



## Advisory Editorial Board:

DOMINIK VAN AAKEN  
FREDERIK AHLEMANN  
BASTIAN AMBERG  
THOMAS BÄHLINGER  
CHRISTOPH BODE  
ROLF BRÜHL  
JOACHIM BÜSCHKEN  
CATHERINE CLEOPHAS  
LEONHARD DOBUSCH  
RALF ELSAS  
DAVID FLORYSIK  
GUNTHER FRIEDL  
WOLFGANG GÜTTEL  
CHRISTIAN HOFMANN  
KATJA HUTTER  
LUTZ JOHANNING  
STEPHAN KAISER  
NADINE KAMMERLANDER  
ALFRED KIESER  
NATALIA KLIEWER  
DODO ZU KNYPHAUSEN-AUFSEß  
SABINE T. KÖSZEGI  
ARJAN KOZICA  
CHRISTIAN KOZIOL  
TOBIAS KRETSCHMER  
HANS-ULRICH KÜPPER  
REINER LEIDL  
ANTON MEYER  
MICHAEL MEYER  
GORDON MÜLLER-SEITZ  
J. PETER MURMANN  
BURKHARD PEDELL  
MARCEL PROKOPCZUK  
TANJA RABL  
SASCHA RAITHEL  
NICOLE RATZINGER-SAKEL  
ASTRID REICHEL  
KATJA ROST  
MARKO SARSTEDT  
DEBORAH SCHANZ  
ANDREAS G. SCHERER  
STEFAN SCHMID  
UTE SCHMIEL  
CHRISTIAN SCHMITZ  
PHILIPP SCHRECK  
GEORG SCHREYÖGG  
LARS SCHWEIZER  
DAVID SEIDL  
THORSTEN SELLHORN  
ANDREAS SUCHANEK  
ORESTIS TERZIDIS  
ANJA TUSCHKE  
STEPHAN WAGNER  
BARBARA E. WEIßENBERGER  
ISABELL M. WELPE  
HANNES WINNER  
CLAUDIA B. WÖHLE  
THOMAS WRONA  
THOMAS ZWICK

# JUNIOR MANAGEMENT SCIENCE

- Maximilian Piehler**, Prospect Theory and Stock Returns During Bubbles 262
- Astrid Schulte-Huermann**, Impact of Weather on the Stock Market Returns of Different Industries in Germany 295
- Andreas Johannes Dambaur**, Kultur als Determinante des Risikoverhaltens von Individuen – Empirische Evidenz von der PGA TOUR 312
- Laura Schlittchen**, Das Phänomen Influencing – Beeinflussung von Fitness Influencern auf die Sportaktivität 349
- Lara Hilgert**, The Effect of Perceived CEO Overconfidence on Consumers' Product and Brand Evaluations 371
- Jule Neckermann**, Over-Confidence Bias in strategischen Entscheidungsprozessen: Entstehung, Konsequenzen und Lösungsansätze 392



## Prospect Theory and Stock Returns During Bubbles

Maximilian Piehler

*Ludwig-Maximilians-Universität München*

### Abstract

I test the hypothesis that investors evaluate stocks based on the prospect theory value of the distribution of past returns. Because some investors tilt towards stocks with high prospect theory value, these stocks become overvalued and earn low subsequent returns. During bubbles this effect should be stronger, due to rising limits to arbitrage and increased participation of individual investors. I do not find strong support for this prediction in the cross section of returns in U.S. stock markets. In contrast to other variables known to explain returns however, prospect theory value does not lose its predictive power during bubbles. Investors with prospect theory preferences seem to choose stocks whose returns optimally combine low standard deviation with high skewness.

*Keywords:* Prospect Theory; bubbles; limits to arbitrage; individual investors.

### 1. Introduction

#### 1.1. Motivation

The decision whether to buy or sell a stock is risky. In order to understand the buying and selling behaviour in the stock market it is therefore important to investigate how investors make risky choices. In other words, how do investors evaluate a given stock and how can this be integrated into pricing models? In the past century, many researchers modelled investor's choices via expected utility. They assumed that investors maximize their expected utility, which is the sum of the utility of each outcome weighted by its probability. It relies on certain assumptions formalized by Neumann (1944). Expected utility has been a useful tool for research in investor behaviour and asset pricing. Nonetheless, the pricing models based on its assumptions do not explain stock prices with complete precision. Several anomalies seem to persist in the stock market, despite multi-factor pricing models. For example, Jegadeesh and Titman (1993) find that stocks that performed well during the past year subsequently outperform stock that did poorly during the past year. This effect is generally called momentum. De Bondt and Thaler (1985) report on long-term reversal: a stock's returns over the last two to five years on average negatively predict the subsequent return. Similarly, there is a short-term reversal in stock returns i.e. the return of the last month negatively predicts the following return, discovered by Lehmann (1990). More recently, Ang et al. (2006) find that the volatility of

idiosyncratic returns has a negative effect on returns, while Novy-Marx (2013) shows how measures of firm profitability have a positive effect on returns, with a predictive power comparable to that of the book-to-market ratio. In light of these return patterns which are hard to make sense of using classical measures of risk and return, some researchers have focused on behavioural explanations. Maybe it is wrong to assume that investors make rational investment choices in the stock market. Kahneman and Tversky (1979) and Tversky and Kahnemann (1992) offer an alternative model of preferences and choice under risk: prospect theory. They specifically develop prospect theory to address violations of the underlying axioms of expected utility. Unlike expected utility, prospect theory is a descriptive model built on experimental evidence. In order to test the accuracy of this theory, it is necessary to apply it in real world settings. Do investors in the stock market behave according to prospect theory? Barberis et al. (2016) derive testable predictions of prospect theory for average returns in the stock market. I closely follow their work and first test their prediction in my sample. However, I argue that the magnitude of the predicted effect is time-varying; it should be different for bubble and non-bubble periods. But first I want to turn to prospect theory and the original prediction made by Barberis et al. (2016).

#### 1.2. Approach

Prospect theory has not been tested in the stock market very often, maybe because there are some specifications that

are hard to translate into empirical work. These specification issues come down to seemingly simple questions that arise when thinking about the choice making process in prospect theory. Generally, individuals evaluate a risky choice based on a certain set of outcomes and associated probabilities. Before getting to the evaluation though, individuals first need to form a representation of the risky choice they are facing. This representation should be formulated in terms of gains and losses, since prospect theory is defined over gains and losses and not absolute wealth. In contrast to the valuation process of a given set of gains and losses, there is no formal description of the representation process in prospect theory. In a short review I therefore give an overview of the research concerning two related questions: (1) What is the reference point investors use to measure gains and losses?, and (2) Where do investors get the information on returns and how could the presentation form influence their representation of stock returns?. A debate in the research on reference points is whether individuals look to the status quo, where the reference point is initial wealth, or the recently held expectations about an outcome to determine gains and losses. Although a reference point based on expectations might make more intuitive sense, I find research supporting both views. Regarding the sources of information and the impact of presentation form on the perceived return, it becomes clear that individuals could form representations that do not reflect the actually observed information. Nolte and Schneider (2018) for example find that the shape of price paths influences perceived attractiveness of an investment, even keeping the underlying return distribution equal. Accounting for these phenomena and integrating them into an empirical framework indeed is an issue when applying prospect theory. Once the process of representation is finished, Tversky and Kahnemann (1992) provide detailed formulae and parameter estimates to calculate the associated utility i.e. the prospect theory value.

My first assumption is that investors evaluate the attractiveness of a stock based on its returns in the past. From theoretical point of view looking at past returns is not ideal, since a stock's value should depend on future cash flows of the underlying firm, not to mention the anomalies where past returns negatively predict future returns. But in practice many investors might confront issues concerning the availability of resources, be it in the form of time, knowledge or data. Taking aside professional investors, which are not the investor group central to the analysis, the assumption of past return distributions as basis for the subsequent evaluation seems reasonable. Therefore I compute the prospect theory value for each stock using its past return distribution.

For the development of predictions I rely on the model of asset prices by Barberis et al. (2016). In their model, two types of investors populate the market. Rational investors construct the tangency portfolio that is based on the mean-variance framework of Markowitz (1952). Individual investors have a preference for stocks with high prospect theory value and do not like stocks with low prospect theory value. They therefore adjust the tangency portfolio positively (negatively) for high (low) prospect theory value stocks. This

setup results in the model's prediction that stocks with high prospect theory value will earn low subsequent returns. By introducing a small extension to this model, I make a second prediction that the subsequent returns on stocks with high prospect theory value will be lower during bubble periods compared to non-bubble periods. The argument for the stronger effect during bubbles is twofold.

The first part of the argument is that the limits to arbitrage for rational investors should rise, conceding individual investors more influence on prices. There are several papers supporting this view. The most important concept is delayed arbitrage, introduced by Abreu and Brunnermeier (2003). They show that rational investors i.e. arbitrageurs do not necessarily have interest in correcting mispricing during bubbles, because they try to "ride" the bubble and exit right before it collapses. In addition, Scheinkman and Xiong (2003) find that overconfident investors with heterogeneous beliefs about fundamental value can buy stocks exceeding their own valuation. Finally, a combination of these heterogeneous beliefs about fundamental value and the workings of the market for borrowing stock reported by D'Avolio (2002) lead to higher short selling costs during bubbles. The second part of the argument is that the fraction of individual investors increases during bubbles, which leads to a stronger overvaluation of high prospect theory value stocks and lower subsequent returns. Barber and Odean (2008) find that individual investor often buy stocks that grab their attention. Therefore the media coverage of a bubble or stock market rally could incentivise new irrational investor to buy the set of stocks that drive the bubble.

In order to test the original prediction by Barberis et al. (2016) and my own prediction it is necessary to define the exact format of past return distribution and what constitutes a bubble in stock markets. Coming back to the debate over reference points for gains and losses, it is best to test three different benchmarks to measure past returns against: raw returns and returns in excess of both the risk-free rate and the market. While raw returns relate to initial wealth, market returns could be a proxy for investor's return expectations. The information source should determine the frequency and the look back window for the returns in the distribution. The sample I use could be loosely split in half pre-internet and half internet era. In the first half it is reasonable to adopt the convention that investors use investment booklets as the main information source<sup>1</sup>. These booklets often show monthly price charts with a lookback period of three to five years. Past return distribution therefore refers to monthly returns over the past five years in my framework. The important change in the internet era is the sheer amount of information available, although price charts still seem to be prominent<sup>2</sup>. Hence I decide not to switch the definition of return distributions, although it is an interesting thought. Regarding the definition of bubbles I rely on a "rational bubble" model and an econometric method developed in a series of papers by Phillips et al.

<sup>1</sup>See Alexander et al., 1998, Table 3, p.307.

<sup>2</sup>See Glaser et al., 2019, Appendix A Table 1, p.7-8.

(2015a,b) and Phillips et al. (2011). They use right-tailed unit root tests to infer explosive price behaviour by working with the stochastic processes of the underlying dividends and fundamentals.

The empirical analysis begins with the detection and time stamping of bubbles in the S&P 500, which is assumed to be a good proxy for U.S. stock markets in general. I detect the existence of a bubble and the time stamping process results in two phases of explosive prices, from February to August 1987 and from December 1995 to September 2000. These time periods coincide with the bubble leading up to Black Monday and the tech bubble at the end of the last century. The main empirical test consists of Fama and MacBeth (1973, just Fama-MacBeth hereafter) regressions, in order to allow good comparability of my results with the analysis of Barberis et al. (2016). First, the prospect theory value is computed for every month using a stock's returns in excess of the market over the past five years. Then the prospect theory value, together with control variables known to predict returns, serves as the independent variable in a monthly regression on returns. The result of the Fama-MacBeth test is the time-series average of these cross-sectional regressions. I find support for the original prediction that stocks with high prospect theory value earn low returns in U.S. stock markets. In order to test whether this effect is different during bubbles, the sample is split into two sub periods. As a result I cannot confirm that high prospect theory value stocks earn lower returns during bubbles compared to non-bubble periods. Depending on the number of control variables used, the prospect theory value has higher predictive power for returns during bubbles. But the most sophisticated model shows that the effect of the prospect theory value is very similar to non-bubble periods. Portfolio sorts on the prospect theory value point into the same direction. On the surface, the low minus high prospect theory value portfolio earns higher abnormal returns during bubbles. Conducting robustness tests shows that the returns to this strategy remain high even when using different returns, return distribution construction windows, subsamples or excluding low price stocks. However, the returns to the low minus high portfolio are less robust during bubbles, because the returns in the high prospect theory value deciles are not significantly different from zero.

To test the plausibility of the argument that limits to arbitrage are related to the relationship between returns and prospect theory value, I analyse the interaction of prospect theory value with several proxies for the limits to arbitrage. These tests show that the negative effect of prospect theory value on returns is generally stronger for stocks experiencing higher limits to arbitrage. On the other hand they show that rising limits to arbitrage have no impact on the predictive power of the prospect theory value during bubbles. This is at least true for the different proxies for the limits to arbitrage I employ.

In the later part of the paper I take a closer look into the mechanism behind the predictive power of the prospect theory value for returns. Each Fama-MacBeth regression turns off two components of prospect theory by setting their pa-

rameters to one. In addition, the sample is split into bubble and non-bubble periods once again. Loss aversion is found to be the most significant predictor of returns by itself. This indicates that investors with prospect theory preferences should prefer stocks with low standard deviation. In a relevance hierarchy probability weighting comes in second. The prospect theory variable constructed with probability weighting only is not a significant predictor of returns, but in conjunction with loss aversion it contributes positively to the significance. Since probability weighting seems to play at least some role, investors under prospect theory should also prefer skewness. Sorting stocks into deciles based on prospect theory value provides support for these findings. The results indicate that investors with prospect theory preferences like stocks that optimally trade of low standard deviation and high skewness.

