is calculated using a so-called pointwise mutual information (PMI) measure. According to Turney (2002), he achieves a classification accuracy between 66% and 84%, depending on the domain.

Our approach resembles a recently proposed approach by Van de Kauter, Marjan, Breesch, and Hoste (2015). In their lexicon based sentiment analysis, they calculate the resulting sentiment in four different ways. Table 5.4 summarizes them. To further improve the performance, we experimented with a suitable range for the neutral class. The best results are achieved, when the median sentiment values between -0.3 and +0.3 are considered neutral, lower values negative, and greater positive.

TABLE 5.4: Aggregation strategies for the sentence's sentiment.

Aggregation strategy	Explanation	Score calculation example
Sum	All polarity words are summed up.	-1,+2,-1 $\rightarrow$ 0
Median	The median of all polarity words defines the score.	-1,+1,+2 $\rightarrow$ +1
Highest score	Only the highest score (positively or negatively) is considered.	-2,+1,+3 $\rightarrow$ +3
Most common	The sign that occurs most often is taken.	-1,+1,+1 $\rightarrow$ +1

The first dictionary is AFINN 111 from Nielsen (2011b). It consist of 2.477 words, including 15 phrases, and scores ranging from -5 for extremely negative to +5 for extremely positive words. We normalize the scores to range from -1 to +1. The scores are manually labeled by the author according to valence, but not to sub- or objectivity and arousal or dominance. It is biased towards the negative class, as 1598 or 65% of the words belong to it. Nevertheless, Nielsen shows, the dictionary performs better than existing one's, such as ANEW, General Inquirer, or OpinionFinder (Nielsen, 2011a). Table 5.5 lists the results for the English training and test dataset. Precision, recall, a weighted macro-average F1-score (as we focus on the minority class, i.e., negative sentences) and accuracy are measured against the prior manual labeling (see also section 5.3). The results show, that the sum of the scores for all polarity words returns the best F1-scores, and accuracies. We must admit, that 54.66% F1-score and 54.08% accuracy cannot be considered satisfactory, as it means almost every second time the dictionary misclassifies a sentence. A reasonable score would start at about 70% for this measure, as shown by (Andreevskaia and Bergler, 2008). Originally the AFINN word list was developed for a micro-blog sentiment analysis (Nielsen, 2011a). Here, specific words occur more often than in regular press articles. This explains, why it is not perfectly suited to rate press articles. Another reason might be the additional neutral class in our case. Most of the lexicon based studies determine either positive or negative orientations (Taboada et al., 2011; Ding, Liu, and Yu, 2008; Mudinas, Zhang, and Levene, 2012). Additional classes significantly increase the complexity of classification.

For more reliable results, other word lists are tested. We searched for word lists specifically developed for press releases, or newspaper articles. The Lexicoder Sentiment Dictionary (LSD) is specifically designed to rate the sentiment of political texts (Young and Soroka, 2012). It consists of 4.409 words (1.625 positive (37%), 2.784 negative (63%), and is tested by Young and Soroka (2012) on New York times coverage in the fields of economy, environment, crime, and international affairs. Table 5.6 summarizes our results.

All measures are slightly above the 50% bar, except for the accuracy of most common sign in part (d), which is below. The negative class shows the lowest results. As already mentioned, 63% of the words in LSD-dictionary are negative, hence we would have expected the word list to operate best in this class.

We test a final dictionary for the English sample, developed for financial news texts. Loughran and McDonald argue, that most dictionaries misclassify common words in a financial context (Loughran and McDonald, 2011). According to them, words such as cost, tax or capital cannot generally be considered negative. Hence they propose a new word list consisting of 2.690 words (353 positive (13%), 2.337 negative (87%)), specifically targeting the financial sector. Table 5.7 summarizes the results. Almost all scores are close, or above the 60% ratio, it is the best performing word list in our sample. This is due to strong scores for the neutral

TABLE 5.5: The results of our AFINN-dictionary sentiment analysis of our US training dataset.

(a) Aggregated by sum

	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	No. of sentences
Negative	37.95	57.80	45.82		109
Neutral	65.42	47.46	55.01		845
Positive	48.75	65.22	55.79		598
Total	57.07	55.03	**54.66**	**54.08**	1552

(b) Aggregated by median

Negative	37.29	60.55	46.15		109
Neutral	62.97	45.68	52.95		845
Positive	49.25	65.89	56.37		598
Total	55.88	54.51	**53.79**	**53.58**	1552

(c) Aggregated by highest score

Negative	37.72	57.80	45.65		109
Neutral	65.11	47.93	55.21		845
Positive	48.61	64.21	55.33		598
Total	56.83	54.90	**54.59**	**53.96**	1552

(d) Aggregated by most common

Negative	36.55	48.62	41.73		109
Neutral	66.96	44.62	53.55		845
Positive	46.96	68.39	55.68		598
Total	57.12	54.06	**53.54**	**53.13**	1552

class. Similar to the previous dictionaries, the word list struggles with the negative class. Precision and F1-score are between 30 and 40%, which is expected by simple guessing.

For the English sample, we conclude that the unsupervised lexicon based sentiment analysis with existing word lists is not accurate enough for robust research in our case. With accuracy levels of 60% or less, results are below the commonly excepted and recognized level of 70% and above (Andreevskaia and Bergler, 2008). Thus, a more sophisticated approach is necessary. Therefore, supervised learning techniques, as described in the following chapter, are a potential solution. Before doing so, we test a word list with the German sample, to see whether the poor results are potentially language specific.

For the German sample, we use a word list created by the University of Leipzig (Remus, Quasthoff, and Heyer, 2010). It consists of 3.468 words (1.650 positive (48%), 1.818 negative (52%)), with additional word forms. In total, there are 15.649 positive and 15.632 negative

TABLE 5.6: The results of our Lexicoder sentiment dictionary based sentiment analysis of the US training set.

(a) Aggregated by sum

	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	No. of sentences
Negative	25.69	33.94	29.25		109
Neutral	58.96	52.54	55.57		845
Positive	46.92	53.51	50.00		598
Total	51.99	51.61	**51.57**	**50.73**	1552

(b) Aggregated by median

Negative	25.47	37.61	30.37		109
Neutral	55.91	49.82	52.69		845
Positive	49.85	56.86	53.13		598
Total	51.44	51.68	**51.29**	**50.79**	1552

(c) Aggregated by highest score

Negative	25.18	32.11	28.23		109
Neutral	58.81	53.73	56.15		845
Positive	47.16	52.68	49.76		598
Total	51.96	51.80	**51.73**	**50.92**	1552

(d) Aggregated by most common

Negative	25.47	37.61	30.37		109
Neutral	61.97	44.73	51.96		845
Positive	45.67	61.71	52.49		598
Total	53.12	50.77	**50.65**	**49.91**	1552

word forms and inflections. The results are summarized in Table 5.8.

Similar to the English sample, scores are below 60%. The F1-score even is below 50%. Differing from the results for the English sentences, the negative class is best predicted by the word list. Precision scores range up to 70% and more.

For the German sample, we can conclude that conventional lexicon based sentiment analysis does not suit our sample of sentences. Therefore, supervised machine learning method is applied and introduced in the following chapter.

TABLE 5.7: The results of our Loughran- and McDonald-dictonary based sentiment analysis of the US training set.

(a) Aggregated by sum

	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	No. of sentences
Negative	31.42	65.14	42.39		109
Neutral	64.12	76.33	69.69		845
Positive	64.84	37.63	47.62		598
Total	62.10	60.63	**59.27**	**59.59**	1552

(b) Aggregated by median

	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	No. of sentences
Negative	30.77	66.06	41.98		109
Neutral	63.62	75.74	69.15		845
Positive	67.44	39.13	49.52		598
Total	62.78	60.95	**59.68**	**59.91**	1552

(c) Aggregated by highest score

	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	No. of sentences
Negative	30.77	62.39	41.21		109
Neutral	63.98	77.16	69.96		845
Positive	65.49	37.12	47.39		598
Total	62.23	60.70	**59.24**	**59.66**	1552

(d) Aggregated by most common

	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	No. of sentences
Negative	30.77	66.06	41.98		109
Neutral	66.07	76.67	70.11		845
Positive	64.62	42.14	51.01		598
Total	63.03	61.53	**60.78**	**60.48**	1552

5.5 Supervised learning sentiment analysis

In supervised learning, a learner receives classified data to learn from, and makes predictions on unseen data (Mohri, Rostamizadeh, and Talwalkar, 2012). The most applied methods in natural language processing are naive Bayes (NB), and Support Vector Machines (SVM) (Ye, Zhang, and Law, 2009). Since both classifiers are applied to our data, we briefly explain their functionality in this paragraph. Figure 5.2 describes the six steps process of supervised machine learning (see Kotsiantis, Zaharakis, and Pintelas (2007)). In Chapter 4 and 5.2 we already covered steps 1-3. The fourth step, the definition of the training set, is part of the next paragraph.

TABLE 5.8: The results of our German sentiment-dictionary based sentiment analysis of the German training set.

(a) Aggregated by sum

	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	No. of sentences
Negative	71.00	40.73	51.76		523
Neutral	33.41	49.55	39.91		557
Positive	31.57	63.27	42.12		226
Total	48.15	48.39	**45.04**	**57.00**	1306

(b) Aggregated by median

Negative	71.00	40.73	51.76		523
Neutral	32.08	47.58	38.32		557
Positive	34.00	68.14	45.36		226
Total	48.00	48.39	**44.92**	**57.00**	1306

(c) Aggregated by highest score

Negative	68.88	45.70	54.94		523
Neutral	34.14	48.11	39.94		557
Positive	31.54	62.39	41.90		226
Total	47.60	49.6	**46.29**	**58.01**	1306

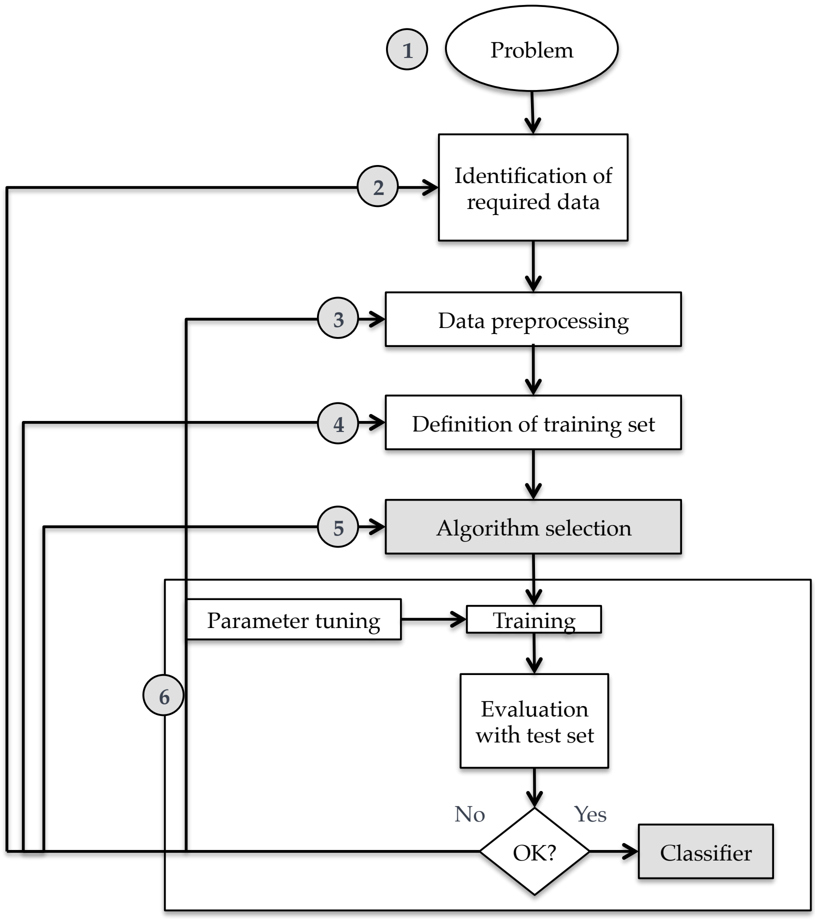
(d) Aggregated by most common

Negative	73.79	29.06	41.70		523
Neutral	34.42	47.40	39.88		557
Positive	26.40	70.80	38.46		226
Total	48.80	44.10	**40.36**	**53.45**	1306

Definition of training set

To run an algorithm on our data, we need to select data to train and test the algorithm with. For the size of these sub dataset, there exists no definite percentage value from the entire set or rule of thumb to apply. The size depends on the classification method, the complexity of the applied classifier, and the dataset in focus. For large datasets (i.e., > 100.000 observations), a common approach is to take 10% of the data for training and testing, thereof 2/3 as training data, and 1/3 for validation as testing data. This approach is called "simple split" (Olson and Delen, 2008), and assumes that the two subsets are of the same kind, i.e., have the same properties. Even if they are selected randomly, this gives rise to criticism. Hence, researchers propose to use k-fold cross validation instead. Here, the randomly selected 10% of the dataset are split into k mutually exclusive subsets, with the same size. K-1 of these subsets are used to train the classifier, and the remaining one to test the classifier. An iterative approach is chosen, until each of these k subsets is tested on. We then measure the cross validation accuracy

FIGURE 5.2: The process of supervised machine learning. Graph from Kotsiantis, Zaharakis, and Pintelas (2007).



(CVA) as

$$CVA = \frac{1}{k}\sum_{i=1}^{k} A_i, \tag{5.9}$$

with A_i to be the accuracy, as in Equation 5.1. A common value for k is 3 (Olson and Delen, 2008), as in our approach.

Algorithm selection

Though naive Bayes algorithms are considered simple (Kotsiantis, Zaharakis, and Pintelas, 2007), they can outperform more complex solutions (Hill, Lewicki, and Lewicki, 2006). The algorithm in naive Bayes is based on the Bayesian theorem, and named after the English statistician and philosopher Thomas Bayes (1701-1761). In statistics, two events can either be independent or dependent from another. The probability of independent events A and B, that both happen, can be written by multiplying the individual probabilities

$$P(A \cap B) = P(A) \times P(B). \tag{5.10}$$

An example would be to flip two coins. Each event is independent from the other. For dependent events Equation 5.10 would not be correct, as the outcome of one event depends on the other. For example, the sentence including "business failure" is more likely to be associated with a failed startup than an operating one. With Bayes theorem, we can calculate the probability of two dependent events

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(A \cap B)}{P(B)}, \tag{5.11}$$

with $P(A|B)$ and $P(B|A)$ as the probability of event A given that event B occurred, or vice versa, and P(A) and P(B) as the probability that A or B occurs. $P(A|B)$ is also called the posterior probability, as it is the likelihood of an event A to occur, when an event B has already occurred. $P(B|A)$ is called the likelihood, P(B) the marginal likelihood, and P(A) is the prior probability, as it is measured based on data from the past. To predict future outcomes with Equation 5.11, the naive Bayes machine learning algorithm estimates the probability of an object, belonging to a certain class based on prior observations. It is called "naive", since it is based on strong simplifying assumptions. For instance, it assumes all features (e.g., the words in sentences) are equally important for the classification and there occurrences are independent from another (Lantz, 2013). Despite this violation in many cases, the algorithm performs quite well (Domingos and Pazzani, 1997; Rish, 2001; Narayanan, Arora, and Bhatia, 2013). There are plenty discussions among researchers to explain this phenomenon. Some argue that for a correct classification it is of no difference if the classifier assigns an object with 51% or 99% to a class (Domingos and Pazzani, 1997).

The calculation of all different probabilities in Equation 5.11 is very difficult, especially for large training datasets. Therefore, naive Bayes algorithms take a simpler approach. Naive Bayes assumes independence among events, in particular, it assumes class-conditional independence (Lantz, 2013). This means, as long as we consider events of the same class, we can regard them as independent and apply formula 5.10. To make this more tangible, we present an example. The algorithm builds a frequency table as it is trained. This table is created with the help from the training dataset. To visualize what the algorithm does, Table 5.9 shows such a table for our example. Three words are selected (innovation, product, and failure) and their frequencies are displayed across our dimensions (positive, neutral, and negative). As the algorithm gets trained, the number of words and frequencies increases. If we then want to classify a sentence, we simply apply Equation 5.11 and add the class-conditional independence assumption by also considering Equation 5.10.

For simplicity, we calculate the probability of a sentence to be positive and includes the words innovation and product, but not failure. Table 5.9 summarizes the individual occurences of the features F1 to F3. In the Bayesian notation this means

$$P(Pos|F1 \cap F2 \cap \neg F3) = \frac{P(F_1|Pos)P(F_2|Pos)P(\neg F_3|Pos)P(Pos)}{P(F1)P(F2)\neg P(F3)}. \tag{5.12}$$

Equation 5.12 returns the likelihood. To make it a probability, it has to be divided by the likelihoods of the neutral and negative sentences

$$PB(Pos|F1 \cap F2 \cap \neg F3) = \frac{P(Pos|F1 \cap F2 \cap \neg F3)}{\sum\limits_{i=1}^{3} P(C_i|F1 \cap F2 \cap \neg F3)}. \tag{5.13}$$

Then, we can calculate the probability using data from the frequency Table 5.9.

TABLE 5.9: Frequency table of our classification example.

	Innovation (F1)		Product (F2)		Failure (F3)		
Likelihood	Yes	No	Yes	No	Yes	No	Total
Pos	$5/20$	$15/20$	$4/20$	$16/20$	$1/20$	$19/20$	20
Neu	$3/40$	$37/40$	$12/40$	$28/40$	$4/40$	$36/40$	40
Neg	$1/40$	$39/40$	$6/40$	$34/40$	$9/40$	$31/40$	40
Total	$9/100$	$91/100$	$22/100$	$78/100$	$13/100$	$87/100$	100

For our example sentence we obtain

$$PB(Pos|F1 \cap F2 \cap \neg F3) = \frac{0.048}{0.048 + 0.020 + 0.003} = 0.672. \tag{5.14}$$

The probability of such a sentence to be positive is 67.2%. As the probability sums up to 1 or 100%, it is also the most likeliest class, hence it will be assigned with a positive sentiment value of 1. This procedure is being repeated by the algorithm, until all sentences are classified. For any given case, we can generalize Equation 5.13, and obtain:

$$PB(C_i|F_1,...,F_n) = \frac{P(C_i)\prod_{k=1}^{n} P(C_i|F_k)}{\sum_{j=1}^{k}\left[P(C_j|F_1,...,F_n) \times P(F_j)\right]}. \tag{5.15}$$

With C_i indicating class i and features F_1 to F_n. Table 5.10 summarizes the strengths and weaknesses of the algorithm.

TABLE 5.10: Strengths and weaknesses of naive Bayes algorithm, adapted from (Lantz, 2013, p. 95).

Strengths	Weaknesses
Simple, fast, and very effective.	Relies on often-faulty assumption of equally important and independent features.
Does well with noisy and missing data.	Not ideal for datasets with large numbers of numeric features.
Requires relatively few examples for training, but also works well with very large numbers of examples.	Estimated probabilities are less reliable than the predicted classes.
Easy to obtain the estimated probability for a prediction.	

The second algorithm is Support Vector Machines (SVM). SVM have rather recently gained popularity, even though their mathematical fundamentals are already known for decades (Lantz, 2013, p. 225). This is due to the algorithm's performance, which several researchers describe as potentially most accurate for text classification (Liu, 2007), as well as the integration into many programming languages. We briefly introduce the fundamental mechanisms of SVM and refer to Vapnik (1998), Campbell, Cristianini, and Shawe-Taylor (1999), and Burges (1998) for further reading.

SVM are called a linear learning system (Liu, 2007, p. 113). They classify data into two classes, with an output value $c_i \in \{1, -1\}$. Since our sentiment analysis has an additional neutral class, we consider sentiment values for sentences by the classifier between -0,3 and +0,3 as neutral. Let's assume a training dataset $\mathbf{T}$ to be of the following form:

$$\mathbf{T} = \{(\mathbf{x}_1, c_1), (\mathbf{x}_2, c_2), ..., (\mathbf{x}_n, c_n)\}. \tag{5.16}$$

T is an n-dimensional vector, composed of tuples of a k-dimensional vector called the input vector $\mathbf{x}_i = (x_{i1}, x_{i2}, ..., x_{ik})$, with $\mathbf{x}_i \in \mathscr{H} \subseteq \mathbb{R}^n$. SVM define the class for each sentence by finding a decision function of the form:

$$f(\mathbf{x}_i) = \langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b. \tag{5.17}$$

Where $\langle \mathbf{w} \cdot \mathbf{x}_i \rangle$ is the dot product in our Euclidean space $\mathscr{H}$, and $b \in \mathbb{R}$ is a scalar. Equation 5.17 shows, each input vector gets a real number assigned. Hence, it allows us to come up with a classification criteria:

$$c_i = \begin{cases} 1 & if \langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b \geq 0.3 \\ 0 & if \langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b < 0.3 \text{ and} > -0.3 \\ -1 & if \langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b \leq -0.3. \end{cases} \tag{5.18}$$

To visualize Equation 5.17 and 5.18, Figure 5.3 shows an example. In a two dimensional space, the hyperplane is a straight line. It divides the space into two parts, i.e., one for each class. SVM now aim to position the hyperplane such that the distance to all data points is maximized. This means, we must maximize the distance of the two hyperplanes $\langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b = 1$ and $\langle \mathbf{w} \cdot \mathbf{x}_j \rangle + b = -1$. With some vector geometry, it can be shown, that the distance d, which is often called margin, can be written as

$$d = \frac{2}{||\mathbf{w}||}. \tag{5.19}$$

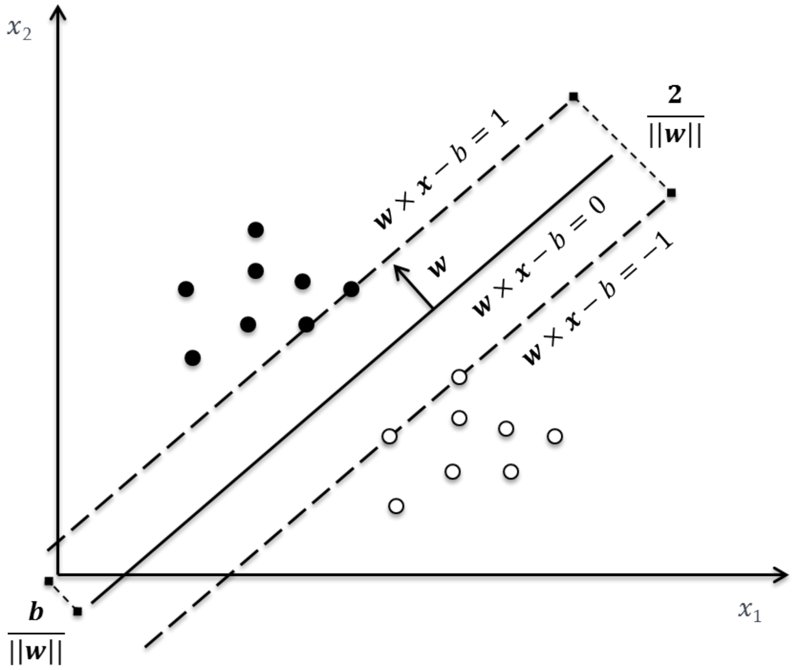
Here, $||\mathbf{w}||$ stands for the norm of $\mathbf{w}$. To maximize the distance d, $||\mathbf{w}||$ has to be minimized, which can be written as

$$min \frac{1}{2} ||\mathbf{w}||^2. \tag{5.20}$$

There is no definite solution to this problem, since many hyperplanes can be found to minimize the overall distance towards the data points. To solve this problem, constraints are added, which only consider those data points that are closest to the hyperplane

$$\begin{aligned} \langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b \geq 1 \quad & \text{for } c_i = 1, \\ \langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b \leq -1 \quad & \text{for } c_i = -1. \end{aligned} \tag{5.21}$$

FIGURE 5.3: Visualization of the SVM classifier in the linear separable case. The hyperplane in the middle is chosen such that it maximizes the distance between the two classes. Graph from (Liu, 2007, p. 99).



An analytical solution can be found by optimization theory tools, such as the Lagrangian multiplier method (Liu, 2007). In the example in Figure 5.3, all data points have been linearly separable. This means, a hyperplane can be drawn, which separates all elements of the two classes from another. Most real world examples involve cases where this is not possible (Kotsiantis, Zaharakis, and Pintelas, 2007), and the algorithm cannot determine a hyperplane, for example due to falsely labeled data. Figure 5.4 displays such a case.

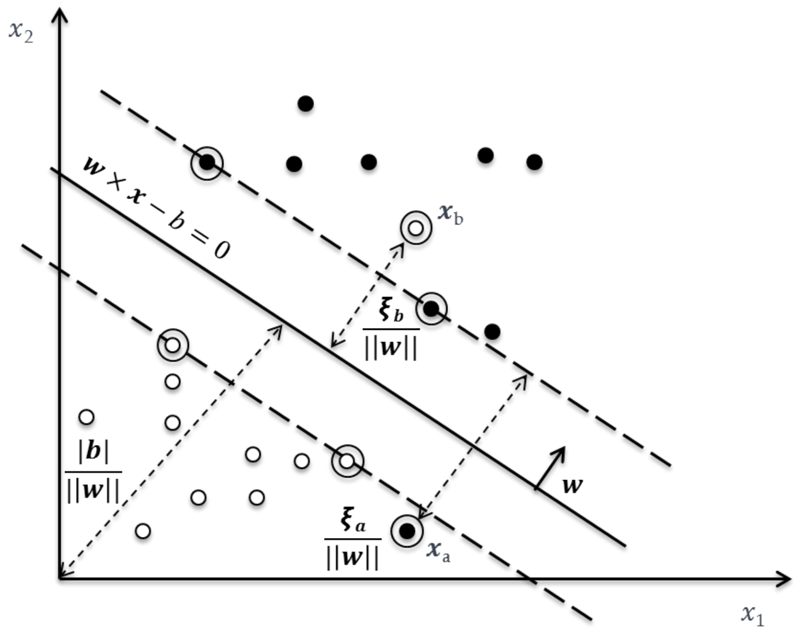
The problem can be solved by introducing so-called soft margins, which allow for errors (such as falsely labeled data)(Veropoulos, Campbell, Cristianini, et al., 1999). These errors are formally described by slack variables, and added to the constraints in Equation 5.21, we obtain

$$\begin{aligned} \langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b &\geq 1 - \xi_i \quad \text{for } c_i = 1, \\ \langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b &\leq -1 + \xi_i \quad \text{for } c_i = -1. \end{aligned} \tag{5.22}$$

With $\xi_i \geq 0$. From Equation 5.22 we can see, that an error occurs when ξ_i becomes greater than 1. Equation 5.20 then becomes

$$min \frac{1}{2} ||\mathbf{w}||^2 + V \sum_{i=1}^{n} \xi_i, \tag{5.23}$$

FIGURE 5.4: SVM in the non-linear separable case. $\boldsymbol{x}_a$ and $\boldsymbol{x}_b$ are error data points. Graph from (Liu, 2007, p. 105).



with a cost value V $\in \mathbb{R}$. The algorithm now tries to minimize the costs associated with each data point that got assigned to a false class.

Another way of dealing with non-separable data points is the so-called kernel trick. Therefore, the data gets mapped onto a higher dimensional space, the transformed feature space (Kotsiantis, Zaharakis, and Pintelas, 2007). With this technique, any consistent training set can be made separable. Therefore, we transform data from the input space $\mathscr{H}$ to the feature space $\mathscr{F}$ via a nonlinear mapping

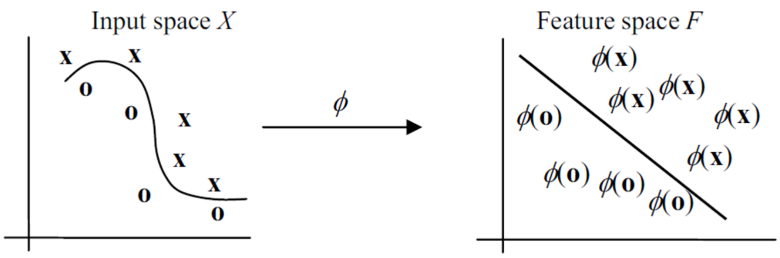
$$\begin{aligned} \phi : \mathscr{H} &\to \mathscr{F}, \\ \mathbf{x}_i &\mapsto \phi(\mathbf{x}_i). \end{aligned} \tag{5.24}$$

As a consequence, our training dataset **T** in Equation 5.16 becomes

$$\mathbf{T'} = \{((\phi(\mathbf{x})_1, c_1), (\phi(\mathbf{x}_2), c_2), ..., (\phi(\mathbf{x}_n), c_n)\}. \tag{5.25}$$

Figure 5.5 visualizes the idea of transforming into a higher-dimensional space, As a result the data becomes linear separable.

FIGURE 5.5: Transformation logic from input to feature space. Graph from (Liu, 2007, p. 109).



Evaluation of the training and the testing data

The final step before we can evaluate our data, is to test and optimize the algorithm's performance. As described in this chapter, k-fold validation is applied. Several reasons exist for a potentially unsatisfactory performance. Most common are (Kotsiantis, Zaharakis, and Pintelas, 2007):

1. There are relevant features missing for the problem,
2. The size of the training dataset is too small,
3. The problem's dimensionality is too high,
4. An inappropriate algorithm was chosen,
5. The dataset is imbalanced.

Point 1-3 relate to our input vector $\mathbf{x}_i$. It consists of features that are relevant for the classification. In our case these features are the words of sentences. We can simply add certain features to the training set to improve its performance (Marsland, 2015). This implies a good knowledge of the necessary features. The size of the training set can also be increased, by evaluating more sentences. And finally, we can reduce the dimensionality by removing rare words, stop words, lemmatization and stemming (see Chapter 5.2). If all this does not lead to a better performance, it might be due to point 4, the inappropriate algorithm.

A topic of increasing interest among natural language processing researchers is dealing with imbalance (Japkowicz and Stephen, 2002). Imbalance of a dataset is given when one class is overrepresented in comparison to other classes (Japkowicz and Stephen, 2002). The algorithm then performs well on the majority class, where lots of examples are included in the training set, but performs poorly on the minority class. There are generally three techniques to reduce the imbalance:

(a) Up-sizing the minority class (randomly or focused)

(b) Downsizing the majority class (randomly or focused)

(c) Changing the misclassifying costs of small and large classes

Point (a) results in re-sampling patterns of the minority class (either randomly or randomly close to the boundary (hyperplane) or the input space). The down-sizing of the majority class in point (b) is done by eliminating data. Again, this can be done completely random or far away from the hyperplane. Point (c) is a completely different approach. It tries to compensate the imbalance by modifying the relative cost associated to misclassifying the majority or minority class. For example, consider an imbalance ratio of 5:1 in favor of the majority class. The cost of misclassifying a minority class sentence will be set 5 times higher than for the majority class.

In Table 5.11 we present results for the SVM classifier, each language separately, before applying any of the above-mentioned techniques.

TABLE 5.11: Performance results of the SVM classifier.

	USA			GER		
Class	Precision	Recall	F1-Score	Precision	Recall	F1-Score
-1	0.611	0.367	0.458	0.655	0.727	0.689
0	0.739	0.699	0.718	0.643	0.639	0.641
+1	0.669	0.779	0.720	0.790	0.754	0.772
Average	0.698	0.698	0.693	0.696	0.706	0.701

Table 5.11 shows that the negative class in the US and neutral class in Germany display the lowest levels of F1-Score. To improve them, we applied the up-sizing strategy. For the US, we randomly selected fifty negatively rated sentences and duplicated them in the training dataset. For Germany we proceed similarly, with the neutral and negative class, and duplicated 25 sentences for each category, accepting a potential negative result for the F1-Score of our positive class. The results of this procedure are presented in Table 5.12

Almost all values exceed the 70% value in F1-Score, only the German neutral class, with 69.8% is slightly below. For the entire classification in the German language, an average of 72.4% is achieved. With this training and testing data, we now turn to the evaluation of our different classifiers and potential performance boosters. This will be part of the next paragraph.

TABLE 5.12: Performance results of the SVM classifier after up-sizing.

	USA			GER		
Class	Precision	Recall	F1-Score	Precision	Recall	F1-Score
-1	0.806	0.674	0.734	0.785	0.703	0.742
0	0.732	0.779	0.754	0.667	0.733	0.698
+1	0.735	0.722	0.728	0.741	0.736	0.738
Average	0.743	0.742	0.741	0.728	0.723	0.724

Achieved results

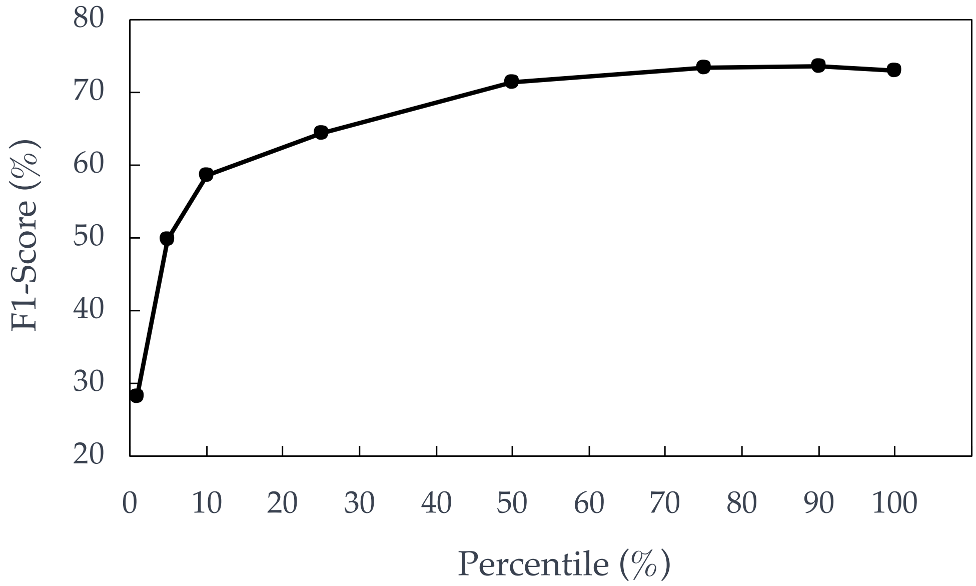
We start by introducing the achieved results of the training and testing procedure for naive Bayes. All results are averaged over the two languages in scope. Table 5.13 summarizes the F1-Score for two different classifiers.

TABLE 5.13: Achieved results with the naive Bayes algorithm.