This study mainly contributes to the work of Barberis et al. (2016) and shows that the relationship between the prospect theory value and returns is strong and predicts returns almost equally well during bubbles and non-bubble periods. The analysis of the power of known return predictors also shows that behavioural explanations, like prospect theory, can be valuable in research on asset pricing. In addition, my paper provides further evidence on the reliability of the tests for explosive price behaviour developed by Phillips et al. (2015a). The paper is organized as follows. In chapter 2 prospect theory and the construction of return distributions is presented. This entails the presentation and discussion of each component relevant to prospect theory. Subsequently chapter 3 introduces the model of asset price by Barberis et al. (2016). There I develop a small extension to the model and form the hypothesis. Chapter 4 presents the evidence backing the hypothesis, together with an overview of the research on the limits to arbitrage. Chapter 5 consists of a discussion of the history of bubbles and different measurement models of explosive behaviour in prices. In addition, it presents the model for detection and time stamping of bubbles by Phillips et al. (2015a) leading up to the application of the method to the time series of the S&P 500. In chapter 6, I analyse the results from several empirical tests designed to provide evidence for my hypothesis. This part of the paper starts off with the data description and summary statistics and then continues over to portfolio sorts and the Fama-MacBeth regressions designed to be the central empirical test. In later parts of chapter 6 I look more closely into the limits to arbitrage argument and the mechanism behind prospect theory. In the final chapter of the paper I summarize the results and main conclusion and give a direction for future research in related areas.

## 2. Prospect theory

### 2.1. Representation and valuation

In the framework of this study, investors evaluate stocks in a two-step process. The first step is the preliminary analysis of the presented stock, in which some representation of outcomes is formed. How do investors mentally represent a

stock? Which sources do they use as a basis for the representation? In order to answer these questions in the context of prospect theory and the stock market, it is important to define what gains and losses are in the stock market, while also keeping in mind the information sources used by investors. I therefore present a short overview of the related research in section 2.3. But first it is useful to take a closer look at the second step, valuation, during which investors evaluate the representation of the stock they formed earlier. For this part of the process, [Tversky and Kahnemann \(1992\)](#) provide a formal framework and parameter estimates.

2.2. Valuation

2.2.1. Valuation methodology

Based on a critique of the expected utility framework, [Kahneman and Tversky \(1979\)](#) develop a descriptive model of decision making under risk. They analyse the preferred choices of individuals confronted with a set of outcomes and associated probabilities. Because these “prospects” are restricted to two non-zero outcomes, the applicability of the original version of prospect theory is limited. Further, individuals sometimes choose dominated prospects. [Tversky and Kahnemann \(1992\)](#) address these problems in their cumulative formulation of prospect theory. I first present the functional form of this theory and then discuss the components and their implications for decision making under risk.

Let  $X$  be a set of uncertain outcomes  $x$ , where all outcomes except one neutral outcome are gains or losses. The function that assigns each uncertain outcome a probability  $p$  is the prospect to be considered by the individual:

$$(x_{-m}, p_{-m}; x_{-m+1}, p_{-m+1}; \dots; x_{-1}, p_{-1}; x_0, p_0; x_1, p_1; \dots; x_n, p_n). \tag{1}$$

The subscripts  $-m$  and  $n$  stand for the most negative and most positive outcome, respectively. Before individuals determine the value of a prospect they transform probabilities into decision weights:

$$\pi_i = \begin{cases} w(p_i + \dots + p_n) - w(p_{i+1} + \dots + p_n) \\ w(p_{-m} + \dots + p_i) - w(p_{-m} + \dots + p_{i-1}) \end{cases} \tag{2}$$

for  $\begin{matrix} 0 \leq i \leq n \\ -m \leq i < 0 \end{matrix}$ ,

according to the weighting functions

$$w^+(p) = p^\gamma / (p^\gamma + (1-p)^\gamma)^{1/\gamma}, w^-(p) = p^\delta / (p^\delta + (1-p)^\delta)^{1/\delta}. \tag{3}$$

As indicated by the parameters  $\gamma$  and  $\delta$  the weighting function is different for gains and losses. The same is true for the value function:

$$v(x) = \begin{cases} x^\alpha & \text{if } x \geq 0 \\ -\lambda(-x)^\alpha & \text{if } x < 0 \end{cases}, \tag{4}$$

An individual then assigns the prospect in (1) the following value:

$$\sum_{i=-m}^n \pi_i v(x_i). \tag{5}$$

Put into words, an individual analyses each pair of outcome and probability according to (2)-(4) and values the prospect as a sum of these parts. I now turn to discuss essential components of prospect theory, namely probability weighting, reference dependence, diminishing sensitivity and loss aversion.

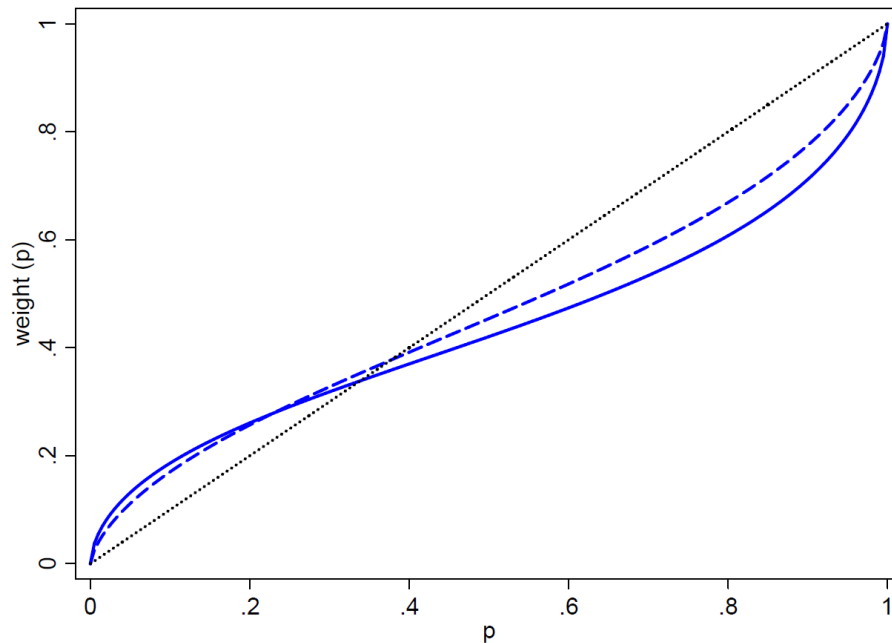
2.2.2. Components of prospect theory and related concepts

*Probability weighting*

Probability weighting refers to the transformation of probabilities into decision weights. These weights are a representation of the perceived impact an outcome has on the evaluation of the prospect. As a consequence decision weights do not equal the objective probability of an outcome. An interesting case in probability weighting is individuals’ preference for lotteries and insurance. [Kahneman and Tversky \(1979\)](#) describe two gambles that illustrate these preferences. In the first gamble, the options are a gain of 5000 with a 0.001 chance or a gain of 5 with certainty. In the second gamble, the options are a loss of 5000 with a 0.001 chance or a loss of 5 with certainty. Notice that the expected value of the options in both gambles is identical. The majority of respondents prefer the very unlikely gain of 5000 in the first gamble, while choosing the certain loss of 5 in the second gamble. These choices show that individuals overweight very small probabilities, causing them to prefer lottery-like outcomes in the domain of gains and insurance in the domain of losses. The weighting function that transforms probabilities into decision weights in prospect theory accounts for this by its inverse s-shaped slope. Figure 1 plots the weighting functions for  $\gamma=0.61$  and  $\delta=0.69$ , as estimated by [Tversky and Kahnemann \(1992\)](#). In the original version the probabilities used as input to the weighting function are exact, individual probabilities. As [Quiggin \(1982\)](#) shows, the application of decision weights to exact probabilities can lead to violations of dominance. He suggests using the entire cumulative distribution of probabilities instead. In cumulative prospect theory, the formulation of decision weights is in accordance with this approach. For any positive outcome, the decision weight is the difference between the probabilities of “the outcome is at least as good as  $x_i$ ” and “the outcome is strictly better than  $x_i$ ”, the weighting function applied to both. For any negative outcome, the decision weight is the difference between the probabilities derived from “the outcome is at least as bad as  $x_i$ ” and “the outcome is strictly worse than  $x_i$ ”, again, the weighting function applied to both. These types of decision weights capture the contribution of a single outcome in context of the entire probability distribution. Consequently the overweighting of small probabilities carries over into an overweighting of the tails of any distribution in cumulative prospect theory.

*Reference dependence*

Reference dependence implies that individuals evaluate the attractiveness of an outcome based on gains and losses in relation to a reference point. In expected utility there is no



**Figure 1:** The weighting function (Based on Tversky and Kahnemann (1992), p.313).

Graphic illustration of the probability weighting function in prospect theory for  $\gamma=0.61$  (solid line) and  $\delta=0.69$  (dashed line).

difference to whether the outcome of a gamble represents a relative increase or decrease in wealth to the individual. Only the final wealth is considered. Intuitively though, the same final wealth can feel different depending on how it was reached, through a gain or a loss. In accordance with this rationale Kahneman and Tversky (1979) find experimental evidence that individuals' choices are dependent on their reference point. It is not clear though, what constitutes such a point. Köszegi, B. and Rabin, M. (2006) for example assume that the reference point is the individual's recently held expectations about an outcome.

#### *Diminishing sensitivity*

The principle of diminishing sensitivity determines the shape of the value function. The value function in prospect theory is concave for gains and convex for losses, as seen in figure 2. This shape is inspired by the proposition that the marginal value of gains and losses decrease with their magnitude. Kahneman and Tversky (1979) motivate this proposition through following problems posed to individuals in their experiments. The options in the first problem are a gain of 6000 with probability 0.25 or gains of 4000 and 2000, each with probability 0.25. The majority of respondents choose option two, implying that the value of the separate gains is larger than the value of the single gain, although both options offer a reward of 6000 in total. Since apparently  $v(6000) < v(4000) + v(2000)$ , the marginal value of smaller gains must be larger and the function for gains convex. The second problem has the same structure, only in the domain of losses, and in turn supports the convexity of the value function for losses. Tversky and Kahnemann (1992) estimate a

median value of 0.88 for  $\alpha$  in equation (4), in accordance with moderate levels of diminishing sensitivity.

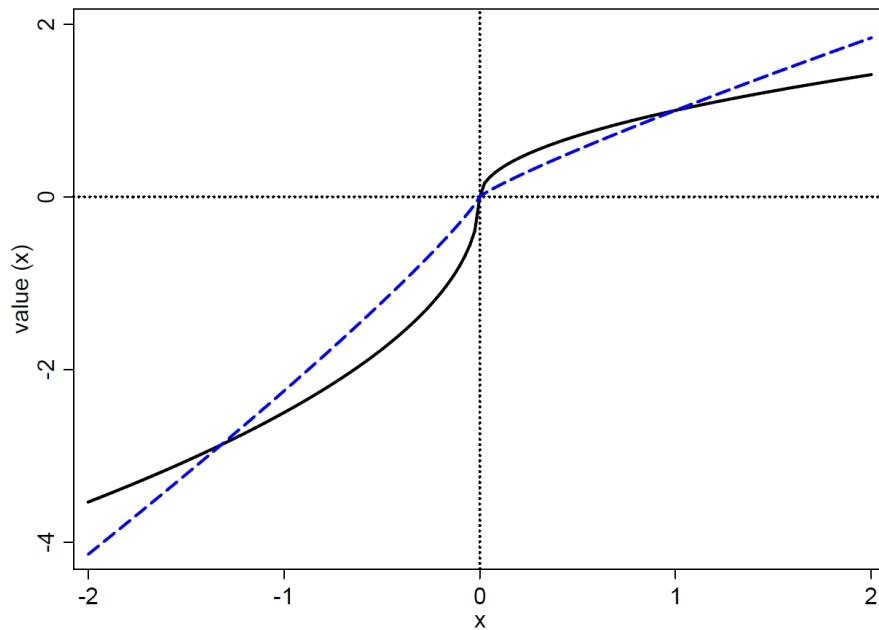
#### *Loss aversion*

Loss aversion is the tendency of individuals to perceive losses stronger than gains of the same size. Therefore the value function for losses is steeper than for gains. This produces a kink at the origin. It illustrates how people are very reluctant to accept gambles with equal probabilities and exactly opposite outcomes. For example, a coin toss with a payment of 50 or -50 seems unattractive, because the individual is much more sensitive to the potential loss. In the value function in (4) the parameter  $\lambda$  determines the degree of loss aversion. Tversky and Kahnemann (1992) estimate a median value of 2.25 for  $\lambda$ , finding evidence of strong loss aversion in individuals' decisions under risk. Loss aversion is a widely accepted concept not only in neuroscience research<sup>3</sup>, but also specifically in the finance literature. Benartzi and Thaler (1995), for example, use investor's loss aversion as a part of their explanation of the equity premium puzzle.

#### *Narrow framing*

One further concept in the analysis of choices under risk, not originally integrated in prospect theory, is sometimes applied together with prospect theory: narrow framing. Tversky and Kahnemann (1981) first describe the effect of "decision framing" on the preferences of test subjects. Decision framing refers to how individuals conceive a prospect i.e. how they perceive the choices they are presented with. The prefer-

<sup>3</sup>See Tom et al., 2007, for a study on the neural basis of loss aversion.



**Figure 2:** The value function (Based on Tversky and Kahnemann (1992), p.311).

Graphic illustration of the value function in Prospect Theory for  $\alpha=0.5$ ,  $\lambda=2.5$  (solid line) and  $\alpha=0.88$ ,  $\lambda=2.25$  (dashed line). The blue dashed line is the parameter specification used in the main analysis of this paper.

ences observed depend on the framing of the choice. Narrow framing is a specific way of framing choices: Each choice is viewed separately from other choices that might have to be made. This means that the current choice is isolated from other choices. Rabin, M. and Weizsäcker, G. (2009) replicate the experimental results of Tversky and Kahnemann (1981) and find in a survey that 89 percent of respondents engaged in narrow framing. The concept of narrow framing is also applicable in the stock market. Some investors might not track the performance of their overall portfolio, but rather the gains and losses of individual stock in the portfolio. Kumar and Lim (2008) use trade clustering as a proxy for narrow framing and show that investors engaging in narrow framing exhibit stronger disposition effects<sup>4</sup>. If investors track gains and losses on the individual stock level, they should indeed be more prone to the disposition effect, because they are risk-averse (risk-seeking) for gains (losses) and there is no offset over the entire portfolio. An intuitive reason why people could engage in narrow framing in the stock market is the complexity of the uncertainty involved. Managing a portfolio requires decision making with the correlations between stocks in mind. This task is cognitively demanding and it is therefore easier to consider only individual stocks.