Classifier	Preprocessing (%)	SelectPercentile (%)	SelectKBest (%)
BernoulliNB	52.8	48.8	52.8
MultinomialNB	72.1	73.3	72.1

All sentences in our training and testing dataset are preprocessed with the described steps in Chapter 5.2. In addition, we use the "feature selection" module from Scikit-learn to boost performance. Feature selection is the process of algorithmically finding features (in this case words) that best predict the sentiment orientation of a sentence (Garreta and Moncecchi, 2013, p. 84). Therefore, it makes use of an evaluation function, which returns a score for every feature. We applied two different feature selection tools, SelectPercentile, and SelectKBest. Both take advantage of a user-specified percentile, or defined number of features, based on their score. A χ^2-test is applied to test for statistical robustness. Regarding the results in Table 5.13, we can conclude, that performance does not improve by applying the feature selection in combination with the BernoulliNB algorithm, but it slightly does for the MultinominalNB, even though, merely with SelectPercentile. As for both tools one must select the percentile or number of features to include, we can iterate through, and obtain a ramp-up curve. Figure 5.6 shows a logarithmic growth of the F-Score with percentile. It reaches its maximum at about 90 %. We can conclude that some features do not improve accuracy and should be omitted.

FIGURE 5.6: Ramp-up curve for SelectPercentile tool in the naive Bayes algorithm.



As naive Bayes is by definition simpler than Support Vector Machines (SVM), we now present results the for the second algorithm.

TABLE 5.14: The achieved results for Support Vector Machines in percentage.

Classifier	Preprocessing	TD-IDF	MCO	TD-IDF + MCO
SVC	75.2	72.0	-	61.2
LinearSVC	74.9	74.2	74.5	72.2

With SVM, we have the choice of two classifiers as well. One is Support Vector Classification (SVC), the other is Linear Support Vector Classification (LinearSVC). The latter is supposed to scale better with large samples, and we except it to return better results. In this case, we try to boost performance by applying four steps:

1. Preprocessing (as with naive Bayes),
2. Term Frequency Inverse Document Frequency (TF-IDF),
3. Multi-Class Optimization (MCO),
4. TF-IDF and MCO.

Number one refers to the above-mentioned preprocessing steps. The second step, called TF-IDF, measures the importance of a single word for classification purposes. It aims to identify words or features, which are strongly associated with one category, such as the word "terrific" with the positive class, for example. It consists of two elements, the first is term frequency

$$TF = TF(n,s) = \frac{f(n,s)}{max\{f(x,s) : x \in s\}}, \tag{5.26}$$

with n as number of occurrences of a word in a sentence s. From Equation 5.26 one can see that it is a measure for relative occurrence of terms. It is normalized by the maximum number of occurrences of a word x in the sentence s. Thus we can say $0 < TF \leq 1$. The second term is Inverse Document Frequency (IDF), and defined as

$$IDF = IDF_i = log\frac{N}{k_i}, \tag{5.27}$$

with N as the number of sentences the entire corpus has, and k_i as the number of sentences, which include a specific term i.

MCO is the third tool. It stands for "Multi-Class Optimization", and aims to improve performance measures especially in cases, where more than two classes exist. In this project, we use the version developed by Crammer and Singer (2001), which appears to be most powerful for multi-class optimization problems.

As a final step, we combine TF-IDF and MCO. Table 5.15 summarizes the results. After the same preprocessing steps as with naive Bayes, SVC slightly outperforms the LinearSVC classifier. Additional tools seem not to increase the classifiers performance. Therefore, we choose to verify the results by using k-fold validation introduced in Chapter 5.5. This means, the results from Table 5.13 and 5.14 are being verified, using three different training and testing sets. We leave out BernoulliNB, since its performance lacks far behind the others.

According to the results, LinearSVC and LinearSVC + MCO achieve the best values. As they both perform in all three tests equally, we have chosen to take the simpler LinearSVC for our analyses.

TABLE 5.15: Achieved results for the Support Vector Machines algorithm.

Classifier	Type	Fold 1	Fold 2	Fold 3	∅
NB	Multinomial + SP	72.9	67.8	66.2	69.0
SVM	SVC	75.3	69.7	72.0	72.3
SVM	SVC + TF-IDF	72.9	67.8	66.0	69.0
SVM	LinearSVC	74.7	71.1	70.5	72.1
SVM	LinearSVC + TF-IDF	74.0	70.1	70.2	71.4
SVM	LinearSVC + MCO	74.7	71.1	70.5	72.1
SVM	LinearSVC + TF-IDF + MCO	73.5	71.4	70.0	71.6

5.6 Matching

One of the aims of this research project is to identify a potential effect of entrepreneurial failure on media judgment of startups (see Chapter 6.5). To do so, our goal is to obtain a dataset for analysis, that replicates as closely as possible, a randomized experiment. This means to obtain similar covariate distributions for the treated and the control groups. It can be achieved by using the popular matching methods available. The following chapters introduce the applied matching procedures and their results. Before we do so, we start with the preprocessing part.

Matching preprocessing

Recent publications from Gary King and colleagues point out a fundamental weakness of the existing matching procedures (King, Lucas, and A Nielsen, 2014; King and Nielsen, 2015). They criticize that most matching methods maximize according to one metric (such as propensity score or Mahalanobis distance), whearas a simultaneous optimization of balance and matched sample size is desired. Table 5.16 summarizes scenarios for different optimization strategies. By optimizing simultaneously towards high sample size and high balance between the treatment and control group, we achieve low model dependence and variance. The goal of matching should, therefore, be to jointly optimize the balance of a dataset and matched sample size. Here, balance can be defined as the similarity of the empirical distributions of the full set of covariates in the matched treated and control groups (Stuart, 2010). All other cases could result in either high variance and, therefore, poor reliability of results or high model dependence leaving researchers with discretion.

King et al. (2014) recommend to do two tests before matching the data. These tests can

TABLE 5.16: The effects of sample size and balance on model dependence and variance.

	Sample size small	Sample size large
Balance high	• Low model dependence • High variance	• Low model dependence • Low variance
Balance low	• High model dependence • High variance	• High model dependence • Low variance

be considered as preprocessing steps before the actual matching. They intent to measure, whether or not the effect of interest is dependent on the way data is matched and the number of observations pruned (in our case startups). As a first step we look at balance. Therefore, we define a distance measure. Since distance measures can be used for various purposes, many different one's exist (see Stuart (2010) for an overivew). Here, we focus on an approach by King, Lucas, and A Nielsen (2014), and introduce the L1 distance measure, as it is discrete, just as our sentiment classification

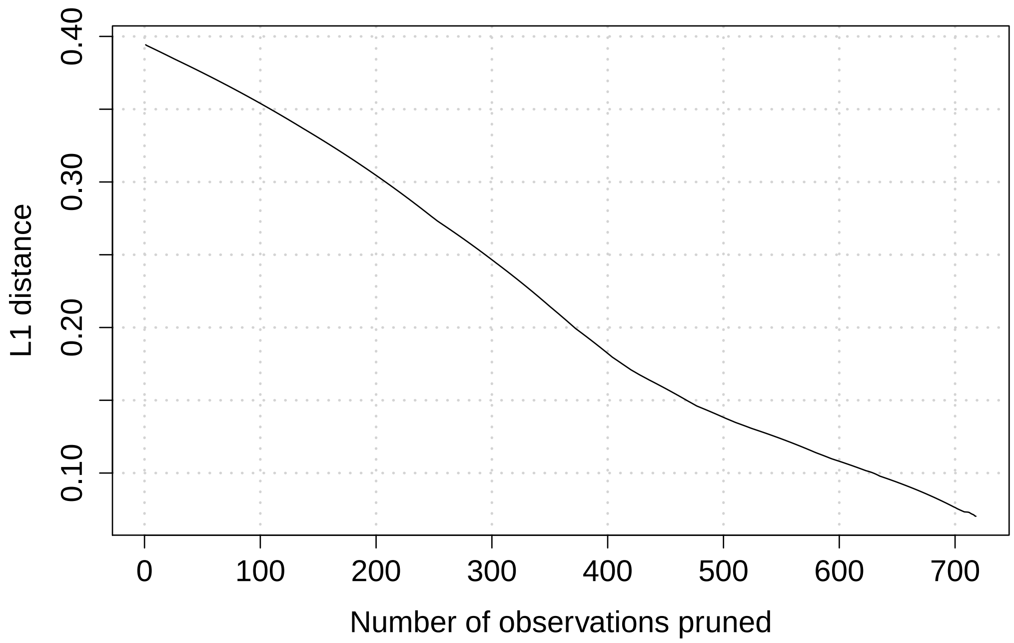
$$L1(H) = \frac{1}{2} \sum_{(l_1 \dots l_k) \in H} \left| f_{l_1 \dots l_k} - g_{l_1 \dots l_k} \right| . \tag{5.28}$$

L1 can be understood as the difference between the relative empirical frequencies of treated and control units, in a bin for each of the variables $l_1 \dots l_k$. It is thus a measure for the distances of covariates of the treated and control set. A close distance means both sets are similar. For further reading, we refer to King, Lucas, and A Nielsen (2014).

Figure 5.7 shows the L1 distance measure as a function of pruned observations. We can see, as observations are being pruned the distance between the resulting startups sample becomes smaller, and so does the imbalance.

While a greater balance is desired, variance might increase at the same time. Thus, in a second step, we look at the model dependence. Figure 5.8 shows the effect of failure on media coverage of startups as a function of pruned observations. The effect stays negative through the entire pruning process. The displayed width of the curve results from different matching procedures which are used to estimate the effect. From Figure 5.8, we can conclude that the sign of the outcome (i.e. the treatment effect) is not influenced by our sample size. In other words, keeping as many startups as possible after the matching process in our dataset will not lead to a different effect than pruning a certain number of them. This allows us to match the data without any constraints,

FIGURE 5.7: L1-distance as a function of pruned observations. As startups are pruned from the dataset, the average L1-distance of the remaining one's decreases.



while at the same time achieving a high balance. We can proceed with the matching procedure.

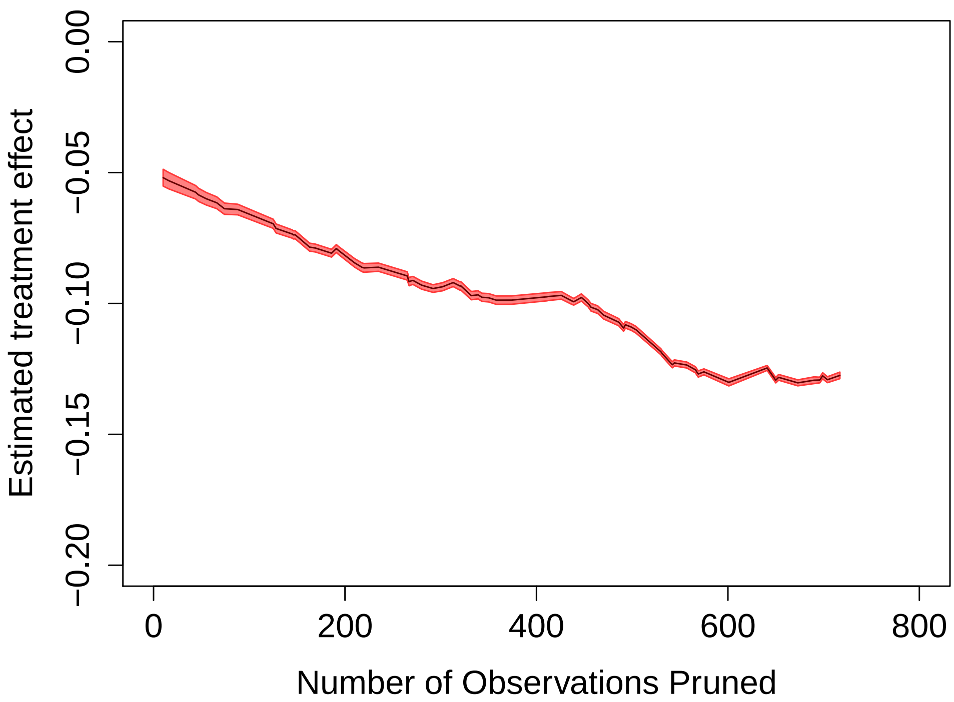
Matching procedure

In a randomized experiment, a coin is flipped to decide whether or not an object is assigned to the control or treatment group. Such an experiment is mostly immune to a potential selection effect, which means that the results could be due to the construction of the sample or the choice of model applied rather then due to a true effect. To reduce this risk, matching methods are applied. In general, matching can be defined as (Rubin, 1973):

> ***Matching.*** *A method of data organization to reduce bias and increase precision in observational data.*

It is used in studies, where a random assignment of a treatment to the subject of interest is not given (Rubin, 1973). These studies are also called observational studies. In the following, we introduce the matching procedure along three main steps:

FIGURE 5.8: Estimated treatment effect as a function of pruned observations. Pruning observations increases the average negative treatment effect.



Step 1 - Ignorability assumption

In order to choose the covariates for our matching process, we make use of the so-called "ignorability assumption". Rosenbaum and Rubin (1983) introduced this concept, which assumes the assignment to the treatment or control group to be effectively random, i.e., unobserved covariates or missing data can be ignored. It can be formally written with the observed covariate vector $\vec{X}$,

$$(Y(0), Y(1)) \perp T | \vec{X} = \vec{x} \; for \; all \; \vec{x}. \tag{5.29}$$

Here Y(i) with $i \in \{0,1\}$ denotes the observed Y(1) or unobserved Y(0) potential outcomes of the treatment T. It assumes that no unobserved differences between the treatment and control groups exist, conditional on the observed covariates.

Step 2 - Implement a matching method

Matching methods can be divided into two groups. There are fixed ratio, i.e., 1-to k (for integer k $\geq$1), or l-to m (for integer l $\geq$1 and m $\geq$1), and variable ratio methods, where the ratio can vary between the individual groups of treated and control units. Since matching can be done in several ways, Table 5.17 summarizes the most important matching types, which are used for this thesis.

TABLE 5.17: Applied matching methods in this thesis.

Method	Description	Ratio type
Exact	Matches treated with control units, which have the same covariates.	variable
Nearest	Matches to the closest in terms of distance measure.	fixed
Optimal	Matches by minimizing the average absolute distance across all matched pairs.	fixed
Genetic	Uses intense genetic search algorithm, to find a set of weights for each covariate, such that the optimal balance is achieved.	variable
Coarsened Exact	Matches on a covariate while maintaining the balance of other covariates.	variable

We did the calculation for all different methods, the quality measure gets introduced in step 3.

Step 3 - Assess the quality of the matching results and iterate

The next step is to assess the quality of the matching results. In the matching environment, quality is derived from the balance of the covariates. Here, balance is associated with the similarity of the empirical distribution of the covariates used to match treatment and control groups. This measure differs from the aforementioned L1-distance measure for the pre-matching tests. In this case we make use of the MatchIt package in R-Studio, which build on suggestions from Ho et al. (2004).

The more similar the covariates distributions are, the lower is the mean difference measure, defined as

$$Mean_{Diff} = \frac{\mu(\vec{X}|T=1) - \mu(\vec{X}|T=0)}{\sigma(\vec{x}|T=0)}, \tag{5.30}$$

with the covariates distribution function $\mu(\vec{X}|T=i) = \frac{1}{n_i}\sum_{T=i} X_j$ and $i \in \{0,1\}$, and the standard deviation of the control group

$$\sigma_{(\vec{x}|T=0)} = \sqrt{\frac{\sum_{i\in\{i:T_i=0\}}(X_i - \mu(\vec{X}|T=0)^2}{n_0 - 1}}. \tag{5.31}$$

The results of the matching process are part of the next paragraph.

Matching results

The key performance indicators are the $Mean_{Diff}$ measure from Equation 5.30, and the number of unmatched startups. Both measures should be as low as possible, to achieve high similarity, and thus comparability, as well as a large sample size. Table 5.18 summarizes results of the second measure found for each matching method.

TABLE 5.18: Results of the matching process

	Treated		Control		Total
Method	Matched	Unmatched	Matched	Unmatched	No.
Exact	249	74	421	528	1272
Nearest Neighbor	323	0	646	303	1272
Optimal	323	0	646	303	1272
Genetic	323	0	462	487	1272
Coarsened Exact	312	11	512	437	1272

The nearest neighbor and optimal methods, have the lowest number of unmatched startups, followed by the genetic method and, therefore, the highest number of matches. This appears intuitively right, as the two other methods are more stringent, since they try to achieve exact matching pairs, with variable ratio.

Table 5.19 summarizes the results for the mean difference measure (see eq. 5.30). Here, exact matching makes obviously no difference between the covariates of treatment and control group, since only perfect matches are considered. Genetic and Coarsened exact matching show a small difference from -0.001, and match therefore better, compared to nearest neighbor and optimal matching. The better mean difference measure for genetic and coarsened exact,

TABLE 5.19: Matching summary for each matching method.

Method	Mean difference	Matching ratio
Exact	-	variable
Nearest	0.1209	1:2
Optimal	0.1209	1:2
Genetic	-0.001	variable
Coarsened Exact	-0.001	variable

are the results of pruned observations (startups) from the dataset. In Chapter 5.6 we show, that the treatment effect is independent of the number of pruned observations and therefore does not change its algebraic sign. In order to make a final decision on which matching method to choose, and therefore, which startups to include for further analysis, we can base the decision only on the lowest number of unmatched startups. This means we choose the nearest neighbor and optimal matching methods. With a ratio of 1:2 the dataset includes 323 treated startups and 646 untreated startups.

Chapter 6

Study I - Media judgment and cultural differences of failed startups

The first study of this thesis addresses the question of how the media judges failure of startups in the US and Germany. As we have seen in Chapter 3.1, the media substantially influence the people's perception, opinion, and judgment. We begin with a short introduction, followed by a theoretical overview on media judgment of failure, and a summary of our data and methodology. Finally, we present and discuss our results.

6.1 Introduction

Cultural values and beliefs are considered key for a nation's entrepreneurial activity (Singer, Amorós, and Moska, 2014). While some societies are considered entrepreneurial friendly (e.g., the US), others are considered less supportive (e.g., the German). So far, researchers have studied the entrepreneurship culture mostly qualitative via surveys or interviews (Singer, Amorós, and Moska, 2014). In this study, we use media reporting as a surrogate for entrepreneurial friendliness. Previous research has shown that media reporting shapes the public's agenda, influences their beliefs and, therefore, the culture of a society (Ghanem, 1997). With a focus on media reporting, especially after a startup has failed, we compare the sentiment of media reporting between the US and Germany, as it reflects the entrepreneurial friendliness of the two societies (Hindle and Klyver, 2007). Due to limited research in this field, we follow a call from Hindle and Klyver (2007) for further investigations. As exemplary societies to study, we select the US and Germany. According to the "Global Entrepreneurship Monitor", they are expected to deviate significantly in dealing with failure (Reynolds et al., 2000). German culture emphasizes risk avoidance and displays low levels of tolerance for failure, which is of a less significance in building a fruitful ground for entrepreneurship (Landier, 2005). On the contrary, the US culture appears to be a role model for entrepreneurial countries, as it is considered risk tolerant, and in surveys, entrepreneurs express low levels of fear of failure (Singer, Amorós, and Moska, 2014).

Several lines of evidence have highlighted the importance and benefits of entrepreneurship to employment, innovation, productivity increase, and income growth (e.g., Carree and Thurik, 2010; Shane, 1996; Lee et al., 2011). Hence, governments seek to support entrepreneurship by starting various initiatives (e.g., Entrepreneurship für Deutschland in Germany, or Partnering to Accelerate Entrepreneurship Initiative (PACE) in the US). Still, vast differences in entrepreneurial activity exist between countries (Singer, Amorós, and Moska, 2014). One of the key determinants mentioned by researchers in this context is national culture. Culture is a set of shared values, beliefs, and expected behavior (e.g., Herbig, 1994; Lonner, Berry, and Hofstede, 1980). It influences people's tolerance of new ideas and inquisitiveness (Wallace and Fogelson, 1961) and is, therefore, able to foster or discourage entrepreneurial activities. Culture gets influenced by media, as it is capable of stimulating fear, impacting the images on the social construction of reality, or promoting stereotypes (McGuire, 1986). Hindle and Klyver (2007) show, that media coverage plays an important role in entrepreneurship, as the volume of media reporting is proportional to the number of entrepreneurs in a country, and can strongly influence a firm's performance (Wartick, 1992). Still, Hindle and Klyver (2007) call for further investigations on how media affects entrepreneurial activity. To the best of our knowledge, no quantitative study exists, which investigates the relationship between media coverage and success of startups. Therefore, we build a panel dataset of more than 1.000 startups from the US and Germany. This study addresses the question of how media judges entrepreneurial failure and compares the two countries in scope.

In the following chapter we introduce the underlying theory and derive three research hypotheses. We subsequently present our panel dataset, and how it is structured. In the methodology and empirical framework chapter, sentiment analysis, startup matching, and our analysis approach are discussed. Finally, we present results, discuss them, and wrap up our findings.

6.2 Theory and hypotheses

The research field of entrepreneurship has gained rather recently popularity with the emergence of innovative startup firms, and the associated relevance of these startups for economies (Carree and Thurik, 2010). However, Mueller and Thomas (2001) argue that policy makers need to base their decisions on how to promote entrepreneurship in part on assumptions based on success stories or anecdotes. From a research point of view, the national culture is considered vital for entrepreneurial activity (see 6.1), hence researchers study its impact by following the cultural dimensions theory of Lonner, Berry, and Hofstede (1980). Some of them argue that individualism and masculinity can be positively associated with entrepreneurship. High levels of uncertainty avoidance and power distance have a negative impact on entrepreneurship (Johnson and Lenartowicz, 1998; Lee and Peterson, 2001; McGrath et al., 1992; Lee and Peterson, 2001; Shane, 1992). Nakata and Sivakumar (1996) suggest a more complicated relationship, arguing that entrepreneurial activity depends on the individual case. However, the consensus amongst researchers exists regarding culture. Cultural values influence entrepreneurial activity in a country (e.g., Hayton, George, and Zahra, 2002; Wennberg, Pathak, and Autio, 2013; Johnson and Lenartowicz, 1998). Especially Western cultural values appear to support entrepreneurship. Here are two main explanations of this circumstance: a) the findings are skewed due to cultural biases in the Western-dominated field of research and b) concepts of new entrepreneurship are mainly developed in western societies (e.g., Davis and McClelland (1962); Schumpeter (1934)). To account for cultural differences in entrepreneurship, other concepts need consideration as well. Depending on the cultural values of society, reputation and status could be better predictors than Hofstede's dimensions. Begley and Tan (2001) studied the influence of social status of entrepreneurs and found it to be a better predictor of entrepreneurial activity in East Asia, than in Western countries. The researchers concluded, that improving the social status of entrepreneurs in the eye of the general public in East Asian countries to be the key driver to increase rates of entrepreneurs. A potential way to achieve this could be through active support from mass media.

The extensive entrepreneurship literature has neglected the influence of media and failed to measure a culture's endorsement of entrepreneurship, only assumptions that the US is more entrepreneurial friendly than Germany, can be made from existing surveys Singer, Amorós, and Moska (2014). Hindle and Klyver (2007), who investigate survey data from the Global Entrepreneurship Monitor, find no significant influence of successful entrepreneurship stories in mass media on nascent entrepreneurship rates or on the number of startups. Instead, the volume of media stories about entrepreneurship correlated with the volume of people managing a young business. One has to leave the entrepreneurship literature to find further indications of how mass media influences the field. Henderson and Robertson (1999) revealed within the group of young adults, that poor knowledge of actual entrepreneurs exists, which results from unfavorable media reporting.

Our focus lies on failed startups, and how media judges them prior to, in the event of, and after failure. In addition, we investigate if media judgment varies between the US and Germany. Moreover, we seek to extend the current understanding of organizational theory, especially social judgment of organization theory. The theory primarily builds on three pillars, legitimacy, reputation, and status, which can be influenced by media (Bitektine, 2011). Here we follow an approach by Wry, Deephouse, and McNamara (2006), and Deephouse (1996), who use media judgment as a surrogate for legitimacy. The media performs social evaluations by granting or withholding legitimacy. Thus, the media can be seen as a mechanism of social control, in terms of society's cultural norms and values (Parsons, 2013). Reputation and status are both results of social evaluations (George et al., 2016). For startups, it is more difficult to actively manage reputation and status, in comparison to large cooperations, which often are equipped with so-called communication departments (Tihanyi, Graffin, and George, 2014). Hence, a favorable media judgment to foster good reputation and status is harder to achieve for startups. However, previous research shows, to become a successful venture, good reputation and status are highly relevant (Bitektine, 2011). The lack of previous studies, made us believe it is a field worth investigating. In the next paragraphs to follow, we derive our research hypotheses.

Numerous studies have illustrated how entrepreneurial failure leads to negative experiences for entrepreneurs, such as stigmatization, and the loss of reputation or image (e.g., Wiesenfeld, Wurthmann, and Hambrick, 2008; Jenkins et al., 2014; Singh, Corner, and Pavlovich, 2015). The same holds for firms. Efrat (2006) shows, how historically adverse events, such as firm failure or bankruptcy, have been stigmatized. Thus, we posit:

H1: *Media judgment of startups after failure is less positive than media judgment of similar startups, which did not fail.*

Hypothesis 1 refers to a potential sudden drop of media sentiment after a startup fails. We are interested if failure is being judged harshly by the media in comparison to operating startups, and how the judgment evolves. More specifically, we aim to identify differences in media judgment between the US and Germany. As German cultural norms emphasize risk avoidance and seem to attach a social stigma to failure (Reynolds et al., 2000), we posit hypothesis 2:

H2: *Media judgment by German media on German startups after failure is less positive than US media judgment by US startups.*

In addition, we are interested in potential effects of an in-group bias, and wonder, if the media of the two countries in scope favors startups of their own country, in comparison to foreign startups. Hypothesis 3, therefore, states:

TABLE 6.1: Overview of the used dataset for Study I.

	Country	No. of Startups	No. of sentences	No. of observations
Treatment group	USA	186	14.873	1686
	GER	137	13.132	697
Control group	USA	604	124.813	4366
	GER	345	52.705	1673
Total		1272	205.523	8422

H3: *Media judges native startups more positively than foreign startups.*

With these three hypotheses, we start our investigation. The next chapter gives a short overview of our dataset.

6.3 Data

In Chapter 4 we give a broad summary of our dataset, and explain how it is build up, as well as descriptive statistics for a better understanding of the data. The presented dataset is used as a base to start from in the following study, and variables are derived from it. Here, we briefly introduce the important numbers for this study. Table 6.1 lists for the treatment and the control group the number of startups, sentences, and observations, i.e. the number of mean sentiment values, aggregated by quarter.

In total there are 205.523 sentences with associated sentiment values. For each of the 1.272 startups we aggregated the mean sentiment value per quarter, from this, we calculate the mean sentiment value for each quarter and obtain 8.422 observations. The aggregation by quarter for each startup ensures, that single sentences less drive our results within a short period of time. By building the mean value over all quarters, we ensure that our results are not solely driven by individual startups, but rather representing the entire 1.272 startups in the dataset.

The next chapter introduces the applied methodology and the associated empirical framework.

6.4 Methodology and empirical framework

In this chapter, we describe the empirical foundation of the thesis, which includes the underlying mathematical concepts, such as the difference-in-differences (DiD) estimation, and the applied regression model. We begin by introducing the DiD estimation.

A comparison of sentiment values

As a first step, we start by building a mathematical framework, to investigate the dataset. Our dependent variable is the sentiment value S, which varies between -1, 0, and +1 per sentence. In mathematical terms, we can derive this more formally. For the set $M_s = \{x | -1 \leq x \geq 1\}$ with $x \in \mathbb{Z}$ of potential sentiment values for a startup, we define

$$S := S_{i,g,t} = \text{Sentiment of startup i, group g, at time t.} \tag{6.1}$$

The group g can either be treatment or control, and time t is any point in time of the observed period. In practice, a startup can only be part of the treatment or control group. In other words, we can not observe the same startup at a given point in time treated and untreated. Building on 6.1, we derive

$$S_{i,g,t} = \alpha + \beta \text{TREAT}_g + x \sum_{j=1}^{n} X_{j,t} + \varepsilon_{i,t}. \tag{6.2}$$

Here, α stands for the average sentiment of an untreated (operating) startup and hence is not indexed. TREAT_g is an indicator variable and equals 1 for treated (failed), and 0 for control startups, the coefficient β thus accounts for fundamental differences between startups of the control and the treatment group. The sum over $X_{j,t}$ represents variations in key observables, such as cumulated funding sum, market, or founder's information, which vary between startups and might also be time variant. All these variables might have an influence on the judgment. The last term $\varepsilon_{i,t}$ represents the error, which accounts for omitted variables. Our underlying assumption is that the coefficient β, which represents the difference in media judgment of treated and untreated startups, is negative and significant.

We could use Equation 6.2 for a simple OLS regression to calculate β. This type of approach is called cross-sectional, since it aims to analyze data at a particular point in time (cross-sectional) (Rindfleisch et al., 2008). It is a standard approach in research but has raised discussions about its explanatory power due to a lack of unobserved variables and the corresponding heterogeneity. For example, some startups might invest more heavily into marketing, which might lead to a higher media attention and better reporting. At the same time, some startups might have higher qualified founders, which gain publicity through better media training. To better carve out the impact of failure onto media judgment of startups, we expand Equation 6.2 and apply a so-called difference-in-differences (DiD) approach. This research design helps to distinguish the components of a potential drop in media sentiment after business failure, which can be due to selection (startups with lower original quality get selected into the treatment group and eventually fail) or treatment (i.e., due to business failure).

Difference-in-differences estimation

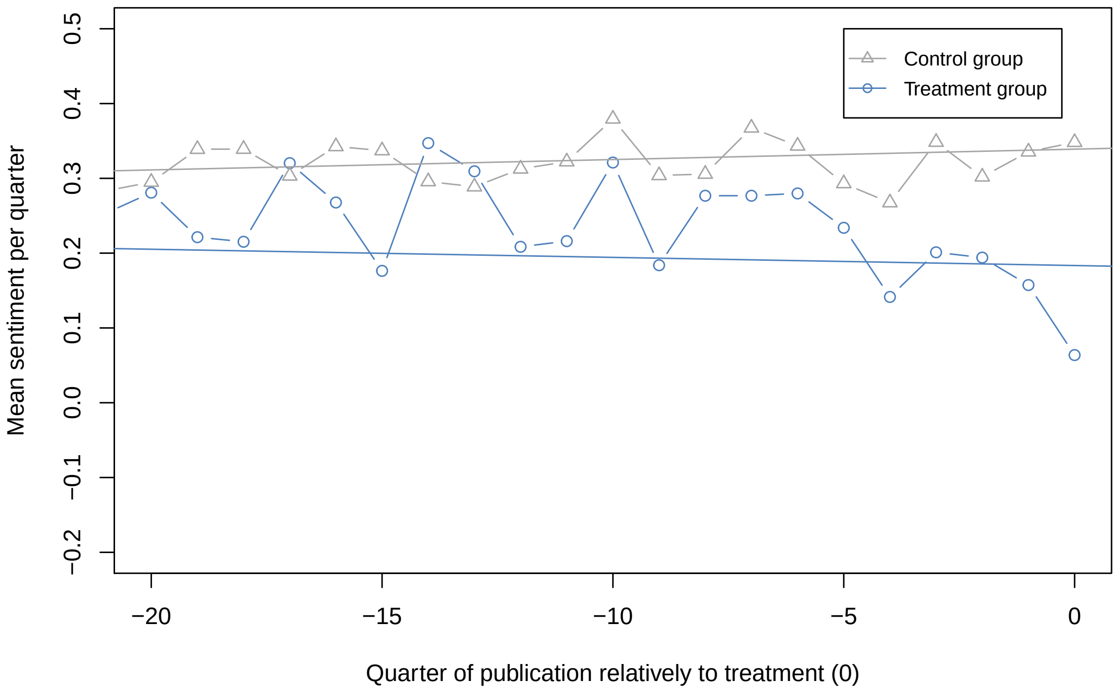
We examine the failure effect on media judgment by applying a DiD approach. The DiD estimation takes advantage of the fact that we can observe the media judgment in the form of the sentiment value, before and after the treatment effect (failure). As there might be differences between treatment and control group even prior to treatment due to for instance the use of different technology or market segment the startup is in, a sole comparison of the after treatment sentiment could be leading to false inferences. The DiD estimation allows disentangling these effects by including differences from the pretreatment sentiment. This means the pretreatment differences indicate the expected differences after the treatment. All further changes serve as a base, from which we can identify the component attributable to the treatment effect, i.e., to failure. This is known as the parallel trend assumption (Pischke, 2005). Figure 6.1 shows the two trends for treatment and control group. Before treatment in quarter 0, the linear fits for the two curves are almost parallel. Their linear equations become:

$$S_{i,treatment,t} = 0.21 - 0.0011t, \tag{6.3a}$$

$$S_{i,control,t} = 0.30 + 0.0005t, \tag{6.3b}$$

with a slight difference in slopes. Compared to the magnitude of the effect (+1 or -1), both slopes are barely 1/10 of a percent, which is why we can assume our parallel trend assumption to be satisfied. The difference in the two equations is mostly driven by the quarters from -4 to zero. During this time, it might already be obvious that startups are about to fail. Potential campaigns might not have been successful, or investors have stopped supporting. The later explains the fall of the media sentiment and hence observed differences.

FIGURE 6.1: Parallel line assumption test for the DiD estimation.



With this in mind, we can assume parallel trends and proceed to enrich Equation 6.4 by two additional dummy variables

$$S_{i,g,t} = \alpha + \beta \mathrm{TREAT}_g + \gamma \mathrm{POST}_{i,t} + \delta (TREAT \cdot POST)_{i,t} + x \sum_{j=1}^{n} X_{j,t} + \varepsilon_{i,t}. \tag{6.4}$$

The first new dummy variable in Equation 6.4 is $\mathrm{POST}_{i,t}$. It is equal to 1 for observations of both groups after treatment. For the control group startups, the treatment date is taken from matched failed startups. With the coefficient γ, we can express changes due to the period after treatment. Following the term, is, an interaction term of the TREAT and POST variable. The corresponding coefficient δ is our quantity of interest. It describes after treatment sentiment of failed startups, δ can thus be called treatment effect. Equation 6.4 allows calculating a potential treatment effect. At the same time, it is very strict, as it controls for observed heterogeneity of startups through observables $\mathrm{X}_{j,t}$. In a more general approach, we can ignore these covariates, because of their balance across groups, and obtain

TABLE 6.2: Results of Equation 6.5 in our DiD approach.

	Pretreatment	Posttreatment	Difference
Treatment group	$\alpha+\beta+\varepsilon_{i,t}$	$\alpha+\beta+\gamma+\delta+\varepsilon_{i,t}$	$\gamma+\delta$
Control group	$\alpha+\varepsilon_{i,t}$	$\alpha+\gamma+\varepsilon_{i,t}$	γ
Difference	β	$\beta+\delta$	δ

$$S_{i,g,t} = \alpha + \beta \mathrm{TREAT}_g + \gamma \mathrm{POST}_t + \delta (TREAT \cdot POST)_{i,t} + \varepsilon_{i,t}. \tag{6.5}$$

An efficient and simple way to determine the treatment effect δ now is to calculate the difference in differences. Table 6.2 shows the results of Equation 6.5 in a DiD logic. Hence, the coefficient δ can be calculated by building the difference of the post treatment, and the pre-treatment values for each group, followed by a subtraction of the results (see Chapter 6.5.1 for results).

Applied regression framework

The DiD estimation is a powerful tool, since it allows to circumvent the normally arising endogeneity problems when comparing heterogeneous objects or individuals (see Meyer (1995) for an overview). For a more rigorous analysis of the dependency interaction between sentiment and entrepreneurial failure, we apply a regression framework. It has three main advantages compared to the DiD approach:

1. Standard errors can be calculated easily,
2. Control variables can be used to reduce variance,
3. It is easy to change the underlying model.

Hence, we proceed with the most simple regression framework, the ordinary least squares (OLS) regression. For a detailed introduction of this regression type, we refer to Wooldridge (2010).

Variable definition and summary

Table 6.3 summarizes the key variables for this analysis, plus the mean value (mean), standard deviation (Std. dev.), and the minimum and the maximum value (Min. and Max.). Our dependent variable is the average sentiment. It contains the sentiment of all sentences in the press mentioning a particular startup, averaged by quarter. This means, if a startup has multiple observations within one-quarter, we calculate the mean for each. Our main independent variables have already been mentioned in Equation 6.5, but are further broken down, e.g., Post1 to Post7_20 are dummy variables for the separate quarters after treatment, and Pre1 to Pre3_20 for the respective quarters before. The difference in differences variable in Equation 6.5 is called DD with DD_3_20 to DD_1 as dummy variables for quarters before, and DD1 to DD7_20 after treatment. In addition, Table 6.3 includes a dummy variable to distinguish countries (1 for Germany, 0 for USA), the logarithmized funding sum and observations, a factor variable for different market groups, an age variable for startups, the work experience in years per founder, the team size as an integer, and the academic background as a dummy variable per founder.

TABLE 6.3: Dataset overview of variables.

Variable	Explanation	Mean	Std. dev.	Min.	Max.
Dependent variable					
Mean Sentiment	Value between -1 and +1 representing the media sentiment of press articles based on sentences, which name the startup in scope. The mean values are aggregated by quarters.	0.267	0.438	-1	+1
Independent variable					
Treatment	Dummy variable, which is 1 for treated startups and 0 for untreated.	0.274	0.446	0	1
Post	A Dummy variable, which is 1 for sentences published after the startup has failed. Operating startups get a fictitious failure date from the matched failed startup.	0.489	0.500	0	1

TABLE 6.3: Dataset overview of variables.

Variable	Explanation	Mean	Std. dev.	Min.	Max.
Dependent variable					
Post1	A Dummy variable, which is 1 for sentences published in the first quarter after the startup has failed (same for operating startups with matched failure date).	0.424	0.201	0	1
Post2	Same as above for the second quarter after the startup has failed.	0.029	0.169	0	1
Post3	Same as above for the third quarter after the startup has failed.	0.025	0.156	0	1
Post4	Same as above for the fourth quarter after the startup has failed.	0.021	0.147	0	1
Post5	Same as above for the fifth quarter after the startup has failed.	0.041	0.198	0	1
Post6	Same as above for the sixth quarter after the startup has failed.	0.029	0.168	0	1
Post7_20	Same as above for the seventh to the twentieth quarter after the startup has failed.	0.237	0.425	0	1
Pre1	Dummy variable, which is 1 for sentences published in the last quarter the startup was operating (same for operating startups with matched failure date).	0.0244	0.154	0	1
Pre2	Same as above for the second last quarter the startup was operating.	0.024	0.156	0	1
Pre3_20	Same as above between the twentieth and third last quarter the startup was operating.	0.3408	0.474	0	1

TABLE 6.3: Dataset overview of variables.

Variable	Explanation	Mean	Std. dev.	Min.	Max.
Dependent variable					
DD	Difference in differences variable, which is 1 for treated startups and sentences published after the startup has failed, and 0 else. Operating startups always have a value of 0.	0.075	0.264	0	1
DD_3_20	The difference in differences variable as above, but solely for the third to the twentieth quarter before failure.	0.08	0.27	0	1
DD_2	The difference in differences variable as above, but solely for the second last quarter before failure.	0.004	0.07	0	1
DD_1	The difference in differences variable as above, but solely for the last quarter before failure.	0.007	0.08	0	1
DD1	The difference in differences variable as above, but solely for the first quarter after failure.	0.014	0.121	0	1
DD2	The difference in differences variable as above, but solely for the second quarter after failure.	0.008	0.089	0	1
DD3	The difference in differences variable as above, but solely for the third quarter after failure.	0.006	0.076	0	1
DD4	The difference in differences variable as above, but solely for the fourth quarter after failure.	0.005	0.070	0	1
DD5	The difference in differences variable as above, but solely for the fifth quarter after failure.	0.009	0.093	0	1

TABLE 6.3: Dataset overview of variables.

Variable	Explanation	Mean	Std. dev.	Min.	Max.
Dependent variable					
DD6	The difference in differences variable as above, but solely for the sixth quarter after failure.	0.005	0.067	0	1
DD7_20	The difference in differences variable as above, but solely for the seventh to twentieth quarter after failure.	0.028	0.166	0	1
Country	Categorial variable, 1 stands for Germany, and 0 for USA.	0.028	0.166	0	1
log(Funding sum)	Logarithmized funding sum in $US, raised by the startup until that quarter.	10.33	7.37	0	19.95
log(Observations per quarter)	Number of observations (sentences) published for each quarter.	1.73	1.37	0	8.08
Market Group	Factor variable to define the market the startup is in. Ten different market groups exist.	Services, Apps/Software, Research, SCG, Social media, Health, Entertainment, Education, Social and others.			
Age	Age of the startup in quarters.	21.61	16.27	0	80.00
Work experience per founder	Sum of work experience in the founder's team, divided by the number of founders.	8.431	8.09	0	49
Team size	Number of founders in the team.	1.283	0.69	1	5
Academic background per founder	Number of founders who have a university degree divided by team size.	0.772	0.400	0	1

6.5 Results

We now report our findings on media judgment of startups along the hypotheses from Chapter 6.2. First, we start with our difference in differences estimation (DiD) in Chapter 6.5.1, followed by more detailed analyses, a country specific DiD estimation in Chapter 6.5.2, cultural differences are investigated in Chapter 6.5.3, and an in-group-bias effect in Chapter 6.5.4.

6.5.1 Difference in differences estimation

Before we delve into the regression analysis part, we present the results of the DiD estimation using summary statistics for the treatment and control group. Table 6.4 reports the mean sentiment values of all sentences for the pre-treatment, and post-treatment period. For the post-treatment period, treated startups have a lower mean sentiment value (0.089) compared to the control group (0.298), reflecting a difference of -0.209. The more depressed sentiments can also be seen in the pre-treatment period (0.234 versus 0.291), reflecting a difference of -0.057. In the treated group itself, the sentiment drops after the startup fails (-0.145), whereas matched startups in the control group experience a slight increase (0.007). Since the idea of the DiD estimation is to only attribute the difference of the two groups to the treatment effect instead of the whole post-treatment or treatment value, the difference in differences effect turns out to be -0.152 (-0.145 - 0.007), which is the difference of differences between post- and pre-treatment differences.

From the results in Table 6.4 we can conclude one central insight. There is a negative DiD effect for the sentiment, suggesting a decrease in the sentiment value attributable to the failure of the startup (with a magnitude of -0.152). We compare this value to the average media sentiment of a startup prior to failure. Assuming a one failed to one operating startup correlation, the mean of pretreatment control and treatment group results in (0.234+0.291)/2, and hence 0.263. Thus, business failure reduces the average media sentiment by 58%. This appears to be a tremendous change, with potential implications for the founder. We will further investigate this finding in Study III (see Chapter 8).

In Chapter 6.2 we proposed hypothesis 1:

> ***H1:*** *Media judgment of startups after failure is less positive than media judgment of similar startups, which did not fail.*

TABLE 6.4: Difference in differences estimation of treatment and control group for the entire dataset.

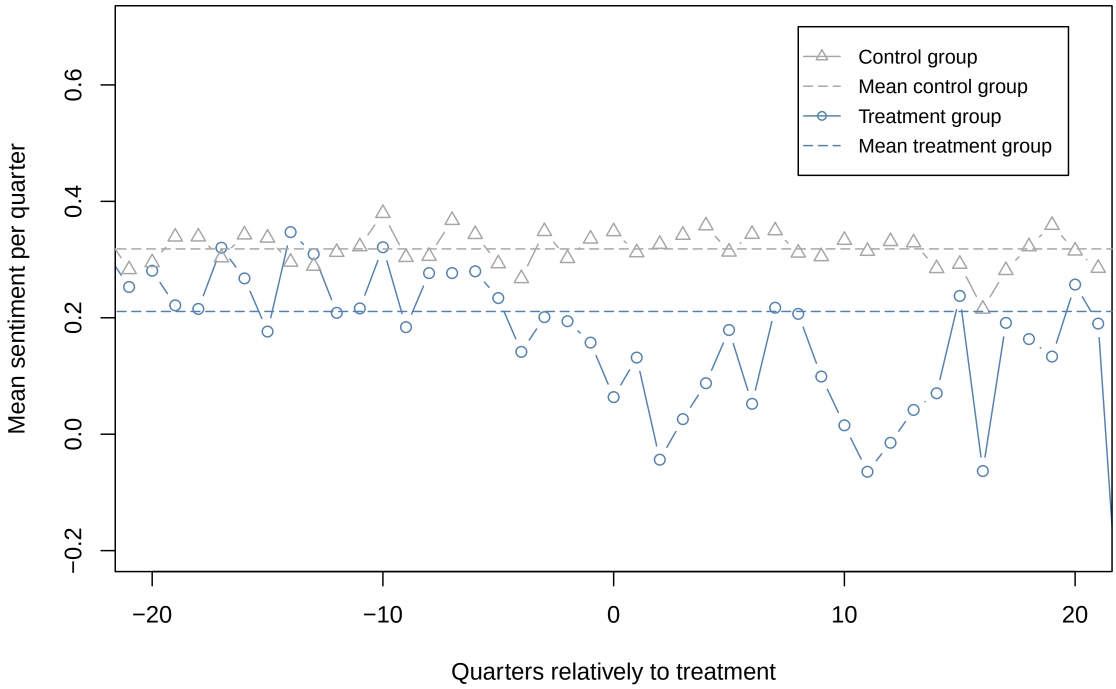
	Pretreatment	Posttreatment	Difference
Treatment group	0.234 (N=2340)	0.089 (N=486)	-0.145 (N=2826)
Control group	0.291 (N=3840)	0.298 (N=3236)	0.007 (N=7076)
Difference	-0.057 (N=6180)	-0.209 (N=3722)	-0.152 (N=9902)

The results of our DiD estimation underline this hypothesis (a drop of 58%). Before we deep dive into a potential explanation for this effect, by applying a robust regression framework, we take a look at the temporal development of the sentiment change. Figure 6.2 presents the sentiment for our two groups (treatment and control) over time. The y-axis shows the mean sentiment per quarter. On the x-axis we see quarters referenced to the treatment quarter 0. Following the timeline, we can see that even before the startup fails (quarter 0) the sentiment seems to fall. It also stays low in the subsequent quarters after the treatment took place. This leads us to two additional insights. First, it appears that the media sentiment anticipates the failure of startups or founders might delay a potential failure intentionally, which gets judged negatively by the media. Second, the low sentiment after failure is not of short duration but remains even in the years after failure.

These findings become more and more solid with the number of observations they are based on. For this purpose, Figure 6.3 shows the number of observations per startup for treatment and control group for each quarter in reference to the treatment quarter 0, so we can directly compare the two graphs. Startups from the control group show a linear to linear quadratic growth in the number of observations, i.e., published sentences per quarter. This increase stops around ten quarters after the treatment, followed by a negative growth of similar magnitude. The treatment group is different. At first, observations grow akin to the control group, and suddenly reach a peak within the treatment quarter, followed by a substantial decline in the number of observations. During the treatment quarter, publications rise to about four times of prior quarters, indicating a high interest of media to report about failure. Both curves show, especially in the two years before and one year after the treatment, a sound basis for our analysis, since there are between three to eight observations per startup and quarter for each of the 969 startups in our dataset. Hence, we can have reason to believe that the observed results are solid and not caused by outliers.

The decline of sentences per quarter for startups from the control group after

FIGURE 6.2: Mean sentiment per startup as a function of quarters in the startup's life, for treatment and control group.



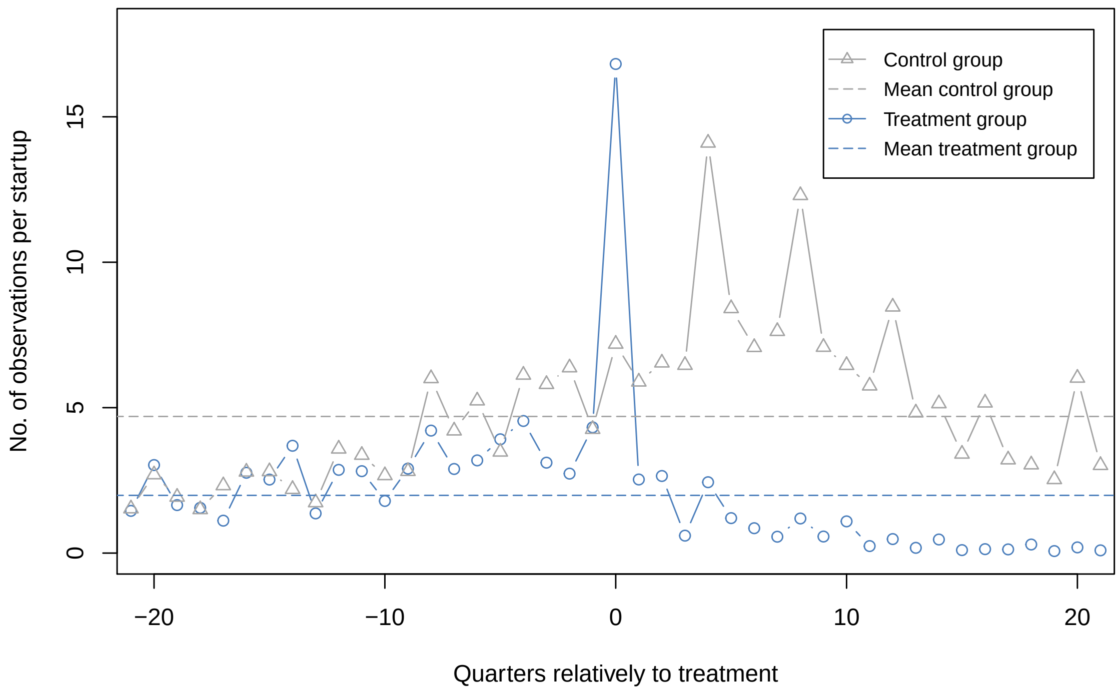
treatment can be explained with our investigated period. As we focus on the years, 1995 to 2015, our time line is finite and so are our observations. Investigated startups of this period reach at least the treatment quarter 0, but with time the number of startups, and hence, the number of observations in our dataset become less.

Figure 6.2 shows the results for all the startups in the dataset, and therefore, US and German ones. Also included are observations from opposite sources, i.e., German media articles about US startups and vice versa. As we are interested in changes in media reporting due to failure, we also like to understand how the sentiment varies between the two countries. Hence, we further deep dive into a state perspective in the next chapter and only consider press articles about startups from their countries of origin.

6.5.2 Country specific difference in differences estimation

Table 6.5 summarizes the DiD estimation split by country, where only news materials from the associated homeland are considered. In the US, treated startups experience a decline in media sentiment after failure (-0.120). At the same time, the sentiment of operating startups

FIGURE 6.3: Observations per startup for treatment and control group.



rises slightly (0.018). The DiD estimator for the US turns out to be -0.138. This implies that the change in media sentiment for US startups is smaller, compared to the entire dataset from Table 6.4. We take a look at Germany. Here, treated startup's sentiment declines also after failure (-0.072), while operating startup's sentiment rises (0.022), just like in the US. However, magnitudes are lower, and the resulting DiD estimator becomes -0.094, which is two-thirds of the decline compared to the US. In other words, media sentiment after failure falls in the US 32% lower, in comparison to Germany (-0.138 to -0.094). The difference to the results for our DiD estimator in Chapter 6.5.1 can be explained with the missing sentences from American newspapers about German startups and vice versa.

The results from Table 6.5 show that in general startups seem to be judged more positively in the US than in Germany. After a startup fails in the US, the average media judgment remains positive. We have calculated a value of (0.214), one magnitude larger compared to German operating startups (0.011 to 0.033). Nevertheless, after failure, the judgment change (the difference between pre- and post-treatment values) experienced by startups is more negative in the US (-0.120 to -0.072). Media judgment of failed startups in Germany turns to the negative, suggesting that this country is not able to tolerate failures. To better understand the observed differences between the countries and potential explanations for changes and

TABLE 6.5: Difference in differences analysis of treatment and control group for the US and Germany.

		Pretreatment	Posttreatment	Difference
USA	Treatment group	0.334 (N=1440)	0.214 (N=246)	-0.120 (N=1686)
	Control group	0.374 (N=2308)	0.392 (N=2058)	0.018 (N=4366)
	Difference	-0.040 (N=3748)	-0.178 (N=2304)	-0.138 (N=6052)
GER	Treatment group	0.004 (N=560)	-0.068 (N=137)	-0.072 (N=697)
	Control group	0.011 (N=788)	0.033 (N=885)	0.022 (N=1673)
	Difference	-0.007 (N=1348)	-0.101 (N=1022)	-0.094 (N=2370)

differences in media judgment, we turn to a regression model. This will be part of the next chapter.

6.5.3 Cultural differences in media judgment of entrepreneurial failure

For a more rigorous analysis of the link between entrepreneurial failure, media judgment and country specific effects, we built a regression framework. In Table 6.6, we present the baseline analysis, which follows Equation 6.5, just like the DiD estimation, employing sentiment as the dependent variable. Column (1) shows a linear OLS model and reports findings from the total dataset, including the treatment and control group. Column two to four add fixed effects for robustness checks (for further details on fixed effects see Angrist and Pischke (2008)), and Equation 6.5 becomes

$$S_{i,g,t} = \alpha + \beta \mathrm{TREAT}_g + \gamma \mathrm{POST}_t + \delta(TREAT \cdot POST)_{i,t} + \lambda_s + \eta_t + \tau_m + \varepsilon_{g,t}. \quad (6.6)$$

To exclude potential effects solely resulting from particular startups or years in our dataset, we included fixed effects in subsequent models. We control for startups with λ_s, and the market they are in with τ_m, as well as year fixed effects with η_t. Our results turn out to be robust, as the difference in differences variable (called DD in our regression model, in order to distinguish it from the DiD estimator) stays strongly significant (at a p = 0.01 level). Its value becomes almost the same as in the DiD estimation from Chapter 6.5.1 (-0.150

compared to -0.152).

TABLE 6.6: Regression model for the DD estimator.

	Dependent variable:			
	Mean_sentiment			
	(1)	(2)	(3)	(4)
Treatment	-0.096^{***}	-0.157	-0.174	-0.175
	(0.013)	(0.384)	(0.384)	(0.384)
Post	-0.012	-0.006	0.011	0.011
	(0.010)	(0.011)	(0.015)	(0.015)
DD	-0.129^{***}	-0.150^{***}	-0.154^{***}	-0.153^{***}
	(0.023)	(0.025)	(0.025)	(0.025)
Constant	0.333^{***}	0.314^{***}	8.337^{**}	8.362^{**}
	(0.008)	(0.100)	(4.192)	(4.192)
Startup ID	No	Yes	Yes	Yes
Year	No	No	Yes	Yes
Market/Industry	No	No	No	Yes
Observations	9,161	9,161	9,161	9,161
R^2	0.021	0.326	0.326	0.327
Adjusted R^2	0.021	0.245	0.246	0.246
Residual Std. Error	0.422 (df = 9157)	0.370 (df = 8181)	0.370 (df = 8180)	0.370 (df = 8179)
F Statistic	66.511^{***} (df = 3; 9157)	4.044^{***} (df = 979; 8181)	4.045^{***} (df = 980; 8180)	4.043^{***} (df = 981; 8179)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

In summary, the findings in Table 6.6 support hypothesis H1 and we can conclude that the media sentiment of a startup drops after failure significantly at the p = 0.01 level by -0.152 on average (see the three models with fixed effects in Table 6.6).

From Figure 6.2 we learn that the drop in sentiment occurs even before the startup fails. This decline starts about a year or four quarters before treatment. Interestingly, the sentiment stays low, even in subsequent years long after the startup even existed. Hence, we further investigate the pre-failure decline and after failure effects in a separate regression.

Table 6.7 and Table 6.8 show a temporal resolution along quarters of the DD variable. In addition, we added time variant variables, such as age of the startup, raised funding sum, and a number of observations per quarter. Ten-time periods have been identified to explain the sentiment evolution, each representing one model in Table 6.7 (Part 1) and Table 6.8 (Part 2). Model 10 shows a negative treatment effect throughout all quarters prior to (except DD_1) until the fifth quarter after treatment. From the seventh to the twentieth, we see this again. Surprisingly, values decrease consistently, reaching a minimum value of -0.343 in the fourth quarter after treatment. In other words, almost one year after the startup went out of business, media still reports negatively about the startup.

TABLE 6.7: Pre-failure and after failure media sentiment effects - part 1.

	Dependent variable:				
	Mean_sentiment				
	(1)	(2)	(3)	(4)	(5)
Treatment	-0.780^{**}	-0.780^{**}	-0.784^{**}	-0.781^{**}	-0.747^{**}
	(0.343)	(0.343)	(0.343)	(0.343)	(0.344)
Pre3_20	−0.004	−0.004	−0.004	−0.001	−0.001
	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)
Pre2		−0.013	−0.013	−0.010	−0.010
		(0.035)	(0.035)	(0.035)	(0.035)
Pre1			0.019	0.022	0.022
			(0.036)	(0.036)	(0.036)
Post1				0.024	0.024
				(0.027)	(0.027)
Post2					0.003
					(0.030)
DD_3_20	0.057^{**}	0.057^{**}	0.059^{**}	0.049^{*}	0.040
	(0.026)	(0.026)	(0.027)	(0.028)	(0.028)
DD_2		0.007	0.009	−0.001	−0.009
		(0.076)	(0.076)	(0.076)	(0.077)
DD_1			0.019	0.008	0.0005
			(0.072)	(0.072)	(0.073)
DD1				−0.078	−0.088
				(0.056)	(0.056)
DD2					−0.095
					(0.069)
Age	0.002	0.002	0.003	0.003	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
log(Funding sum)	−0.001	−0.001	−0.001	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(Observations per quarter)	0.042^{***}	0.042^{***}	0.042^{***}	0.042^{***}	0.042^{***}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Constant	32.472	32.994	35.465	36.211	38.200
	(31.714)	(31.747)	(31.919)	(31.942)	(32.072)
Startup	Yes	Yes	Yes	Yes	Yes
Year	Yes	No	No	No	No
Market/Industry	Yes	Yes	Yes	Yes	Yes
Observations	7,337	7,337	7,337	7,337	7,337
R^2	0.309	0.309	0.309	0.309	0.309
Adjusted R^2	0.227	0.227	0.227	0.227	0.227
Residual Std. Error	0.374 (df = 6562)	0.374 (df = 6560)	0.374 (df = 6558)	0.374 (df = 6556)	0.374 (df = 6554)
F Statistic	3.788^{***} (df = 774; 6562)	3.777^{***} (df = 776; 6560)	3.768^{***} (df = 778; 6558)	3.760^{***} (df = 780; 6556)	3.754^{***} (df = 782; 6554)

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

In the introductory chapter of this thesis, we state that cultural differences between the US and Germany exist, which affect entrepreneurial risk-taking and potentially media reporting about it. So the question is, how do we quantify these potential effects? What do we need for such an analysis? Regarding our dataset, we have different options. First, we can just distinguish media reporting by country. This means we add a country variable to our regression model. Since our focus is an entrepreneurial failure, we are also interested in a potential change in media sentiment after failure. For this purpose, we can interact our country variable with the DD variable. Table 6.9 displays our results. The sample is in three phases. First, there is the entire sample, second, only sentences from a startup's homeland press and third, the cross

TABLE 6.8: Pre-failure and after failure media sentiment effects - part 2.

	Dependent variable:				
	Mean_sentiment				
	(6)	(7)	(8)	(9)	(10)
Treatment	−0.734**	−0.721**	−0.710**	−0.713**	−0.642*
	(0.344)	(0.344)	(0.344)	(0.344)	(0.344)
Pre3_20	0.001	0.003	0.006	0.004	0.024
	(0.014)	(0.014)	(0.015)	(0.015)	(0.017)
Pre2	−0.008	−0.007	−0.003	−0.005	0.014
	(0.035)	(0.035)	(0.035)	(0.036)	(0.037)
Pre1	0.023	0.026	0.030	0.028	0.050
	(0.036)	(0.036)	(0.036)	(0.037)	(0.038)
Post1	0.026	0.028	0.031	0.029	0.050*
	(0.027)	(0.027)	(0.028)	(0.028)	(0.030)
Post2	0.004	0.006	0.009	0.006	0.027
	(0.030)	(0.030)	(0.030)	(0.031)	(0.033)
Post3	0.018	0.020	0.024	0.021	0.041
	(0.032)	(0.032)	(0.032)	(0.032)	(0.034)
Post4		0.033	0.037	0.035	0.056
		(0.035)	(0.035)	(0.035)	(0.037)
Post5			0.031	0.028	0.050*
			(0.026)	(0.026)	(0.029)
Post6				−0.023	−0.002
				(0.030)	(0.032)
Post7_20					0.032*
					(0.018)
DD_3_20	0.026	0.009	−0.005	−0.002	−0.087**
	(0.029)	(0.029)	(0.031)	(0.031)	(0.038)
DD_2	−0.029	−0.046	−0.060	−0.058	−0.152*
	(0.077)	(0.077)	(0.078)	(0.078)	(0.081)
DD_1	−0.014	−0.035	−0.049	−0.047	−0.142*
	(0.073)	(0.073)	(0.074)	(0.074)	(0.077)
DD1	−0.104*	−0.124**	−0.138**	−0.136**	−0.233***
	(0.056)	(0.057)	(0.058)	(0.058)	(0.063)
DD2	−0.111	−0.127*	−0.144**	−0.141**	−0.241***
	(0.069)	(0.069)	(0.070)	(0.071)	(0.075)
DD3	−0.200**	−0.217***	−0.233***	−0.231***	−0.330***
	(0.078)	(0.078)	(0.079)	(0.079)	(0.083)
DD4		−0.230***	−0.244***	−0.242***	−0.343***
		(0.082)	(0.082)	(0.083)	(0.086)
DD5			−0.112*	−0.109	−0.211***
			(0.068)	(0.068)	(0.073)
DD6				0.030	−0.078
				(0.092)	(0.095)
DD7_20					−0.201***
					(0.049)
Age	0.003	0.003	0.003	0.004	0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
log(Funding_sum + 1)	−0.001	−0.001	−0.001	−0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(Observations_per_quarter)	0.042***	0.042***	0.042***	0.042***	0.041***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Constant	38.765	39.335	39.564	41.828	43.034
	(32.111)	(32.287)	(32.312)	(32.455)	(32.509)
Startup	Yes	Yes	Yes	Yes	Yes
Year	Yes	No	No	No	No
Market/Industry	Yes	Yes	Yes	Yes	Yes
Observations	7,337	7,337	7,337	7,337	7,337
R^2	0.310	0.311	0.311	0.311	0.313
Adjusted R^2	0.228	0.228	0.228	0.228	0.230
Residual Std. Error	0.373 (df = 6552)	0.373 (df = 6550)	0.373 (df = 6548)	0.373 (df = 6546)	0.373 (df = 6544)
F Statistic	3.756*** (df = 784; 6552)	3.760*** (df = 786; 6550)	3.755*** (df = 788; 6548)	3.745*** (df = 790; 6546)	3.766*** (df = 792; 6544)

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

data, where there is specific sentence consideration from startups based on which nations they result. For the three different datasets, we have one model, with the country variable, and another one, with the interaction term of country and DD estimator, which results in six models.

Model 1a and b are calculated with the entire dataset. The DD estimator is significant at the p = 0.01 level, with values of -0.132 and -0.154. In addition, our interaction term also appears to be significant, and positive for Germany (0.112). An inclusion of selected articles from the home country startups is in consideration in model 2a and b. The results show German media reports less positively about startups than US media, with a country variable of -0.321 and -0.326 respectively. However, the interaction of country and DD variable shows, that it is the US media, which changes sentiment more drastically after failure (-0.105). The final two columns in Table 6.9 show results for a crossed dataset. US media reports about German startups and vice versa. The country variable remains significant for both models, indicating that German media also reports less positively about US startups, than German ones. In addition the logarithmized observations are negatively significant (-0.019) at a p = 0.05 level. This could mean, media cross reporting focuses on big failures of startups.

In the above regression models, the number of observations per startup has been statistically significant. Figure 6.4 shows the mean number of observations per startup on the y-axis, as a function of quarters relative to treatment on the x-axis. From this graph, we conclude two things. First, German media reports more about failed startups, than US media (3,7 observations per quarter on average compared to 1 in the US). Second, within the treatment quarter, German media reporting rises to 8 times of the overall average. Hence, startups, which eventually fail, seem to get the most attention by the media when they fail.

Interestingly, these result somewhat contradict our hypothesis from Chapter 6.2. We assume the US media to judge failure less harsh than German media. For the overall sentiment value, we can confirm this hypothesis, for the change in sentiment we cannot (Germany +0.112 at a p = 0.01 level). So our hypothesis:

H2: *Media judgment by German media on German startups after failure is less positive than US media judgment by US startups.*

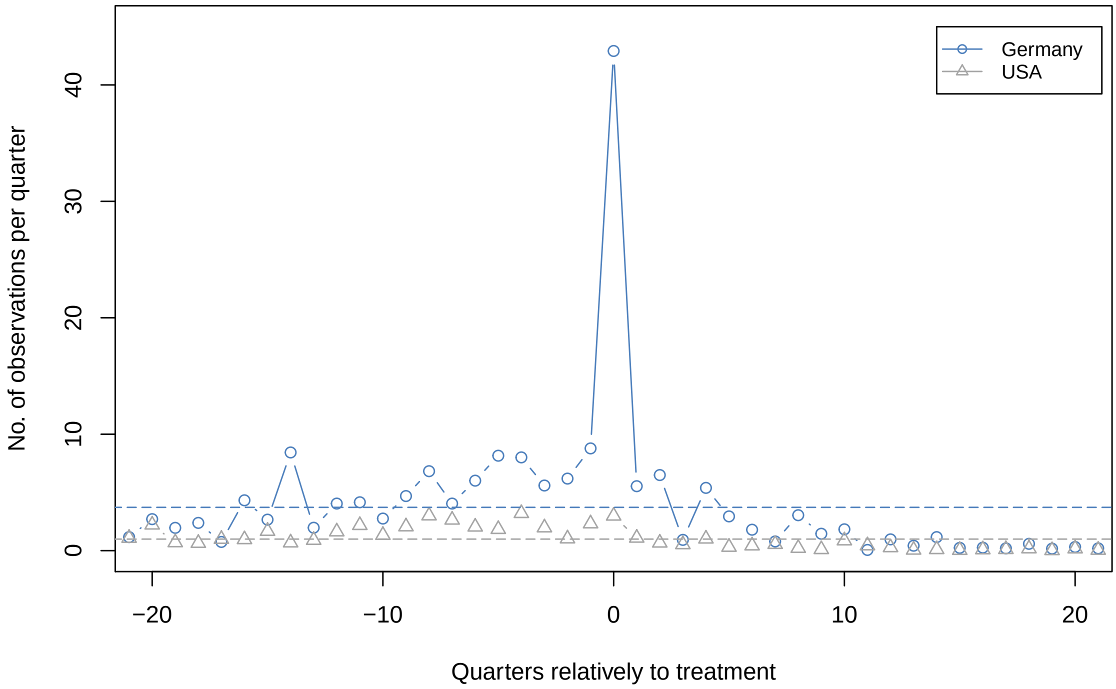
can be confirmed with the addition, that the overall negative change in media sentiment is larger in the US, and reporting about failure, in general, is less common compared to Germany. After failure reporting in the US still remains positive, while it becomes negative in

TABLE 6.9: Regression models with three different datasets. Model 1a, and b are based on the entire dataset, model 2a, and b are based on reporting about native startups, model 3a, and b are based on reporting about foreign startups.

	Dependent variable:					
	Mean_sentiment					
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
	All	All + i	Own	Own + i	Cross	Cross + i
Treatment	−0.055***	−0.056***	−0.045***	−0.045***	−0.014	−0.013
	(0.016)	(0.016)	(0.016)	(0.016)	(0.041)	(0.041)
Post	−0.032***	−0.032***	−0.028**	−0.028**	−0.032	−0.033
	(0.011)	(0.011)	(0.012)	(0.012)	(0.029)	(0.029)
DD	−0.132***	−0.154***	−0.098***	−0.117***	−0.122*	−0.089
	(0.026)	(0.029)	(0.028)	(0.030)	(0.065)	(0.079)
Country (1 = Germany)	−0.195***	−0.200***	−0.321***	−0.326***	0.436***	0.445***
	(0.011)	(0.011)	(0.013)	(0.013)	(0.034)	(0.036)
Age	0.001**	0.001**	0.0003	0.0003	0.0005	0.0005
	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.001)	(0.001)
log(Funding sum)	0.002***	0.002***	−0.0003	−0.0003	−0.0003	−0.0003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
log(Observations per quarter)	0.022***	0.021***	0.025***	0.025***	−0.019**	−0.019**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.009)	(0.009)
DD:Country (1 = Germany)		0.100**		0.105*		−0.073
		(0.050)		(0.058)		(0.097)
Constant	0.317***	0.319***	0.365***	0.367***	−0.053	−0.060
	(0.013)	(0.013)	(0.014)	(0.014)	(0.037)	(0.038)
Observations	7,337	7,337	6,291	6,291	1,046	1,046
R^2	0.068	0.068	0.111	0.111	0.181	0.181
Adjusted R^2	0.067	0.067	0.110	0.110	0.175	0.175
Residual Std. Error	0.410 (df = 7329)	0.410 (df = 7328)	0.399 (df = 6283)	0.398 (df = 6282)	0.394 (df = 1038)	0.395 (df = 1037)
F Statistic	75.909*** (df = 7; 7329)	66.946*** (df = 8; 7328)	111.686*** (df = 7; 6283)	98.169*** (df = 8; 6282)	32.745*** (df = 7; 1038)	28.710*** (df = 8; 1037)

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

FIGURE 6.4: Number of observations per failed startup and quarter in Germany and the US.