## 2.3. Representation

### 2.3.1. Gains and losses in the stock market

Before applying the evaluation methods in prospect theory, investors first need to form a representation of the out-

comes associated with a state of the world. In the context of stock market, it would be interesting to know how investors form a mental representation of a stock. What information are they using as the basis for their decision making? Prospect theory is defined over gains and losses. What is the reference point for measuring gains and losses? In the following section, I briefly discuss important research related to answering those questions.

The question of how to define a reference point in reference dependent models of preference, like prospect theory, is not limited to behavioural finance. There is an ongoing discussion in the related behavioural economics research, focussing on two main concepts for reference points. The first proposed reference point is the status quo. An investor in the stock market should care about returns on the wealth she invested. The status quo then is a return of zero over the considered period, since it reflects the initial wealth. Benartzi and Thaler (1995) for example use a return of zero as the reference point and assume initial wealth is reset to the most recent wealth for the subsequent evaluation period. Köszegi, B. and Rabin, M. (2006) propose a different definition of the reference point. In their model, the reference point is the recently held expectations about an outcome. Imagine an investor who predicts a positive stock market trend in the near future. She would expect an increase in stock prices and a positive return on her investment. Her reference point based on her expectations would be the increasing market return. In a setting where no specific market trend is expected by the investor, she could reasonably form her expectations based

<sup>4</sup>The disposition effect is the tendency to sell winners early and hold on to losers.

on the risk free rate<sup>5</sup>. Although the expectations-based definition of the reference point makes intuitive sense, there is one important assumption: individuals correctly predict the environment they face. Even with a short horizon it is questionable whether investors can form correct expectations of the market return or the risk-free rate and also predict their own reaction to them. In their experiments, [Baillon et al. \(2015\)](#) find support for the status quo being the preferred reference point. However, in an empirical study by [Lin and Meng \(2015\)](#), investors in the U.S. stock market have reference points above the risk-free rate, aligning closest with their portfolios past returns. Since there is no conclusive evidence on either side of the argument, three different reference points form part in my framework: zero or raw returns, returns in excess of the risk-free rate and returns in excess of the market. I do not make a specific distinction whether they are based on status quo or expectations. Returns in excess of the market are the benchmark case, because investors could reasonably compare their performance with the overall performance of the stock market.

### 2.3.2. Information sources

With the reference points in mind, what source are investors using to determine returns? Information sources probably vary greatly over time and investor groups. In the preinternet era investment booklets could be a common guide. In fact, [Alexander et al. \(1998\)](#) present survey results from US mutual fund shareholders showing that investment booklets are the most frequently used source of information. Other sources in their survey include financial publications, family and friends or brokers. [Barberis et al. \(2016\)](#) point to the Value Line Investment Survey as an example of an investment prospectus. Each stock page in this prospectus shows a chart of past prices and a future price target on top, fundamental variables over time in the middle and a written analysis at the bottom<sup>6</sup>. The price chart is the only visual representation of data and immediately catches the reader's eyes. It is not unreasonable to assume that the average individual investor looking at this kind of prospectus would be greatly influenced by the price chart in their decision making. Further, the focus of these charts on past price developments could suggest that past prices contain important information on future prices. Therefore I assume that investors use past prices to form their expectations about the future. Specifically investors are assumed to evaluate the attractiveness of a stock based on past return distributions.

But the stock return information from the prospectus is coded into the price charts. What investor biases could alter the information during the transition from price charts to returns? [Nolte and Schneider \(2018\)](#) test how the shape of price paths influences investment decisions. Although the underlying return distribution of each price path presented

to their test subjects is equal, they find that that some price paths are preferred to others. It seems that the framing of price paths affects the perceived attractiveness of the underlying return distribution. It is also becoming clear that, in the mind of an investor, asking for prices and asking for returns is not essentially the same. Confronting subjects with return instead of price charts can lead to lower expectations (see [Glaser et al. \(2019\)](#)). [Glaser et al. \(2007\)](#) show that subjects expect trend continuation if asked for future price changes, while reverting to mean of the series if asked for future price levels. This evidence suggests that investors do not always end up with accurate stock returns, when gathering information from price charts. As consequence using stock returns observed in the marketplace as a depiction of the returns investors have in mind might not be correct. This bias is not only inherent to my framework, but also to other studies applying prospect theory to the stock market (see e.g. [Barberis et al. \(2016\)](#))<sup>7</sup>. The extent of this bias however depends on the assumption that price charts are the sole information source and on the fraction of unsophisticated investors in the marketplace that cannot accurately discern the difference between prices and returns.

About half of my sample period belongs to the internet era. Surely, this has an impact on the information sources used by investors. First, the channel through which many investors trade has moved to brokerage firms that offer online trading. This means that they receive a smaller amount of professional advice from their brokers, but in turn can avoid higher trading fees. Other new sources of information in the internet era are online investment communities or free financial websites like Yahoo Finance. Regarding the question whether one single display method (price charts versus return charts) is preferred, [Glaser et al. \(2019\)](#) observe no clear pattern for the different sources. All of the online investment information sources have in common that they offer investors much more data on stocks than in previous decades. More available information should enhance investors performance, but only if they can use it efficiently for their investment decisions. [Keller and Staelin \(1987\)](#) find that the decision effectiveness of people decreases with the amount of information. Information overload also causes subjects to opt for simple investment alternatives in an experiment reported by [Agnew and Szykman \(2005\)](#). When searching for information online investors have access to much more frequent data. This frequent feedback on stocks could actually hinder performance, as investors focus narrowly on very recent data and lose track of long term performance over multiple time periods (see [Lurie and Swaminathan \(2009\)](#)).

### 2.3.3. Prospect theory value of return distributions

It seems hard to accurately discern which information, and at what frequency, is used by each individual investor.

<sup>5</sup>In theory, even the status quo could be the reference point based on expectations, if the investor expects to retain initial wealth.

<sup>6</sup>See Appendix C, p. XII, where a sample stock page of the Value Line Investment Survey is presented.

<sup>7</sup>In addition [Benartzi and Thaler \(1995\)](#) as well as [Menkhoff and Schmelting \(2006\)](#) assume that investors evaluate stocks using past return distributions. Although they do not further specify how investors would get information on past returns, there might be a possibility that the same bias is inherent to their studies



In face of the uncomplete understanding on what gains and losses are in the stock market, I make following assumptions: (1) investors evaluate a stock based on its past return distribution, (2) investors track the stock's performance with monthly frequency and (3) they consider returns from the past five years in the benchmark case. The assumption of monthly returns is based on the idea that investors could think that they can see trends emerging at this frequency, which influences their perception of a stock. I use the same lookback period for the return distributions as Barberis et al. (2016) to keep comparability high. However, I also present the results of the main analysis for other periods in Appendix D. The prospect theory value is constructed as follows. For each stock, the past sixty monthly returns are sorted from most negative to most positive. The resulting return distribution is equally weighted. As a consequence recent returns have the same impact on the decision of an investor as the earliest returns in the distribution:

$$(r_{-m}, 1/60; r_{-m+1}, 1/60; \dots; r_{-1}, 1/60; r_0, 1/60; r_1, 1/60; \dots; r_n, 1/60) \quad (6)$$

For convenience, the form and subscripts of equation (6) are equal to those of equation (1). It is easy to see that an equally weighted distribution results in the same probability for every outcome. The prospect theory value of this distribution then is the application of equations (2) until (5), keeping in mind the fixed probability. Rearranged in to one equation, investors determine the prospect theory value according to:

$$TK = \sum_{i=-m}^{-1} v(r_i) [w^-(i + m + 1/60) - w^-(i + m/60)] + \sum_{i=1}^n v(r_i) [w^+(n - i + 1/60) - w^+(n - i/60)] \quad (7)$$

The parameters for the shape of the value and probability weighting functions are taken from Tversky and Kahnemann (1992). Other authors confirm the plausibility of their estimation (see e.g. Abdelloui (2000) or Bruhin et al. (2010)).

### 3. Model of asset prices

#### 3.1. The model by Barberis et al. (2016)

The starting point in the model stock market is the mean-variance framework as proposed by classical Portfolio Theory. There are two types of assets; risky assets  $J$  with  $j \in \{1, \dots, J\}$  and the risk-free asset  $r_f$ . A risky asset  $j$  earns return  $r_j$  and is distributed around the mean  $\mu_j$  with standard deviation  $\sigma_j$ . Assets are traded by two types of investors. Rational investors 'arbitrageurs' buy and sell stock maximizing the Sharpe-Ratio of their portfolio. Put another way, they invest in the tangency portfolio as proposed by Portfolio Theory. This portfolio has return  $r_t$ , and the  $J \times 1$  vector  $w_t$  determines the weight of the risky assets  $J$ . On the other hand, the individual investors exhibit prospect theory preferences.

While these investors also start by building the efficient tangency portfolio, they chose to adjust their holdings positively for stocks with high prospect theory value and negatively for stocks with low prospect theory value. The asset weights in their portfolio  $w_p$  are given by:

$$w_p = w_t + kw_{TK} \quad (8)$$

for some  $k > 0$ , with  $w_{TK}$  as the vector of weights in high (low) prospect theory stocks. The weight of the  $j^{th}$  stock in the  $J \times 1$  vector of  $w_{TK}, w_{TK}^j$  is:

$$w_{TK}^j = TK_j - \overline{TK} \quad (9)$$

The variable  $TK$  refers to the prospect theory value of a stock's return distribution as in equation (7). In the benchmark case, the return distributions represent sixty monthly returns in excess of the market. The individual investors in this model do not sort their portfolio on absolute TK values of stock, but rather on higher than average values. Accordingly,  $\overline{TK}$  is the average TK value of all stocks in  $J$ , formally  $\overline{TK} = \sum_{j=1}^J TK_j / J$ .

The fraction of rational investors in the stock market is  $\pi$ . Conversely,  $1 - \pi$  is the fraction of irrational investors that adjust the tangency portfolio for stocks with high (low) TK value. Later in this chapter I will argue that  $\pi$  decreases during bubble periods. Given the fractions of the two types of investors in the population, the market portfolio  $w_m$  is:

$$w_m = \pi w_t + (1 - \pi)(w_t + kw_{TK}) = w_t + \eta w_{TK} \quad (10)$$

where  $\eta = (1 - \pi)k$ .

As stated above, the model of Barberis et al. (2016) starts with a mean-variance framework. Specifically, the basis is the matrix solution for the case of multiple risky assets and a risk-free asset. Guided by this solution, they develop a model of asset prices fitting the specifications in equations (8) through (10). In the resulting model, the mean return of assets  $\mu_j$  is given by:

$$\begin{aligned} \frac{\mu_j - r_f}{\mu_m - r_f} &= \beta_j - \frac{\eta w_{TK}^j s_j^2}{\sigma_m^2 (1 - \eta \beta_{TK})} \\ &= \beta_j - \frac{\eta (TK_j - \overline{TK}) s_j^2}{\sigma_m^2 (1 - \eta \beta_{TK})} \end{aligned} \quad (11)$$

I present a commented proof of their model in the Appendix A. The  $\beta$ -factors are the product of regression of the respective excess returns on the market excess returns. The variable  $s_j^2$  is the covariance of the residuals from the regression of  $j$ 's excess returns. In equation (11) the mean excess return equals  $\beta_j$  of stock  $j$ , adjusted negatively for a term which is increasing in the prospect theory value  $TK_j$  of the stock. This leads to the original prediction that stocks with high prospect theory value earn low subsequent returns.

3.2. Model extension and hypothesis

For the main part of my analysis, I test how this prediction changes during bubble periods in the stock market. I extend the model of Barberis et al. (2016) by partly reinterpreting the variable  $\eta$  appearing in equation (11):

$$\eta = (1 - \pi)k \tag{12}$$

The term for the fraction of irrational investors  $(1 - \pi)$  remains unaltered in its interpretation. The factor  $k$  however subsumes the extent to which rational investors face limits to arbitrage. I argue that both of these factors increase during bubbles. In the following paragraph I turn to the reinterpretation in more detail and show the effect of the proposed increase in  $(1 - \pi)$  and  $k$  on returns in equation (11).

As indicated in equation (8) individual investors split their holdings between the tangency portfolio with weight  $w_t$  and prospect theory portfolio with weight  $w_{TK}$ . The vector of weights  $w_{TK}$  is scaled by some factor  $k > 0$ . There is no further definition of this scaling factor. It subsumes all effects that increase or decrease the weight in the prospect theory portfolio. Limits to arbitrage for rational investors could be a factor that increases the weight in the prospect theory portfolio. During normal market conditions any mispricing introduced by individual investors should be swiftly corrected by contrary arbitrage positions of rational investors. Because of arbitrage the influence of individual investor's preferences on prices and the weight of the prospect theory portfolio in general are kept at low. If the limits to arbitrage rise during bubbles, rational investors have less interest to correct mispricing which gives the individual investor's preferences space to manifest themselves. So my argument is that arbitrage provides a boundary to the power of individual investors over stock prices. In summary, an increase in the limits to arbitrage increases  $k$  which in turn leads to a stronger weight in the prospect theory portfolio<sup>8</sup>. The product of the fraction of individual investors  $(1 - \pi)$  and the factor  $k$  determines the value of  $\eta$ , as in equation (12). Because  $k$  scales the fraction of individual investors by the extent of limits to arbitrage,  $\eta$  represents something akin to the effective participation of individual investors in the stock market. If  $k$  increases while keeping  $(1 - \pi)$  constant, the impact of their preferences on prices still increases: the effective participation of individual investors is larger. During bubble periods, both the fraction of individual investors and the limits to arbitrage should increase. Therefore the effective participation of individual investors should increase as well. The evidence supporting this proposition is presented in chapter four. First however it is necessary to take a closer look at the impact an increase in  $\eta$  has on the model in equation (11), which also leads up to my main hypothesis. If  $\eta$

<sup>8</sup>Although I relate the factor  $k$  to the limit to arbitrage, it is a variable that subsumes any factor that increases the weight in the prospect theory portfolio. Because a high level of mispricing is one of the defining characteristics of bubbles, I come to the conclusion that any hindrance to correct this mispricing should lead to a higher weight in the prospect theory portfolio.

increases the numerator in the fraction on the right hand side increases, while the denominator becomes smaller. The fraction as a whole becomes larger, which increases the deviation from  $\beta_j$ . Put differently; an increase in either  $k$  or  $(1 - \pi)$  i.e. the effective participation of individual investors  $\eta$  increases the mispricing and the deviation from rational prices based on the risk factor  $\beta_j$ . The formal proof of this intuition is the partial derivative of equation (11) with regards to  $\eta$ :

$$\frac{\partial EX}{\partial \eta} = -\frac{(TK_j - TK)s_j^2}{\sigma_m^2(1 - \eta\beta_{TK})^2} \tag{13}$$

where  $EX$  is an abbreviation for the mean excess return. The proof for equation (13) is in Appendix B. Under the assumption that the stock is high prospect theory value  $TK_j > \overline{TK}$ , an increase in  $\eta$  decreases excess returns. If  $TK_j < \overline{TK}$  then the the excess return is rising in  $\eta$ . In the original prediction, stocks with high (low) prospect theory values earn low (high) subsequent returns. An increase in the effective participation rate of individual investors seems to work in the same direction. If  $\eta$  increases the effects should just be larger. My main hypothesis is that stocks with high (low) prospect theory values earn lower (higher) subsequent returns during bubble periods compared to non-bubble periods. Formally, the model representing this prediction is:

$$\frac{\mu_j - r_f}{\mu_m - r_f} = \beta_j - \frac{\eta_{non-bubble}(TK_j - \overline{TK})s_j^2}{\sigma_m^2(1 - \eta_{non-bubble}\beta_{TK})} \tag{14}$$

$$\frac{\mu_j - r_f}{\mu_m - r_f} = \beta_j - \frac{\eta_{bubble}(TK_j - \overline{TK})s_j^2}{\sigma_m^2(1 - \eta_{bubble}\beta_{TK})} \tag{15}$$

where  $\eta_{bubble} > \eta_{non-bubble}$ . This prediction is based on the proposition that the limits to arbitrage and/or the fraction of individual investors in the stock market increase. In the following chapter, I review the evidence for these propositions.