Germany.

6.5.4 In-group bias in media judgment

As a next step, we break our regression in Table 6.9 down to country level. Hence, Table 6.10 does the same analysis as above but includes only US media articles. In column 1a and b, we can see, a negative significant DD variable (-0.111 and -0.109 at a p = 0.01 level), and country variable (-0.067 for both at a p = 0.01 level), suggesting that US media judges German startups more negatively than their own. Column 2 shows findings for the "own" model, meaning US media judges US startups. Similar to the country specific DiD estimation in Table 6.5, our DD variable is negative and significant (-0.099 at a p = 0.01 level). Interestingly, the logarithmized observations are positively significant (0.036 at a p = 0.01 level), indicating that the US media favors startups to write positive stories about. This changes as we take a look at how US media reports about German startups. In model 3 we can see, as the number of observations increases, sentiment becomes more negative (significant at a p = 0.01 level). Also, the DD variable is negatively significant (at a p = 0.05

level), and its value is far greater (-0.163).

TABLE 6.10: US media reporting comparison.

	Dependent variable:			
	Mean_sentiment			
	(1a)	(1b)	(2)	(3)
	All	All + i	Own	Cross
Treatment	-0.052^{***}	-0.052^{***}	-0.053^{***}	-0.021
	(0.017)	(0.017)	(0.019)	(0.044)
Post	-0.029^{**}	-0.029^{**}	-0.029^{**}	-0.020
	(0.012)	(0.012)	(0.013)	(0.029)
DD	-0.111^{***}	-0.109^{***}	-0.099^{***}	-0.163^{**}
	(0.029)	(0.030)	(0.031)	(0.081)
Country (1 = Germany)	-0.067^{***}	-0.067^{***}		
	(0.015)	(0.015)		
Age	0.0002	0.0002	-0.00001	0.0004
	(0.0003)	(0.0003)	(0.0004)	(0.001)
log(Funding sum)	-0.001	-0.001	-0.001	0.0001
	(0.001)	(0.001)	(0.001)	(0.002)
log(Observations per quarter)	0.026^{***}	0.026^{***}	0.036^{***}	-0.026^{***}
	(0.004)	(0.004)	(0.004)	(0.009)
DD:Country (1 = Germany)		-0.012		
		(0.074)		
Constant	0.372^{***}	0.371^{***}	0.360^{***}	0.391^{***}
	(0.014)	(0.014)	(0.015)	(0.033)
Observations	5,801	5,801	4,964	837
R^2	0.026	0.026	0.031	0.021
Adjusted R^2	0.024	0.024	0.030	0.014
Residual Std. Error	0.389 (df = 5793)	0.389 (df = 5792)	0.391 (df = 4957)	0.364 (df = 830)
F Statistic	21.673^{***} (df = 7; 5793)	18.964^{***} (df = 8; 5792)	26.439^{***} (df = 6; 4957)	2.963^{***} (df = 6; 830)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 6.11 gives a report of the analysis for the German media. For ease of comparison, the model is identical to Table 6.10. The first two columns indicate with significance at the p = 0.1 level, and no significance, German media seems to report more negatively after failure (-0.097 and -0.114), but with significance at a p = 0.01 level more positively about German startups (0.175 and 0.171). Own reporting, i.e., German media reports about German startups, shows no significance, but a negative DD variable. Only the number of observations is significant at a p = 0.1 level. In the cross model 3, we find no evidence for an in-group bias of German media reporting.

We posit in Chapter 6.2, that the media is more critical with foreign startups, as a result of the in-group bias. Our results support this assumption for the US. In Germany media seems to favor German startups, but does not statistically significant judge US startups more negatively.

TABLE 6.11: German media reporting comparison.

	Dependent variable:			
	Mean_sentiment			
	(1a)	(1b)	(2)	(3)
	All	All + i	Own	Cross
Treatment	0.008	0.008	0.011	0.009
	(0.033)	(0.033)	(0.035)	(0.107)
Post	−0.018	−0.018	−0.005	−0.154
	(0.026)	(0.026)	(0.027)	(0.101)
DD	−0.097*	−0.114	−0.099	0.013
	(0.058)	(0.081)	(0.067)	(0.154)
Country (1 = Germany)	0.175***	0.171***		
	(0.034)	(0.037)		
Age	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.003)
log(Funding sum)	−0.00005	−0.00005	0.0003	0.0001
	(0.002)	(0.002)	(0.002)	(0.005)
log(Observations per quarter)	−0.011	−0.011	−0.015*	0.047
	(0.008)	(0.008)	(0.008)	(0.036)
DD:Country (1 = Germany)		0.027		
		(0.092)		
Constant	−0.089**	−0.085**	0.087***	−0.125
	(0.038)	(0.040)	(0.028)	(0.093)
Observations	1,536	1,536	1,327	209
R^2	0.026	0.026	0.005	0.028
Adjusted R^2	0.021	0.020	0.0002	−0.001
Residual Std. Error	0.431 (df = 1528)	0.431 (df = 1527)	0.419 (df = 1320)	0.497 (df = 202)
F Statistic	5.717*** (df = 7; 1528)	5.010*** (df = 8; 1527)	1.048 (df = 6; 1320)	0.954 (df = 6; 202)
Note:				*p<0.1; **p<0.05; ***p<0.01

This means we have found evidence to support our assumption from Chapter 6.2:

H3: *Media judges native startups more positively than foreign startups.*

6.6 Discussion

How does media judgment of startup failure differ between the US and Germany? Following our research agenda from Chapter 6.2, we posit a decline in media sentiment after startup failure. Indeed, failure leads to a less positive media sentiment in both countries. How can we explain these findings? First of all, media reports positively in both countries about running startups, which is in line with Hofstede's theory and empirical studies, suggesting both countries to welcome entrepreneurial activity. The marked decline, even before the startup is out of business, results in a relatively neutral media sentiment (0.089), and shows that media appreciates entrepreneurial risk taking. If it does not, we would expect the sentiment to become negative. Failure is thus not despised, supporting founders to seek self-actualization, and independence from authorities, which is once again backed by Hofstede's measured values for individualism versus collectivism in these countries (see Chapter 3.4).

On a country level, the results are less positively in Germany compared to the US. The sentiment turns negative after failure, suggesting it to be less excepted and tolerated by society. Negative reporting, accompanied by high levels of publications immediately after failure support this assumption. Entrepreneurs hence experience a wave of bad press in addition to their business closure; this might trigger stigmatization of entrepreneurs as result of an attribution process (see Chapter 3.1), as it contributes to a collective impression of society and thus judgment.

In the US, the change (before to after failure) in sentiment is larger but remains on a relatively high level, i.e., more positively than for German operating startups. Failure in the US seems to be less deviant from common rules and expectations and is therefore not negatively judged. As a consequence, it does not serve as a base for stigmatization. Supporting this, is the fact, that reporting on startups after failure almost vanishes, which would be essential for stigmatization. Media thus takes the failure off the public agenda in the US, whereas it enhances it in Germany. With the theories introduced in Chapter 3.1 on how media is able to influence the public agenda and salience of topics, we have to assume failure to be much more present in people's mind in Germany compared to the US. This presents might hinder individuals to become entrepreneurs, due to the huge amount of negative reporting. In addition, it might prevent failed entrepreneurs to start anew, as negative media coverage sometimes is accompanied by criticism, which might lead to stigma, expressed by the general public.

The last hypothesis from Chapter 6.2 deals with the in-group bias. Our findings support an in-group bias in the US, where startups experience a fair judging by media as opposed to the German. As the in-group bias is used by individuals to reduce uncertainty, we would have expected it to be more prominent in Germany (Fischer and Derham, 2016). Therefore, it can be the result of a selection bias, as potentially media is more likely attracted by grand failures than regular business closures. In Germany, we found no evidence to support an in-group bias, even though previous research suggests it to be more likely to occur in societies with greater levels of collectivism (Yamagishi, Jin, and Miller, 1998). This inconsistency may be due to the low number of observations in this analysis and should be studied further before drawing final conclusions.

6.7 Conclusion

The purpose of this study was to measure entrepreneurial friendliness in the US and Germany. We sought to understand how media judges startups, especially in the event of failure. If this judgment varies between the two countries in scope, and if so, to what extent. The above-mentioned findings support the broadly held assumption, that the US is more entrepreneurial friendly than Germany, as even in the event of failure media still reports relatively positive about startups. Failed startups are not mentioned in the press, long after they went out of business, which is mostly in the interest of its founders. In Germany failure is judged slightly negatively, and hence differs from the US. Despite the negative values, there is no evidence pointing towards a cruel judgment. Therefore, we can still consider Germany as entrepreneurial friendly, even though media seems to be more resentful.

This study extends entrepreneurial failure research, as it includes the role of media, which has been mostly neglected so far, even though prior research called for a further investigation (Hudson, 2008). By adding two countries to the study, we can identify cultural differences in treating entrepreneurial failure. This study thus sets a starting point for further investigations in this field. Questions such as why do journalist in Germany judge failure negatively, while this is not observed in the US, need to be answered. An additional step to investigate would be how it influences entrepreneurial activity? Are potential founders detained from a career as entrepreneurs, due to adverse media reporting after failure in Germany? These three questions show, there is still much to explore.

Chapter 7

Study II - Startup specific differences in media judgment of failure

The second study of this thesis sought to investigate the relationship between media judgment, and startup-specific characteristics, such as its novelty, or distinctiveness of the business model, the raised funding sum, the role of serial entrepreneurship, prior work experience, and age of the startup. As the majority of startups fail, we are especially interested in the way media judges along these dimensions in the event of failure.

7.1 Introduction

The way media judges startup business failure is a key determinant for an entrepreneurial friendly environment. Nowadays the media is not simply conveying facts, it also conveys a feeling and tone, which gets absorbed by the general public (McCombs and Ghanem, 2001). As a consequence, the media effect a firm's performance (Wartick, 1992). This occurs indirectly, by impacting a firm's reputation. A startup's reputation, especially in the eyes of the general public, is therefore strongly influenced by the way journalists report about it. A favorable reputation is key for acquiring benefits, such as setting a higher price level for products, attract more qualified employees, or better access to investors or capital (Fombrun and Shanley, 1990). Hence, startups should aim for positive publicity. This study aims to identify general trends in media judgment, which are derived from startup specific characteristics after failure. Do certain startups experience more positive or negative media judgment, primarily through a very innovative business model, or huge amounts of funding? We apply ordinary least squares regression models to our dataset on failed and operating US and German startups.

In this study, we follow a call by Hindle and Klyver (2007), who suggest a more sophisticated investigation of the link between media coverage and entrepreneurship. Our focus is set on how media judges startups, especially in the event of failure. Previous research in the field of entrepreneurship has mainly focused on topics such as learning from failure (Cope, 2011; McGrath, 1999; Shepherd, 2003a; Stokes and Blackburn, 2002), effects of failure on entrepreneurs (Ucbasaran et al., 2012; Singh, Corner, and Pavlovich, 2007; Brunstein and Gollwitzer, 1996; Politis and Gabrielsson, 2009), or stigma of failure (Jenkins et al., 2014; Landier, 2005; Simmons and Wiklund, 2011; Simmons, Wiklund, and Levie, 2014). Most founders have difficulties in learning from failure (Cannon and Edmondson, 2005), and suffer from failure, as it can be a traumatic experience for them (Cope, 2011), often associated with stigmatization of different types. While some entrepreneurs experience discrimination in form of lower future job opportunities, or limited access to financial resources (Cope, 2011), others deterred from entrepreneurship (Hindle and Rushworth, 2000). Surprisingly, little is known about the relationship between media coverage and entrepreneurship (Hindle and Rushworth, 2000). More specifically, to our knowledge no research in this field has investigated the link between startup specific attributes and how they impact media reporting.

We aim to shed light in this unexplored field by taking a meso perspective. Therefore, we measure media judgment via the sentiment expressed in press releases about startups, and are interested in what drives the judgment. The so-called sentiment analysis uses a support vector machine learning algorithm to detect expressed sentiments in press articles. With the help of an OLS regression model, we quantify our findings, and compare them to the existing organizational theory, namely the social judgment of organizations theory. This theory positions legitimacy, reputation, and status as central concepts for evaluator's social judgment (Bitektine, 2011). Our approach uses media judgment as an indirect measure of these three. Withholding legitimacy, low reputation and status are thus reflected in negative media judgment about startups. In addition, we make use of Hofstede's cultural dimensions theory, and stigma research, to interpret our findings.

This study is structured as follows; the next chapter derives our research hypotheses, Chapter 7.3 gives a brief overview of our dataset. Chapter 7.4 introduces our methodical approach, the key measures used, and the chosen empirical framework. In Chapter 7.5 we present results in four steps, followed by a discussion and conclusion in Chapter 7.6 and 7.7.

7.2 Theory and hypotheses

The literature often uses novelty, atypicality or distinctiveness synonymously. For reasons of simplicity we stick to novelty, and use it in the following.

According to Hofstede's cultural dimensions theory, media is expected to judge novel, former unknown business ideas in countries with high levels of uncertainty avoidance, less positively. It is considered a deviation from social norms, and therefore to be rejected (Lonner, Berry, and Hofstede, 1980). Especially in case of failure we expect this to occur. Hence, our first hypothesis is formulated as:

> *H1: Novelty of business models negatively moderates a startup's media judgment after failure.*

Since the emergence of the venture capitalist's industry, startups are able to receive financial support from external investors. These investments are often placed in a three steps process, to reduce an investor's risk (Podolny, 2001): The seed stage, the venture capital stage, and the expansion stage. Throughout these stages one operating startup in our dataset raised up to USD 657 M, a failed one up to USD 374 M. These huge amounts of money, might be taken into consideration, as media reports and judges a startup's failure. We posit:

> *H2: The amount of founding negatively moderates the media judgment of startups after failure.*

Researchers have often pointed out the importance of experience and knowledge, for being a successful entrepreneur (see Agarwal et al. (2004) and Chatterji (2009) for an overview). As our data shows, a significant number (29%) of all founders have no prior work experience, and hence, might face more difficulties than experienced founders. We, therefore, expect founders with prior work experience to be more successful in general. In addition, we expect experienced founders in case of failure to be less criticised by media, as it is less likely to be blamed on the founder's team, but on external effects. Our hypothesis three is therefore:

> H3: Prior work experience of the team before founding positively moderates the media judgment after failure.

Another focus is how media judges serial entrepreneurs. Much research in this field has been conducted, after MacMillan's pioneering study, which pointed out the importance of serial entrepreneurs, in order to understand entrepreneurship (MacMillan, 1986). Two types of serial entrepreneurs exist. First, there are serial entrepreneurs, who exit a business before it is fails. This could be motivated by other opportunities such as starting a potentially more profitable business, or being unable to grow the existing one further. Then there are serial entrepreneurs who start a new venture, after the previous one has failed. These entrepreneurs often consider external factors as being responsible for their venture failure, rather then their own entrepreneurial skill set (Eggers and Song, 2015). The two types of serial entrepreneurship have in common that they left a startup to found a new one. We are interested in how media judges their startups after failure, and hypothesize:

> *H4: Prior founding experience negatively moderates the media judgment after failure.*

Storey (2016) found age to be among five other critical factors for a small firm's growth. A startup aims for fast and strong growth. We expect the media to take into consideration the age of a startup. Failure after several years could be accompanied by delaying filing for insolvency, and hence, judged harsh by the media. Therefore, we posit:

> H5: Media judgment of startups after failure, will be less positive for startups which existed longer, than for those who failed earlier.

7.3 Data

For Study II we build on the data gathered for Study I (see Chapter 4). In addition, we included the category list from CrunchBase, which describes each startup with at least one and up to 11 keywords, e.g., the startup 3D Vision Systems is associated with the categories hardware and software (see Figure 7.1).

These key words have been supplied by the CrunchBase community, e.g., startup founders or investors. We use this categorization as a starting point for our novelty or distinctiveness measure. As these categories exist only for those startups, which are originally taken from CrunchBase, not all Startups from Study I are included in this analysis. Table 7.1 presents a summary of the 839 included startups.

In addition to these startups, we use funding data from CrunchBase. Within numerous funding rounds, startups aim to collect financial support from investors. A correlation

FIGURE 7.1: Number of category labels per startup included in the dataset.

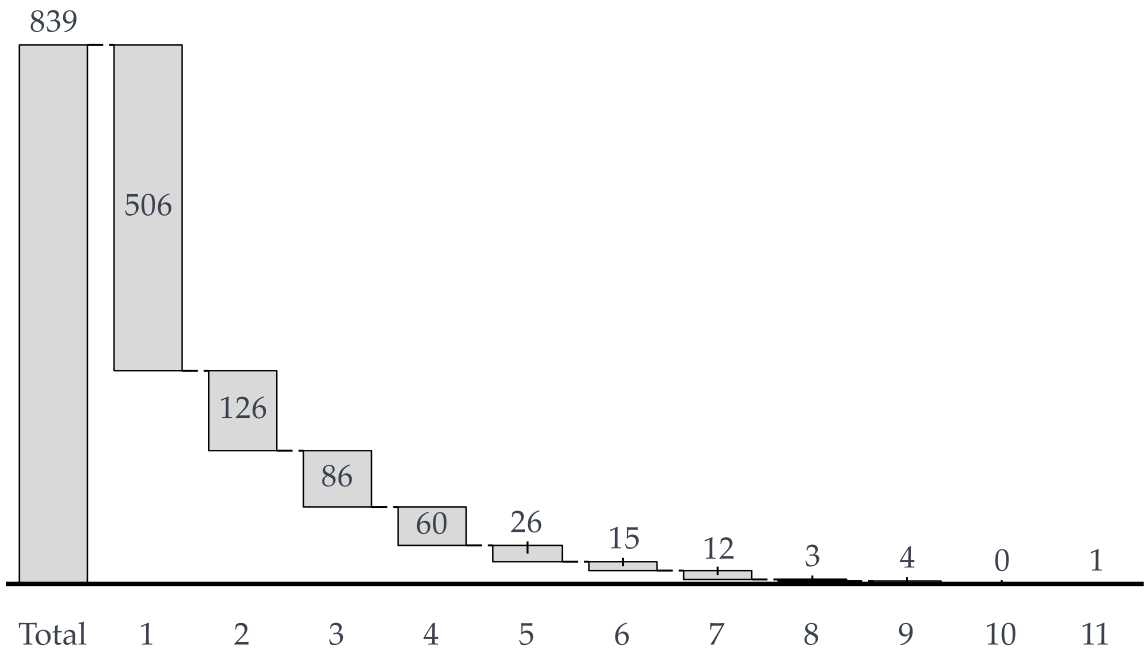


TABLE 7.1: Number of startups per country and operating status included in the dataset.

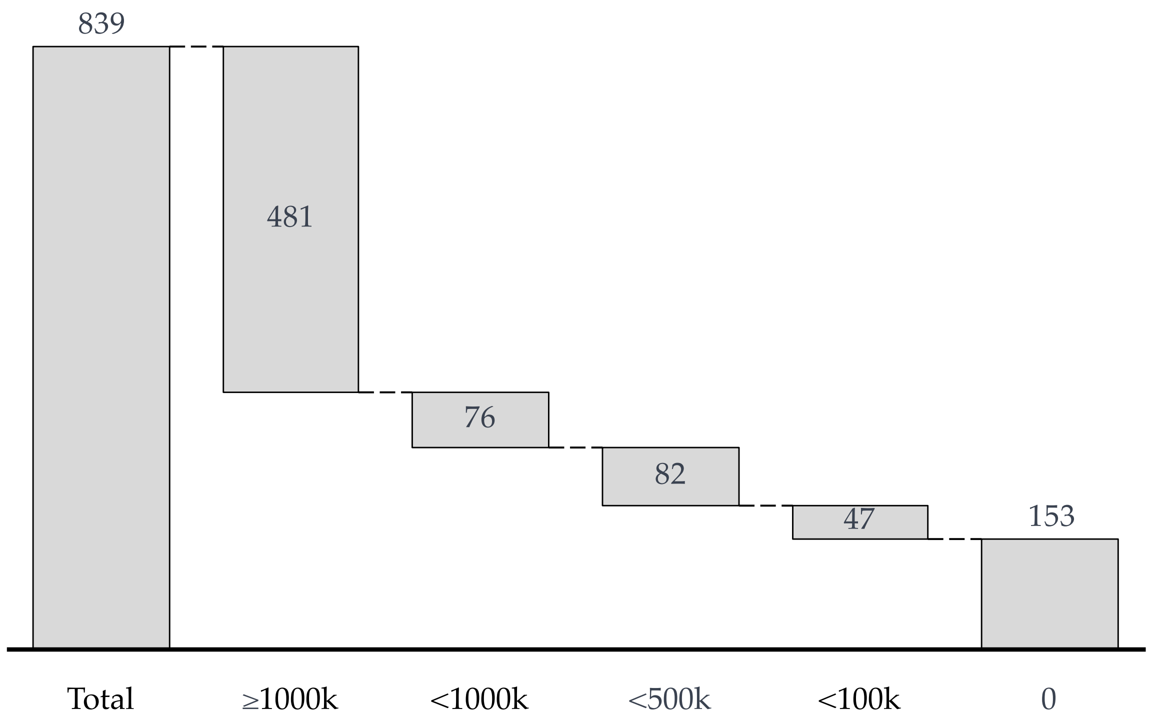
Country	USA	Germany	Sum
Operating	393	157	550
Failed	171	118	289
Sum	564	275	839

between media judgment and positive funding outcomes for startups is expected. Therefore, we include the time dependent variable "funding sum" into our dataset, which reflects for each observation, the accumulated funding sum to that point in time. Figure 7.2 displays the number of startups as a function of their total raised funding sum. We can see that 153 startups did not receive funding at all, 47 received up to USD 100.000, 82 up to USD 500.000, 76 up to USD 1.000.000, and 481 received more than 1.000.000.

The distribution of funding sums can be considered representative, as random selection and matching techniques to resemble a randomized experiment as close as possible are used. Also, the data is fairly evenly distributed across a wide range, which is a perfect precondition to identify a potential correlation to the media judgment.

Our last analysis in this study focuses on work experience prior to founding a startup.

FIGURE 7.2: Number of startups per funding sum.

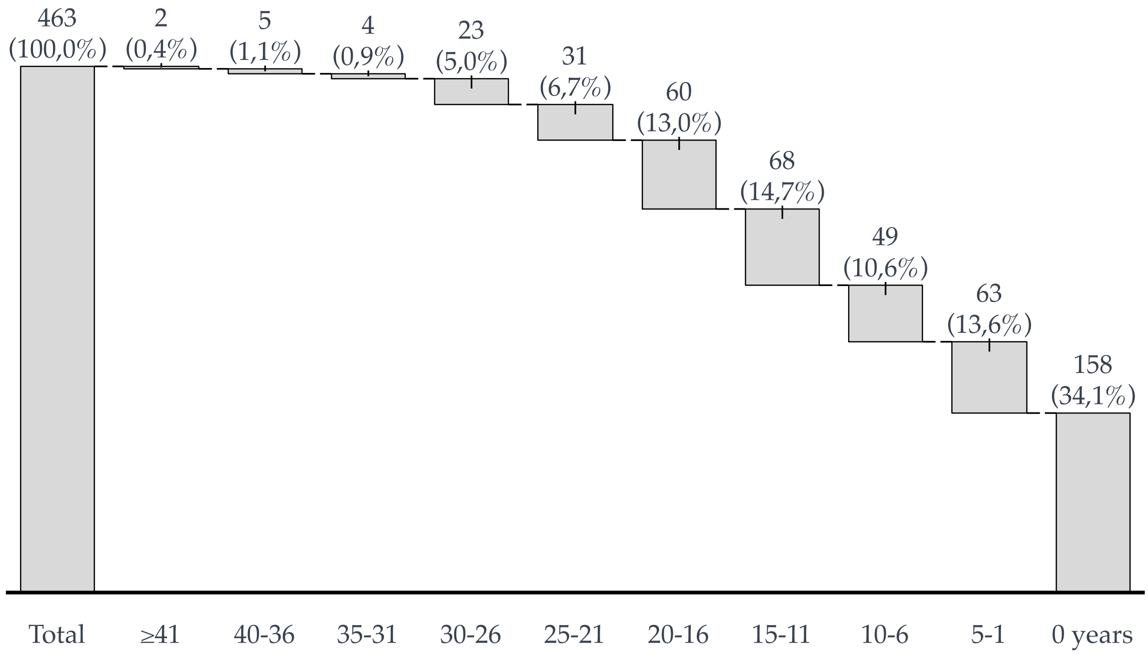


Therefore, publicly available data from social network services, where members of the founder team share their resume are used. One of the thereof gained variables is "cumulated work experience". It is the sum of the founder team's prior work experience in years. It can be seen as an experience or knowledge measure. As we have discussed in Chapter 7.1, we expect more experienced founder teams to receive more positive media judgment. Figure 7.3 shows, 41.7% of the founder team's have 10+ years of prior work experience, while 34.1% have no prior work experience.

7.4 Methodology and empirical framework

We aim to analyze media judgment of failed startups, as a function of a startup's novelty, collected funding sum, and overall experience of the founder's team. For the first analysis, a measure for novelty is needed. The following chapter hence introduces this measure and the general approach.

FIGURE 7.3: Cumulated work experience per founder's team.



7.4.1 Novelty

The measure for novelty needs to fulfill two requirements. First, it has to be measurable, i.e., a numeric value which allows comparing outcomes is necessary. Second, the measure has to describe a startup's newness or deviation from existing cultural beliefs. Previous research by Goldberg, Hannan, and Kovács (2016) uses two similarity measures to develop a distance measure, which fits these requirements. Their measure builds on two fundamental works by Jaccard (1901) and Shepard et al. (1987). Jaccard developed a similarity measure, which is defined as

$$J(m,n) = \frac{|m \bigcap n|}{|m \bigcup n|} = \frac{|m \bigcap n|}{|m|+|n|+m \bigcap n|}, \tag{7.1}$$

with m and n as categories of the set of categories C. The term $|m \bigcap n|$ stands for the number of startups of the intersection, in other words, which are associated with both categories (e.g., software development and online marketing). $|m \bigcup n|$ represents the number of startups associated with category m, n, or both. J takes values in the range [0,1], where 0 means perfect dissimilarity (i.e., no categorical overlap between the two), and 1 perfect similarity (i.e., both startups are labeled with the same categories). This measure allows to compare two labels of our category space C to each other. As a result, we obtain a similarity value. For example, the label finance occurs six times and technology occurs nine times. Five of these startups are labeled with both. The similarity of finance and technology is therefore given by

$5/(6+9-5) = 0.5$. As we are aiming for a distance measure between two labels, and hence startups, we need to derive the relationship between distance and similarity. Shepard manages to do so in the field of cognitive psychology (Shepard et al., 1987). His measure, backed by other research (Chater and Vitányi, 2003; Tenenbaum and Griffiths, 2001), is defined as

$$sim(m,n) = e^{-\lambda d(m,n)}, \quad \lambda > 0. \tag{7.2}$$

It posits a negative exponential relationship between similarity and distance, with a sizing or normalization factor for the similarity measure of λ. Solving Equation 7.2 with the help of Equation 7.1 for d, we obtain

$$d(m,n) = \begin{cases} -\frac{ln(J(m,n))}{\lambda} & \text{if } m \neq n, \\ 0 & \text{if } m = n. \end{cases} \tag{7.3}$$

From Equation 7.3 we can deduce that J(m,n) has to be greater than zero by definition of the ln-function. From this distance measure it is possible to develop a measure for novelty. Therefore, we follow an approach by Kovács and Hannan (2015), sketch their main ideas, and begin with a definition for novelty:

Novelty. A function of the average pair-wise distance, between the labels it gets assigned to.

The average pair-wise distance is given by

$$D(x) = \sum_{i \in C_x} \sum_{j \in C_x} k(i,x)k(j,x)d(i,j), \tag{7.4}$$

with the function $k(i,x)$, which is equal to one, when startup x is labeled with the category i, and 0 else. C_x are the labels associated with startup x. In case of a small distance between the categories of a startup in Equation 7.4, we would assume it have low novelty, and vice versa. Therefore, novelty is inverse proportional to D(x), and can be written as

$$N(x) = \begin{cases} 1 - \left(\frac{1}{1 + \frac{D(x)}{|C_x| - 1}} \right) & \text{if } |C_x| > 1, \\ 0 & \text{if } |C_x| = 1, \end{cases} \tag{7.5}$$

with $|C_x|$ as the number of different categories a startup x is associated with. From Equation 7.5 we can see, as the pair-wise distance for a given number of labels $|C_x|$ goes up, the same holds for our novelty measure N(x).

We have calculated $N(x)$ for the startups in our dataset with a total of 839 startups from the CrunchBase database in the respective timeline.

7.4.2 Empirical framework

Before we delve into the applied regression framework, we introduce the variables. In Table 7.2 the dependent variable, mean sentiment, includes for each quarter and startup the mean sentiment value of all observations (i.e., sentences). Our key explanatory variables are novelty, funding sum, and cumulated work experience. Similar to Study I, we include the DD variables, as well as control variables, such as country, observations, and age.

The following regression model is used to estimate media judgment as a function of a startup's novelty, cumulated funding sum, the team's work experience, prior founding experience, and age of the startup

$$S_{i,g,t} = \alpha + Z_{i,t} + \kappa TREAT_i + \lambda POST_{i,t} + \mu DD_{i,t} \cdot M_{i,t} + \phi \sum X_i + \tau \sum Y_{i,t} \ + \varepsilon_{i,t}, \quad (7.6)$$

here we follow the same notation for the sentiment as in Study I (see Chapter 6.4). $Z_{i,t}$ is the sum of five components. It consists of $NOVE_i$ as the in Chapter 7.5.1 introduced novelty of a startup i, $FUND_{i,t}$ which stands for the cumulated funding of a startup i at time t, $WORK_i$ is the founder's team work experience prior to founding, $PRIO_i$ represents prior founding experience of the founder's team, and $AGE_{i,t}$ the age of a startup i at time t. In addition, Equation 7.6 includes the difference in differences estimator $DD_{i,t}$ multiplied by the moderator term $M_{i,t}$. This term represents one of the five terms included in $Z_{i,t}$, which are subsequently included in our regression model in Table 7.3. In addition, Equation 7.6 includes the sums over time variant and invariant control variables.

Equation 7.6 is estimated using an ordinary least squares model without fixed effects.

TABLE 7.2: Summary of all variables in the dataset.

Variable	Explanation	Mean	Std. dev.	Min.	Max.
Dependent variable					
Mean Sentiment	A value between -1 and +1 representing the media sentiment of press articles based on sentences which name the startup in scope. The mean values are aggregated by quarters.	0.30	0.61	-1	1
Independent variable					
$NOVE_i$	Novelty of a startup.	0.70	0.39	0	1
$\log(FUND_{i,t})$	Logarithmized funding sum in $US raised by the startup until that quarter.	10.9	7.3	1	20.4
$\log(WORK_i)$	Logarithmized cumulated work experience of founder's team prior to startup in scope.	1.8	1.2	1	3.7
$PRIO_i$	Prior founding experience, number of founders with founding experience divided by the number of founders in the team.	0.08	0.26	0.00	1.00
$AGE_{i,t}$	Age of the startup in quarters.	21.61	16.27	0	80.00
Treatment	Dummy variable, which is 1 for treated startups and 0 for untreated.	0.274	0.446	0	1
Post	Dummy variable, which is 1 for sentences published after the startup has failed. Operating startups get a fictitious failure date from the matched failed startup.	0.489	0.500	0	1
DD	Difference in differences variable, which is 1 for treated startups and sentences published after the startup has failed, and 0 else. Operating startups always have a value of 0.	0.075	0.264	0	1
Country	Categorial variable, 1 stands for Germany, and 0 for USA.	0.028	0.166	0	1
log(Observations per quarter)	Number of observations (sentences) published for each quarter.	1.73	1.37	0	8.08

7.5 Results

In this chapter we investigate five different aspects. Starting with an analysis of media judgment and startup novelty, followed by funding sum, work experience, serial entrepreneurship, and age of a startup.

7.5.1 Media judgment and novelty of startups

The first hypothesis in this study states, that media judgment of failed startups will be more negative with the degree of novelty of the startup. To test this assumption, we analyze our data using a linear OLS-regression model, from Chapter 7.4.2. The results are presented in Table 7.3. In column (1) the estimator for the interaction term of the DD and the novelty variable appears to be statistically significant at a $p = 0.05$ level. Its value of -0.182 indicates a negative impact of a startup's novelty on the after failure media judgment. In other words, the more novel or atypical a startup is, the more harsh will media judge it in the event of failure and from there on. This effect holds for both countries. In the US though, the overall media judgment of failed novel startups remains positive, as the sum of all covariates becomes 0.119, whereas in Germany it is negative (-0.153). This is similar to what we observed in our first study (see Chapter 6). However, we find no significance for a general tendency of a more negative media judgment of novel startups with the independent variable $NOVE_i$. This might indicate that novel startups are only getting judged differently by the media in case of failure.

These findings support our hypothesis 1:

H1: *Novelty of business models negatively moderates a startup's media judgment after failure.*

The next paragraph presents the results of the influence of raised funding on media judgment.

7.5.2 Media judgment and funding of startups

Hypothesis 2 states, that the media judgment of a failed startups will be more negative with the amount of funding they receive from investors, since losing huge amounts of money might not be tolerated by societies and media. Column (2) in Table 7.3 presents the results. We find a significant negative impact of -0.020 at the $p = 0.01$ level. This means, with every USD 1.000.000 of funding the startup collects, the media judgment becomes less positive by the above mentioned amount. In other words, the more funding a startup raises throughout its journey, the less positive will media judge it in case of failure, and from there on. As funding sums reach up to triple digit million dollar values, the magnitude of this effect is large in comparison to the other variables included in the regression model. Similar to the novelty analysis, we find no significance for our independent variable log($FUND_i$). Funding might thus not be considered as relevant by the media, when it reports about running startups.

These results support hypothesis 2 from Chapter 7.1:

***H2:** The amount of founding negatively moderates the media judgment of startups after failure.*

7.5.3 Media judgment and work experience of startup teams

The third hypothesis states, that the media judgment of failed startups will be better for teams with experience in comparison to teams with low prior work experience. To test this hypothesis, we implement the cumulated work experience variable in our regression model (Table 7.3, column (3)). The interaction term with our DD variable has a positive sign, however, is not significant. Therefore, hypothesis 3 cannot be supported by our findings. Instead, the independent variable log($WORK_i$) is significant at a p = 0.01 level, indicating a more negative judgment with years of work experience. This means a less positive judgment of startups whose founders gained prior work experience.