4. Review of the evidence

4.1. Limits to arbitrage

4.1.1. Arbitrage

Under the assumption of efficient markets all public information is incorporated into stock prices. In addition rational investors do not face constraints and can trade freely. If individual investors entered this market and drove stock price away from the fundamental value, rational investors immediately try to profit and correct prices i.e. conduct arbitrage. Since there are no constraints on trading, arbitrage is risk-free and requires no capital and any mispricing after price-shocks should quickly dissipate. What if the assumption on no constraints to trading is lifted? Rational investor now might not have the same incentive to correct mispricing, because they are facing some risks. The existence of risk associated with arbitrage means that rational investors face limits to arbitrage, which make persistent mispricing of

stocks possible. Barberis and Thaler (2003) identify three main risks and costs associated with taking arbitrage positions that could impose limits to arbitrage: fundamental risk, noise trader risk and implementation costs.

#### 4.1.2. Risks associated with arbitrage

##### *Fundamental risk*

A risk factor facing the arbitrageur is changes in fundamental value of stocks. If a rational investor determines that a stock is undervalued, she would try to profit from the mispricing and buy the stock at a low price. Imagine that the fundamental value of the cash flows suddenly decreases, because unexpected negative company information arrives. The fundamental value now has dropped below the buying price and the arbitrageur has incurred a loss. Although shorting is a possibility, it might be hard to find a good substitute. Gromb and Vayanos (2010) explain fundamental risk in two parts, connected to cross-asset arbitrage and intertemporal arbitrage. Cross-asset arbitrage is rational investors exploiting price discrepancies between two different assets at one point in time, and intertemporal arbitrage is exploitation of price differences for one asset at two points in time. They find the impact of demand shocks on prices not only depends on the ability to hedge risk but also the risk aversion of arbitrageurs. Therefore fundamental risk could play an important role in the decision making and investment process of arbitrageurs.

##### *Noise trader risk*

Noise traders are unsophisticated investors that react to noise not related to actual cash flow information. Arbitrageurs face the risk that noise traders drive prices further away from fundamental value for a longer period. As Shleifer and Vishny (1997) point out, many rational investors do not invest their own money but rather their clients' money. The client may have short time horizons and demand positive returns. With this delegated portfolio management, clients could force the arbitrageur to liquidate positions early. If the arbitrageur has to liquidate before noise trader's beliefs bounce back to original levels, they incur a loss<sup>9</sup>. These problems do not only occur with delegation from clients, but also from other parties such as lenders, that make arbitrage not only risky but costly.

##### *Implementation costs*

Since arbitrage often involves shorting stocks, it is important to consider which costs are associated with such a transaction. In order to sell stocks short, an arbitrageur has to borrow the stock by giving some collateral, in most cases cash. Gromb and Vayanos (2010) point out that a low interest rate earned on the collateral compared to the market is costly for the arbitrageur. Further, the desired stock has to be shortable in the first place. D'Avolio (2002) analyses the market for borrowing stock and finds that most of the stocks in CRSP

are shortable. Those that aren't can be found in the category of high illiquid stocks with low market capitalization. In addition, he shows that high loan fees seem to be rare on average, but high for some stocks. Duffie et al. (2002) develop a model where borrowers and lenders in the stock market bargain over the terms of their agreement. This leads to friction cost if the lender has problems finding lendable shares. Costs for conducting arbitrage transaction can also come from insufficient capital. Arbitrageurs do not invest their own money, so they either have to lend it directly or attract investor capital. Gromb and Vayanos (2002) show that financially constrained arbitrageurs are sometimes not able to hold their arbitrage positions. If the difference between the arbitrageur's position and the price increases, some lenders might first require higher margins and then terminate the loan, if possible. This is similar to the case of Shleifer and Vishny (1997), where investors pull out capital when they see a negative arbitrage position.

#### 4.1.3. Measures for the limits to arbitrage

##### *Proxies for the limits to arbitrage*

Because the limits to arbitrage are difficult to measure by a single variable, researchers have tried to find adequate approximations instead<sup>10</sup>. Mashruwala et al. (2006) rely on idiosyncratic risk as a measure of the risk an arbitrage position entails. Their argument is that only a part of a stock's volatility can be hedged, specifically the part that is known to the arbitrageur, which is assumed to be the volatility explainable by the standard market model. The other part of the stock's volatility, coming from the residuals in the standard market model regression, is an approximation for the absence of close substitutes in the hedging process. This absence is a limit to arbitrage and a risk that the arbitrageur faces. Brav et al. (2009) also use idiosyncratic risk as a proxy for limits to arbitrage, but take the residual variance from a four-factor model, instead of the standard one-factor model. Size is also considered a factor related to the limits of arbitrage. Because the stocks of firms with large market capitalization tend to be easier to trade, borrow or short, size is an inverse estimate for the limits to arbitrage. In addition to size, Ali et al. (2003) use several measures of transaction costs to study the limits to arbitrage. They pick stock price as a measure of direct transaction costs, because it is inversely related to the bid-ask spread. Similar to Mashruwala et al. (2006), they employ the dollar trading volume as an indirect measure of transaction cost. A higher dollar trading volume suggests that an arbitrage position could be executed without delays and for low cost. Stoll (2000) also finds that the ask-bid price spread is negatively related to measures of trading activity, such as the average daily dollar trading volume. Theoretically there should be fewer frictions when implementing arbitrage strategies for stocks with high trading volume.

<sup>9</sup>See De Long et al., 1990, p.705.

<sup>10</sup>See Chu et al., 2017, p.5-6, for an overview of different approximations of limits to arbitrage.

Therefore the dollar trading volume is an inverse proxy for the limits to arbitrage.

#### *Proxies for short-sale constraints*

An arbitrageur has to have the ability to sell stock short, if she wants to hedge against adverse effects to her arbitrage position. As discussed in the paragraph on implementation costs, frictions in the search process for shorable stocks can be a limit to arbitrage. [Chen et al. \(2002\)](#) measure short-sale constraints indirectly by looking at the breadth of ownership, the number of investors with long positions in a stock. This is based on an argument by [Miller \(1977\)](#) that if short-sale constraints exist, stock prices are anchored around the valuation of optimists and pessimists do not have any impact on prices. Therefore a low breadth of ownership, a sign that mainly optimists trade a stock, should be an approximation of short-sale constraints. [Nagel \(2005\)](#) employs another measure related to the ownership, institutional ownership. He argues that borrowing stock for shorting is easier for stocks with high institutional ownership, related to the findings by [D'Avolio \(2002\)](#). Therefore high institutional ownership should indicate a lower degree of short-sale constraints. Similarly, [Hirshleifer et al. \(2011\)](#) use institutional holdings as proxy for availability of loanable shares. If the loanable shares are scarce, short-selling is difficult. Finding loanable shares should be easier for stock with high institutional holdings. In a second strand of the literature, short interest, the fraction of shares sold short to shares outstanding, is a constraint on the arbitrageur side. While the proxies using institutional holdings variables focus on the supply side, short interest can be seen as the demand side of the short-sale constraint issue ([Asquith et al. \(2005\)](#)). In this framework, a high short interest leads to more arbitrageurs competing for the shorable shares, increasing cost and building a short-sale constraint.

#### 4.1.4. Limits to arbitrage and implementation costs during bubbles

For the main part of my analysis it is important to think about how the extent of limits to arbitrage could change during bubble-periods compared to non-bubble periods. I make a three-fold argument based on synchronization issues ([Abreu and Brunnermeier \(2003\)](#)), resale options ([Scheinkman and Xiong \(2003\)](#)) and implementation costs ([D'Avolio \(2002\)](#)). [Abreu and Brunnermeier \(2003\)](#) present a model of how bubbles in asset prices can form and persist due to synchronization issues among arbitrageurs. It is a description of how rational investor behaviour could result in a failure to correct mispricing. In the model, assets prices are assumed to deviate from fundamental value at some point. One after another, the arbitrageurs become aware of the mispricing. They have prior knowledge of how bubbles evolve. Therefore they want to profit from the mispricing as long as the bubble grows and exit right before it collapses. The bursting of the bubble occurs when a sufficient amount of arbitrageurs decides to sell off the positions that profit from the mispricing. But they have no knowledge when other ar-

bitrageurs became aware of the bubble or when other will try to cancel out their respective positions. Arbitrageurs have to try to time the market, because only a fraction of arbitrageurs can exit before the bubble collapses. The incentive to correct any mispricing in asset prices is therefore delayed until the point when a sufficient amount of arbitrageurs think the bubble will burst soon. This is a strong argument for the increased limits to arbitrage during bubble periods. [Griffin et al. \(2011\)](#) support the model of delayed arbitrage by rational investors during the dotcom-bubble. They find that institutional investors pull capital in the months before the bubble bursts, while individual investors increase their exposure. It seems that arbitrageurs drive the collapse of the bubble and delay the correction of mispricing. [Scheinkman and Xiong \(2003\)](#) offer a different approach to explain the continued overvaluation of assets during bubbles. In their model an investor not only buys an asset, but also the option for resale to others. The investors are not fully rational and exhibit overconfidence. Every investor is confident that her view of fundamental value is correct. This variability in beliefs leads to a situation where the resale option for the asset can be exercised with other overconfident investors, which see higher fundamental value in the asset. Therefore investors buy assets with prices in excess of their own evaluation of fundamental value. The resulting mispricing is a consequence of the irrational behaviour of investors. If the overconfident investors make up a significant part of the market, the limits to arbitrage rise. Another requirement for successful arbitrage is the ability to sell short. [D'Avolio \(2002\)](#) studies the characteristics of the market for lending and borrowing stock i.e. the processes underlying short selling. In the sample, most stocks are shorable for relatively low fees. Only few stocks demand very high loan fees. However, short-selling becomes more difficult when investor opinions on the value of a stock diverge. Imagine a short seller with a stock valuation lower than the average investor in the market. Since other investors are more optimistic, the lender could cancel the loan of the short seller and receive a higher fee. The short seller can now either close out the position by buying back the stock or chose to pay a higher fee. In both situations the short seller faces more costs associated with her position. It is reasonable to assume that, during bubble periods, some investors overvalue more than other, similar to the overconfidence model of [Scheinkman and Xiong \(2003\)](#). Therefore the short selling cost should rise during bubble periods, increasing the limits of arbitrage. In summary, any of the three explanations makes for a theoretical argument to why limits to arbitrage should rise during bubbles. Since evidence for the delayed arbitrage model by [Abreu and Brunnermeier \(2003\)](#) comes from the time period of the tech bubble and this accounts for the majority of bubble months I detect in my sample later, it is the most appealing explanation. However, the delayed arbitrage model cannot be measured with the proxies for limits to arbitrage or short-sale constraints above. Analysing the change in trading behaviour of arbitrageurs during the bubble like [Griffin et al. \(2011\)](#) seems to be the only way

to test the delayed arbitrage model. The explanation based on Scheinkman and Xiong (2003) runs into similar issues. While it makes intuitive sense that overconfident investors continuously exercise resale options, it is hard to test. Explanations based on investor behaviour in general run into this issue. In these models, investors do not necessarily face higher limits to arbitrage, because they consciously decide to deviate from the own fundamental valuation of stocks. Only later, when the price pressure of this type of investor is high, arbitrageurs face higher limits to arbitrage.

#### 4.2. Individual investor's participation during bubbles

The second proposition related to my hypothesis is that the fraction of irrational investors might increase during bubble periods compared to non-bubble periods. How could irrational investors be drawn in to participating in the stock market? Maybe their decision to invest in stocks depends on attention-grabbing news media coverage. If ever more positive information of returns to be earned arrives, more irrational investors might consider participating. Barber and Odean (2008) find that individual investors buy stocks that grab their attention, because they are faced with a search problem. There are many stocks they could buy, but they can only choose between those that they notice. Consequently they buy stocks that recently grabbed their attention. Institutional i.e. rational investors on the other hand do not confront a search problem as large as individual investors. They can narrow their search for stocks by demanding certain criteria (for example sector or certain multiples). Therefore institutional investors do not buy attention grabbing stocks to the same extent. During bubble periods individual investors receive continuous positive information about a set of stocks. They are incentivised to buy stocks driving the bubble and contribute to the overvaluation. Effectively, they have a greater influence on prices than during non-bubble periods. The evidence on higher participation of irrational investors during bubbles is scarce. Although Brennan (2004) observes an increase in individual investors owning stock during the time of the tech bubble, the reason was more the democratization of equity markets in the U.S. and not attention-grabbing stock information. Since one of the bubbles my test detects is the tech bubble, I might have to accept the notion that this one time increase in individual investors in the U.S. could drive results and not something inherent to all bubbles. However this would only be true if the results indicated a long term change even after the tech bubble. As an argument, it does not seem unreasonable that more individual investors will think about buying stocks during bubbles, because they are fed information about high returns.