Therefore, our hypothesis 3:

***H3:** Prior work experience of the team before founding positively moderates the media judgment after failure.*

cannot be supported or rejected.

7.5.4 Media judgment and serial entrepreneurship

Next, we explore how prior founding experience affects media judgment after failure. This analysis is reported in column 4 of Table 7.3. The estimate of the interaction term between DD and $PRIO_i$ variable turns out to be positively significant (0.139) at a p = 0.1 level. This suggests a positive influence of prior founding experience on media judgment after failure. At

TABLE 7.3: Regression analysis of media judgment and startup-specific characteristics.

	Dependent variable:				
	Mean_sentiment				
	(1)	(2)	(3)	(4)	(5)
Treatment	−0.046*	−0.047**	−0.045*	−0.038	−0.042*
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Post	−0.012	−0.014	−0.012	−0.012	−0.014
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
DD	−0.015	0.147	−0.067	−0.175***	0.491**
	(0.062)	(0.092)	(0.077)	(0.044)	(0.196)
Country (1 = Germany)	−0.272***	−0.272***	−0.272***	−0.270***	−0.270***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
$NOVE_i$	0.004	−0.005	−0.007	−0.008	−0.003
	(0.019)	(0.018)	(0.018)	(0.018)	(0.018)
$\log(FUND_i)$	0.001	0.002	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\log(WORK_i)$	−0.012**	−0.012**	−0.011*	−0.011*	−0.012**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
$PRIO_i$	−0.076**	−0.089***	−0.083***	−0.115***	−0.086***
	(0.031)	(0.031)	(0.031)	(0.036)	(0.031)
$\log(AGE_{i,t})$	−0.023**	−0.024**	−0.025**	−0.025**	−0.020*
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
log(Observations per quarter)	0.029***	0.029***	0.030***	0.030***	0.029***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
DD:$NOVE_i$	−0.182**				
	(0.075)				
DD:$\log(FUND_i)$		−0.020***			
		(0.006)			
DD:$\log(WORK_i)$			−0.032		
			(0.033)		
DD:$PRIO_i$				0.139*	
				(0.073)	
DD:$\log(AGE_{i,t})$					−0.210***
					(0.065)
Constant	0.427***	0.432***	0.436***	0.437***	0.422***
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
Observations	3,478	3,478	3,478	3,478	3,478
R^2	0.102	0.103	0.101	0.101	0.103
Adjusted R^2	0.099	0.100	0.098	0.098	0.100
Residual Std. Error (df = 3466)	0.386	0.386	0.386	0.386	0.386
F Statistic (df = 11; 3466)	35.734***	36.266***	35.237***	35.508***	36.207***
Note:					*p<0.1; **p<0.05; ***p<0.01

the same time, the estimated coefficient of $PRIO_i$ (without interaction) is -0.115 at a p = 0.01 level. These two results indicate a negative influence of prior founding experience on media judgment, which appears to get more then compensated by the positive influence after failure. These results contradict our hypothesis 4, where we expect media to judge failure after starting anew more negatively:

H4: *Prior founding experience negatively moderates the media judgment after failure.*

In fact, our results indicate that the opposite seems to be true.

7.5.5 Media judgment and age of startups

The last hypothesis states, that the media judgment will be less positive after failure, the longer the startup has existed. Table 7.3 includes the $AGE_{i,t}$ variable. The interaction term with our DD variable in column (4) shows a negative coefficient of -0.210 at a p-level of 0.01, indicating a negative effect of age and startup failure onto the media judgment. In other words, the older a startup is and fails, the more negative will the media judgment of this failure be. The estimate of -0.210 is the largest in this comparison. For further discussion, we refer to Chapter 7.6. Regarding hypothesis 5:

H5: *Media judgment of startups after failure, will be less positive for startups, which existed longer, than for those, who failed earlier.*

our findings support this assumption.

7.6 Discussion

What are potential implications of our results for startups? We begin with our main findings. Novel business ideas are indeed judged negatively after failure. From a cultural point of view, it is tempting to conclude, that society does not tolerate failure. An additional risk taking by the founder's team in choosing a novel business idea with high levels of uncertainty, due to scarce prior experience in this field, is even less acceptable. The media judgment could be a reflection of this thought, especially in countries with a high level of uncertainty avoidance, such as Germany. From a journalist's perspective one could argue differently. As he or she inevitably has not all the information about the reasons for a startup's failure, the journalist has to draw conclusions from limited knowledge. In this case, the journalist might compare the current to similar previous startups. The accessibility bias argues, that those information which are easier to retrieve from memory dominate the judgment (Iyengar, 1990). As more negative events are being remembered more accurately (Kensinger, 2007), journalist might tend to attribute the startup's failure to it's novelty. But one should be cautious with drawing final conclusions. As risk taking always accompanies entrepreneurship, media could have also rewarded the additional risk by a more positive judgment.

The second main finding is, that media judges failure of highly funded startups more negatively. From an organizational and cultural theory perspective, one could ascribe this finding to a deprivation of legitimacy and stigmatization of a startup, and hence, its founders. Legitimacy, as the "right to exist" (see Chapter 3.2), is withdrawn since the startup has not acted in accordance with normative rules and expectations (Devers et al., 2009; Dimaggio and Powell, 1983). It has lost tremendous amounts of money from investors, which relied on the startup and it's team. This negative judgment due to non-accordance can also be observed for successful entrepreneurs, as it is not the result of failure. Negative judgment also leads to stigmatization, which can have fatal consequences for the founders. As individuals bearing a stigma suffer in various dimensions (see Chapter 3.3), including health and well-being (Baumeister and Leary, 1995). Again, we have to be cautious, since not all startups publish their exact funding sums and some data was missing. Nevertheless, we are confident that our results are solid, as they are significant at a p-level of 0.05. An alternative explanation for negative media judgment of failed startups with high levels of funding could simply be an audience's tendency for more negative news coverage (Iyengar, Norpoth, and Hahn, 2004). Startups, which have received millions of dollars in funding are more likely to be known by audiences. Therefore, a failure of them may have the potential to attract lots of readers, which is in the interest of journalists.

The third main finding in this study is the negative impact of age on media judgment of failed startups. In a similar way as for the funding sum, a failure after years could be attributed to the founder's team. More specifically, journalist could judge a potential delayed filing for insolvency negatively. As a result of management mistakes, the startup might have not been competitive, and instead of closing down the business, management may have ignored the trend of an obvious failure. This behavior could be motivated by a fear of failure, and the associated stigmatization. In addition, entrepreneurs tend to connect personally to their startup, termination of this relationship is, therefore, accompanied by feelings of personal loss, which the entrepreneur tries to avoid (Shepherd, 2003a).

Most surprisingly is the media judgment of failed startups and their serial entrepreneurs. As the results indicate, those startups with serial entrepreneurs are judged more positively after failure, then teams without prior founding experience. One potential explanation for this result is, that experienced entrepreneurs rather fail due to external circumstances, which are less controllable by the startup, e.g., arising competition, or disruptive innovations. Inexperienced teams on the other hand, might commit more mistakes, and are consequently judged harshly by the media. The media, hence, would favor serial entrepreneurship over the potential negative bearing of not sticking to a venture. Malach-Pines et al. (2005) found that regarding entrepreneurs as social heros (high status), is reflected in the entrepreneurial activity of a society. It might thus be, that known entrepreneurs experience more positive media judgment.

Our research in this field underlies a limitation, as we can not tell whether the startup prior to the one in scope was successful or not. As the results, in general, indicate a positive influence of serial entrepreneurship on media judgment, one could also understand the positive influence of serial entrepreneurship after the failure as merely to the experience component of serial entrepreneurship. Contradicting this interpretation is the fact, that the cumulated work experience after failure has no significant impact on media judgment, and work experience, in general, is negatively significant at a p-level of 0.01. Hence, we recommend to treat these results with a healthy amount of caution and suggest further research for validation.

7.7 Conclusion

This study analyzes media judgment of failed startups as a function of its novelty, its received funding, the founder team's work experience, prior founding experience, and age. It aimed to add to the growing literature on startup failure (e.g., Politis and Gabrielsson, 2009; Simmons and Wiklund, 2011; Ucbasaran et al., 2010) by including the media perspective. For this purpose we build on prior research and derive five hypotheses, which are tested on a large scale dataset. Our key findings include a favor of media for typical business ideas over novel (distinctive) one's. Also, a more critical response of the media to startups which have been heavily invested in and failed. In addition, a more positive media judgment of serial entrepreneurs and a less positive with age. Our work extends the field of entrepreneurial failure and social judgment of organizations theory, and differs from prior work, which focuses on fields such as learning from failure (Cope, 2011; Shepherd, 2003b), attitudes towards failre (Politis and Gabrielsson, 2009), or causes of failure (Bruno and Leidecker, 1988). The field of social judgment of organizations considers media legitimacy as one of the most explored fields (Bitektine, 2011), but has to our knowledge not yet studied its impact for startups.

In this study, we have applied one approach for exploring the relationship between media judgment and startup-specific characteristics. It is an important building block in understanding how media impacts entrepreneurship, and how cultural effects influence this relationship. As we see, there is still much to explore.

Chapter 8

Study III - Entrepreneurial failure and implications for founders

The third study of this thesis seeks to investigate founder's motivations to start anew after failure. For this purpose, we measure influences of media judgment, prior founding experience, prior work experience, raised funding for the failed startup, and the novelty of its business model.

8.1 Introduction

Extensive research demonstrates the value of entrepreneurship for job creation, innovation, and economic development (see Lee, Peng, and Barney (2007); Carree and Thurik (2010)). As starting a venture is inevitably related to risk taking, a significant number of startups fail. Numbers range between 50% (Headd, 2003) and 92% (Marmer et al., 2011) within the first years, depending on the selection of startups. From an economic point of view, researchers regard a termination of a venture as necessary for wealth creation, and compare it to, e.g., a dynamic ecosystem, where death replaces senescent organisms (Coelho and McClure, 2005). However, if entrepreneurs decide not to found again after failure, gained knowledge, skills or networks might become obsolete. Thus, there are costs associated with failure. Depending on the source to consult, between 18% and 25% start again after their first venture closed or failed (Westhead et al., 2005; Westhead and Wright, 1998; Wagner, 2002). Failed entrepreneurs tend to blame their lack of success to the external environment (Eggers and Song, 2015). This could be, for instance, negative media reporting, resulting in negative side effects, such as less customers, or investors, who stop their support for the startup. To the best of our knowledge, the relationship between media judgment of failed entrepreneurs, and its impact on their likelihood to start a new startup, has not been investigated so far.

8.2 Theory and hypotheses

The implications of a venture failure for founders can lead to disadvantages in various dimensions. There are negative emotions, such as, pain, remorse, shame, humiliation, anger, guilt, and blame (see Cardon and McGrath (1999); Cope (2011); Harris and Sutton (1986); Shepherd (2003a); Singh, Corner, and Pavlovich (2007)). Researchers try to explain these emotions by comparing a venture failure to the loss of someone or something important, as entrepreneurs tend to connect personally to their venture (Shepherd, 2003b). In addition, Singh, Corner, and Pavlovich (2007) and Cope (2011) found evidence for losses associated with the entrepreneur's social network. They relate to marriage breakdowns and the termination of close relationships after failure. On a professional side, as failure can be a traumatic experience for an entrepreneur, it may deter subsequent venture activity (Politis and Gabrielsson, 2009). This is associated with stigma, in the form of a collective judgment of one's social network (Cope, 2011). Unfortunately, little is known about stigma from an entrepreneur's perspective, as most research focuses on the socio-economic level (Hampel and Tracey, 2016; Landier, 2005; Simmons, Wiklund, and Levie, 2014). We aim to support research in this field and seek to contribute to the social judgment of organizations theory, by investigating what drives entrepreneurs to start again, and what is hindering them to become serial entrepreneurs.

Serial entrepreneurs have sold or closed a business before starting a new one. They benefit in important ways from their past experience. For instance, do serial entrepreneurs often find it easier to raise funding for their ventures, as they are expected to be more skilled and socially connected in comparison to novice entrepreneurs (Hsu, 2007; Zhang, 2011). Also, they benefit from valuable learnings and experience, which they have gained in prior venture activities (Cope, 2011; Coelho and McClure, 2005; Gruber, MacMillan, and Thompson, 2008). A lack of experience is considered as potentially the most important factor for venture failure (Shepherd, 2003b). Also, do serial entrepreneurs often gain industry specific knowledge, which is considered a key determinant of success (Eggers and Song, 2015). However, there are also drawbacks from being a serial entrepreneur. Eggers and Song (2015) found, that serial entrepreneurs tend to blame external factors for their venture failure, and might not benefit from valuable learnings of the failure. In addition, they tend to change industry sectors for subsequent ventures, which increases opportunity costs of gained knowledge. Especially failed entrepreneurs act under the pretext of blaming external factors are less likely to change their leadership style in their next venture and may continue with deficits in strategic planning, and decision making (Eggers and Song, 2015). Also, learnings can be of no advantage, as they might be inferred improperly from a small number of experiences (Rerup, 2009). An empirical study by Cannon and Edmondson (2001) concludes that learning from failure is also very much dependent on the individual itself, since it requires an agreement and acknowledgment of personal mistakes to learn from failure.

The social judgment of organizations theory suggests, that media is capable of assigning or withholding legitimacy (Bitektine, 2011). It plays an important role in the evaluation process an audience performs before judging an entrepreneur. We, therefore, expect it to influence an entrepreneur's decision to start, again after a previous venture has failed, and posit:

> *H1: The more negative media judgment of a startup is after failure, the less likely will a founder be to found again.*

Media judgment may certainly be taken into consideration by entrepreneurs when they deliberate about whether or not to start again. Serial entrepreneurs have already decided to start anew. We are interested if novice entrepreneurs resign from entrepreneurship after their first venture attempt has failed, and compare them to serial entrepreneurs, who have started a venture before. Hence we posit:

> *H2: Entrepreneurs who founded before, are more likely to found again after failure, than novice entrepreneurs.*

Being a successful entrepreneur requires many different skills, experiences, and knowledge (Markman and Baron, 2003). Especially industry-specific experiences are considered key for running a successful venture (Agarwal et al., 2004; Chatterji, 2009). More experienced entrepreneurs are, therefore, more likely to be successful. In other words, a potential failure of entrepreneurs with lots of prior work experience, should be less expected by the entrepreneurs. Hence, we expect entrepreneurs with more work experience to be less likely to found again, than less experienced founders. Our hypothesis is:

> *H3: The less experience founders have, the more likely will they found again after failure.*

In Study II we identify a negative correlation between media judgment after a startup has failed, and the amount of funding it receives from investors. A potential explanation is media's mechanism of social control, introduced by Parsons (2013). According to this approach, media can be viewed as a monitoring service, which reports the illegitimate behavior of startups based on commonly agreed norms and regulations (Ingram and Rao, 2004). Deviation from this set of rules is consequently pointed out by the media. Hence, similar to hypothesis 1 we expect those founders whose startup has raised high funding amounts, to be less likely to found again after failure, and posit:

> *H4: The more funding a startup raises, the less likely will the founders be, to start again after failure.*

Entrepreneurs seek to identify unexploited market niches. Thereby they strive for products and services, with a unique selling proposition (USP). This is often accompanied with innovations, and a novel business model behind it. As results of Study II indicate, media judges novel business models more harshly after failure than traditional one's. We expect founders of more novel startups to be less likely to found again after failure, as the more negative media, and a potential deviation from cultural norms, reduce their legitimacy. We posit:

> *H5: The more novel a startup is, the less likely will a founder be to found again after failure.*

The next chapter will introduce the data underlying this study.

8.3 Data

The founder's data is from publicly available professional network profiles. In the following chapter, the preprocessing and derivation of new variables is explained.

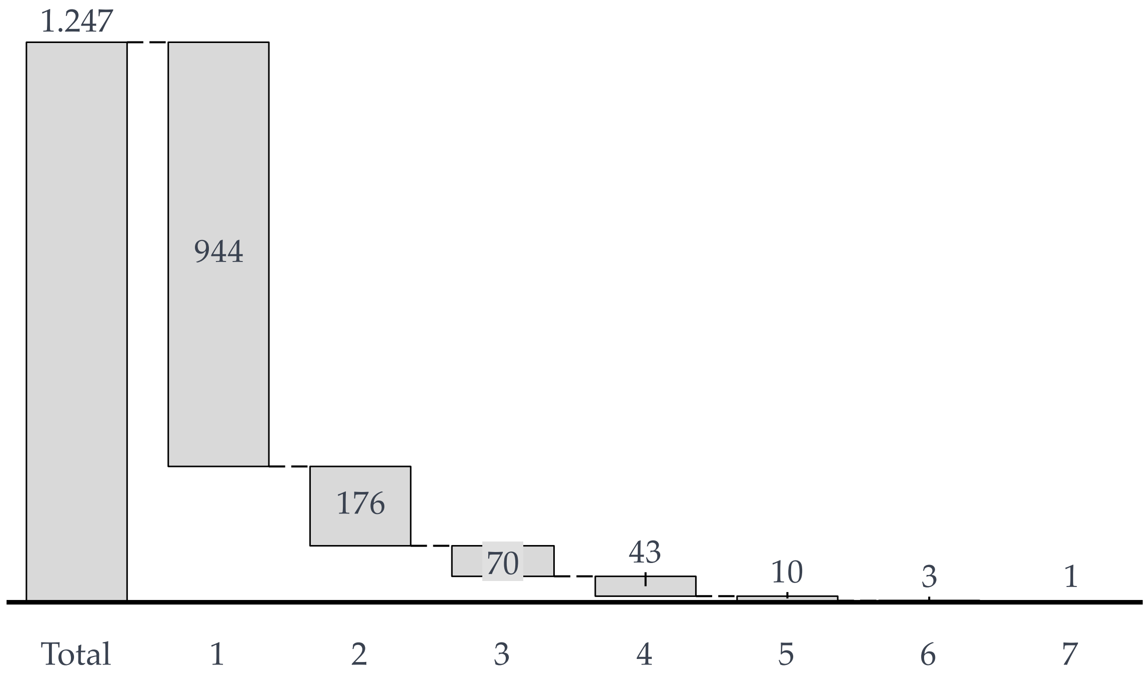
8.3.1 Collection and formatting of founder's data

As described in Chapter 3.3, the professional development of a person can be strongly influenced by stigmatization. This stigmatization can be the result of a negative experience from the past. Entrepreneurial failure satisfies the prerequisites for stigmatization, and therefore, might influence the professional career of a founder. This is reflected in the research hypothesis H1 (see above). To address this research question, publicly available data of the professional development of founders is collected.

Since the CrunchBase dataset does not list the names of the founders, they have to be searched manually via online search engines. We obtain one to seven founders per startup. The chart in

Figure 8.1 shows the distribution of the number of founders per startup. On average there are 1.4 founders per startup.

FIGURE 8.1: Number of founders per startup.



With the names of the startups and founders, we search for their professional network profile to obtain information concerning our four variables of interest (see Table 8.1). Some founders also decide not to mention the startup in their profile, one reason could be that they try to avoid the association with its failure. These founders are neglected in the dataset.

When the manual selection of the right founders is finished, the following step is to code the founder's public profile manually. Since this thesis focuses on the influences of failure on founder's professional development, variables are created accordingly. They are displayed in Table 8.1.

With these additional variables it is possible to analyze the likelihood to found again and its dependence on prior work or founding experience in Chapter 8.5. Table 8.2 summarizes the resulting founders data.

8.3.2 Descriptive statistics

To get a better understanding of the underlying data this section provides descriptive statistics of our data. Table 8.2 lists the number of startups and founders included in the data. In total

TABLE 8.1: Independent variables created from the professional network profile data of founders.

Variable name	Description	Type
Work experience	The number of prior years of work experience before founding a startup.	Integer
Founded before	The number of startups founded prior to startup in focus.	Categorial variable
Founded again	Identifier for founded again after startup in focus has failed.	Categorial variable
Team size	Total number of founders per startup.	Integer

we search for founders from the 1272 startups in our dataset, and obtained 1782 names. From them, we are able to identify 1364 publicly available profiles, 699 of them are complete, representing 489 startups in total.

TABLE 8.2: Overview of number of startups, as well as founders in the dataset.

	USA	Germany	Total
Total No. of startups	790	482	1272
Thereof identified founders	919	863	1782
Thereof with publicly available profile	734	630	1364
Thereof with complete data	413	286	699
Associated number of startups	307	182	489

With regards to our research objective, we search specifically for failed founders, and their likelihood to found again. Table 8.3 shows the 258 failed founders in our dataset with complete information on their professional background, split by country and the decision to found again or not. The numbers show, more than 50% of the failed founders decide to start again, with a 10.6% higher rate in the US (61.4% compared to 50.8% in Germany).

With these first impressions of our data we now turn to a regression model.

TABLE 8.3: Overview failed founders per country and found again decision.

	Founded again			
	Yes	No	Rate	Total
USA	81	51	61.4%	132
GER	64	62	50.8%	126
Total/Average	143	115	58.5%	258

8.4 Methodology and empirical framework

Our dependent variable "Founded again" is binary, and takes the values 0 and 1, where the latter indicates, that the founder decides to found again after a venture failure. A binary dependent variable suits perfectly to a so-called probit regression model. In this paragraph, we will introduce the basic ideas behind this model, by following Wooldridge (2010), and Cameron and Trivedi (2005).

As we are interested in a founder's likelihood to found again after a venture's failure and its determinants, we focus on failed entrepreneurs. Our dataset includes 258 failed and 451 unfailed founders, with all variables available. To make a proper estimation of what influences the found again chances after failure, we have to ensure that the underlying sample is not biased. For instance, the sample of failed founders could be dominated by startups with more novel business models, which are judged harshly by the media after failure (see Study II). Without a correction in form of a sample selection process, one could falsely conclude that failed founders do not intend to found again due to negative media judgment. Therefore, we extend the probit regression model with a prior sample selection process. This two steps approach is introduced in the next paragraphs.

First, we need to define a binary response model (probit), which returns values such that

$$FA_i = \begin{cases} 1 & \text{if a founder founds again,} \\ 0 & \text{if a founder does not found again.} \end{cases} \tag{8.1}$$

We approach this by introducing the latent variable (unobservable)

$$FA_i^* = \vec{\beta} \cdot \vec{x}_i + \varepsilon_i. \tag{8.2}$$

Here, $\vec{\beta} \cdot \vec{x}_i$ is the scalar product of the parameter vector $\vec{\beta}$ and a covariate vector $\vec{x}_i$. The error term ε_i is assumed to be normally distributed with a variance σ^2, i.e., $\varepsilon_i\ N(0, \sigma^2)$. With Equation 8.2, we can extend Equation 8.1, and obtain

$$FA_i = \begin{cases} 1 & FA_i^* > 0, \\ 0 & FA_i^* \leq 0. \end{cases} \tag{8.3}$$

Instead of searching for the exact value for FA_i^*, our main interest is the sign. To calculate the probability of a failed founder i to found again, we can write

$$P(FA_i = 1|\vec{x}_i) = P(FA_i^* > 0|\vec{x}_i) = P(\varepsilon_i \leq \vec{\beta} \cdot \vec{x}_i|\vec{x}_i). \tag{8.4}$$

With the assumption of the error term ε_i to be normally distributed, we can rewrite Equation 8.4 it, and obtain

$$P(FA_i = 1|\vec{x}_i) = \int_{-\infty}^{\vec{\beta} \cdot \vec{x}_i} \phi(\vec{\beta}\vec{x}_i) dx_i = \Phi(\vec{\beta}\vec{x}_i). \tag{8.5}$$

Equation 8.5 can be seen as the likelihood function of a founder to found again, with a covariates vector of $\vec{x}_i$. As we mentioned above, this function might be biased, due to the selection of failed founders. What we are really looking for is $P(FA_i = 1|\vec{x}_i|OS) = 1$, with the selection function for the operating status, in this case returning the operating status

$$OS(\vec{x}) = \begin{cases} 1 & , \vec{\gamma}\vec{x} + \eta > 0, \\ 0 & , \vec{\gamma}\vec{x} + \eta \leq 0. \end{cases} \tag{8.6}$$

$OS(\vec{x})$ equals 1 for failed, and 0 for operating startups. η is the associated error term, $\vec{x}$ is assumed to be always observable, with the parameter vector $\vec{\gamma}$. FA can only be observed, if $OS(\vec{x}) = 1$. In order to obtain $P(FA_i = 1|\vec{x}_i)|OS = 1)$, some mathematical relations lead to

$$P(FA_i = 1|\vec{x}|OS = 1) = \frac{1}{\phi(\vec{\gamma}\vec{x})} \int_{\vec{\gamma}\vec{x}}^{\infty} \phi \left[\frac{\vec{\beta}\vec{x}_i + \rho\eta}{\sqrt{(1-\rho^2)}} \right]. \tag{8.7}$$

The parameter ρ is the correlation between the two error terms ε and η. Its value ranges from -1 to +1, the closer the value is to zero, the weaker the correlation. In Equation 8.7, we assume the error terms to be independent of the explanatory variables.

Solving Equation 8.7 is a two step process. First, we solve Equation 8.6 for γ using a probit model for $OS(\vec{x})$ on $\vec{x}$, then we estimate $\vec{\beta}$ and ρ with the help of Equation 8.7. The results of the estimation will be presented in the next chapter. Before we do so, we introduce and motivate the probit regression model as in Equation 8.6. In order to explain the operating status we used the age, number of funding rounds, total number of sentences about the startup, and the prior work experience of the founder's team. It seems reasonable to include the age variable, since the likelihood of a startup to fail should increase with its age. The number of funding rounds indicates how attractive a startup is from an investor's perspective, and can be seen as a quality measure. It is thus useful to include. Similar function the number of sentences, the more attention a startup receives should, in general, reflect the public's interest, and hence can be viewed as a quality measure. Prior experience has been shown to be critical for success (Agarwal et al., 2004), thus we include it as well. Our probit model can be written as

$$OS(Age, NFR, NS, WE) = \begin{cases} 1 & , \alpha Age + \kappa NFR + \lambda NS + \tau WE + \eta > 0, \\ 0 & , \alpha Age + \kappa NFR + \lambda NS + \tau WE + \eta \leq 0. \end{cases} \tag{8.8}$$

With Age as the age variable, NFR as the number of funding rounds, NS as the number of sentences, and WE for work experience of the founding team. The estimations are made with the sampleSelection package available for R-Studio (see Ott Toomet and Arne Henningsen (2008) for further information). Table 8.4 shows the results of our estimation.

TABLE 8.4: Probit model

	Dependent variable:
	Operating status
Age	-0.084^{***}
	(0.010)
NFR	-0.200^{***}
	(0.048)
NS	-0.337^{***}
	(0.036)
WE	-0.235^{***}
	(0.063)
η	-2.349^{***}
	(0.227)
Observations	699
Note:	*p<0.1; **p<0.05; ***p<0.01

All estimators are significant at a p = 0.01 level, which indicates their relevance for predicting the operating status of a startup. The next step is to estimate Equation 8.7. This will be part of the following results chapter.

8.5 Results

We know report our findings in two ways: A bivariate probit model with a prior selection process, in this case, we use a maximum likelihood selection model. In addition we calculate the marginal effects, to quantify the results. The corresponding values are displayed in each column, below the estimated effect, and its standard error. The number of observations equals the number of failed founders (258). Table 8.5 builds up five regression models in total. We begin to investigate the likelihood to found again after failure by taking into consideration the observed change in media judgment, i.e., before and after failure. Then, we focus on prior founding activity and the chances to found again. Our third regression considers prior work experience, followed by the raised amount of funding by the founder's team. Finally, we conclude with the startup's novelty.

8.5.1 Found again chances and media judgment

For our analysis of the link between chances to found again after failure, and the media judgment experienced during a previous entrepreneurial endeavor, which resulted in failure, Table 8.5 column (1) presents the results.

The "Sentiment change" variable is defined as the difference between prior and after failure media judgment, it becomes -0.482 at a p = 0.01 level, and -0.881 at a p = 0.1 level in the final model (column 4). In order to better interpret these values, we have calculated the marginal effects to -0.098 and -0.159 repectively. Since our values for the sentiment range between -1 and +1, a negative significant estimator can be interpreted as a higher likelihood for founders to start again, if the change in media judgment (before to after failure) is greater. In other words, the more the media criticize a startup's failure, the more likely will the founder be to found again. In our case up to 9.8% or 15.9% respectively. This is rather astonishing, since according to the Global Entrepreneurship Monitor, less failure tolerance should result in lower levels of entrepreneurship. We will discuss this in more detail in Chapter 8.6. Referring to our hypothesis, where we assume positive media judgment to foster founders to start anew, we do not find evidence to support it. Apparently, the opposite seems to be true, and founders seem to develop a "now more than ever" mentality after failure.

TABLE 8.5: Probit model with selection - analysis of found again chances.

	Dependent variable:				
	Found again				
	(1)	(2)	(3)	(4)	(5)
Country Dummy	−0.194	−0.304	−0.334*	−0.208	−0.283
	(0.164)	(0.196)	(0.203)	(0.276)	(0.390)
	-0.039	-0.068	-0.068	-0.042	-0.051
Team size	0.081	0.071	0.065	0.041	−0.096
	(0.073)	(0.079)	(0.080)	(0.120)	(0.204)
	0.017	0.014	0.013	0.008	-0.017
Age	−0.012	0.002	−0.002	−0.011	−0.072*
	(0.020)	(0.020)	(0.021)	(0.028)	(0.038)
	-0.002	0.017	-0.017	-0.002	-0.020
Sentiment change	−0.482**	−0.344	−0.351	−0.520	−0.881*
	(0.219)	(0.262)	(0.263)	(0.346)	(0.483)
	-0.098	-0.070	0.071	-0.105	-0.159
Founded before		1.475***	1.473***	2.248***	2.092***
		(0.256)	(0.256)	(0.530)	(0.700)
		0.300	0.300	0.452	0.378
Work experience			−0.010	−0.014	−0.018
			(0.014)	(0.016)	(0.018)
			-0.002	-0.003	-0.003
log(Total funding)				0.002	−0.016
				(0.062)	(0.090)
				0.000	-0.003
Novelty					−0.598*
					(0.325)
					-0.108
Constant	0.324	−0.086	−0.005	0.021	0.460
	(0.226)	(0.283)	(0.302)	(0.477)	(0.735)
	-0.039				
Observations	258	258	258	258	258
Log Likelihood	−516.304	−491.160	−490.868	−414.753	−310.729
ρ	−0.329 (0.215)	−0.181 (0.249)	−0.116 (0.273)	−0.145 (0.353)	0.340 (0.340)
Note:					*p<0.1; **p<0.05; ***p<0.01

8.5.2 Found again chances of serial entrepreneurs

Our second hypothesis for this study states, that entrepreneurs who found a startup before and leave it for whatever reason, are more likely to found again after failing with another startup. The "Founded before" variable shows a positive significance at a $p = 0.01$ level for all four models (column 2 to 5). The marginal effects indicate an increased likelihood to found again after failure, when the founder has founded before of 30.0 to 45.2%, depending on the model. This effect is between 2.37 (0.378 over -0.159 in Model 5) and 4.3 (0.452 over -0.105 in model (4)) times larger, than the change in media sentiment. In model 2, the effect from change in media sentiment is no longer significant, as we add the "Founded before" variable, suggesting a rather small effect of media on the likelihood to found again.

8.5.3 Found again chances and experience

As experience is one of the key metrics for successful entrepreneurs, we posit a correlation between work experience before founding, and the likelihood to found again after failure, in hypothesis 3. From the results in Table 8.5 we find no support for this hypothesis. Our independent variable "Work experience" varies between -0.002 and -0.003 with no statistical significance. This is rather astonishing, as we expect experienced founders, with lots of priorly gained industry experience, to display more confidence in their own abilities. A failure should, therefore, be less expected by them, resulting in a lower likelihood to found again.

8.5.4 Found again chances and raised funding

When startups receive funding, they manage to convince others of their idea. In Study II we see, that media judgment after failure becomes more and more negative, with the amount of funding a startup receives. Media takes a monitoring role and withdraws legitimacy from those who do not act in accordance with commonly accepted rules and norms (see Chapter 8.2). Hence, we expect those founders, who manage to raise lots of funding before they fail, to be less likely to found again. Our results do not support this assumption. The estimator log(Total funding) shows no significance and its marginal effects are close to zero (column 4).

8.5.5 Found again chances and startup novelty

More novel business models are the result of a creative idea. Therefore, founders of a novel startup could creatively consider new business ideas after failure in order to found again.

Contradicting this thought is the fact, that novel business ideas, which fail, are judged harsh by the media, and this might reduce the likelihood to found again, since it withdraws legitimacy (see Chapter 7). Our independent variable novelty is negative (-0.108), and significant at a p = 0.1 level (see column 5). This means, the more novel a startup is, the less likely are its founders to start again after failure. Hence, this finding supports hypothesis 5.

8.6 Discussion

What are the implications of our findings for founders or people, who consider founding a startup as a career choice? Following our hypotheses, our results indicate no correlation of negative media judgment and a reduced likelihood to found again. Surprisingly the opposite seems to be true. Media's mechanism of social control, i.e., by withdrawing legitimacy via critical media reporting, seems not to affect the founders. They seem to develop a "now more than ever" mentality, as their likelihood to found again increases with more negative media reporting. This study supports a call from Hudson (2008), who suggests further research in the filed of legitimacy, as organizations persist, even if an evaluating audience is not granting legitimacy. Apparently, the same holds for startup founders, which is a novel finding in the field of entrepreneurship. A potential explanation for this finding could be overconfidence (see Chapter 2.3). Overconfident entrepreneurs expect negative events to rather occur to others then themselves. These entrepreneurs are more likely to start anew. Unfortunately, we can not control for confidence in our dataset, so we can not further verify this potential explanation. Attribution theory suggests that negative media judgment could lead to high stigma of these entrepreneurs. The observed "now more than ever" mentality could, therefore, be interpreted, as an attempt of the founders to change the public's attribution to a positive direction.

Our second hypothesis stated, that serial entrepreneurs are more likely to found again after failure. As we have seen in Table 8.5, this is the case. Entrepreneurs gain valuable knowledge, while building a company, which is beneficial for future founding activities and their success (Agarwal et al., 2004). Prior experience might trigger restart intentions. Interestingly, as we add the independent variable for work experience, also the country variable becomes significant. This shows, that differences between the two countries exist. In the US it seems more likely to found again after failure. Work experience itself, is not found to effect the likelihood to found again. Apparently, young founders with low experience do not differ in this dimensions, from more experienced individuals, who decide to found a company.

In Study II we see, that the more funding a startup raises, the worse is its media judgment after failure. Hence, we expect high funding amounts to influence the decision to found again. With the underlying data, we do not find evidence to support this assumption. In other words, it makes no difference for the founders, if their startup failure is connected to an investor's loss of huge amounts of money or not. This differs from the media perspective, which clearly differentiates between "small" and "big" failures. One way of interpreting this is to assume that personal attributes of the founders are more relevant for the decision to start anew.

A negative influence on the likelihood to start again has the novelty of a startup's business model. Founders, who are creative in developing a new business model, are less likely to found again. This finding is similar to the finding of Study II, which indicates a more negative media judgment after failure of more novel business models. A potential explanation for founders with innovative business models to be less likely to start again could be, that for them the failure seems less expected. Since venture capitalists prefer startups with novel business models to invest in, founders with such business models may expect to be more attractive for them, than traditional one's (Aldrich and Fiol, 1994). Failure might thus impact their self-concept (see Chapter 3.3) more than that of other founders. As they are less prepared, they might not develop self-protection strategies, and hence suffer more from failure associated stigma (Janoff-Bulman and Frieze, 1983; Jones, 1984). As a consequence, these founders might reorientate their professional career.

8.7 Conclusion

The present study is designed to identify determinants, which raise or lower the chances to found again after failure. As we see, negative media judgment and prior founding experience raise the likelihood, and a startup's novelty lowers the likelihood to found again. Most surprisingly is the fact that founders seem to develop a "now more than ever" mentality, as more negative media reporting results in a higher likelihood to found again. This implies, that the traditional concept of legitimacy in the field of organizational theory is not applicable here, or has to be adapted. It therefore supports a call from Hudson (2008) to change our current perception of organizational theory and legitimacy, who argue that even without legitimacy organizations can continue to exist. Our findings suggest, that rather personal or startup-specific determinants seem to matter. Media seems to be less important.

Our study can be seen as a first step in this field. Further investigations should consider a potential bias in the selection of founders. As some founders obviously do not mention their connection to a former failed startup in their professional network profile, we had to exclude them from this analysis. To what extent this influences our results, should be the goal of future research in this field. Another limitation is the size of the dataset. In total, we are able to identify 258 founders with complete profiles, for this analysis. A dataset of this size, can easily lead to false inferences. Our results should, therefore, be treated with a healthy degree of caution.

Practical implications of our analysis can be drawn. It appears as if engagement in entrepreneurial activities is strongly linked to the personal characteristics of the individual. In other words, media may influence founders less than their own skills, drive and motivation. This is a good sign for societies, which display higher levels of uncertainty avoidance, and less failure tolerance than others. Their policy makers can focus on programs to develop entrepreneurial skills, rather than aiming for a more failure tolerant societies, which often is more difficult to achieve.

Chapter 9

Discussion

The overall aim of this thesis is threefold. First, we want to quantify entrepreneurial friendliness by using media sentiment of press articles as a surrogate. Our focus is the failure event, and its effect on media judgment of startups. In order to classify our results, we compare the US to Germany, as they are considered to verify substantially in their entrepreneurial friendliness. Second, our intention is to understand differences in media judgment, due to startup-specific characteristics. Third, our last study investigates the influence of media judgment, startup- and team-specific characteristics, on the likelihood to found again, after failure. We do so by applying supervised machine learning sentiment analysis to sentences, extracted from press articles, naming selected failed and operating startups in the US and Germany within 1995 and 2015. Our results show significant differences in media judgment between the two countries, before and after a startups fails. This chapter summarizes all findings from three studies, and discusses their potential explanations, implications and gives suggestions for future research in this field. An overview of or hypotheses and the corresponding results, is given in Table 9.1.

Startups are temporary organizations used to search for a repeatable and scalable business model (Blank, 2012). They aim for fast growth through innovative products or services. Organizational theory states, startups need legitimacy in order to get accepted by greater audiences, such as customer, employees or investors (see Chapter 3.2). Some of these audiences react immediately to the release of new information about a firm's performance (Bitektine, 2011). Media is thus equipped with a mechanism of social control, as it is capable of assigning or withholding legitimacy (Hamilton, 1995). By measuring the sentiment or judgment of startups in press articles, we can deduce its legitimacy, and draw conclusions about entrepreneurial friendliness. With our analysis in Study I, we see, media judgment of startup failure is less positive compared to similar operating startups. However, the observed reduction differs between the US and Germany. While failure is still judged positively in the US, it is judged negatively in Germany. This means, US media considers startup failure as

TABLE 9.1: Overview of the included studies, their hypotheses, and our findings.

Study	Hypotheses	Supported by findings
Media judgment and cultural differences of failed startups	Media judgment of startups after failure is less positive than media judgment of similar startups, which did not fail.	Yes
	Media judgment by German media on German startups after failure is less positive than US media judgment by US startups.	Yes
	Media judges native startups more positively than foreign startups.	In part, only true for US
Startup specific differences in media judgment of failed startups	Novelty of business models negatively moderates a startup's media judgment after failure.	Yes
	The amount of founding negatively moderates the media judgment of startups after failure.	Yes
	Prior work experience of the team before founding positively moderates the media judgment after failure.	No evidence
	Prior founding experience negatively moderates the media judgment after failure.	No, opposite is true
	Media judgment of startups after failure, will be less positive for startups, which existed longer, than for those, who failed earlier.	Yes
Entrepreneurial failure and implications for founders	The more negative media judgment of a startup is after failure, the less likely will a founder be to found again.	No, opposite is true
	Entrepreneurs who founded before, are more likely to found again after failure, than novice entrepreneurs.	Yes
	The less experience founders have, the more likely will they found again after failure.	No evidence
	The more funding a startup raises, the less likely will the founders be, to start again after failure.	No evidence
	The more novel a startup is, the less likely will a founder be to found again after failure.	Yes

legitimate, and seems to appreciate entrepreneurial risk taking. Entrepreneurs can be less afraid of the negative effects often accompanied with a startup's failure, such as devaluation (Wiesenfeld, Wurthmann, and Hambrick, 2008), blame (Moulya and Sankaranb, 2000), or disadvantages for future career options (Cannella, Fraser, and Lee, 1995). Extending this point with the results of Study II, we can even assert that adding the negative impact of a startup's novelty and high levels of raised funding will not change the positive judgment to the negative in the US. A possible explanation for these results is to attribute it to the fairly high confidence of the US society in dealing with uncertainty. Lonner, Berry, and Hofstede (1980) found an uncertainty avoidance level of 46 on a scale from 1 to 100. This value is considered below average and smaller in comparison to most European and Asian countries (e.g., Germany 65, France 86, Japan 92, or South Korea 85). Countries with low levels of uncertainty avoidance value practice over principles, and are more forgiven when risk taking turns out badly. In addition, the US displays a high level of individualism vs. collectivism. Its value of 91 on the same scale is one of the highest measured, which becomes obvious when we compare it to other countries (e.g., Germany 67, France 71, Japan 46, or South Korea 18). Societies with high levels of individualism are loosely-knit, and people tend to look after themselves, and their families mainly. Members of these societies are self-reliant and favor initiative. Our results support this view, as there seems to be no intention to criticize founders for being initiative and taking risk in order to be more self-reliant. The results obtained for Germany are different. Here we have seen a negative media judgment after failure, indicating a withdrawal of legitimacy. Entrepreneurs are thus more likely to experience the negative influences of failure. Its score of 65 in uncertainty avoidance is indicating a slight preference for it. In combination with an above average value of 67 for individualism, Germany seems to support initiative, but judges failure negatively.

With this in mind, we except both societies to differently influence the self-concept of founders (see Chapter 3.3). Whereas in the US, positive media judgment after failure might encourage the founder to start anew, negative media judgment in Germany should lower the probability. However, the results of Study III do not support that assumption. No significant difference between the two countries is observed. Surprisingly, the data indicates a different result. According to our regression Table 8.5, the chances to found again seem to grow with the negativity of the sentiment change a startup experiences after failure. It suggests a "now more than ever" mentality of founders. Founders might regard gained learnings and experiences as more relevant, than a potentially negative judgment of a general public. This in turn can be explained with fairly high values of individualism in both countries. In these societies, other people's views are regarded as less relevant. There are, however, other possible explanations. For instance, rising found again chances with negative media judgment could be explained with a selection effect. As we have taken our data on founders from publicly available profiles, some of them might not mention the former failure. Due to self-protection, they might wish not to be associated with these startups, since they might expect negative consequences, such as stigmatization (Jenkins et al., 2014). We, therefore, recommend further research in this field.

Chapter 10

Conclusion

The US appears to be a breeding ground for highly innovative and successful startups. Some of them are among the most valuable companies on earth. Scholars explain this phenomenon in part with the entrepreneurial friendly culture in the States. Surveys (Singer, Amorós, and Moska, 2014), and structured interviews (McGrath, 1999) seem to support this assumption. Germany is considered less entrepreneurial friendly (Singer, Amorós, and Moska, 2014). However, is this really true? No large scale study exists, which aims to quantify how entrepreneurial friendliness the two countries are, especially how these societies treat those, who take risk and fail with their startup. We shed some light by combining social judgment of organizations theory, stigmatization theory, and cultural theories such as Hofstede's dimensions and in-group bias. We seek to add to the emerging field of entrepreneurship research, especially of entrepreneurial friendliness of societies. Therefore, we investigate the media judgment of startups, with special focus on failure. Our data driven approach, is able to quantify cultural differences in startup judgment, and perception, where so far only surveys and qualitative studies exist. Our findings support the wide spread believe of cultural differences in the entrepreneurial friendliness between the two countries, and explain in part, why the US is considered ahead of Germany in their startup eco-system.

The thesis goes some way towards enhancing our understanding of the above-mentioned theories in the context of entrepreneurship. Social judgment of organizations theory considers legitimacy of organizations as key for the acceptance by a general public, and to become successful (Bitektine, 2011). The enclosed studies show, that fundamental differences between the US and Germany exist. According to them, failure in the US is regarded legitimate, since media judgment about failed startups remains positive. No blame for failing with a startup is transmitted by the media. Legitimacy is not merely granted through positive media judgment, it is also not prominently positioned through failure stories in the media. In other words, media legitimacy is granted in two ways. First, through positive media judgment, and second, through less reporting. The opposite can be concluded for Germany.

More than one-third of media reporting occurs after a startup fails. Media's function of social control shows up through frequent reporting about failure as illegitimate.

Further research is necessary to explain the newly arousing questions through the three studies. For instance, why failed entrepreneurs seem to be more likely to found again, when media reporting about their failure is more negative. Another question is why the data shows no in-group bias for Germany. We expect Germany to be more likely to display the bias, as the US, since the German society shows a greater level of uncertainty avoidance, which is connected to a favor for the in-groups (see Chapter 3.4). These questions show, there is still much to explore.

Appendix A

Precleaning of articles (delete pictures and tables)

```
' DELETE PICTURES AND TABLES
'
'
Sub Delete_Pictures()
Dim MyFile As String
Dim MyObj As Object, MySource As Object, file As Variant
myPath = "C:\Users\Matthias Jacobi\Downloads\GEROPERATINGUSSOURCES" & "\"
myExtension = "*.DOC"
Application.DisplayAlerts = False
MyFile = Dir(myPath & myExtension)
Do While MyFile <> ""
Set oDoc = Documents.Open(FileName:=myPath & MyFile) 'Visible:=False
a = oDoc.InlineShapes.Count
Do While ActiveDocument.InlineShapes.Count > 0
oDoc.InlineShapes(1).Delete
Loop
Do While ActiveDocument.Tables.Count > 0
oDoc.Tables(1).Delete
Loop
oDoc.Close SaveChanges:=True
MyFile = Dir
Loop
Application.DisplayAlerts = True
End Sub
```

Appendix B

Transfer articles into panel data structure

```
Sub devide_articles()
'Dim oExcel As Excel.Application
'Dim ExcelWB As Workbook
Dim MyFile As String
Dim MyObj As Object, MySource As Object, file As Variant
Dim NmbrOfArt, i As Integer
Dim k As Long
Dim NmbrOfArtStr, StartWord, EndWord As String
Dim c As Range
Dim blnFound As Boolean
Dim oExcel As Object
ExcelPath = "C:\...\01_Lexis Nexis Articles\02_USA\01_Failed Startups" & "\"
WordPath = "C:\...\01_Lexis Nexis Articles\02_USA\01_Failed Startups" & "\"

Set oExcel = CreateObject("Excel.Application")

start_with_File = "USA_11i Solutions_1-16.DOC"
'!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!"


'ExcExcel open
Excelname = "USA F USA S sentences3.xlsm" '!!!!!!!!!!!!!
Set ExcelWB
= oExcel.Workbooks.Open(FileName:=ExcelPath & Excelname, ReadOnly:=False)
oExcel.Visible = True
MsgBox ("set excell book")
```

```
Application.ScreenUpdating = False
'Word path
myExtension = "*.DOC"

MyFile = Dir(WordPath & myExtension)
If start_with_File <> 0 Then
Do Until MyFile = start_with_File
MyFile = Dir
Loop
'MyFile = Dir
End If
'For i = 1 To 4
'Next i
Do While MyFile <> ""
last_row = ExcelWB.Sheets("Articles").Range("j1").Value
k = last_row
Application.DisplayAlerts = False
Set oDoc = Documents.Open(FileName:=WordPath & MyFile) 'Visible:=False
'nmbr of art. def.
If InStr(1, MyFile, "#") > 0 Then
point1 = InStrRev(MyFile, "-") + 1
point2 = InStr(1, MyFile, "#")
NmbrOfArtStr = Mid(MyFile, point1, point2 - point1)
NmbrOfArt = CInt(NmbrOfArtStr)
Else
point1 = InStrRev(MyFile, "-") + 1
point2 = InStr(1, MyFile, ".DOC")
NmbrOfArtStr = Mid(MyFile, point1, point2 - point1)
NmbrOfArt = CInt(NmbrOfArtStr)
End If
Do Until NmbrOfArt <= 200
NmbrOfArt = NmbrOfArt - 200
Loop
'company def.
point1 = InStr(1, MyFile, "_") + 1
point2 = InStr(5, MyFile, "_")
Company = Mid(MyFile, point1, point2 - point1)
```

```
'add a tag at the end, to define the end of the last document
oDoc.Range.Select
lngEnd = Selection.End
Selection.MoveDown (wdParagraph)
teg = "Dokument " & CStr(NmbrOfArt + 1) & " von " & CStr(NmbrOfArt)
Selection.TypeText (teg)
For i = 1 To NmbrOfArt
2:
'define searching words (Dokument 1 von 200)
StartWord = "Dokument " & CStr(i) & " von " & CStr(NmbrOfArt)
EndWord = "Dokument " & CStr(i + 1) & " von " & CStr(NmbrOfArt)
'find the Source
oDoc.Range.Select
Selection.Find.Text = StartWord
blnFound = Selection.Find.Execute
StartTxt = Selection.End
lngStart = 0
lngEnd = 0
Do Until lngEnd - lngStart > 1
lngStart = Selection.End
Selection.MoveDown (wdParagraph)
lngEnd = Selection.End
Loop
Source = oDoc.Range(lngStart, lngEnd)
ExcelWB.Sheets("Articles").Range("F" & i + last_row) = Source
1:
'find the Date
lngStart = lngEnd
Do Until lngEnd - lngStart > 1
lngStart = Selection.End
Selection.MoveDown (wdParagraph)
lngEnd = Selection.End
Loop
qw = Len("Small Business Trends - USA")
Dt = oDoc.Range(lngStart, lngEnd)
If InStr(1, Dt, "http") > 0 Then GoTo 1
ExcelWB.Sheets("Articles").Range("e" & i + last_row) = Dt
```

```
'find the Title
lngStart = lngEnd
Do Until lngEnd - lngStart > 1
lngStart = Selection.End
Selection.MoveDown (wdParagraph)
lngEnd = Selection.End
Loop
Title = oDoc.Range(lngStart, lngEnd)
ExcelWB.Sheets("Articles").Range("g" & i + last_row) = Title

'copy Text and Text length
'find start of text
'lngStart = lngEnd
oDoc.Range.Select
Selection.Find.Text = StartWord
blnFound = Selection.Find.Execute
lngStart = Selection.End
'find end of text
oDoc.Range.Select
Selection.Find.Text = EndWord
blnFound = Selection.Find.Execute
EndTxt = Selection.Start
'debug if there is no next dokument heading
If blnFound = False Then
EndWord2 = "Dokument " & CStr(i + 2) & " von " & CStr(NmbrOfArt)
oDoc.Range.Select
Selection.Find.Text = EndWord2
blnFound = Selection.Find.Execute
Selection.MoveUp
Selection.Paragraphs.Add
Selection.TypeText (EndWord)
'GoTo 2
oDoc.Range.Select
Selection.Find.Text = EndWord
blnFound = Selection.Find.Execute
End If
lngEnd = Selection.Start
'find end of pfrase with "LENGTH:"
```

```
oDoc.Range(lngStart, lngEnd).Select
Selection.Find.Text = "LENGTH:"
blnFound = Selection.Find.Execute
If blnFound = False Then
oDoc.Range(lngStart, lngEnd).Select
Selection.Find.Text = "LÄNGE:"
blnFound = Selection.Find.Execute
End If
DateStart = Selection.End
Selection.MoveDown (wdParagraph)
lngStart = Selection.End
Length = oDoc.Range(DateStart, lngStart)
txt = oDoc.Range(lngStart, lngEnd)

ExcelWB.Sheets("Articles").Range("c" & i + last_row) = StartWord
ExcelWB.Sheets("Articles").Range("b" & i + last_row) = MyFile
ExcelWB.Sheets("Articles").Range("d" & i + last_row) = Company
ExcelWB.Sheets("Articles").Range("h" & i + last_row) = Length
fgh = StartTxt + 0.2 * (EndTxt - StartTxt)
If lngStart > StartTxt + 0.2 * (EndTxt - StartTxt) Then
'Length at the end of the article
txt = oDoc.Range(StartTxt, EndTxt)
End If
ExcelWB.Sheets("Articles").Range("i" & i + last_row) = txt
If IsNumeric(ExcelWB.Sheets("Articles").Range("A" & i + last_row - 1))
= False Then
k = 1
Else
k = ExcelWB.Sheets("Articles").Range("A" & i + last_row - 1) + 1
End If
ExcelWB.Sheets("Articles").Range("A" & i + last_row) = k
Next i
'ExcelWB.Sheets("Articles").Range("A1").CurrentRegion.Rows.Count

'oExcel.ExcelWB.Sheets("Articles").Range("A1").CurrentRegion.Rows.Count
oDoc.Close SaveChanges:=wdDoNotSaveChanges
Application.DisplayAlerts = True
MyFile = Dir
```

```
Loop
Application.ScreenUpdating = True
Application.DisplayAlerts = False
ExcelWB.Save
Application.DisplayAlerts = True
End Sub
```

Appendix C

Remove duplicated articles

```
#!/usr/bin/env python
# -*- coding: utf-8 -*-

import xlrd, xlwt
import os

#os.chdir('C:/.../01_Sentiment Analysis/01_Data and Macros/04_Datasets')

def copylil(lil):
"""Returns a true copy of list in list."""
returnl = []
for l in lil:
newl = l[:]
returnl.append(newl)
return returnl

wb = xlrd.open_workbook('USA O GER S.xls')
sheet_names_list = wb.sheet_names()
first_sheet_name = sheet_names_list[0]

sh = wb.sheet_by_name(first_sheet_name)

ggblilfile = []
for row_number in xrange(sh.nrows):
lilfile.append(sh.row_values(row_number))

string = len(lilfile) * '0'
checklist = [x for x in string]
```

```
lilfile_copy = copylil(lilfile)
result_rows = []
deleted_rows = []

rown = 0 # Die Reihennummer wird gezählt mit dieser Variable
for row1 in lilfile:
counter = 0
itemn = 0
for row2 in lilfile_copy:
if row1[7:9] == row2[7:9] and checklist[itemn] == '0':
checklist[itemn] = '1'
counter = counter + 1
if counter == 1:
result_rows.append(row2)
if counter > 1:
deleted_rows.append(row2)
itemn = itemn + 1
rown = rown + 1 # Reihennummer um eins erhöht
w = xlwt.Workbook()
ws = w.add_sheet(first_sheet_name)
i = 0
for line in result_rows:
j = 0
for element in line:
ws.write(i, j, element)
j = j + 1
i = i + 1
ws1 = w.add_sheet('Deleted_Lines')
i = 0
for line in deleted_rows:
j = 0
for element in line:
ws1.write(i, j, element)
j = j + 1
i = i + 1
w.save('USA O GER S 1.xls')
```

Appendix D

Articles to sentences

```
import csv
import os
from nltk.tokenize import sent_tokenize, word_tokenize
import nltk
from nltk.tree import Tree
# ---------------- BASICS ---------------------------------- #

os.chdir('C:/.../04_Datasets/05_Articles')

# ---------------- DEFINE FUNCTIONS ------------------------ #
def get_ne_from_tree(ne_tree):
ne_in_sent = []
for subtree in ne_tree:
if type(subtree) == Tree: # If subtree is a noun chunk, i.e. NE != "O"
ne_label = subtree.label()
ne_string = " ".join([token for token, pos in subtree.leaves()])
ne_in_sent.append((ne_string, ne_label))

if subtree.label() in ['NN','NNP']:
ne_label = subtree.label()
ne_string = " ".join([token for token, pos in subtree.leaves()])
ne_in_sent.append((ne_string, ne_label))

return ne_in_sent

def is_named_entity(company_string, sentence):
tokens = nltk.word_tokenize(sentence)
pos_tags = nltk.pos_tag(tokens)
```

```
chunked = nltk.ne_chunk(pos_tags)
named_ent = get_ne_from_tree(chunked)

for tuplex in named_ent:
if tuplex[0].lower() == company_string.lower():
#print("Hier hat es geklappt")
return True

else:
#print("Hier hat es nicht geklappt")
print(named_ent)
print(company_string)
print (sentence)
print(chunked)
return False

def main():
csvfile = open("USA O USA S 6.csv", newline='')
reader = csv.DictReader(csvfile, delimiter=',')
output = open("USA O USA S 6 sentences.csv", 'w', newline='')
writer = csv.writer(output, delimiter=',')
data = [["Article Index Nmbr","Document name","File name","Company name",
"Date","Source","Title","Length","Sentence"]]

for article in reader:
sent_tokenize_list = sent_tokenize(article["Article"])
sent_tokenize_list2 = [x.lower() for x in sent_tokenize_list]

#sent_tokenize_list = sent_tokenize_list.lower()

sentences = []
for idx, sentence in enumerate(sent_tokenize_list2):
word_tokenize_list = word_tokenize(sentence)
#word_tokenize_list_lower = [x.lower() for x in word_tokenize_list]
company_string = article["Company name"].lower()


if company_string in ' '.join(word_tokenize_list):
```

```
#if is_named_entity(company_string, sent_tokenize_list[idx]):
sentences.append(sent_tokenize_list[idx])

for sentence in sentences:
data.append([article["Article Index Nmbr"],
article["Document name"],
article["File name"],
article["Company name"],
article["Date"],
article["Source"],
article["Title"],
article["Length"],
sentence])

writer.writerows(data)

main()
```

Appendix E

Journal overview for literature review

TABLE E.1: Journals used for the literature review.

No.	Psychology	Management	Entrepreneurship
1	Trends in Cognitive Sciences	Academy of Management Journal	"Research Policy. A Journal Devoted to Research Policy, Research Management and Planning"
2	Annual Review of Psychology	Academy of Management Perspectives	Journal of Business Venturing
3	Psychological Bulletin	Academy of Management Review	Entrepreneurship Theory and Practice
4	Personality and Social Psychology Review	Accounting, Organizations and Society	Journal of Small Business Management JSBM
5	Annual Review of Clinical Psychology	Accounting Review	Strategic Entrepreneurship Journal
6	Psychological Review	Administrative Science Quarterly	Small Business Economics
7	Personnel Psychology	American Economic Review	Entrepreneurship and Regional Development
8	Neuroscience and Biobehavioral Reviews	California Management Review	Journal of Enterprising Culture
9	Journal of Applied Psychology	Contemporary Accounting Research	Journal of International Entrepreneurship
10	Journal of Personality and Social Psychology	Econometrica	Journal of Developmental Entrepreneurship
11	Perspectives on Psychological Science	Entrepreneurship Theory and Practice	Venture Capital : An International Journal of Entrepreneurial Finance
12	Clinical Psychology Review	Harvard Business Review	"Journal of Small Business Management and Entrepreneurship"
13	Psychological Science in the Public Interest, Supplement	Human Resource Management	Journal of Small Business Strategy
14	Research on Language and Social Interaction	Information Systems Research	Journal of Entrepreneurship Education
15	Advances in Experimental Social Psychology	Journal of Accounting and Economics	International Small Business Journal
16	Psychological Methods	Journal of Accounting Research	Journal of Small Business and Entrepreneurship
17	Psychological Science	Journal of Applied Psychology	Journal of Entrepreneurship
18	Journal of Experimental Psychology: General	Journal of Business Ethics	"International Journal of Entrepreneurship and Small Business"
19	Journal of the American Academy of Child and Adolescent Psychiatry	Journal of Business Venturing	Family Business Review
20	Educational Psychologist	Journal of Consumer Psychology	"The International Journal of Entrepreneurship and Innovation"
21	Journal of Consulting and Clinical Psychology	Journal of Consumer Research	"International Entrepreneurship and Management Journal"
22	Cognitive Psychology	Journal of Finance	"Zeitschrift für KMU & Entrepreneurship (ehemals: Zeitschrift für Klein- und Mittelunternehmen/ Internationales Gewerbearchiv)"
23	Developmental Science	Journal of Financial and Quantitative Analysis	Technovation
24	Child Development	Journal of Financial Economics	"International Journal of Entrepreneurship and Innovation Management"
25	Current Directions in Psychological Science	Journal of International Business Studies	"Journal of Entrepreneurial Finance and Business Ventures (ehemals: Journal of Entrepreneurial and Small Business Finance)"
26	Learning and Instruction	Journal of Management Studies	Frontiers of Entrepreneurship Research
27	Journal of Abnormal Psychology	Journal of Marketing	"International Journal of Entrepreneurial Behaviour and Research [Online-Ressource]"
28	Neuropsychology Review	Journal of Marketing Research	Entrepreneurship, Innovation and Change
29	Learning and Memory	Journal of Operations Management	"Journal of Small Business and Enterprise Development"
30	Journal of Organizational Behavior	Journal of Political Economy	
31	Journal of the Learning Sciences	Journal of the American Statistical Association	
32	Journal of Child Psychology and Psychiatry and Allied Disciplines	Management Science	
33	Cognition	Marketing Science	
34	Development and Psychopathology	MIS Quarterly	
35	Journal of Research in Crime and Delinquency	MIT Sloan Management Review	
36	Organizational Behavior and Human Decision Processes	Operations Research	
37	Psychological Medicine	Organization Science	
38	Social Cognitive and Affective Neuroscience	Organization Studies	
39	Journal of Educational Psychology	Organizational Behavior and Human Decision Processes	
40	Educational Psychology Review	Production and Operations Management	
41	Clinical Child and Family Psychology Review	Quarterly Journal of Economics	
42	Personality and Social Psychology Bulletin	RAND Journal of Economics	
43	Reading Research Quarterly	Review of Accounting Studies	
44	Psychometrika	Review of Financial Studies	
45	Research in Organizational Behavior	Strategic Management Journal	
46	Journal of Health and Social Behavior		
47	Developmental Psychology		
48	Depression and Anxiety		
49	Emotion		
50	Journal of Memory and Language		

Appendix F

Literature review summary

Reference (Authors, Year)	Dependent/Independent variable or empirical setting	Main topic	Most important ideas/findings
Begley, TM & Tan, WL (2001)	§ Developed theory of face to cultural factors relevant for entrepreneurship § Tested theory in sample of respondants in six East Asian and four Anglo-Saxon countries	Cultural differences in judgment	§ Importance of face and hierarchy in many Asien societies emphasizes relevance of social status judgments to career making-decisions § In Anglo countries, people who view entrepreneurship as higher in social status, regardless of its status in the culture are more likely to express interest in entrepreneurship
Cardon, MS & Foo, MD & Shepherd, D & Wiklund, J (2012)	Short literature review	Emotions	§ Entrepreneurial action leads to passion (not the other way around as previously stated) § No known work when entrepreneurs positively exit a venture
Cardon, MS & Stevens, CE & Potter, DR (2011)	§ Newspaper articles from 1999-2001 § 389 accounts from failure	Blame of entrepreneurs	§ Failure is blamed to misfortune and mistakes evenly in a U.S. study § Regional differences exist, e.g., New York blames more likely mistakes, San Francisco misfortune is more likely to be blamed
Cope, J (2011)	§ Interpretative phenomenological research on eight entrepreneurs	Learning from failure	§ Failure learnings increase the level of preparedness of the entrepreneur for future enterprising activities § Failure can stimulate profound changes in self-awareness
Hayward, MLA & Forster, WR & Sarasvathy, SD & Fredrickson, BL (2010)	Combination of cognitive perspectives on confidence	Overconfidence	§ More confident entrepreneurs will experience greater social support from founding team members during venture activities and after failure
Holland, DV & Shepherd, DA (2013)	Conjoint experiment	Persistance	§ Values and adversity are determinants for an entrepreneur's persistence decision (that is to stay in business) § Adversity affects the motivational factors in the decision to persist § Decision makers are more likely to select a risky opportunity that may help them overcome or minimize a loss rather than accept a sure loss
Isenberg, D (2011)	Article in Harvard Business Review	The cult of failure	§ In "hyperentrepreneurial" countries such as Israel, Taiwan, and Iceland early business failures are common

Lee, SH & Peng, MW & Barney, JB (2007)	Real options theory	The role of bankruptcy laws for entrepreneurial activity	§ In Japan business failure is considered a very shameful deed
Lee, SH & Yamakawa, Y & Peng, MW & Barney, JB (2011)	Database from 29 countries from 1990-2008	The role of bankruptcy laws for entrepreneurial activity	§ Lenient, entrepreneur-friendly bankruptcy laws are significantly correlated with the level of entrepreneurship development as measured by the rate of new firm entry § Lowering exit and entry barriers will positively affect entrepreneurship
McGrath, RG (1999)	Real options reasoning	Wealth creation of entrepreneurship	§ High failure rates for entrepreneurial businesses do not really matter, provided that the cost of failing is contained and that the businesses that do succeed enjoy substantial growth
Patzelt, H & Shepherd, DA (2011)	Survey, 2.700 US citizens	Negative emotions of an entrepreneurial career	§ Self-employed experience fewer negative emotions § This effect is magnified if individuals use problem–focused and/or emotion-focused coping tools § Coping enables individuals to deal with negative emotions that arise when "important goals have been harmed, lost, or threatened" (Folkman and Maskowitz 2004) § Coping refers to "the thoughts and behaviors used to manage the internal and external demands of situations that are appraised as successful" (Folkman and Maskowitz 2004)
Schutjens, V (2006)	Empirical proposition testing of 79 businesses that have closed within 5 years after start-up	Restart intentions	§ Business climate in Silicon Valley encourages risk taking and tolerates failure (Saxenian 1994, Lee et al. 2000) § Policymakers in other regions are aiming to reduce the stigma attached to failure (Waasdorp 2001, Armour and Cumming 2004) § Older entrepreneurs (>45 years) are less likely to start again than younger one's § Perception of final stage of business is important for restart likelihood. People with severe emotional problems associated with closure do more often start a new one (Cope 2003) § Hardly any study concentrated on the decision to start a business after a former has closed, that is the decision to become a serial entrepreneur (Schutjens 2006) § Most entrepreneurs who fail maintain their entrepreneurial intentions and a considerable group succeeds as serial entrepreneurs
Shepherd, DA (2003)	Psychological literature on grief to explore the emotion of business failure	Grief	§ Emotional relationship between the self-employed and their businesses, e.g., loyalty to customers and market, personal growth and the need to prove oneself. A loss of business is likely to generate a negative emotional response (grief)

			§ Negative emotions interfere with individuals‘ attention in the processing of information (Mogg 1990, Wells 1994) § Emotional events receive higher priority in processing information than those that are neutral (Ellis 1971, 1989) § Situations with great demands on attention and information processing have been found to cause individuals to be more prone to emotional interference (Mathews 1990, Wells 1994) § For self-employed founding a new business might enhance recovery from grief over the loss of a previous business (Shuchter 1986, Archer 1999) § Addressing questions such as „What do you do for a living?“, „How is your business going?“ (secondary sources of stress) can reduce the negative emotions associated with thoughts of the events surrounding the loss of a business (restoration orientation)
Shepherd, DA (2009)	Derived meso- and multi-level model for grief recovery	Grief recovery from the loss of a family business	§ More emotional intelligent individuals are better able to recognize and use their grief to process information about the loss and to help family members to do the same § A family member’s grief recovery time is defined as the period it takes before his or her thoughts about the events surrounding, and leading up to the loss of the family business, no longer generate a negative emotional response § The more emotional capable a family is, the speedier will be its grief recovery § Nadeau (1998) explores the healing process of a family that has had a family member die and uses interviews and analysis to demonstrate the importance of families coming to terms with their grief, and making sense of their loss.
Shepherd, DA & Cardon, MS (2009)	A developed emotion framework of project failure	Negative emotional reactions to project failure and the self-compassion to learn from the experienece	§ The importance of a project to an individual is partly dependent on the extent to which it satisfies the psychological need for competence and once it has been lost, this need remains unsatisfied (thwarted). A psychological need for competence ‘is satisfied when feedback provides information to the individual about their high performance at a task’ and a psychological need for competence is thwarted when feedback provides information of poor performance (Deci and Ryan, 2000). § A negative emotional reaction is when an event causes an individual’s core affect to become negative in response to a failure event (Seo 2004) § These negative emotions can lead organizational members to overestimate the likelihood of negative outcomes, and to underestimate the likelihood of positive outcomes for subsequent projects, (Nygren, 1996) as well as become more risk averse (Lerner and Keltner, 2001)