In order to test my hypothesis and the propositions made, it is necessary to detect and time-stamp bubbles in my sample. In the following chapter I therefore discuss what constitutes a bubble and how to measure it, followed by presenting the econometric test for bubbles I apply to my sample.

## 5. Bubbles

### 5.1. A short history of bubbles

When hearing about bubbles, a sharp rise in prices coupled with investor euphoria and a subsequent bounce back of prices to pre-bubble levels comes to mind. Although bubbles seem like a new phenomenon, many date them back until the tulip mania that occurred in the Netherlands leading up to 1637. There, prices for rare species of tulips increased rapidly until, in some cases, they crashed to 10 percent of their maximum<sup>11</sup>. Two other commonly identified bubbles before the start of the 20th century are the Mississippi bubble and the South Sea Bubble. Both shared similar characteristics, they occurred when private enterprises sought to refinance government debt. The Mississippi bubble was closely connected to the French *Compagnie des Indes*, which successively took over large projects like collecting taxes or running trade monopolies outside Europe, financed by the issuances of shares. In the beginning stages of the *Compagnie des Indes* shares sold at 500 livres. From the start of the expansion if its operations in April 1719 to the height of the bubble share prices rose to 10,000 livres in February 1720 until falling sharply after May 1720. The South Sea Bubble happened during the same time period, but the English South Sea company had the refinancing of government debt as its only goal. The acquisition of government debt was again finance through share issues. From January until July 1720 the share prices rose from close to 100 pounds to 900 pounds, before crashing down in September. Although both of these developments are often seen as bubbles, Garber (1990) argues that there were rational explanations for the respective price declines<sup>12</sup>. The next period of explosive prices probably is more well-known today: the bubble leading up to the great market crash in 1929. White (1990) attributes the bull markets in the 1920s to the stable growth after the post-war recession and large companies' trend towards equity financing of new plants and equipment. In addition, commercial banks started to make investment for individuals easier through investment trusts. While stock prices and dividends commoved during most of the 1920s, prices started to run away from 1928 onwards up to the crash<sup>13</sup>. For the following decades there is now consensus whether bubbles existed in the financial markets. One of these shorter and more contested bubbles is the one before Black Monday in October 1987. Some argue that it was computer high frequency trades that caused the crash but there is no broadly accepted version of why a bubble could have existed. The most recently recorded bubble in stock prices is the tech bubble in the late 1990s. Driven by the promise of the internet economy, investors heavily bought in to stocks of technology firms, although the fundamentals did not justify the high valuations. Shiller (2000) describes how investors often willingly ignored concerning facts and

<sup>11</sup>See Garber, 1990, p.37.

<sup>12</sup>See Garber, 1990, for an overview of the three bubbles and alternative explanations.

<sup>13</sup>See White, 1990, Figure 3, p.73.

were blinded by positive news coverage. Ofek and Richardson (2002) show that the internet sector was highly overvalued, where the valuations of a quarter of the firms suggest a required excess return on capital of almost 50 percent over a 10 year period<sup>14</sup>.

## 5.2. Detecting bubbles in asset prices

Because of the enormous impact bubbles can have on the economy and the livelihoods of many people, economists have tried to develop methods for testing explosiveness in asset prices. Gürkaynak, R. S. (2008) provides a survey of the literature on econometric methods. These methods rely on the rational bubble model, where the price is not only based on cash flow streams from dividends and fundamentals, but also has a bubble component. Investors incorporate this component into their valuation and accept the corresponding prices. The variance bound test by Shiller (1981) is based on a model of dividend cash flows. With perfect knowledge of the future dividend streams, the asset price would reflect the present value of the resulting cash flows. Since investors don't know the future, the price reflects the present value of expected dividends plus the forecast error. Consequently, the variance of this price together with the variance of the forecast error should be equal to the variance of prices with perfect knowledge. If they exceed this variance bound, a possible reason is the existence of a bubble in prices. There are several issues with this method. First, it should be noted that it was originally not supposed to be a specific test for bubbles, but rather market efficiency. However, Marsh and Merton (1986) conclude that the variance bound test does not even test for market rationality, when changing the assumptions underlying the stochastic process for dividends. In addition it can only be applied as an ex post test, because the calculation of the variance bound requires information on prices formed on perfect knowledge. West (1987) develops another test that is supposed to indirectly prove the existence of bubbles. The starting point is a model where asset prices are determined by the discounted dividend streams. The basic idea is to compare a set of parameters from different regressions, one run on an actual time-series and the other one run on simulated data. The first set of parameters is obtained from a regression of the stock price on lagged dividends. The second set comes from a pair of equations, determining the dividend process and discount rate. Because both regressions estimate the same model, there is no bubble if the parameters are equal. If there is a bubble the underlying model is slightly different from before. Now asset prices are determined by discounted dividends and a bubble component. The parameter estimates from the regression with actual data are then biased. However, the second set of parameters obtained by simulation is still consistent. The test by West (1987) compares the two sets of parameters. If they are different after controlling for sampling errors, this is indirect evidence of a bubble. Problems for this method arise

in the specification of the stochastic process for dividends. Both of the test presented until now rely on ruling out any other option than bubbles, instead of testing for them directly. Diba and Grossman (1988) design a more direct test of bubbles and lay the foundation for the test by Phillips et al. (2015a,b) that I use to detect bubbles. Prices in their model are determined by dividends and unobservable fundamentals. Essentially, they conclude that if no bubble exists, the processes for dividends and prices should be stationary after the equal amount of differencing. The explosive process for bubbles in their model breaks this relationship. Therefore a bubble exists if the same amount of differencing that makes the process for dividends stationary, does not make prices stationary, which can be tested with unit-root tests. Evans (1991) doubts the power of such tests. He shows that the specification of the explosive bubble process is problematic: for periodically collapsing bubbles, unit root tests have low power of detecting an explosive process. Diba and Grossman (1988) assume that bubbles cannot restart after bursting and therefore their test would fail to detect explosive prices with periodically collapsing bubbles. Although the idea behind the test is promising, reliably testing for bubbles requires a different method, where unit root tests do not lose detection power. This is the issue that Phillips et al. (2015a,b) address with their method that I present in the next section.

## 5.3. The Phillips et al. (2015a) approach for bubble detection

### 5.3.1. Rational bubble model

In a recent series of papers Phillips et al. (2011) and Phillips et al. (2015a,b) develop a new approach based on unit root testing that aims to correct for these fallacies. In order to understand how unit root tests can help to identify explosiveness in asset prices, it is useful to start with a rational bubble model, similar to those discussed in the previous section:

$$P_t = \sum_{i=0}^{\infty} \left( \frac{1}{1+r_f} \right)^i \mathbb{E}_t (D_{t+i} + U_{t+i}) + B_t \quad (16)$$

where  $P_t$  is the after-dividend price of the asset,  $D_t$  is the payoff from dividends,  $U_t$  represents the unobservable fundamentals and  $\mathbb{E}_t$  is an expectation operator. Accordingly the price of an asset depends on two components. A fundamental component that is based on expected cash flows from dividends  $D_t$  and fundamentals  $U_t$ , dubbed  $P_t^f$  here, and a bubble component  $B_t$ . This results in a simplified form of equation (16):

$$P_t = P_t^f + B_t \quad (17)$$

The stochastic process underlying the bubble component has a submartingale property<sup>15</sup>:

$$\mathbb{E}_t (B_{t+1}) = (1+r_f)B_t \quad (18)$$

<sup>14</sup>See Ofek and Richardson, 2002, p.270.

<sup>15</sup>See Phillips et al., 2011, Figure 2, p.204, for an illustration of this kind of process.

Therefore the expected value for the bubble component in  $t + 1$  is an increasing function of the bubble component in  $t$ . This is in accordance with the self-reinforcing price expectations during bubbles. Under the assumption that  $B_t = 0$  asset prices are driven by the characteristics of  $D_t$  and  $U_t$ . If now the stochastic processes for dividends and unobservable fundamentals are at most integrated of order one, then the fundamental component  $P_t^f$  in general is at most integrated of order one, then the fundamental component  $P_t^f$  in general is at most integrated of order one. In this case asset prices follow a random walk. If there is a bubble in asset prices i.e. ( $B_t \geq 0$ ), then the explosive process described in equation (18) would dominate the behaviour of asset prices. Therefore it is possible to infer the existence of a bubble by testing whether prices follow a random walk or an explosive autoregressive process. One disadvantage of testing prices is the possibility that the detected explosiveness is driven by the fundamental component rather than a bubble. Using the price-dividend ratio controls for a false positive interpretation of test results, assuming that the unobservable fundamentals are still at most driven by a I(1) process. Under the null hypothesis asset prices follow the process in equation (19):

$$y_t = dT^{-\eta} + \theta y_{t-1} + \varepsilon_t, \varepsilon_t \sim iid(0, \sigma^2), \theta = 1 \quad (19)$$

where  $d$  is a constant,  $T$  is the sample size and  $\eta$  is a control for the size of the intercept as the sample size approaches infinity. Phillips et al. (2015a) focus on the case of  $\eta \geq 1/2$ , for which the martingale component dominates the drift and  $y_t$  is close to a pure random walk. Phillips et al. (2014) report on the sensitivity of right-tailed unit root testing to specification of the null hypothesis. Instead of testing the null for stationarity against a random walk with intercept or trend, a random walk is the null against the explosive alternative. This corresponds to analysing the right tail of the distribution in unit root testing. To test for a unit root, Phillips et al. (2015a) apply Augmented Dickey-Fuller (ADF hereafter) tests on different subsamples using rolling regressions with sample start  $r_1$  and end  $r_2$  using following form:

$$y_t = \hat{\alpha}_{r_1, r_2} + \hat{\beta}_{r_1, r_2} y_{t-1} + \sum_{i=1}^k \hat{\psi}_{r_1, r_2}^i \Delta y_{t-i} + \hat{\varepsilon}_t \quad (20)$$

where  $k$  is the number of lags and the ADF statistic is the t-statistic of the  $\beta$ -coefficient. The recursive approach is chosen based on the critique by Evans (1991) that unit root tests incorrectly detect stationary behaviour in a model with periodically collapsing bubbles.

### 5.3.2. SADF and GSADF test

The supremum augmented Dickey-Fuller test (SADF hereafter) is the first method that Phillips et al. (2011) develop, which applies unit root tests to an expanding sample sequence. The test statistic is obtained by running the regression in equation (20) on this sequence. Define the total

sample as running on an interval of  $[0, 1]$ ,  $r_0$  as the minimum window size for the regression and  $r_1$  and  $r_2$  as the starting point and endpoint of the subsample, respectively. The first regression runs on the minimum window size  $r_0$ . Leaving the starting point  $r_1$  fixed, the sample size increases by one observation at a time until the total sample size is reached. The supremum value of the resulting ADF test statistic sequence is the SADF test:

$$SADF(r_0) = \sup_{r_2 \in [r_0, 1]} ADF_0^{r_2} \quad (21)$$

Imagine for example a series of 100 observations and a fixed minimum window size  $r_0 = 0.1$ . The first ADF test would then run a regression on observations 1 through 10, the second on observations 1 through 11 etc. until the last regression included all 100 observations. Figure 3 illustrates this procedure. Several papers explore this recursive approach for other unit root tests, such as the Bhargava (1986) statistic, the Busetti and Taylor (2004) statistic, the Kim (2000) statistic or the Chow-type Dickey-Fuller statistic<sup>16</sup>.

The extension of this test is the general supremum augmented Dickey-Fuller test (GSADF hereafter) by Phillips et al. (2015a). This test is based on the same concept as the SADF test, but the number of subsamples used is much larger. Instead of only varying the endpoint  $r_2$ , the starting point  $r_1$  is flexible as well. This double-recursive procedure is meant to make the detection of periodically collapsing bubbles more precise. Figure 4 illustrates the formation of subsamples the test is performed on. The GSADF test then is the supremum value of the resulting ADF test statistic sequence:

$$GSADF(r_0) = \sup_{\substack{r_2 \in [r_0, 1] \\ r_1 \in [0, r_2 - r_0]}} ADF_{r_1}^{r_2} \quad (22)$$

The recommendation for the minimum window size  $r_0$  is dependent on the size of the full sample  $T$  according to the rule  $r_0 = 0.01 + 1.8\sqrt{T}$ . On the one hand this ensures that the initial window is small enough to capture explosive price behaviour early in the sample if  $T$  is large. On the other hand the initial window is big enough for adequate estimation if  $T$  is small.

### 5.3.3. Time stamping bubbles

For empirical application it is important to know not only whether bubbles exist in a sample, but also when they originate and terminate. Since real-time stamping of bubbles is an important concern in practice, Phillips et al. (2015a) use a backward-looking implementation of the SADF test, called the backward sup augmented Dickey-Fuller test (BSADF hereafter). The endpoint  $r_2$  is held fixed, while the starting point  $r_1$  expands backwards towards the beginning of the full sample. The BSADF is then again the supremum value

<sup>16</sup>See Homm and Breitung, 2012, for a comparative analysis of these test statistics.

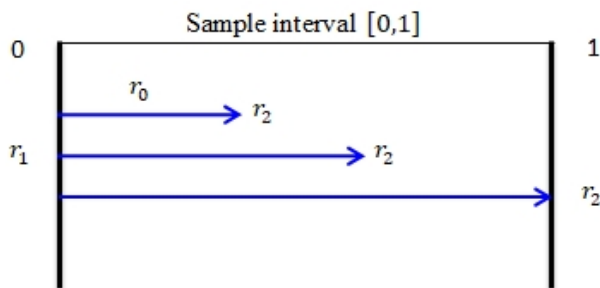


Figure 3: The SADF Test (Based on Phillips et al., 2015a, p.1049).

Set  $r_1 = 0$  and  $r_2 \in [r_0, 1]$ . Use initial window  $[0, r_2]$  and expand  $r_2$ .