			§ Negative setbacks are critical to the development of resilience (Sutcliffe and Vogus, 2003) § Psychological well-being refers to the extent to which an individual experiences self-acceptance, positive relations with other, autonomy, environmental mastery, purpose in life, and personal growth (Ryff 1989) § There are two primary forms of emotion regulation, one focused on manipulating the inputs to the emotional system, such as by preventing the triggering of the emotion or diminishing the level of emotion triggered (antecedent-focused emotion regulation), and one focused on manipulating the outputs of the emotional system, such as by suppressing the emotional response tendencies, once the emotion has already been generated (response-focused emotion regulation) § Perceived social support from others is associated with positive well-being of individuals § Responses to a threat of a person's self-meaning or psychological well-being: Self-compassion and self-kindness § Mindfulness: holding painful thoughts and feelings in balanced awareness rather than over-identifying with them
Shepherd, DA & Covin, JG & Kuratko, DF (2009)	Discussion of two approaches of grief management	Complement of social cognitive theory with psychological theories on grief	§ A robust sense of coping self-efficacy is accompanied by benign appraisal of potential threats, weaker stress reactions to them, less ruminative preoccupation with them, better behavioral management of threats, and faster recovery of well-being from any experienced distress over them.
Shepherd, DA & Haynie, JM (2011)	Derived framework	Impression management strategies in response to the negative attributions associated with the stigma of venture failure	§ The stigmatized use a variety of impression management strategies, presumably in the hope of eliminating or minimizing the stigma of bankruptcy. The strategies were: (1) concealing , (2) defining in a positive light, (3) denying responsibility, (4) withdrawing (Sutton and Callahan 1987) § People often define their self-worth by how they perform at work (Kreiner and Ashforth, 2004) § Individuals with a negative self-view preferred evaluators who had unfavorable impressions of them (Swann, 1992)
Singh, S & Corner, PD & Pavlovich, K (2015)	Lived experience of 12 entrepreneurs	Stigmatization	§ Possible negative discrimination with regard to future employment opportunities and access to future resources both financial and human (Cope 2011, Shepherd 2011, Singh 2007) § Entrepreneurs expect negative judgment from family members and prospective employers to perceive them unfavorably

			§ Distancing to friends and society begins prior to the business failure as a self-protection mechanism § Entrepreneurs are getting treated differently by family after failure and the social circle changes, e.g. friendships were not there anymore
Townsend, DM & Busenitz, LW & Arthurs, JD (2010)	Empirical research with 316 nascent entrepreneurs	Research on why do some individuals decide to start new businesses and others do not	§ Results indicate, the longer it takes for individuals to act on the intention to become an entrepreneur, the less likely they are to decide to create a new venture § Results indicate that individuals decide to start new ventures because they are confident in their abilities to act entrepreneurially
Ucbasaran, D & Shepherd, DA & Lockett, A & Lyon, SJ (2013)	Literature review	An entrepreneurs perspective on business failure and on the consequences	§ Experiencing failure can have adverse motivational effects by generating a sense of “helplessness,” thus diminishing individuals’ beliefs in their ability to undertake specific tasks successfully in the future and leading to rumination, that hinders task performance (Bandura 1991, Cardon & McGrath 1999, Shepherd 2003) § In their study, failure was attributed to lack of effort (as opposed to ability) and, as a result, led students to redouble their efforts
Ucbasaran, D & Westhead, P & Wright, M & Flores, M (2010)	Survey data from 576 entrepreneurs in Great Britain	Comparative optimism and the influence on business failure	§ Experience with business failure was associated with entrepreneurs who are less likely to report comparative optimism § Entrepreneurs have a greater tendency to be over-optimistic than non-entrepreneurs § Definition of comparative optimism: The tendency of people to report that they are less likely than others to experience negative events (Helweg-Larsen and Shepherd 2001)
van Gelder, JL & de Vries, RE & Frese, M & Goutbeek, JP (2007)	Sample of 71 operational and 20 failed business owners from Suva, the capital of Fiji, using discriminant analysis	Different business strategies of operational and failed entrepreneurs	§ Operational business owners are more likely to employ a detailed and long-term planning strategy, whereas failed business owners more often pursue a reactive strategy § Operational business owners are more likely to set more specific and more difficult goals
Welpe, IM & Sporrle, M & Grichnik, D & Michl, T & Audretsch, DB (2012)	Questionaire based experiment and examination of four independent variables	Direct and moderating effects of emotions on exploitation tendencies	§ Fear reduces a positive evaluation of a potential new engagement into a new venture § Joy increases the positive impact of evaluation of exploitation

Wennberg, K & Pathak, S & Autio, E (2013)	Multi-level methodology with data from the Global Entrepreneurship Monitor (GEM) and Global Leadership and Organizational Behavior Effectiveness study (GLOBE) for 42 countries	The influence of cultural practices in the self-efficacy of entrepreneurs	§ Intention-based theories of entrepreneurial entry suggests that individuals consider not only their own ability to succeed and the possibility of failure, but also how this action is consistent with prevailing cultural norms and practices (Krueger and Carsrud 1993) § National culture is often seen as central for entrepreneurship (Hayton et al. 2002) § Some countries are considered models of an "entrepreneurial society", whereas others are perceived as "less entrepreneurial" (Freytag and Thurik 2007) § Influence of national cultural context on individual entrepreneurship is underexplored § The positive effect of self-efficacy on entrepreneurial entry is more pronounced in cultural landscapes that favor institutional collectivism and have higher performance orientation § Negative effects of individual's fear of failure on entry are somewhat smaller in settings with high levels of institutional collectivism § In collectivistic societies, the room for deviation is lesser since pursuing entrepreneurship may represent a potential challenge to established societal norms § Individuals exhibiting similar perceptions may behave differently depending on the cultural context in which they are embedded
Wiesenfeld, BM (2008)	Elaborated model to explain why failure leads to devaluation of elites	Devaluation via blameworthyness	§ Devaluation of executives and directors takes place when their companies perform poorly § They tend to be fired, not rehired elsewhere and if rehired with less capacities
Zacharakis, AL & Meyer, GD & DeCastro, J (1999)	Structured interviews	Entrepreneurial failure from the entrepreneur's and the VC's viewpoint	§ Entrepreneurs admitted that internal factors played a major role in their own venture's problems (58 percent of the time) § Entrepreneurs tended to attribute the failure of other new ventures to those in charge of these ventures at 89 percent of the time § VCs attribute problems to fierce competition rather than to poor management decisions

Appendix G

Literature review summary psychology papers

Reference	Dependent/Independent variable or empirical setting	Main topic	Most important ideas/findings
Savitsky, K & Epley, N & Gilovich, T (2001)	Four studies (scenario methodology, laboratory study, and two times testing of findings)	Expected judgment after failures, shortcomings, and mishaps	§ Fear of harsh judgment is triggered in part by tendency to be inordinately focused on misfortune. People fail to consider the wider range of situational factors that tend to moderate onlookers' impressions.
Ellis, S & Mendel, R & Nir, M (2006)	Laboratory experiment	Learning from successful and failed experiences	§ After successful events reviews of wrong actions are most effective § After failed events any kind of review is effective
Ellis, S & Davidi, I (2005)	Quasi-field experiment	Learning from successful and failed experiences	§ Performance of soldiers in doing navigation exercises improved when debriefed on their failures and successes, compared with others who reviewed their failed events only
Elliot, A & Church M (1997)	Field experiment	A hierarchical model of approach and avoidance achievement motivation	§ Mastery goals facilitated intrinsic motivation. Performance goals enhanced graded performance. Performance avoidance goals proved inimical to intrinsic motivation and graded performance.
Naquin, C & Tynan, R (2003)	Two studies (real teams and controlled scenarios)	The team halo effect	§ The nature of the causal attribution process used to diagnose failure scenarios leads to individuals being more likely to be identified as the cause of team failure than the team as a collective
Lockwood, P & Jordan, C & Kunda, Z (2002)	Three studies (one and two examined the impact of role models on motivation, three was a real-life experiment)	Motivation by positive or negative role models: Regulatory focus determines who will best inspire us	§ Individuals are motivated by role models who encourage strategies that fit their regulatory concerns. Participants of the study were more likely to generate real-life role models that matched their chronic goals.
Brunstein, J & Gollwitzer, P (1996)	Two experiments	Effects of failure on subsequent performance and the importance of self-defining goals	§ Enhanced performance after failure observed, it is relevant for a student's self-defined goals. Impaired performance on a subsequent task characterized as being irrelevant to the same self-definition
Schoorman, F & Holahan P (1996)	Study with 257 students	Psychological antecedents of escalation behavior	§ Although responsibility and negative decision consequences contribute to the escalation, they are not necessary conditions for escalations to occur
Keltner, D & Ellsworth, C & Edwards, K	Five experiments with students	Effects of sadness and anger on social perception	§ Sad students perceived situationally caused events as more likely and situational forces more responsible for an ambigious event than angry students
Lyubomirsky, S & Nolen-Hoeksema, S	Three studies with dysphoric and nondysphoric participants	Effects of self-focused rumination on negative thinking and interpersonal problem solving	§ Dysphoric participants induced to ruminatively self-focus on their feelings and personal characteristics would endorse more negatively biased interpretations of hypothetical situations
Coyne, J & Racioppo, M	Process studies of interventions	Coping research and clinical intervention research.	§ Process studies of interventions designed to improve coping, provide an alternative to fruitless and potentially misleading correlational studies using checklists
Spencer, J & Schutte, A (2004)	Two experiments with children	Perseverative biases	§ Perseverative biases arise from a single behavioral system

Baumgardner, A & Brownlee, E (1987)	Two experiments with socially anxious and nonanxious participants	Strategic failure in social interaction	§ Individuals who are particularly doubtful about their ability to perform up to par will sometimes fail strategically at the outset of social interaction as a means to create lower and safer standards

Appendix H

Curriculum vitae and summary

Curriculum vitae:

For reasons of data protection, the curriculum vitae is not included in the online version.

Summary:

Among the five most valuable firms in the world, there are two former startups. These two, Facebook and Google, have one thing in common: They are both from the US. Other successful startups support the assumption, that the US is especially entrepreneurial friendly, and the place to be for founders. However, is this actually true? Most research findings to this question are based on surveys or interviews. In contrast to these approaches, we aim to quantify entrepreneurial friendliness in the US and compare it to Germany, which is considered less entrepreneurial friendly. A focus is set on the media judgment of startups, as a surrogate of entrepreneurial friendliness, and changes of this sentiment after a startup fails. The media is capable of assigning or withholding organizational legitimacy, and therefore influences startups and founders in their daily work. By employing a difference-in-difference approach, in combination with regression models, we find significant differences in media judgment of startups in the two countries, with permanent positive judgment in the US (even after failure), and a rather neutral judgment in Germany, which changes to the negative after failure. Our findings also reveal, it is not common to report about failure in the US, whereas Germans seem to be fairly interested in detailed reporting about failures. For both countries we find startups with a novel business model to be judged more harshly after failure, than startups with traditional business models. In addition, startups which receive high levels of funding through investors, are also judged more harshly after failure. In a cross-comparison, we find the US to exhibit an in-group bias, meaning a favor in media judgment for their own startups over German startups. Surprisingly, we find a "now more than ever" mentality among failed founders, who have been negatively criticised by the media. This thesis enhances our understanding of entrepreneurial friendliness in the US and Germany, and adds to the social judgment of organizations theory, as it quantifies media legitimacy of startups, and the influence of failure on legitimacy of startups in two different countries.

Bibliography

Aberson, Christopher L., Michael Healy, and Victoria Romero (2000). “Ingroup bias and self-esteem: A meta-analysis”. In: *Personality and social psychology review* 4.2, pp. 157–173.

Abrams, Dominic Ed and Michael A. Hogg (1990). *Social identity theory: Constructive and critical advances*. Springer-Verlag Publishing.

Acs, Zoltan J. and David B. Audretsch, eds. (2010). *Handbook of entrepreneurship research*. Springer.

Acs, Zoltan J. et al. (2005). “Global Entrepreneurship Monitor. Executive Report 2004”. In: *Babson Collage and London Business School. Babson Park, MA. and Londres, UK*.

Agarwal, Rajshree et al. (2004). “Knowledge transfer through inheritance: Spin-out generation, development, and survival”. In: *Academy of Management journal* 47.4, pp. 501–522.

Aldrich, Howard E. and C. Marlene Fiol (1994). “Fools rush in? The institutional context of industry creation”. In: *Academy of Management Review* 19.4, pp. 645–670.

Andreevskaia, Alina and Sabine Bergler (2008). “When Specialists and Generalists Work Together: Overcoming Domain Dependence in Sentiment Tagging”. In: *ACL*, pp. 290–298.

Angrist, Joshua D. and Jörn-Steffen Pischke (2008). *Mostly harmless econometrics: An empiricist’s companion*. princeton university press.

Armour, John and Douglas Cumming (2008). “Bankruptcy law and entrepreneurship”. In: *American Law and Economics Review* 10.2, pp. 303–350.

Arora, Ashish and Anand Nandkumar (2011). “Cash-out or flameout! Opportunity cost and entrepreneurial strategy: theory, and evidence from the information security industry”. In: *Management Science* 57.10, pp. 1844–1860. ISSN: 0025-1909.

Audretsch, David B. and Michael Fritsch (2002). “Growth regimes over time and space”. In: *Regional Studies* 36.2, pp. 113–124.

Balahur, Alexandra et al. (2013). “Sentiment analysis in the news”. In: *arXiv preprint arXiv:1309.6202*.

Baum, J. Robert and Edwin A. Locke (2004). “The relationship of entrepreneurial traits, skill, and motivation to subsequent venture growth”. In: *Journal of Applied Psychology* 89.4, p. 587. ISSN: 0021-9010.

Baum, Joel A. C. and Christine Oliver (1991). “Institutional linkages and organizational mortality”. In: *Administrative science quarterly*, pp. 187–218.

Baumeister, Roy F. and Mark R. Leary (1995). "The need to belong: desire for interpersonal attachments as a fundamental human motivation". In: *Psychological Bulletin* 117.3, p. 497. ISSN: 1939-1455.

Begley, Thomas M. and Wee-Liang Tan (2001). "The Socio-Cultural Environment for Entrepreneurship: A Comparison Between East Asian and Anglo-Saxon Countries". In: *Journal of International Business Studies* 32.3, pp. 537–553. ISSN: 0047-2506. DOI: 10.1057/palgrave.jibs.8490983.

Berglas, S. and E. E. Jones (1978). "Control of attributions about the self through self-handicapping strategies: The appeal of alcohol and the role of underachievement". In: *Personality and Social Psychology Bulletin* 4.2, pp. 200–206.

Berscheid, Ellen and Elaine Walster (1974). "Physical attractiveness". In: *Advances in experimental social psychology* 7, pp. 157–215.

Bettencourt, B. Ann et al. (2001). "Status differences and in-group bias: a meta-analytic examination of the effects of status stability, status legitimacy, and group permeability". In: *Psychological Bulletin* 127.4, p. 520. ISSN: 1939-1455.

Billig, Michael G. and Henri Tajfel (1973). "Social categorization and similarity in intergroup behaviour". In: *European Journal of Social Psychology* 3.1, pp. 27–52.

Bird, Barbara and Mariann Jelinek (1988). "The operation of entrepreneurial intentions". In: *Entrepreneurship Theory and Practice* 13.2, pp. 21–29. ISSN: 10422587.

Bird, Steven, Ewan Klein, and Edward Loper (2009). *Natural language processing with Python*. 1st ed. Beijing and Cambridge [Mass.]: O'Reilly. ISBN: 0596516495.

Bitektine, Alex (2011). "Toward a Theory of Social Judgments of Organizations: The Case of Legitimacy, Reputation, and Status". In: *Academy of Management Review* 36.1, pp. 151–179. DOI: 10.5465/amr.2009.0382.

Blanchflower, David G., Andrew Oswald, and Alois Stutzer (2001). "Latent entrepreneurship across nations". In: *European Economic Review* 45.4-6, pp. 680–691. ISSN: 00142921. DOI: 10.1016/S0014-2921(01)00137-4.

Blank, Steve (2012). *The startup owner's manual: The step-by-step guide for building a great company*. BookBaby.

Boulding, William, Ruskin Morgan, and Richard Staelin (1997). "Pulling the plug to stop the new product drain". In: *Journal of Marketing Research*, pp. 164–176.

Braddock, Jomills Henry and James M. McPartland (1987). "How minorities continue to be excluded from equal employment opportunities: Research on labor market and institutional barriers". In: *Journal of Social Issues* 43.1, pp. 5–39.

Bradley, Don E. and James A. Roberts (2004). "Self-employment and job satisfaction: investigating the role of self-efficacy, depression, and seniority". In: *Journal of Small Business Management* 42.1, pp. 37–58. ISSN: 0047-2778.

Bradley, Gifford W. (1978). "Self-serving biases in the attribution process: A reexamination of the fact or fiction question". In: *Journal of Personality and Social Psychology* 36.1, p. 56. ISSN: 0022-3514.

Brewer, Marilynn B. (1979). "In-group bias in the minimal intergroup situation: A cognitive-motivational analysis". In: *Psychological Bulletin* 86.2, p. 307. ISSN: 1939-1455.

Briar, Scott (1966). "Welfare from below: Recipients' views of the public welfare system". In: *California Law Review*, pp. 370–385.

Brigham, John C. (1974). "Views of black and white children concerning the distribution of personality characteristics". In: *Journal of personality*.

Brosius, Hans-Bernd and Hans Mathias Kepplinger (1990). "The agenda-setting function of television news static and dynamic views". In: *Communication Research* 17.2, pp. 183–211.

— (1992). "Beyond agenda-setting: The influence of partisanship and television reporting on the electorate's voting intentions". In: *Journalism & Mass Communication Quarterly* 69.4, pp. 893–901.

Broverman, Inge K. et al. (1994). "Sex-role stereotypes: A current appraisal". In: *Caring voices and women's moral frames: Gilligan's view*, pp. 191–210.

Brown, Noel and Craig Deegan (1998). "The public disclosure of environmental performance information—a dual test of media agenda setting theory and legitimacy theory". In: *Accounting and business research* 29.1, pp. 21–41.

Bruno, Albert V. and Joel K. Leidecker (1988). "Causes of new venture failure: 1960s vs. 1980s". In: *Business Horizons* 31.6, pp. 51–56. ISSN: 00076813. DOI: 10.1016/0007-6813(88)90024-9.

Brunstein, Joachim C. and Peter M. Gollwitzer (1996). "Effects of failure on subsequent performance: The importance of self-defining goals". In: *Journal of Personality and Social Psychology* 70.2, pp. 395–407. ISSN: 0022-3514. DOI: 10.1037/0022-3514.70.2.395.

Burges, Christopher J. C. (1998). "A tutorial on support vector machines for pattern recognition". In: *Data mining and knowledge discovery* 2.2, pp. 121–167.

Cameron, A. Colin and Pravin K. Trivedi (2005). *Microeconometrics: methods and applications*. Cambridge university press.

Campbell, Colin, Nello Cristianini, and J. Shawe-Taylor (1999). "Dynamically adapting kernels in support vector machines". In: *Advances in neural information processing systems* 11, pp. 204–210.

Cannella, Albert A., Donald R. Fraser, and D. Scott Lee (1995). "Firm failure and managerial labor markets evidence from Texas banking". In: *Journal of Financial Economics* 38.2, pp. 185–210.

Cannon, Mark D. and Amy C. Edmondson (2001). "Confronting failure: Antecedents and consequences of shared beliefs about failure in organizational work groups". In: *Journal of Organizational Behavior* 22.2, pp. 161–177.

— (2005). "Failing to learn and learning to fail (intelligently): How great organizations put failure to work to innovate and improve". In: *Long Range Planning* 38.3, pp. 299–319.

Cardon, Melissa S. and R. G. McGrath (1999). "When the going gets tough... Toward a psychology of entrepreneurial failure and re-motivation". In: *Frontiers of entrepreneurship research* 29.4, pp. 58–72.

Cardon, Melissa S., Christopher E. Stevens, and D. Ryland Potter (2011). "Misfortunes or mistakes?" In: *Journal of Business Venturing* 26.1, pp. 79–92. ISSN: 08839026. DOI: 10.1016/j.jbusvent.2009.06.004.

Cardon, Melissa S. et al. (2005). "A tale of passion: New insights into entrepreneurship from a parenthood metaphor". In: *Journal of Business Venturing* 20.1, pp. 23–45. ISSN: 08839026.

Cardon, Melissa S. et al. (2009). "The nature and experience of entrepreneurial passion". In: *Academy of Management Review* 34.3, pp. 511–532.

Carlsson, Bo (1999). "Small business, entrepreneurship, and industrial dynamics". In: *Are Small Firms Important? Their Role and Impact*. Springer, pp. 99–110.

Carlsson, Bo et al. (1992). *The rise of small business: causes and consequences*. Industrial Institute for Economic and Social Research.

Carree, Martin A. and A. Roy Thurik (2010). "The impact of entrepreneurship on economic growth". In: *Handbook of entrepreneurship research*. Ed. by Zoltan J. Acs and David B. Audretsch. Springer, pp. 557–594.

Carroll, E. Craig and Maxwell E. McCombs (2003). "Agenda-setting Effects of Business News on the Public's Images and Opinions about Major Corporations". In: *Corporate Reputation Review* 6.1, pp. 36–46. ISSN: 1479-1889. DOI: 10.1057/palgrave.crr.1540188. URL: http://dx.doi.org/10.1057/palgrave.crr.1540188.

Cassar, Gavin and Justin Craig (2009). "An investigation of hindsight bias in nascent venture activity". In: *Journal of Business Venturing* 24.2, pp. 149–164. ISSN: 08839026.

Casson, Mark (1982). *The entrepreneur: An economic theory*. Rowman & Littlefield.

Chater, Nick and Paul M. B. Vitányi (2003). "The generalized universal law of generalization". In: *Journal of Mathematical Psychology* 47.3, pp. 346–369.

Chatterji, Aaron K. (2009). "Spawned with a silver spoon? Entrepreneurial performance and innovation in the medical device industry". In: *Strategic Management Journal* 30.2, pp. 185–206.

Coelho, Philip R. P. and James E. McClure (2005). "Learning from failure". In: *American Journal of Business* 20.1, p. 1.

Cooley, Charles Horton (1956). *Social organization: Human nature and the social order*. Free Press.

Cooper, Arnolc C., Carolyn Y. Woo, and William C. Dunkelberg (1988). "Entrepreneur's perceived chances for success". In:

Cope, Jason (2011). "Entrepreneurial learning from failure: An interpretative phenomenological analysis". In: *Journal of Business Venturing* 26.6, pp. 604–623. ISSN: 08839026. DOI: 10.1016/j.jbusvent.2010.06.002.

Crammer, Koby and Yoram Singer (2001). "On the algorithmic implementation of multiclass kernel-based vector machines". In: *Journal of machine learning research* 2.Dec, pp. 265–292.

Crocker, Jennifer and Brenda Major (1989). "Social stigma and self-esteem: The self-protective properties of stigma". In: *Psychological review* 96.4, p. 608.

Crocker, Jennifer et al. (1998). "The handbook of social psychology". In: *The handbook of social psychology*, pp. 504–553.

Daft, Richard L. and Karl E. Weick (1984). "Toward a model of organizations as interpretation systems". In: *Academy of Management Review* 9.2, pp. 284–295.

Dave, Kushal, Steve Lawrence, and David M. Pennock (2003). "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews". In: *Proceedings of the 12th international conference on World Wide Web*, pp. 519–528.

Davidsson, Per (1995). "Culture, structure and regional levels of entrepreneurship". In: *Entrepreneurship & Regional Development* 7.1, pp. 41–62. ISSN: 0898-5626.

Davidsson, Per and Johan Wiklund (1997). "Values, beliefs and regional variations in new firm formation rates". In: *Journal of Economic psychology* 18.2, pp. 179–199.

Davis, Robert C. and David C. McClelland (1962). *The achieving society*.

Deaux, Kay (1976). "Sex: A perspective on the attribution process". In: *New directions in attribution research* 1, pp. 335–352.

Deephouse, David L. (1996). "Does isomorphism legitimate?" In: *Academy of Management journal* 39.4, pp. 1024–1039.

Deephouse, David L. and Suzanne M. Carter (2005). "An examination of differences between organizational legitimacy and organizational reputation". In: *Journal of Management Studies* 42.2, pp. 329–360. ISSN: 00222380.

Dervin, Brenda (1983). *An overview of sense-making research: Concepts, methods, and results to date*. The Author.

Devers, Cynthia E. et al. (2009). "A General Theory of Organizational Stigma". In: *Organization Science* 20.1, pp. 154–171. ISSN: 1047-7039. DOI: 10.1287/orsc.1080.0367.

Dillon, P. (1998). "Failure is just part of the culture of innovation: Accept it and become stronger". In: *Fast Company* 18, pp. 1–2.

Dimaggio, P. J. and W. W. Powell (1983). “The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields”. In: *American Sociological Review* 48.2, pp. 147–160.

Ding, Xiaowen, Bing Liu, and Philip S. Yu (2008). “A holistic lexicon-based approach to opinion mining”. In: *Proceedings of the 2008 international conference on web search and data mining*, pp. 231–240.

Dion, Karen K. (1972). “Physical attractiveness and evaluation of children’s transgressions”. In: *Journal of Personality and Social Psychology* 24.2, p. 207. ISSN: 0022-3514.

Domingos, Pedro and Michael Pazzani (1997). “On the optimality of the simple Bayesian classifier under zero-one loss”. In: *Machine learning* 29.2-3, pp. 103–130.

Douglas, Evan J. and Dean A. Shepherd (2002). “Self-employment as a career choice: Attitudes, entrepreneurial intentions, and utility maximization”. In: *Entrepreneurship Theory and Practice* 26.3, pp. 81–90. ISSN: 10422587.

Efrat, Rafael (2006). “The evolution of bankruptcy stigma”. In: *Theoretical inquiries in Law* 7.2, pp. 365–393.

Eggers, J. P. and Lin Song (2015). “Dealing with Failure: Serial Entrepreneurs and the Costs of Changing Industries Between Ventures”. In: *Academy of Management journal* 58.6, pp. 1785–1803.

Ellis, Shmuel and Inbar Davidi (2005). “After-event reviews: drawing lessons from successful and failed experience”. In: *The Journal of applied psychology* 90.5, pp. 857–871. ISSN: 0021-9010. DOI: 10.1037/0021-9010.90.5.857.

Ellis, Shmuel, Rachel Mendel, and Michal Nir (2006). “Learning from successful and failed experience: the moderating role of kind of after-event review”. In: *The Journal of applied psychology* 91.3, pp. 669–680. ISSN: 0021-9010. DOI: 10.1037/0021-9010.91.3.669.

Everett, Jim and John Watson (1998). “Small business failure and external risk factors”. In: *Small Business Economics* 11.4, pp. 371–390.

Fan, Wei and Michelle J. White (2003). “Personal bankruptcy and the level of entrepreneurial activity”. In: *The Journal of Law and Economics* 46.2, pp. 543–567.

Farina, Amerigo et al. (1976). “Some interpersonal consequences of being mentally ill or mentally retarded”. In: *American Journal of Mental Deficiency*.

Feather, N. T. (1968). “Change in confidence following success or failure as a predictor of subsequent performance”. In: *Journal of Personality and Social Psychology* 9.1, p. 38. ISSN: 0022-3514.

Filion, Louis Jacques (1991). “Vision and relations: elements for an entrepreneurial meta-model”. In: *International Small Business Journal* 9.2, pp. 26–40.

Fischer, Ronald and Crysta Derham (2016). “Is in-group bias culture-dependent? A meta-analysis across 18 societies”. In: *SpringerPlus* 5.1, p. 1.

Fombrun, Charles and Mark Shanley (1990). “What’s in a name? Reputation building and corporate strategy”. In: *Academy of Management journal* 33.2, pp. 233–258.

Funkhouser, G. Ray (1973). “The Issues of the Sixties: An Exploratory Study in the Dynamics of Public Opinion”. In: *Public Opinion Quarterly* 37.1, p. 62. ISSN: 0033362X. DOI: 10.1086/268060.

Gandel, Stephen (2016). *These Are the 10 Most Valuable Companies in the Fortune 500*. Market Intelligence. URL: http://fortune.com (visited on 2016).

Garreta, Raul and Guillermo Moncecchi (2013). *Learning scikit-learn: machine learning in python*. Packt Publishing Ltd.

George, Gerard et al. (2016). “Reputation and Status: Expanding the Role of Social Evaluations in Management Research”. In: *Academy of Management journal* 59.1, pp. 1–13. DOI: 10.5465/amj.2016.4001.

Gergen, Kenneth J. (1971). “The concept of self”. In: *Oxford, England: Holt, Rinehart & Winston*.

Gergen, Kenneth J. and Edward E. Jones (1963). “Mental illness, predictability, and affective consequences as stimulus factors in person perception”. In: *The Journal of Abnormal and Social Psychology* 67.2, p. 95.

Ghanem, Salma I. (1997). “Filling in the tapestry: The second level of agenda setting”. In: *Communication and democracy: Exploring the intellectual frontiers in agenda-setting theory*, pp. 3–14.

Gilbert, Daniel T. and Patrick S. Malone (1995). “The correspondence bias”. In: *Psychological Bulletin* 117.1, pp. 21–38. ISSN: 1939-1455. DOI: 10.1037/0033-2909.117.1.21.

Gioia, Dennis A. and Kumar Chittipeddi (1991). “Sensemaking and sensegiving in strategic change initiation”. In: *Strategic Management Journal* 12.6, pp. 433–448.

Gmür, Markus (2003). “Co-citation analysis and the search for invisible colleges: A methodological evaluation”. In: *Scientometrics* 57.1, pp. 27–57.

Golan, Guy and Wayne Wanta (2001). “Second-level agenda setting in the New Hampshire primary: A comparison of coverage in three newspapers and public perceptions of candidates”. In: *Journalism & Mass Communication Quarterly* 78.2, pp. 247–259.

Goldberg, Amir, Michael T. Hannan, and Balázs Kovács (2016). “What does it mean to span cultural boundaries? Variety and atypicality in cultural consumption”. In: *American Sociological Review* 81.2, pp. 215–241.

Gruber, Marc, Ian C. MacMillan, and James D. Thompson (2008). “Look before you leap: Market opportunity identification in emerging technology firms”. In: *Management Science* 54.9, pp. 1652–1665. ISSN: 0025-1909.

Hamilton, James T. (1995). “Pollution as news: media and stock market reactions to the toxics release inventory data”. In: *Journal of environmental Economics and Management* 28.1, pp. 98–113.

Hampel, Christian E. and Paul Tracey (2016). "How Organizations Move From Stigma to Legitimacy: The Case of Cook's Travel Agency in Victorian Britain". In: *Academy of Management journal*, amj–2015.

Harris, Stanley G. and Robert I. Sutton (1986). "Functions of parting ceremonies in dying organizations". In: *Academy of Management journal* 29.1, pp. 5–30.

Harter, Susan (1986). "Processes underlying the construction, maintenance, and enhancement of the self-concept in children". In: *Psychological perspectives on the self* 3.13, pp. 7–181.

Hartsough, W. Ross and Alan F. Fontana (1970). "Persistence of ethnic stereotypes and the relative importance of positive and negative stereotyping for association preferences". In: *Psychological reports* 27.3, pp. 723–731.

Hatzivassiloglou, Vasileios and Kathleen R. McKeown (1997). "Predicting the semantic orientation of adjectives". In: *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics*, pp. 174–181.

Hayton, James C., Gerard George, and Shaker A. Zahra (2002). "National culture and entrepreneurship: A review of behavioral research". In: *Entrepreneurship Theory and Practice* 26.4, p. 33. ISSN: 10422587.

Hayward, Mathew L.A. et al. (2010). "Beyond hubris: How highly confident entrepreneurs rebound to venture again". In: *Journal of Business Venturing* 25.6, pp. 569–578. ISSN: 08839026. DOI: 10.1016/j.jbusvent.2009.03.002.

Headd, Brian (2003). "Redefining business success: Distinguishing between closure and failure". In: *Small Business Economics* 21.1, pp. 51–61.

Heider, Fritz et al. (1958). "The psychology of interpersonal relations". In: *The Journal of Marketing* 56, p. 322.

Heilbrun, Alfred B. (1976). "Measurement of masculine and feminine sex role identities as independent dimensions". In: *Journal of consulting and clinical psychology* 44.2, p. 183.

Henderson, Roger and Martyn Robertson (1999). "Who wants to be an entrepreneur? Young adult attitudes to entrepreneurship as a career". In: *Education+ Training* 41.5, pp. 236–245.

Herbig, Paul A. (1994). *The innovation matrix: Culture and structure prerequisites to innovation*. Praeger Pub Text.

Herek, Gregory M. (1984). "Attitudes toward lesbians and gay men: A factor-analytic study". In: *Journal of Homosexuality* 10.1-2, pp. 39–51.

Hewstone, Miles, Mark Rubin, and Hazel Willis (2002). "Intergroup bias". In: *Annual review of psychology* 53.1, pp. 575–604.

Hill, Thomas, Pawel Lewicki, and Paweł Lewicki (2006). *Statistics: methods and applications: a comprehensive reference for science, industry, and data mining*. StatSoft, Inc.

Hindle, Kevin and Kim Klyver (2007). "Exploring the relationship between media coverage and participation in entrepreneurship: Initial global evidence and research implications".

In: *International Entrepreneurship and Management Journal* 3.2, pp. 217–242. ISSN: 1554-7191. DOI: 10.1007/s11365-006-0018-8.

Hindle, Kevin and Susan Rushworth (2000). *Yellow Pages® global entrepreneurship monitor: Australia 2000*. Swinburne University of Technology.

Ho, Daniel E. et al. (2004). "MatchIt: Matching software for causal inference". In: *Version 0.8. Used with permission*.

Hofstede, Geert H., Gert Jan Hofstede, and Michael Minkov (1991). *Cultures and organizations: Software of the mind*. Vol. 2. Citeseer.

— (2010). *Cultures and organizations: Software of the mind : intercultural cooperation and its importance for survival*. Rev. and expanded 3rd ed. New York: McGraw-Hill. ISBN: 978-0071664189.

Hogg, Michael A. (2000). "Subjective uncertainty reduction through self-categorization: A motivational theory of social identity processes". In: *European review of social psychology* 11.1, pp. 223–255.

— (2007). "Uncertainty–identity theory". In: *Advances in experimental social psychology* 39, pp. 69–126.

Hsu, David H. (2007). "Experienced entrepreneurial founders, organizational capital, and venture capital funding". In: *Research Policy* 36.5, pp. 722–741.

Hu, Minqing and Bing Liu (2004). "Mining and summarizing customer reviews". In: *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 168–177.

Huberman, Bernardo A., Christoph H. Loch, and Ayse Önçüler (2004). "Status as a valued resource". In: *Social Psychology Quarterly* 67.1, pp. 103–114.

Hudson, Bryant Ashley (2008). "Against all odds: A consideration of core-stigmatized organizations". In: *Academy of Management Review* 33.1, pp. 252–266.

Hybels, Ralph Cushman (1994). *Legitimation, population density, and founding rates: the institutionalization of commercial biotechnology in the US, 1971-1989*. Cornell University.

Indurkhya, Nitin and Frederick Damerau (2010). *Handbook of natural language processing*. 2nd edition. Chapman & Hall/CRC machine learning & pattern recognition series. Boca Raton, FL: Chapman & Hall/CRC. ISBN: 1420085921.

Ingram, Paul and Hayagreeva Rao (2004). "Store Wars: The Enactment and Repeal of Anti-Chain-Store Legislation in America 1". In: *American journal of sociology* 110.2, pp. 446–487.

Iyengar, Shanto (1990). "The accessibility bias in politics: Television news and public opinion". In: *International Journal of Public Opinion Research* 2.1, pp. 1–15.