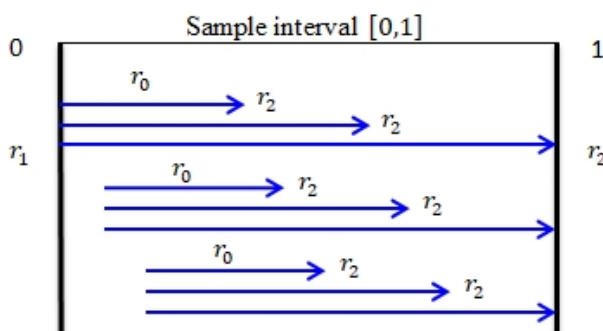


Figure 4: The GSADF Test (Based on Phillips et al., 2015a, p.1049).

Set  $r_1 = [0, r_2 - r_0]$  and  $r_2 \in [r_0, 1]$ . Use moving window  $[r_1, r_2]$ .

of the resulting statistics sequence:

$$BSADF_{r_2}(r_0) = \sup_{r_1 \in [0, r_2 - r_0]} \{ADF_{r_1}^{r_2}\}. \tag{23}$$

Similar to the double-recursive sampling technique of the GSADF test, varying the endpoint  $r_2$  results in a statistic sequence of BSADF tests. This sequence consists of one BSADF statistic for every observation of the original sample provided the minimum window size restriction is fulfilled. Working from the first chronological test statistic to the last, bubble periods are defined according to the following equations:

$$\hat{r}_e = \inf_{r_2 \in [r_0, 1]} \{r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T}\} \text{ and} \tag{24}$$

$$\hat{r}_f = \inf_{r_2 \in [\hat{r}_e + \delta \log(T)/T, 1]} \{r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T}\} \tag{25}$$

where  $scv_{r_2}^{\beta_T}$  is the critical value for the ADF statistic based on  $Tr_2$  observations. Here it is important to notice that a critical value is needed for every month in the sample. The origination of a bubble  $\hat{r}_e$  is then the first chronological observation that passes the respective critical value. In order to avoid spurious explosive behaviour to be identified as a bubble a minimum duration based on sample size is required. After origination the explosive behaviour should persist at

least for  $\delta \log(T)$  observations. This rule is somewhat arbitrary, since  $\delta$  is a parameter that can be chosen freely according to the frequency of the data and assumed minimum bubble duration. For example, if the data frequency is monthly, sample size  $T = 480$  and bubbles should last a half a year at minimum, then  $\delta \approx 2.2$ . The end of a bubble  $\hat{r}_f$  is the first observation after  $\hat{r}_e + \delta \log(T)$  that falls below the critical value. This procedure can be used to time-stamp multiple bubbles that periodically expand and collapse, not only in stock prices, but also in many other assets traded on exchanges.

### 5.3.4. Other applications

Since its inception the BSADF test has been applied to detect bubbles in several assets. Cheung et al. (2015) search for bubbles in bitcoin prices, which seem to be driven by speculative behaviour. They find several smaller bubbles and one large bubble, collapsing the Mt. Gox exchange in 2014. Caspi et al. (2018) set out to detect explosiveness in oil prices. They test the nominal oil price against prior levels and the general price level in the U.S. and find periods of explosiveness in the 1970s and 1980s, as well as before the financial crisis in 2008. Escobari and Jafarinejad (2016) detect bubbles in real estate investment trust and are able to capture, among others, the reflection of the housing bubble before 2008. Hu and Oxley (2017) search for explosiveness in currency exchange rates and conclude that there were no bubbles in the



G10 countries during 1990 and 2014. In summary, this evidence suggests that the approach by Phillips et al. (2015a) can be readily applied to detect bubbles in a broad range of settings. I now turn to the application of this test to the U.S. stock markets.

#### 5.4. Bubbles in the S&P 500

##### 5.4.1. Data

The tests for the existence of bubbles and their time-stamping are performed on the S&P 500. This index covers approximately 80% of all available market capitalization of American equities<sup>17</sup>. The data for the S&P 500 is taken from Robert Shiller's website at Yale<sup>18</sup>. It consists of the monthly time-series of the index price, dividend and earnings. Additionally a consumer price index published by the U.S. Bureau of Labour Statistics allows for the conversion to real values. The data sampled stretch over 480 monthly observations from January 1978 up to December 2017. As indicated by the "rational bubble" model in section 5.3.1, using the price-dividend ratio controls for false positive interpretation of results. To obtain the price dividend ratio, I first compute the percentage dividend yield, which is real dividend divided by real price. The price dividend ratio then is the reciprocal value of the percentage dividend yield. The tests for bubbles run test regressions on an index of this price dividend ratio, starting with value 100 in January 1978.

##### 5.4.2. SADF test of the S&P 500

Before applying the unit root test to the S&P 500 data, I obtain critical values under the null hypothesis by simulating the stochastic process specified in equation (19). The parameters controlling the size of the intercept are set to one i.e.  $d = \eta = 1$ . As discussed earlier, the data-generating process is then close to a pure random walk. The Monte-Carlo simulations for the finite sample critical values run over 480 observations with 1,000 replications. The SADF test statistic is stored for every replication. The basis for the 90%, 95% and 99% right-tailed critical values is the statistics sequence resulting from all replications. As suggested by Phillips et al. (2015a), the lag order is small and fixed, at  $k = 1$ , and the minimum window size for the recursive rolling regressions in the SADF test is 44 observations following their rule. I do not calculate the GSADF test statistic because the increased number of subsamples makes simulations of critical values much more extensive computationally. The alternative hypothesis of explosive behaviour is tested by running an SADF test on the S&P 500 data. Table 1 shows the finite sample critical values and the SADF test results. With a test statistic of 2.428 the SADF test surpasses its critical values at all significance levels. There is strong evidence for the existence of bubble(s) in the indexed price-dividend ratio of the S&P 500. These results are similar to the results reported by aforementioned authors.

<sup>17</sup>See the S&P 500 website, <https://eu.spindices.com/indices/equity/sp-500>.

<sup>18</sup>See "U.S. Stock Markets 1871-Present and CAPE Ratio", [www.econ.yale.edu/~shiller/data.htm](http://www.econ.yale.edu/~shiller/data.htm).

##### 5.4.3. Time stamping of bubbles in the S&P 500

After establishing that one or more bubbles exist in the S&P 500, I turn to the BSADF test (equations (23) to (25)) to determine the exact number of bubbles and the time stamping of each bubble. First critical values for the BSADF tests are obtained by Monte-Carlo simulations with 1,000 replications. The calculation is more complex than for finite sample critical values, since the simulations run on each subsample in the sequence  $Tr_2$ . In the end, there is a critical value for every observation; each based on the number of months passed until that point in the sample. Just as for the SADF, the minimum window size is 44 observations and the order of lags is fixed at  $k = 1$ . I then compare this sequence of simulated critical values to the BSADF test statistic sequence for the price-dividend ratio of the S&P500. In accordance with the crossing-time equations in (24) and (25), the starting point of a bubble is the first chronological observation in the sequence, where the BSADF test statistic passes its corresponding 95% critical value. The assumption is that a bubble has to last at least six months in order to be considered as such. Therefore if the observation six months after the test statistic first passed the critical value remains above the critical value, this time period in the S&P 500 is considered a bubble. Any observation that drops below its critical value afterwards is the end of the bubble. The arbitrary definition of a six month minimum bubble period (or any other chosen duration) seems to be one of the biggest issues for correctly specifying the BSADF test in empirical applications. What if, for example month five and seven after the start of the bubble surpass their respective critical values, but month six does not? The dating algorithm in equations (24) and (25) would not capture this time period as a bubble. Although I did not run into this problem, I attempt to use averaged values over a time span in order to avoid short dips (e.g. month six in the example) in the BSADF test statistic sequence. However, the definition of average values inevitably becomes arbitrary as well: it is not clear how many months should be used for averaging. The conclusion is that economic intuition has to be applied. It is unreasonable to assume that the bubble suddenly stops for one or two months. Ex post it is therefore best to fill in these gaps, if they are just short breaks. If one wants to apply the BSADF test as a real time detector for bubbles, this is a bigger issue since it is not clear whether the drop is the end of the bubble or just a short break. Despite these problems, the time-stamping of bubbles seems to be surprisingly accurate. Figure 5 shows the BSADF test statistic sequence plotted against the 95% critical value sequence for the entire sample. The grey areas are the bubble periods identified by the BSADF test. They coincide with two bubbles corresponding to the tech bubble in the late 20<sup>th</sup> century and the lead-up to the crash of Black Monday in 1987. In fact, the starting point of the bubble is 1995m12, exactly one month later as reported by Phillips et al. (2015a). The ending point of the tech bubble indicated by my tests is 2000m9, which is eleven months earlier. Due to the characteristics of the time stamping algorithm described earlier, there are two short

**Table 1:** The SADF test with critical values for the S&P 500 (Based on Phillips et al., 2015a, p.1066).

The table shows the results of running the SADF test specified in equation (21) on the indexed price-dividend ratio of the S&P 500. The sample stretches over 480 monthly observations, from January 1978 to December 2017. The finite sample critical values are obtained via Monte-Carlo simulation with 1,000 replications.

SADF	Finite sample critical values		
	90%	95%	99%
2.428	1.227	1.568	1.998

breaks between explosive periods that are assumed to be part of the tech bubble. In addition there are two visible spikes of the BSADF tests statistic. The first spike indicates explosive prices leading up to the market crash in October starting in 1987m2. This is the second and much shorter bubble I detect in the S&P 500. The second spike coincides with the financial crisis and passes the critical value sequence from 2008m10 to 2009m3. The detection of a crisis period as explosive is an issue that Phillips and Shi (2018) try to address by implementing a wild bootstrap for the critical values. Phillips et al. (2015a) again find similar periods of explosive prices, but the start and end are different. From their online appendix<sup>19</sup>, it becomes clear that they use a variable lag depending on Bayesian information criteria, while I use a fixed lag. In my own informal tests I find that the specification of the lags in the augmented Dickey-Fuller tests influences results. Therefore the differences could be due to using fixed versus variable lags. Figure 6 shows the real price and price-dividend ratio of the S&P500 for my sample as well as the bubble periods indicated earlier. The BSADF test runs regression of the time-series of the indexed price-dividend ratio (the dotted line in Figure 6). As discussed when presenting the rational bubble model (equation (16)), this is necessary for the correct specification of bubbles. The detected bubble periods coincide well with price increases in the real price index of the S&P500, but only when dividends are not keeping up with this increase. For example, the BSADF test does not detect a bubble in the period from 2016 onwards, although the price of the S&P500 reaches record highs. This is because the price dividend ratio remains more or less constant. Of course the BSADF test is always limited to the specific timer series analysed. The detection of bubbles, say in housing prices, requires a separate implementation of the test. In summary, the bubble periods I detect coincide with the tech bubble and the lead up to Black Monday. They run from December 1995 to September 2000 and from February 1987 to August 1987, respectively. The strategy for the empirical analysis is to split the sample into two different sub samples. The first sample contains all months that were not stamped as a bubble and the second sample contains all months that were stamped as a bubble. This allows me to analyse the changes in the relationships between returns, the prospect theory value and other important control variables. It is therefore not only a

test of my hypothesis that stocks with high prospect theory value earn lower returns during bubbles, but also gives an indication how the predictive power of other variables varies with market conditions.

## 6. Prospect theory and stock returns during bubbles

### 6.1. Empirical approach

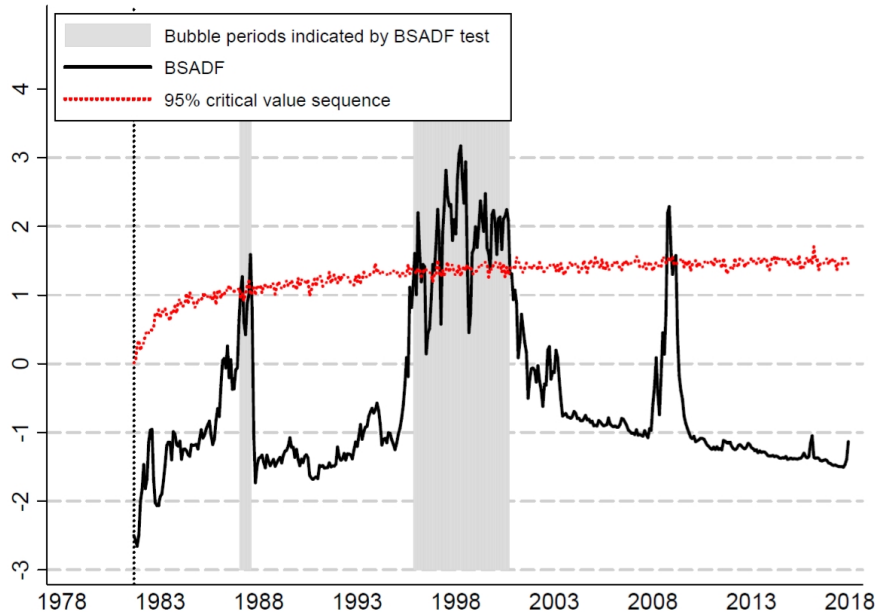
Barberis et al. (2016) show that stock with high (low) prospect theory value earn low (high) subsequent returns. Earlier, I made the prediction that this relationship should be more pronounced during bubble periods compared to non-bubble periods. In order to these two predictions in my sample, I rely on several tests. The first tests are portfolio sorts with a low high prospect theory value strategy that should earn abnormal returns if the predictions hold. These include checks for robustness, where the parameters for constructing the prospect theory value or the underlying sample are modified. In order to test the main predictions in the cross-section of returns, I use the Fama-Macbeth approach. This lets me test while controlling for known return predictors. The plausibility of the rising limits to arbitrage argument is evaluated using the same method, but the regressions include an interaction term of prospect theory value with different approximations for the limits to arbitrage, as presented in section 4.1.3. In the last section I look at the underlying components of prospect theory likely to cause the observed relationship with returns. The empirical analysis closely follows Barberis et al. (2016).

### 6.2. Data

#### 6.2.1. Databases

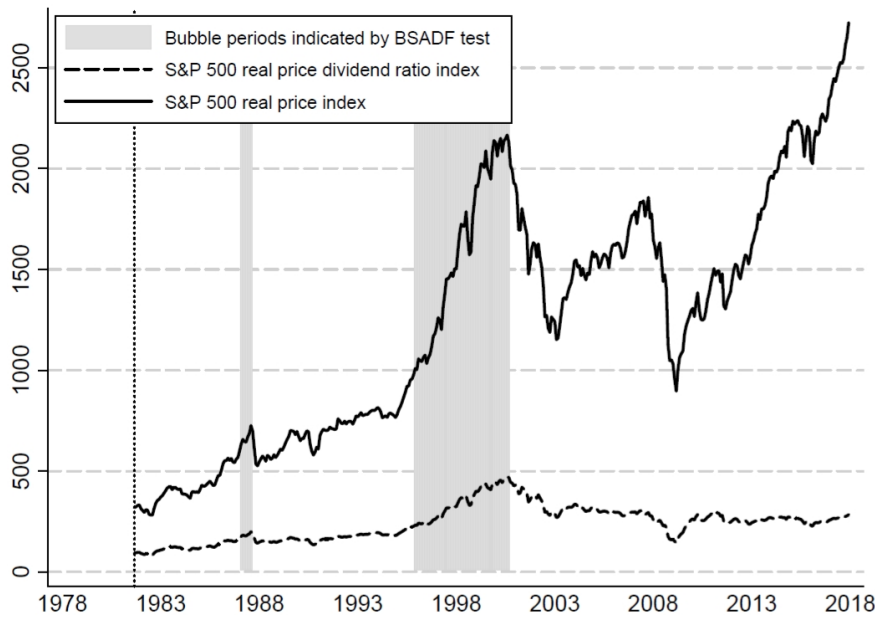
The data for the empirical analysis is provided by Wharton Research Data Services. The CRSP/WRDS Merge database is the source for monthly security information of all companies listed on US stock exchanges, namely security identifiers, exchange codes and prices. The fundamental information on these companies is obtained from the Compustat database. Besides balance sheet items it also contains information related to security markets like common shares outstanding, all sampled with yearly frequency. The Fama and French database is the basis for the monthly market returns and risk-free rate. The sample for these three datasets starts in January 1978 and ends in December 2017, corresponding to the time period used for the time-stamping of bubbles.

<sup>19</sup>See <https://sites.google.com/site/shupingshi/GSADF.zip?attredirects=0>.



**Figure 5:** The BSADF test for bubbles in the S&P 500 (Based on Phillips et al., 2015a, p.1066).

The graph shows the results of the BSADF test procedure. The bold line represents the test statistic for the indexed price-dividend ratio of the S&P 500, the red-dotted line are the 95% critical values obtained via Monte-Carlo simulations with 1,000 replications for each point in time. The sample of the S&P 500 starts in January 1978 and ends in December 2017. The dotted black line represents the end of the minimum window for estimation.



**Figure 6:** Bubbles in the S&P 500 (Based on Phillips et al., 2015a, p.1066).

The graph shows the results of the BSADF test procedure. The bold line is the price index of the S&P 500, adjusted for inflation. The dashed line is the index of the price-dividend ratio for the S&P 500, again with prices adjusted for inflation. The sample of the S&P 500 starts in January 1978 and ends in December 2017. The dotted black line represents the end of the minimum window for estimation.

In addition, I need daily data on prices and trading volume, which comes from CRSP/WRDS Merged daily. The daily data sample starts in January 1983 and ends in December 2017, as available. After merging the four datasets, I exclude stocks from smaller exchanges because of missing price in-

formation. The stocks in the final dataset are listed on NYSE, NYSE Amex, NASDAQ-NMS, NASDAQ OMX BX or NYSE Arca. Finally, stocks are required to have a history long enough as to make the calculation of the prospect theory value possible i.e. a history of at least 60 months.

### 6.2.2. Data Summary

Table 2 summarizes the variable for prospect theory value, dubbed TK, and the control variables I use in the empirical analysis. In addition, Table 3 shows the correlations of all variable. TK is the prospect theory value of a stock's return distribution, computed following equation (7). In the benchmark case, the return distribution consists of the past 60 monthly returns. Size is the log of the market value of equity in month  $t-1$ , obtained by multiplying the common shares outstanding at the end of year  $t-1$  with the end of month price. Beta is a stock's beta as in Fama and French (1992). Each June of a given year stocks are sorted into 100 portfolios based on size and pre beta, a stock's market beta over the last six years. Then the return of each portfolio in the following year is used in a regression on the market return in the current and prior month. Beta is the sum of the coefficients from this regression. Bm is the log book-to-market ratio computed as the difference of the logs of book value of equity and market value of equity. The book value of equity is the shareholder's equity minus the preferred stock value. Following Daniel and Titman (2006), shareholder's equity is approximated by stockholder's equity. If the stockholder's equity is missing, I replace it with common equity plus preferred stock par value or assets minus liabilities plus book minority interest, as available. The preferred stock value is redemption, liquidation or carrying value in that order. After subtracting the stockholder's equity with the preferred stock value, book deferred taxes are added to get the final book value of equity. Mom is the control for the momentum effect reported by Jegadeesh and Titman (1993), the cumulative return from month  $t-12$  to  $t-2$ . Illiq is Amihud (2002) measure for illiquidity, the average ratio of absolute return to dollar trading volume in month  $t-1$ . Further controls include rev, for the short term reversal, an ltrev, for the long term reversal. Rev is a stock's return in  $t-1$ , while ltrev is the cumulative return from  $t-60$  to  $t-13$ . Max is the maximum daily return in month  $t-1$ , while Min is the negative of the lowest return in month  $t-1$ . Skew is the sample skewness of returns over the past five years. Because probability weighting is an important factor for the prospect theory value, skewness should be included in the analysis (see Barberis and Huang (2008)). The computation for the variables starts in January 1983 to match the minimum window required by the time-stamping of the bubbles and stock requirements mentioned earlier. On average, the prospect theory value of a stock's return distribution is negative, which is in accordance with the result in Barberis et al. (2016). Menkhoff and Schmeling (2006) also find that the prospect theory value of simulated return distributions tends to be negative. Table 3 presents the correlations between TK and the control variables employed. Beta is negatively correlated to the prospect theory value of a stock. Probability weighting could be an intuitive explanation for the positive relation between the prospect theory value of a stock and the sample skewness. A skewed return distribution should be preferred, since tails are assigned higher weights in prospect theory. TK is also

positively correlated with the controls for past return anomalies, namely momentum, short term reversal and long term reversal. The correlation with ltrev is the strongest out of all control variables. It is also the variable that is constructed with the return information closest to TK. The reason behind the positive correlation with mom and rev should be similar, although the correlation coefficient on rev seems particularly large considering it represents only one month of shared return information. Finally, TK is positively related to size, the log market value of equity, and negatively related to the log book-to-market ratio Bm. Looking at correlations is only a first rough approximation of the relationship between the variables, therefore I now turn to the first step laid out in the empirical approach, portfolio analysis of stocks sorted on the prospect theory values.

### 6.3. Portfolio sorts on prospect theory value

Sorting stocks into portfolios on prospect theory values make it possible to exploit the prediction that high prospect theory value stocks should earn low returns. The portfolio with the highest TK values should earn the lowest returns and the portfolio with the lowest TK values should earn the highest returns. If the prediction is correct, going short and long in these portfolios, respectively, should generate positive excess returns. This approach of building long-short portfolios is popular in the finance literature, especially as a first assessment of the power of anomalies. For example, momentum (Jegadeesh and Titman (1993)) is exploited by forming such a portfolio, only with past winners and losers. It is also a well-known strategy for asset pricing tests (e.g. Fama and French (1996)), although Berk (2000) points out that sorting into many groups could weaken otherwise strong explanatory power of pricing models. I begin building the TK portfolios by sorting stocks into deciles based on their prospect theory values starting in January 1983 and ending in December 2017. For each of the ten portfolios formed on TK, the average equal weighted return over the next month is computed. The returns in the benchmark case are in excess of the market. The long-short portfolio takes a long position in the low TK portfolio (decile one) and a short position in the high TK portfolio (decile ten). This portfolio should generate positive excess returns, if the prediction is accurate. Table 4 shows the results of the portfolio sorts on prospect theory value. Decile one offers the highest average returns while decile ten exhibit's the lowest returns. Consequently, the low high TK portfolio earns positive returns. The excess returns to this strategy are highly significant. Overall, all decile portfolios are significant at least at the 10% level. The significance of the low TK and high TK decile is most important, since the return on the portfolio rely on both. While I use equal weighted returns, Barberis et al. (2016) also look at value weighted returns. They present similar results and show that the low-high TK portfolio even generates excess returns with value-weighting. Further, they show that the predictive power of prospect theory values for returns lasts for several months, although the level of excess returns declines very fast after the first month. My second prediction

**Table 2:** Summary statistics.

The table shows mean, standard deviation and median for each of the variables used in the empirical analysis over the period starting in January 1983 and ending December 2017. TK is the prospect theory value of a stock's return calculated using a distribution of the past 60 returns. Size is the logarithm of the market value of equity. Beta is a stock's beta based on 100 portfolio sorted on size and pre-beta as in Fama and French (1992). Bm is the log book-to-market ratio, calculated as the difference of the logarithm of the book value of equity and the logarithm of the market value of equity. Mom is the momentum of a stock; the cumulative return from month t-12 to t-2. Illiq is the illiquidity measure of Amihud (2002), scaled by  $10^5$ . Rev is the return in month t-1. Ltrev is the cumulative return of a stock from month t-60 to t-13. Max is the highest daily return in month t-1, while Min is the negative of the lowest daily return. Skew is the sample skewness of returns over the past five years.

Variables	Mean	SD	Median
TK	-0.078	0.032	-0.075
Size	5.902	2.115	5.759
Beta	1.058	0.297	1.046
Bm	-0.575	0.821	-0.500
Mom	0.099	0.394	0.081
Illiq	0.122	0.610	0.002
Rev	0.010	0.123	0.004
Ltrev	0.412	0.753	0.344
Max	0.065	0.593	0.042
Min	0.052	0.052	0.038
Skew	0.0534	0.893	0.146

**Table 3:** Correlation matrix.

The table shows the correlation between TK and the control variables presented in Table 2. The sample starts in January 1983 and ends in December 2017.

	TK	Beta	Size	Bm	Mom	Illiq	Rev	Ltrev	Skew	Max	Min
TK	1.00										
Beta	-0.22	1.00									
Size	0.35	0.01	1.00								
Bm	-0.15	-0.07	-0.32	1.00							
Mom	0.21	0.05	0.08	-0.09	1.00						
Illiq	-0.26	-0.01	-0.33	0.16	-0.10	1.00					
Rev	0.06	0.02	0.03	0.03	-0.02	-0.01	1.00				
Ltrev	0.44	0.16	0.15	-0.34	-0.08	-0.14	-0.02	1.00			
Skew	0.21	0.19	-0.27	0.10	0.24	0.09	0.10	0.22	1.00		
Max	-0.06	0.03	-0.03	0.01	-0.02	0.07	0.07	-0.02	0.02	1.00	
Min	-0.36	0.17	-0.25	0.02	-0.05	0.31	-0.31	0.00	0.05	0.08	1.00

was that high (low) prospect theory value stocks should earn lower (higher) returns during bubble periods compared to non-bubble periods. Therefore I do the same portfolio sorting exercise comparing the date-stamped bubble periods to the non-bubble periods. Table 5 presents the results of this test.

The results for non-bubble periods indicate no large changes of the returns to the low high TK strategy. The first decile still has the highest average excess return and the tenth the lowest average excess return. Returns in the first decile are higher, while the magnitude does not change in the tenth decile. Overall, the low-high TK portfolio earns less compared to the test using the full sample. During bubble periods the returns on the same portfolio are on average more than a percent higher. This is in accordance with my main hypothesis. However, most of the decile portfolio are

not significantly different from zero anymore. The increase in returns mainly comes from decile one: while returns in decile ten are almost on the same level, returns in decile one increase compared to non-bubble periods. This could indicate that some investors during bubble periods only invest in stock with high prospect theory value, causing low prospect theory value stocks to become relatively undervalued and earn higher returns. Generally the returns on the lower half of TK decile portfolios are higher during bubble periods. In order to confirm the results from the portfolio sorts, I first conduct robustness checks employing different specifications related to the construction of the prospect theory value. Table 6 presents the returns on the low-high TK portfolio for these tests, for both bubble and non-bubble periods. First, the construction window of the return distributions used for the computation of prospect theory values changes from

**Table 4:** Portfolio sorts on TK.

The table reports the average monthly excess return of portfolios formed on the prospect theory value. Stocks are assigned to deciles based on their TK value each month, ranging from low TK (decile one) to high TK (decile ten). The sample starts in January 1983 and ends in December 2017. The numbers in parentheses show the t-statistic.

Decile	Excess return	
1	0.0184	(5.45)
2	0.0137	(4.93)
3	0.0121	(4.88)
4	0.0097	(4.31)
5	0.0097	(4.50)
6	0.0079	(3.86)
7	0.0074	(3.77)
8	0.0062	(3.48)
9	0.0058	(3.51)
10	0.0036	(2.04)
Low-high TK	0.0148	(5.97)
N	420	

**Table 5:** Portfolio sorts during bubbles.

The table reports the average monthly excess return of portfolios formed on the prospect theory value. Stocks are assigned to deciles based on their TK value each month, ranging from low TK (decile one) to high TK (decile ten). Bubble periods were date-stamped following PSY (2015). The sample starts in January 1983 and ends in December 2017.