Iyengar, Shanto and Donald R. Kinder (2010). *News that matters: Television and American opinion*. University of Chicago Press.

Iyengar, Shanto, Helmut Norpoth, and Kyu S. Hahn (2004). "Consumer demand for election news: The horserace sells". In: *Journal of Politics* 66.1, pp. 157–175.

Jaccard, Paul (1901). *Etude comparative de la distribution florale dans une portion des Alpes et du Jura*. Impr. Corbaz.

Janoff-Bulman, Ronnie and Irene Hanson Frieze (1983). "A theoretical perspective for understanding reactions to victimization". In: *Journal of Social Issues* 39.2, pp. 1–17.

Japkowicz, Nathalie and Shaju Stephen (2002). "The class imbalance problem: A systematic study". In: *Intelligent data analysis* 6.5, pp. 429–449.

Jenkins, Anna et al. (2014). "Stigmatization of failed entrepreneurs: prevalence and solutions". In: *Frontiers of entrepreneurship research* 34.5, p. 2.

Jensen, Michael C. (1983). "Organization theory and methodology". In: *Accounting review*, pp. 319–339.

Jensen, Michael C. and Aradhana Roy (2008). "Staging exchange partner choices: When do status and reputation matter?" In: *Academy of Management journal* 51.3, pp. 495–516.

Johnson, James P. and Tomasz Lenartowicz (1998). "Culture, freedom and economic growth: do cultural values explain economic growth?" In: *Journal of World Business* 33.4, pp. 332–356.

Jonas, Klaus and Miles Hewstone (2007). *Sozialpsychologie*. Springer.

Jones, Candace, William S. Hesterly, and Stephen P. Borgatti (1997). "A general theory of network governance: Exchange conditions and social mechanisms". In: *Academy of Management Review* 22.4, pp. 911–945.

Jones, Edward E. (1984). *Social stigma: The psychology of marked relationships*. WH Freeman.

Jones, Edward E. and Keith E. Davis (1965). "From acts to dispositions the attribution process in person perception". In: *Advances in experimental social psychology* 2, pp. 219–266.

Jones, Gareth R. et al. (2010). *Organizational theory, design, and change*. Pearson Upper Saddle River.

Jones, Geoffrey and Rohit Daniel Wadhwani (2006). *Entrepreneurship and business history: Renewing the research agenda*. Division of Research, Harvard Business School.

Kelley, Harold H. (1973). "The processes of causal attribution". In: *American Psychologist* 28.2, pp. 107–128. ISSN: 0003-066X. DOI: 10.1037/h0034225.

Kelley, Harold H. and John L. Michela (1980). "Attribution theory and research". In: *Annual review of psychology* 31.1, pp. 457–501.

Kensinger, Elizabeth A. (2007). "Negative emotion enhances memory accuracy behavioral and neuroimaging evidence". In: *Current Directions in Psychological Science* 16.4, pp. 213–218.

King, Gary et al. (2014). "The Balance-Sample Size Frontier in Matching Methods for Causal Inference". In: *PS: Political Science and Politics6* 42, S11–S22.

King, Gary, Christopher Lucas, and Richard A Nielsen (2014). "The Balance-Sample Size Frontier in Matching Methods for Causal Inference". In: *American Journal of Political Science*.

King, Gary and Richard Nielsen (2015). "Why Propensity Scores Should Not Be Used for Matching". In: *Copy at http://j. mp/1sexgVw Export BibTex Tagged XML Download Paper* 481.

Kiousis, Spiro, Philemon Bantimaroudis, and Hyun Ban (1999). "Candidate image attributes experiments on the substantive dimension of second level agenda setting". In: *Communication Research* 26.4, pp. 414–428.

Klapper, Joseph T. (1960). "The effects of mass communication". In: *Free Press*.

Klassen, Robert D. and Curtis P. McLaughlin (1996). "The impact of environmental management on firm performance". In: *Management Science* 42.8, pp. 1199–1214. ISSN: 0025-1909.

Komisar, Randy (2000). "Goodbye Career, Hello Success". In: *Harvard Business Review* 78.2, pp. 160–174.

Konar, Shameek and Mark A. Cohen (1997). "Information as regulation: The effect of community right to know laws on toxic emissions". In: *Journal of environmental Economics and Management* 32.1, pp. 109–124.

Kostova, Tatiana and Kendall Roth (2002). "Adoption of an organizational practice by subsidiaries of multinational corporations: Institutional and relational effects". In: *Academy of Management journal* 45.1, pp. 215–233.

Kotsiantis, Sotiris B., I. Zaharakis, and P. Pintelas (2007). *Supervised machine learning: A review of classification techniques*.

Kovács, Balázs and Michael T. Hannan (2015). "Conceptual spaces and the consequences of category spanning". In: *Sociological science* 2, pp. 252–286.

Kurzban, Robert and Mark R. Leary (2001). "Evolutionary origins of stigmatization: the functions of social exclusion". In: *Psychological Bulletin* 127.2, p. 187. ISSN: 1939-1455.

La Hayward, Mathew and Donald C. Hambrick (1997). "Explaining the premiums paid for large acquisitions: Evidence of CEO hubris". In: *Administrative science quarterly*, pp. 103–127.

Landier, Augustin (2005). "Entrepreneurship and the Stigma of Failure". In: *Available at SSRN 850446*.

Lantz, Brett (2013). *Machine learning with R*. Packt Publishing Ltd.

Lee, Sang M. and Suzanne J. Peterson (2001). "Culture, entrepreneurial orientation, and global competitiveness". In: *Journal of World Business* 35.4, pp. 401–416.

Lee, Seung-Hyun, Mike W. Peng, and Jay B. Barney (2007). "Bankruptcy law and entrepreneurship development: A real options perspective". In: *Academy of Management Review* 32.1, pp. 257–272.

Lee, Seung-Hyun et al. (2011). “How do bankruptcy laws affect entrepreneurship development around the world?” In: *Journal of Business Venturing* 26.5, pp. 505–520. ISSN: 08839026. DOI: `10.1016/j.jbusvent.2010.05.001`.

Levitt, Eugene E. and Klassen, Jr, Albert D (1976). “Public attitudes toward homosexuality: Part of the 1970 national survey by the Institute for Sex Research”. In: *Journal of Homosexuality* 1.1, pp. 29–43.

Liang, Yuxian Eugene and Soe-Tsyr Daphne Yuan (2016). “Predicting investor funding behavior using crunchbase social network features”. In: *Internet Research* 26.1, pp. 74–100. ISSN: 1066-2243. DOI: `10.1108/IntR-09-2014-0231`.

Liu, Bing (2007). *Web data mining: exploring hyperlinks, contents, and usage data*. Springer Science & Business Media.

— (2012). “Sentiment analysis and opinion mining”. In: *Morgan & Claypool Publishers*.

Lockwood, Penelope, Christian H. Jordan, and Ziva Kunda (2002). “Motivation by positive or negative role models: Regulatory focus determines who will best inspire us”. In: *Journal of Personality and Social Psychology* 83.4, pp. 854–864. ISSN: 0022-3514. DOI: `10.1037//0022-3514.83.4.854`.

Lonner, Walter J., John W. Berry, and Geert H. Hofstede (1980). “Culture’s Consequences: International Differences in Work-Related Values”. In: *University of Illinois at Urbana-Champaign’s Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.

Loughran, Tim and Bill McDonald (2011). “When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks”. In: *The Journal of Finance* 66.1, pp. 35–65.

MacMillan, Ian C. (1986). *To really learn about entrepreneurship, let’s study habitual entrepreneurs*. Wharton School of the University of Pennsylvania, Snider Entrepreneurial Center.

Macnamara, Jim R. (2003). “Mass media effects: a review of 50 years of media effects research”. In:

Maidique, Modesto A. and Billie Jo Zirger (1985). “The new product learning cycle”. In: *Research Policy* 14.6, pp. 299–313.

Malach-Pines, Ayala et al. (2005). “Entrepreneurs as cultural heroes: A cross-cultural, interdisciplinary perspective”. In: *Journal of Managerial Psychology* 20.6, pp. 541–555.

March, James G. and Zur Shapira (1987). “Managerial perspectives on risk and risk taking”. In: *Management Science* 33.11, pp. 1404–1418. ISSN: 0025-1909.

Markman, Gideon D. and Robert A. Baron (2003). “Person–entrepreneurship fit: why some people are more successful as entrepreneurs than others”. In: *Human resource management review* 13.2, pp. 281–301.

Marmer, Max et al. (2011). “Startup genome report extra: Premature scaling”. In: *Startup Genome* 10.

Marra, Alessandro et al. (2015). "A network analysis using metadata to investigate innovation in clean-tech – Implications for energy policy". In: *Energy Policy* 86, pp. 17–26. ISSN: 03014215. DOI: `10.1016/j.enpol.2015.06.025`.

Marsh, Herbert W. (1986). "Global self-esteem: Its relation to specific facets of self-concept and their importance". In: *Journal of Personality and Social Psychology* 51.6, p. 1224. ISSN: 0022-3514.

Marsland, Stephen (2015). *Machine learning: an algorithmic perspective*. CRC press.

Maurer, John G. (1971). *Readings in organization theory: Open-system approaches*. Random House (NY).

McCarthy, John D., Clark McPhail, and Jackie Smith (1996). "Images of protest: Dimensions of selection bias in media coverage of Washington demonstrations, 1982 and 1991". In: *American Sociological Review*, pp. 478–499.

McCombs, Maxwell E. and Salma I. Ghanem (2001). "The convergence of agenda setting and framing". In: *Framing public life: Perspectives on media and our understanding of the social world*, pp. 67–81.

McCombs, Maxwell E. and Donald L. Shaw (1972). "The Agenda-Setting Function of Mass Media". In: *Public Opinion Quarterly* 36.2, p. 176. ISSN: 0033362X. DOI: `10.1086/267990`.

McCombs, Maxwell E. et al. (1997). "Candidate images in Spanish elections: Second-level agenda-setting effects". In: *Journalism & Mass Communication Quarterly* 74.4, pp. 703–717.

McDonald, Daniel G. (2004). "Twentieth-century media effects research". In:

McGrath, Rita Gunther (1995). "Advantage from adversity: learning from disappointment in internal corporate ventures". In: *Journal of Business Venturing* 10.2, pp. 121–142. ISSN: 08839026.

— (1999). "Falling Forward: Real Options Reasoning and Entrepreneurial Failure". In: *The Academy of Management Review* 24.1, p. 13. ISSN: 03637425. DOI: `10.2307/259034`.

McGrath, Rita Gunther et al. (1992). "Does culture endure, or is it malleable? Issues for entrepreneurial economic development". In: *Journal of Business Venturing* 7.6, pp. 441–458. ISSN: 08839026.

McGuire, William J. (1986). "The myth of massive media impact: Savagings and salvagings". In: *Public communication and behavior* 1, pp. 173–257.

Mead, George Herbert (1934). *Mind, self and society*. Vol. 111. Chicago University of Chicago Press.

Meyer, Bruce D. (1995). "Natural and Quasi-Experiments in Economics". In: *Journal of Business & Economic Statistics* 13.2, p. 151. ISSN: 07350015. DOI: `10.2307/1392369`.

Miller, Dale T. and Michael Ross (1975). "Self-serving biases in the attribution of causality: Fact or fiction?" In: *Psychological Bulletin* 82.2, p. 213. ISSN: 1939-1455.

Minniti, Maria and William D. Bygrave (2001). “A dynamic model of entrepreneurial learning”. In: *Entrepreneurship: Theory and practice* 25.3, p. 5.

Mohri, Mehryar, Afshin Rostamizadeh, and Ameet Talwalkar (2012). *Foundations of machine learning*. MIT press.

Moulya, V. Suchitra and Jayaram Sankaranb (2000). “The tall poppy syndrome in New Zealand: An exploratory investigation”. In: *Transcending boundaries: Integrating people, processes and systems* 285.

Mudinas, Andrius, Dell Zhang, and Mark Levene (2012). “Combining lexicon and learning based approaches for concept-level sentiment analysis”. In: *Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining*, p. 5.

Mueller, Stephen L. and Anisya S. Thomas (2001). “Culture and entrepreneurial potential: A nine country study of locus of control and innovativeness”. In: *Journal of Business Venturing* 16.1, pp. 51–75. ISSN: 08839026.

Mullen, Brian, Rupert Brown, and Colleen Smith (1992). “Ingroup bias as a function of salience, relevance, and status: An integration”. In: *European Journal of Social Psychology* 22.2, pp. 103–122.

Mummendey, Amélie and Sabine Otten (1998). “Positive–negative asymmetry in social discrimination”. In: *European review of social psychology* 9.1, pp. 107–143.

Muoghalu, Michael I., H. David Robison, and John L. Glascock (1990). “Hazardous waste lawsuits, stockholder returns, and deterrence”. In: *Southern Economic Journal*, pp. 357–370.

Nakata, Cheryl and Kumar Sivakumar (1996). “National culture and new product development: An integrative review”. In: *The Journal of Marketing*, pp. 61–72.

Narayanan, Vivek, Ishan Arora, and Arjun Bhatia (2013). “Fast and accurate sentiment classification using an enhanced Naive Bayes model”. In: *International Conference on Intelligent Data Engineering and Automated Learning*, pp. 194–201.

Nielsen, Finn Årup (2011a). “A new ANEW: Evaluation of a word list for sentiment analysis in microblogs”. In: *arXiv preprint arXiv:1103.2903*.

— (2011b). “Afinn”. In: *Richard Petersens Plads, Building* 321.

Niven, David (2001). “Bias in the News Partisanship and Negativity in Media Coverage of Presidents George Bush and Bill Clinton”. In: *The Harvard International Journal of Press/Politics* 6.3, pp. 31–46.

Noelle-Neumann, Elisabeth and Rainer Mathes (1987). “Theevent as event’and theevent as news’: The significance ofconsonance’for media effects research”. In: *European Journal of Communication* 2.4, pp. 391–414.

O’Hare, Neil et al. (2009). “Topic-dependent sentiment analysis of financial blogs”. In: *Proceedings of the 1st international CIKM workshop on Topic-sentiment analysis for mass opinion*, pp. 9–16.

Ogneva, Maria (2010). "How companies can use sentiment analysis to improve their business". In: *Mashable*.

Olson, David L. and Dursun Delen (2008). *Advanced data mining techniques*. Springer Science & Business Media.

Ott Toomet and Arne Henningsen (2008). "Sample Selection Models in R: Package sampleSelection". In: *Journal of Statistical Software* 27.7. URL: `http://www.jstatsoft.org/v27/i07/`.

Özgür, Arzucan, Levent Özgür, and Tunga Güngör (2005). "Text categorization with class-based and corpus-based keyword selection". In: *International Symposium on Computer and Information Sciences*, pp. 606–615.

Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan (2002). "Thumbs up?: sentiment classification using machine learning techniques". In: *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pp. 79–86.

Parker, Simon C. (2013). "Do serial entrepreneurs run successively better-performing businesses?" In: *Journal of Business Venturing* 28.5, pp. 652–666. ISSN: 08839026. DOI: `10.1016/j.jbusvent.2012.08.001`.

Parsons, Talcott (2013). *Social system*. Routledge.

Patzelt, Holger and Dean A. Shepherd (2011). "Negative emotions of an entrepreneurial career: Self-employment and regulatory coping behaviors". In: *Journal of Business Venturing* 26.2, pp. 226–238. ISSN: 08839026. DOI: `10.1016/j.jbusvent.2009.08.002`.

Peng, Mike W. and Oded Shenkar (2002). "Joint venture dissolution as corporate divorce". In: *The Academy of Management Executive* 16.2, pp. 92–105.

Perkins, Jacob (2014). *Python 3 Text Processing with NLTK 3 Cookbook*. Packt Publishing Ltd.

Pischke, J. S. (2005). "Empirical Methods in Applied Economics: Lecture Notes". In: *Downloaded July* 24, p. 2015.

Podolny, Joel M. (1993). "A status-based model of market competition". In: *American journal of sociology* 98.4, pp. 829–872.

— (2001). "Networks as the Pipes and Prisms of the Market 1". In: *American journal of sociology* 107.1, pp. 33–60.

Politis, Diamanto (2005). "The Process of Entrepreneurial Learning: A Conceptual Framework". In: *Entrepreneurship Theory and Practice* 29.4, pp. 399–424. ISSN: 10422587. DOI: `10.1111/j.1540-6520.2005.00091.x`.

Politis, Diamanto and Jonas Gabrielsson (2009). "Entrepreneurs' attitudes towards failure". In: *International Journal of Entrepreneurial Behavior & Research* 15.4, pp. 364–383. ISSN: 1355-2554. DOI: `10.1108/13552550910967921`.

Pollock, Timothy G. and Violina P. Rindova (2003). "Media legitimation effects in the market for initial public offerings". In: *Academy of Management journal* 46.5, pp. 631–642.

Rai, Shailendra Kumar (2008). “Indian entrepreneurs: an empirical investigation of entrepreneur’s age and firm entry, type of ownership and risk behavior”. In: *Journal of Services Research* 8.1, p. 213.

Rao, Hayagreeva (2004). “Institutional activism in the early American automobile industry”. In: *Journal of Business Venturing* 19.3, pp. 359–384. ISSN: 08839026.

Remus, R., U. Quasthoff, and G. Heyer (2010). “SentiWS – a Publicly Available German-language Resource for Sentiment Analysis”. In: *Proceedings of the 7th International Language Resources and Evaluation (LREC’10)*, pp. 1168–1171.

Rerup, Claus (2009). “Attentional triangulation: Learning from unexpected rare crises”. In: *Organization Science* 20.5, pp. 876–893. ISSN: 1047-7039.

Reynolds, Paul D. (1999). “Creative destruction: source or symptom of economic growth”. In: *Entrepreneurship, small and medium-sized enterprises and the macroeconomy*, pp. 97–136.

Reynolds, Paul D. et al. (2000). “Global entrepreneurship monitor”. In: *Executive Report*.

Rindfleisch, Aric et al. (2008). “Cross-Sectional Versus Longitudinal Survey Research: Concepts, Findings, and Guidelines”. In: *Journal of Marketing Research* 45.3, pp. 261–279. DOI: 10.1509/jmkr.45.3.261. URL: http://dx.doi.org/10.1509/jmkr.45.3.261.

Rindova, Violina P., Timothy G. Pollock, and Mathew La Hayward (2006). “Celebrity firms: The social construction of market popularity”. In: *Academy of Management Review* 31.1, pp. 50–71.

Rish, Irina (2001). “An empirical study of the naive Bayes classifier”. In: *IJCAI 2001 workshop on empirical methods in artificial intelligence*. Vol. 3, pp. 41–46.

Rosenbaum, Paul R. and Donald B. Rubin (1983). “The central role of the propensity score in observational studies for causal effects”. In: *Biometrika* 70.1, pp. 41–55. ISSN: 0006-3444. DOI: 10.1093/biomet/70.1.41.

Rosenberg, Morris (1965). “Society and the adolescent self-image”. In:

— (1979). “Conceiving the self”. In: *Clinical Rehabilitation* 22.2, pp. 179–187.

Rubin, Donald B. (1973). “Matching to remove bias in observational studies”. In: *Biometrics*, pp. 159–183.

Samuels, Frederick (1973). *Group images*. Rowman & Littlefield.

Savitsky, Kenneth, Nicholas Epley, and Thomas Gilovich (2001). “Do others judge us as harshly as we think? Overestimating the impact of our failures, shortcomings, and mishaps”. In: *Journal of Personality and Social Psychology* 81.1, pp. 44–56. ISSN: 0022-3514. DOI: 10.1037//0022-3514.81.1.44.

Schachter, Stanley and Jerome Singer (1962). “Cognitive, social, and physiological determinants of emotional state”. In: *Psychological review* 69.5, p. 379.

Schindehutte, Minet, Michael Morris, and Jeffrey Allen (2006). "Beyond achievement: Entrepreneurship as extreme experience". In: *Small Business Economics* 27.4-5, pp. 349–368.

Schumpeter, Joseph Alois (1934). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle*. Vol. 55. Transaction publishers.

Schutjens, Veronique and Erik Stam (2006). "Starting Anew: Entrepreneurial Intentions and Realizations Subsequent to Business Closure". In: *ERIM REPORT SERIES RESEARCH IN MANAGEMENT*.

Schwab, Klaus, Xavier Sala-i-Martin, et al. (2010). "The global competitiveness report 2015-2016". In:

Seki, Yohei et al. (2007). "Overview of Opinion Analysis Pilot Task at NTCIR-6". In: *NTCIR*.

Shafritz, Jay M., J. Steven Ott, and Yong Suk Jang (2015). *Classics of organization theory*. Cengage Learning.

Shane, Scott A. (1992). "Why do some societies invent more than others?" In: *Journal of Business Venturing* 7.1, pp. 29–46. ISSN: 08839026.

— (1993). "Cultural influences on national rates of innovation". In: *Journal of Business Venturing* 8.1, pp. 59–73. ISSN: 08839026.

— (1996). "Explaining variation in rates of entrepreneurship in the United States: 1899-1988". In: *Journal of Management* 22.5, pp. 747–781. ISSN: 0149-2063.

— (2009). "Why encouraging more people to become entrepreneurs is bad public policy". In: *Small Business Economics* 33.2, pp. 141–149.

Shepard, Roger N. et al. (1987). "Toward a universal law of generalization for psychological science". In: *Science* 237.4820, pp. 1317–1323.

Shepherd, Dean A. (2003a). "Learning from business failure: Propositions of grief recovery for the self-employed". In:

— (2003b). "Learning from business failure: Propositions of grief recovery for the self-employed". In: *Academy of Management Review* 28.2, pp. 318–328.

— (2009). "Grief recovery from the loss of a family business: A multi- and meso-level theory". In: *Journal of Business Venturing* 24.1, pp. 81–97. ISSN: 08839026. DOI: 10.1016/j.jbusvent.2007.09.003.

Shepherd, Dean A. and J. Michael Haynie (2011). "Venture failure, stigma, and impression management: A self-verification, self-determination view". In: *Strategic Entrepreneurship Journal* 5.2, pp. 178–197. ISSN: 19324391. DOI: 10.1002/sej.113.

Shepherd, Dean A. and Holger Patzelt (2017). "Researching Entrepreneurial Failures". In: *Trailblazing in Entrepreneurship*. Springer, pp. 63–102.

Short, J. (2009). "The Art of Writing a Review Article". In: *Journal of Management* 35.6, pp. 1312–1317. ISSN: 0149-2063. DOI: 10.1177/0149206309337489.

Simmons, Sharon A. and Johan Wiklund (2011). “Stigma and entrepreneurial failure: Implications for Entrepreneurs’ career choices”. In: *Working Paper*.

Simmons, Sharon A., Johan Wiklund, and Jonathan Levie (2014). “Stigma and business failure: implications for entrepreneurs’ career choices”. In: *Small Business Economics* 42.3, pp. 485–505.

Singer, Slavica, José Ernesto Amorós, and Daniel Moska (2014). “Global Entrepreneurship Monitor 2014 Global Report”. In: URL: `http://www.gemconsortium.org/report`.

Singh, Smita, Patricia Doyle Corner, and Kathryn Pavlovich (2007). “Coping with entrepreneurial failure”. In: *Journal of Management & Organization* 13.04, pp. 331–344.

— (2015). “Failed, not finished: A narrative approach to understanding venture failure stigmatization”. In: *Journal of Business Venturing* 30.1, pp. 150–166. ISSN: 08839026. DOI: `10.1016/j.jbusvent.2014.07.005`.

Sitkin, Sim B. (1992). “Learning through failure: the strategy ofsmall losses”. In: *Research in organizational behavior* 14, pp. 231–266.

Small, Henry (1973). “Co-citation in the scientific literature: A new measure of the relationship between two documents”. In: *Journal of the American Society for information Science* 24.4, pp. 265–269.

Smilor, Raymond W. (1997). “Entrepreneurship: Reflections on a subversive activity”. In: *Journal of Business Venturing* 12.5, pp. 341–346. ISSN: 08839026.

Snyder, Charles Richard, Raymond L. Higgins, and Rita J. Stucky (1983). *Excuses: Masquerades in search of grace*. John Wiley & Sons.

Sokolova, Marina, Nathalie Japkowicz, and Stan Szpakowicz (2006). “Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation”. In: *Australasian Joint Conference on Artificial Intelligence*, pp. 1015–1021.

Star, Alvin D. and Michael Z. Massel (1981). “Survival rates for retailers”. In: *Journal of Retailing* 57.2, pp. 87–99.

Staw, Barry M. and Jerry Ross (1989). “Understanding behavior in escalation situations”. In: *Science* 246.4927, pp. 216–221.

Stokes, David J. and Robert Blackburn (2002). “Learning the hard way: the lessons of owner-managers who have closed their businesses”. In: *Journal of small business and enterprise development* 9.1, pp. 17–27.

Storey, David J. (2016). *Understanding the small business sector*. Routledge.

Stuart, Elizabeth A. (2010). “Matching methods for causal inference: A review and a look forward”. In: *Statistical science : a review journal of the Institute of Mathematical Statistics* 25.1, pp. 1–21. ISSN: 0883-4237. DOI: `10.1214/09-STS313`.

Taboada, Maite et al. (2011). “Lexicon-Based Methods for Sentiment Analysis”. In: *Computational Linguistics* 37.2, pp. 267–307. ISSN: 0891-2017.

Tajfel, Henri and John C. Turner (1979). "An integrative theory of intergroup conflict". In: *The social psychology of intergroup relations* 33.47, p. 74.

Tajfel, Henri et al. (1971). "Social categorization and intergroup behaviour". In: *European Journal of Social Psychology* 1.2, pp. 149–178.

Taylor, Charles R., Ju Yung Lee, and Barbara B. Stern (1995). "Portrayals of African, Hispanic, and Asian Americans in magazine advertising". In: *American Behavioral Scientist* 38.4, pp. 608–621.

Taylor, L., R. Schroeder, and E. Meyer (2014). "Emerging practices and perspectives on Big Data analysis in economics: Bigger and better or more of the same?" In: *Big Data & Society* 1.2. ISSN: 2053-9517. DOI: 10.1177/2053951714536877.

Taylor, Shelley E. and Jonathon D. Brown (1988). "Illusion and well-being: a social psychological perspective on mental health". In: *Psychological Bulletin* 103.2, p. 193. ISSN: 1939-1455.

Tenenbaum, Joshua B. and Thomas L. Griffiths (2001). "Generalization, similarity, and Bayesian inference". In: *Behavioral and brain sciences* 24.04, pp. 629–640.

Tesser, Abraham and Jennifer Campbell (1980). "Self-definition: The impact of the relative performance and similarity of others". In: *Social Psychology Quarterly*, pp. 341–347.

Thomas, James B., Shawn M. Clark, and Dennis A. Gioia (1993). "Strategic sensemaking and organizational performance: Linkages among scanning, interpretation, action, and outcomes". In: *Academy of Management journal* 36.2, pp. 239–270.

Thompson, Cynthia A., Richard E. Kopelman, and Chester A. Schriesheim (1992). "Putting all one's eggs in the same basket: A comparison of commitment and satisfaction among self-and organizationally employed men". In: *Journal of Applied Psychology* 77.5, p. 738. ISSN: 0021-9010.

Tihanyi, Laszlo, Scott D. Graffin, and Gerard George (2014). "Rethinking governance in management research". In: *Academy of Management journal* 57.6, pp. 1535–1543.

Treiman, D. J. and H. I. Hartmann (1981). *Women, Work, Wages: Equal Pay for Jobs of Equal Value*.

Tripathy, Abinash, Ankit Agrawal, and Santanu Kumar Rath (2016). "Classification of sentiment reviews using n-gram machine learning approach". In: *Expert Systems with Applications* 57, pp. 117–126. ISSN: 09574174. DOI: 10.1016/j.eswa.2016.03.028.

Turney, Peter D. (2002). "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews". In: *Proceedings of the 40th annual meeting on association for computational linguistics*, pp. 417–424.

Ucbasaran, Deniz, Paul Westhead, and Mike Wright (2009). "The extent and nature of opportunity identification by experienced entrepreneurs". In: *Journal of Business Venturing* 24.2, pp. 99–115. ISSN: 08839026. DOI: 10.1016/j.jbusvent.2008.01.008.

Ucbasaran, Deniz et al. (2010). “The nature of entrepreneurial experience, business failure and comparative optimism”. In: *Journal of Business Venturing* 25.6, pp. 541–555. ISSN: 08839026. DOI: 10.1016/j.jbusvent.2009.04.001.

Ucbasaran, Deniz et al. (2012). “Life After Business Failure: The Process and Consequences of Business Failure for Entrepreneurs”. In: *Journal of Management* 39.1, pp. 163–202. ISSN: 0149-2063. DOI: 10.1177/0149206312457823.

Van de Kauter, Marjan, Diane Breesch, and Véronique Hoste (2015). “Fine-grained analysis of explicit and implicit sentiment in financial news articles”. In: *Expert Systems with Applications* 42.11, pp. 4999–5010. ISSN: 09574174. DOI: 10.1016/j.eswa.2015.02.007.

Vapnik, Vladimir (1998). *Statistical learning theory. 1998*.

Venkataraman, Sankaran (1997). “The distinctive domain of entrepreneurship research”. In: *Advances in entrepreneurship, firm emergence and growth* 3.1, pp. 119–138.

Verba, Sidney and Kay L. Scholozman (1979). “Injury to insult: Unemployment, class, and political response”. In:

Veropoulos, Konstantinos, Colin Campbell, Nello Cristianini, et al. (1999). “Controlling the sensitivity of support vector machines”. In: *Proceedings of the international joint conference on AI*, pp. 55–60.

Vianen, Annelies Em (2000). “Person-organization fit: The match between newcomers’ and recruiters’ preferences for organizational cultures”. In: *Personnel psychology* 53.1, pp. 113–149.

Wagner, Joachim (2002). “Taking a second chance: Entrepreneurial restarters in Germany”. In:

Wallace, Anthony F. C. and Raymond D. Fogelson (1961). “Culture and personality”. In: *Biennial Review of Anthropology*, pp. 42–78.

Wartick, Steven L. (1992). “The relationship between intense media exposure and change in corporate reputation”. In: *Business & Society* 31.1, pp. 33–49.

Watson, John and Jim E. Everett (1996). “Do Small Businesses Have High Failure Rates? Evidence from Australian Retailers”. In: *Journal of Small Business Management* 34. ISSN: 0047-2778.

Watts, Mark D. et al. (1999). “Elite cues and media bias in presidential campaigns explaining public perceptions of a liberal press”. In: *Communication Research* 26.2, pp. 144–175.

Weick, Karl E. (1979). “The social psychology of organizing (Topics in social psychology series)”. In:

— (1988). “Enacted sensemaking in crisis situations”. In: *Journal of Management Studies* 25.4, pp. 305–317. ISSN: 00222380.

Weigelt, Keith and Colin Camerer (1988). “Reputation and corporate strategy: A review of recent theory and applications”. In: *Strategic Management Journal* 9.5, pp. 443–454.

Wejnert, Barbara (2002). “Integrating models of diffusion of innovations: A conceptual framework”. In: *Annual review of sociology* 28.1, pp. 297–326.

Welpe, Isabell M. et al. (2012). “Emotions and Opportunities: The Interplay of Opportunity Evaluation, Fear, Joy, and Anger as Antecedent of Entrepreneurial Exploitation”. In: *Entrepreneurship Theory and Practice* 36.1, pp. 69–96. ISSN: 10422587. DOI: 10.1111/j.1540-6520.2011.00481.x.

Wennberg, Karl, Saurav Pathak, and Erkko Autio (2013). “How culture moulds the effects of self-efficacy and fear of failure on entrepreneurship”. In: *Entrepreneurship & Regional Development* 25.9-10, pp. 756–780. ISSN: 0898-5626. DOI: 10.1080/08985626.2013.862975.

Wennberg, Karl et al. (2010). “Reconceptualizing entrepreneurial exit: Divergent exit routes and their drivers”. In: *Journal of Business Venturing* 25.4, pp. 361–375. ISSN: 08839026.

Westhead, Paul and Mike Wright (1998). “pNovice, serial, and portfolio founders: Are they different? q”. In: *Journal of Business Venturing* 13.3. ISSN: 08839026.

Westhead, Paul et al. (2005). “Novice, serial and portfolio entrepreneur behaviour and contributions”. In: *Small Business Economics* 25.2, pp. 109–132.

Wiesenfeld, Batia M., Kurt A. Wurthmann, and Donald C. Hambrick (2008). “The stigmatization and devaluation of elites associated with corporate failures: A process model”. In: *Academy of Management Review* 33.1, pp. 231–251.

Wilson, Theresa, Janyce Wiebe, and Paul Hoffmann (2005). “Recognizing contextual polarity in phrase-level sentiment analysis”. In: *Proceedings of the conference on human language technology and empirical methods in natural language processing*, pp. 347–354.

Woodward, Mark (2013). *Epidemiology: study design and data analysis*. CRC press.

Wooldridge, Jeffrey M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wry, Tyler, David L. Deephouse, and Gerry McNamara (2006). “Substantive and evaluative media reputations among and within cognitive strategic groups”. In: *Corporate Reputation Review* 9.4, pp. 225–242. ISSN: 1479-1889.

Wylie, Ruth C. (1979). *The self-concept: Vol. 2. Theory and research on selected topics*.

Yamagishi, Toshio, Nobuhito Jin, and Allan S. Miller (1998). “In-group bias and culture of collectivism”. In: *Asian Journal of Social Psychology* 1.3, pp. 315–328.

Ye, Qiang, Ziqiong Zhang, and Rob Law (2009). “Sentiment classification of online reviews to travel destinations by supervised machine learning approaches”. In: *Expert Systems with Applications* 36.3, pp. 6527–6535. ISSN: 09574174.

Yi, Jeonghee et al. (2003). “Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques”. In: *Data Mining, 2003. ICDM 2003. Third IEEE International Conference on*, pp. 427–434.

Young, Lori and Stuart Soroka (2012). "Affective news: The automated coding of sentiment in political texts". In: *Political Communication* 29.2, pp. 205–231.

Zacharakis, Andrew L., G. Dale Meyer, and Julio DeCastro (1999). "Differing Perceptions of New Venture Failure: A Matched Exploratory Study of Venture Capitalists and Entrepreneurs*". In:

Zhang, Junfu (2011). "The advantage of experienced start-up founders in venture capital acquisition: Evidence from serial entrepreneurs". In: *Small Business Economics* 36.2, pp. 187–208.

Zimmerman, Monica A. and Gerald J. Zeitz (2002). "Beyond survival: Achieving new venture growth by building legitimacy". In: *Academy of Management Review* 27.3, pp. 414–431.

Zuckerman, Miron (1979). "Attribution of success and failure revisited, or: The motivational bias is alive and well in attribution theory". In: *Journal of personality* 47.2, pp. 245–287.