Decile	Excess return			
	Non-Bubble		Bubble	
1	0.0168	(4.50)	0.0268	(3.53)
2	0.0129	(4.18)	0.0181	(2.89)
3	0.0114	(4.13)	0.0155	(2.97)
4	0.095	(3.79)	0.0110	(2.15)
5	0.0094	(3.93)	0.0117	(2.29)
6	0.0077	(3.40)	0.0089	(1.91)
7	0.0074	(3.36)	0.0076	(1.78)
8	0.0065	(3.24)	0.0044	(1.29)
9	0.0062	(3.29)	0.0041	(1.22)
10	0.0039	(2.02)	0.0021	(0.46)
Low-high TK	0.0130	(4.73)	0.0247	(4.43)
N	355		65	

the benchmark of 60 months to 48 and 36 months. For bubble periods I expect lower returns to the low-high TK the longer the construction window, because longer equal weighted return distributions cannot capture the explosive nature of prices. During non-bubble periods the returns to the low-high TK portfolio are slightly higher compared to the construction window of 60 months and are still highly significant. Like for the benchmark case, the returns on the portfolio are higher during bubbles. Taking the results from Table 5 into consideration, the relationship between longer construction windows and lower returns on the low-high TK portfolio during bubbles seems to hold as predicted. The

second technical robustness check for the computation of the prospect theory values deals with the returns used. Instead of excess returns over the market, I now use returns in excess over the risk-free rate and raw returns. Using returns in excess of the risk-free rate leads to a similar decrease in returns on the low-high TK portfolio as using raw returns. The impact of different specifications of returns is generally small. Dividing the sample into two sub-periods generates more interesting results. Since the bubble periods are already a sub periods, it is not reasonable to analyse them together with non-bubble periods for this test. The returns on the portfolio are lower for the period before the new millennium. In both

**Table 6:** Robustness tests (Based on Barberis et al., 2016, p.3088).

The table reports the average returns of a low-high prospect theory value portfolio, formed each month by going long (short) in stocks in the lowest (highest) TK decile. In the panel construction window the TK values are computed with different window sizes for the return distributions. The panel returns tests the robustness for raw returns and returns in excess of the risk-free rate. Time period splits the sample into two sub periods. In the last panel, I exclude stocks that fall under 5 Dollars in the month before the computation of TK values.

		Low-high TK			
		Non-Bubble		Bubble	
Construction window	48 months	0.0117	(4.26)	0.0250	(4.44)
	36 months	0.0112	(4.07)	0.0255	(4.57)
Returns	Raw returns	0.0131	(4.11)	0.0258	(4.12)
	Returns in excess of the risk-free rate	0.0133	(4.20)	0.0260	(4.13)
Time periods	1983-2000	0.0103	(2.92)		
	2001-2017	0.0150	(3.74)		
Exclude low price stock	Price $\geq$ 5\$	0.0109	(6.43)	0.0153	(3.17)

periods, the returns are lower than the average in Table 4 because I excluded bubble periods for this robustness test. The next robustness test shows that low price stocks could be part of the explanation for the return on the portfolio. The magnitude of returns drops after excluding stocks that fall below a price of five dollars in the months before computation of the prospect theory value. Low priced stocks tend to be more illiquid and less subject to the forces to arbitrage. They are also more risky and could have a more skewed distribution. Both of these characteristics attract investors with prospect theory preferences; therefore low price stocks should have higher prospect theory values. Individual investors in general seem to prefer lottery like stocks. Kumar (2009) analyses the propensity to gamble that different types of investors exhibit. He also finds that investors consider a stock a lottery especially if it has a low price. In summary, the exclusion of low price stocks should have a negative effect on the returns of the low-high TK portfolio. Since the returns on the portfolio are still high and significant after excluding low price stocks, prospect theory preferences seem to influence prices in general and not only for a small part of the stock market. All robustness tests confirm the initial observation that there are significant abnormal returns to a low-high TK portfolio during non-bubble periods. In addition, the returns to the portfolio are larger during bubbles, but suffer from lower significance. How do the returns to the portfolio develop during the tech bubble? Figure 7 illustrates the returns on the low-high TK portfolio for every month date-stamped as a bubble, from December 1995 up to September 2000. It is noticeable that the returns to the portfolio are extreme for most of the tech bubble.

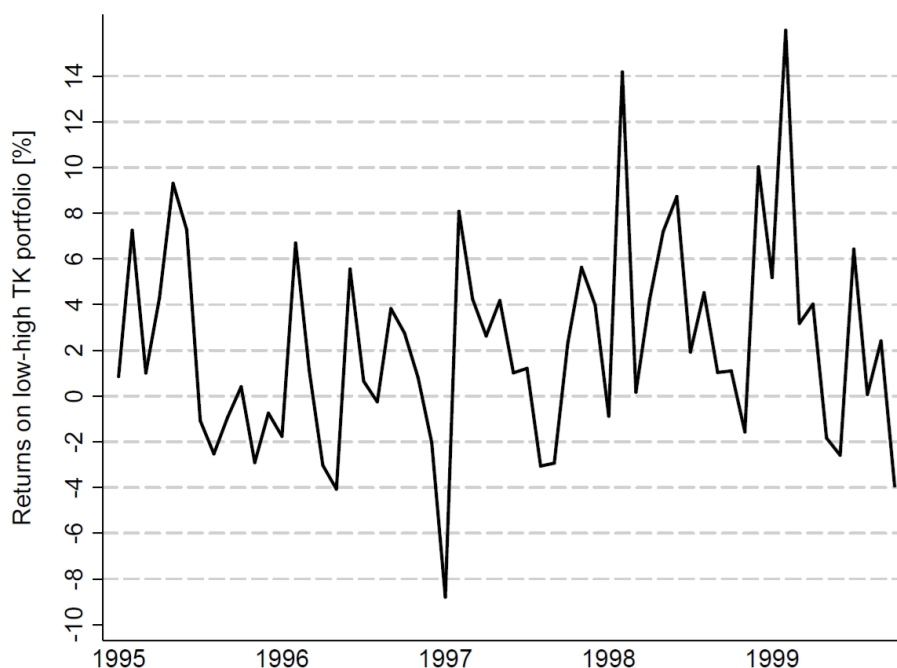
There is one month in 1997 when they drop below minus eight percent, while they are especially positive in months from 1998 onwards. This fits earlier results and the magnitude of returns seems to coincide with the progress of the bubble itself, which hits its highest levels in the course of

1998<sup>20</sup>. In the analysis for the portfolio sorts during bubbles in Table 5, I concluded that the returns on the low-high TK portfolio seem to be driven mainly by the positive returns of the lower TK deciles. These returns also seem to be highly volatile. The high average returns seem to come from fewer extreme values rather than a constant trend. Also the low significance in the high TK deciles could be a problem for the implementation of the strategy during bubbles. Portfolio sorts are not appropriate to test a relationship between prospect theory value and returns. This leads up to the question whether the prospect theory value of a stock in fact has predictive power for returns in the cross section.

#### 6.4. Prospect theory and stock returns in the cross section

In this section, I test my two predictions using the Fama-MacBeth approach. They develop a method for testing the relationship between risk and return using three steps. In the first step, they run a time-series regression to obtain beta coefficient estimates. They then use these estimates to a run cross-sectional regression each month. The coefficients and standard errors are then obtained as the time-series average. This approach is frequently applied in asset pricing research. For my analysis only the last two steps are necessary, since I am not trying to measure the relationship between return and risk and therefore do not need market beta-regression coefficients. The Fama-MacBeth approach is very useful, because it allows capturing the cross-sectional variation of factors. I can therefore test my predictions while controlling for variables known to explain returns. Starting in January 1983 and ending in December 2017, I run a regression of the prospect theory value TK and my control variables mentioned in Table 2 on the percentage return in that month. The coefficients

<sup>20</sup>See Figure 5, where the BSADF test statistic reaches its highest level in 1998.



**Figure 7:** The low-high TK portfolio during bubbles.

The figure shows the returns of a low-high prospect theory value portfolio, formed each month by going long (short) in stocks in the lowest (highest) TK decile. The sample runs from 1995m12 up to 2000m9, which falls into bubble periods detected applying a method by Phillips et al. (2015b). The returns are monthly percentage returns.

and standard errors reported are the time-series average of these monthly regressions. Panel A of Table 7 reports the results of this exercise for the months not stamped as a bubble using the method by Phillips et al. (2015a), while Panel B reports the results for the months stamped as a bubble period using the same method. In each panel, I run the regressions with five different specifications. Model (1) is the first test of my predictions using only TK as an independent variable. Model (2) brings in controls for market beta, size, book-to-market ratio, momentum effect and illiquidity. Model (3) and Model (4) include the control for short term reversal, rev, and for long term reversal, ltrev, together with the maximum and minimum daily return variable, respectively. I do this to highlight the effect especially rev has on the predictive power of TK. Model (5) includes the control for skewness of past returns. The results in Panel A support the original prediction by Barberis et al. (2016). The coefficient on TK is negative, indicating that an increase in TK measured at the beginning of the month leads to a decrease in returns. This effect is not only sizeable, but also economically significant. After including the controls for size, beta, the book-to-market ratio and the momentum effect the coefficient on TK remains significant and loses only little in magnitude. I also find that including the control for the short term reversal effect, Rev, has a strong impact on the magnitude and significance of TK. Comparing the coefficients of TK in Model (2) and Model (3) shows that the introduction of rev leads to a decrease in magnitude with the t-statistic shrinking. Why should the predictive power of TK for returns be affected most by the

return in month  $t-1$ ? Barberis et al. (2016) argue that investors get the most recent monthly return from a different source than returns that are further in the past. If this is correct, investors might attribute the previous month's return more weight than the other returns in the distribution. In that case, a high return in the previous month would overshadow possible low returns in the past and investors would be incentivised to buy the stock, causing it to become overvalued and earn low returns. This would also be an alternative explanation of the short term reversal effect. The most negative daily return in month  $t-1$  has a similar and strong negative effect on subsequent returns. The strong effect of the control variable rev could lead to the conclusion that TK's predictive power is due to reversal, but Barberis et al. (2016) find that leaving out month  $t-1$  in the construction of TK has no large impact. Skewness should influence the results to some degree, because probability weighting favours extreme outcomes with low probability. Therefore a skewed distribution with a heavy weigh on the right tail should be preferred by investors with prospect theory preferences. As expected the inclusion of the skewness control in model (5) leads to a slight decrease of the significance on TK. I now turn to the results for the Fama-MacBeth regressions during bubbles in panel B. This is the test of my main prediction that high (low) prospect theory values stocks earn lower (higher) returns in bubble periods compared to non-bubble periods. The results suggest that, in the cross-section, the prospect theory value does not have more predictive power for returns during bubbles. The significance of TK is very similar to non-bubble



**Table 7:** Fama-MacBeth regressions (Based on Barberis et al., 2016, p.3090).

The Table shows the results of the Fama-MacBeth approach. In each cross-sectional regression, percentage return is the dependent variable. TK is a stock's prospect theory value, measured at the beginning of the month using a distribution consisting of 60 monthly returns (see equation (7)). Beta is a stock's market beta, computed with the returns of 100 portfolios formed on size and pre-beta in June of each year, following Fama and French (1992). Size is the log equity market value at month t-1; the product of common shares outstanding in December of year t-1 and monthly stock price. Bm is the log book-to-market ratio calculated as the difference of the log book value and the log equity market value in December of year t-1. Book value is computed following Daniel and Titman (2006). Mom is a stock's cumulative return from month t-12 until t-1. Illiq is the illiquidity measure of Amihud (2002), scaled by 105. Rev is a stock's return in month t-1. Ltrev is a stock's cumulative return from month t-60 until t-13. Max is the highest daily return in month t-1, while Min is the negative of the lowest daily return. Skew is the sample skewness of monthly returns over the past five years. TK, Rev, Mom, Max and Min are scaled by factor 100. The entire sample starts in January 1983 and ends in December 2017. Panel A shows the results of the Fama-MacBeth regressions for the months (N=355) not stamped as a bubble using a method by Phillips et al. (2015a), while Panel B shows the results for months (N=65) stamped as a bubble using the same method. The numbers in brackets show the t-statistics.

Panel	A. Non-Bubble					B. Bubble				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
TK	-0.121*** (-4.68)	-0.082*** (-4.49)	-0.057** (-3.31)	-0.080*** (-3.68)	-0.092*** (-3.45)	-0.251*** (-4.67)	-0.226*** (-5.51)	-0.192*** (-4.77)	-0.190** (-3.36)	-0.243** (-2.94)
Beta		0.247 (0.95)	0.306 (1.20)	0.516* (2.15)	0.411 (1.58)		0.675 (1.24)	0.763 (1.46)	0.758 (1.90)	0.573 (1.45)
Size		-0.115*** (-4.44)	-0.113*** (-4.39)	-0.134*** (-4.81)	-0.116*** (-4.07)		-0.126 (-1.25)	-0.113 (-1.15)	-0.128 (-1.37)	-0.073 (-0.76)
Bm		0.189*** (3.83)	0.204*** (4.17)	0.170*** (3.99)	0.160*** (3.60)		0.194 (1.17)	0.208 (1.28)	0.170 (1.44)	0.169 (1.50)
Mom		0.002 (1.42)	0.001 (0.64)	0.002 (1.04)	-0.001 (0.57)		0.007 (1.82)	0.006 (1.58)	0.007 (1.68)	0.005 (1.05)
Illiq		0.226*** (2.85)	0.232** (2.95)	0.339*** (4.09)	0.344*** (3.33)		0.217 (1.34)	0.254 (1.57)	0.279 (1.57)	0.264 (1.32)
Rev			-0.049*** (-13.90)	-0.056*** (-14.17)	-0.058*** (-12.62)			-0.054*** (-5.84)	-0.066*** (-6.60)	-0.073*** (-6.38)
Ltrev				-0.075 (-1.01)	-0.071 (-0.84)				-0.049 (-0.23)	0.128 (0.49)
Max				0.007 (1.94)	0.015** (2.90)				0.017 (1.48)	0.016 (1.12)
Min				-0.052*** (-8.30)	-0.053*** (-6.76)				-0.045** (-3.79)	-0.049** (-3.74)
Skew					0.117* (2.56)					0.187 (1.56)
R-squared	0.015	0.044	0.050	0.057	0.062	0.012	0.049	0.056	0.065	0.073

t-statistics in parentheses \*p<0.05, \*\*p<0.01, \*\*\*p<0.001

periods when not controlling for other return predictors in model (1). In some models it is actually greater than during non-bubble periods, but for the most sophisticated specification TK exhibits slightly weaker predictive power for returns. This leads to the conclusion that the predictive power of TK is roughly equal during bubble and non-bubble periods. There is no strong evidence that my prediction holds on average. Another interesting finding is that the coefficients of TK are greater in magnitude during bubbles. But comparing the size of coefficients between the two different sub samples is not a statistical test, even if the independent variables are standardized. In general it is important to note that a precise interpretation of the predictive power of variables in a Fama-MacBeth regression is only valid under the assump-

tion of linear relationship. In my case the variables are scaled differently so the interpretation has to be restricted to the t-statistics. Nonetheless I can compare the effects of the same variables between the different panels. Bubble periods are extreme market periods characterised by overvaluation, so the predictive power for some variables should be expected to change. While the book-to-market ratio, for example, has significant positive predictive power for returns during non-bubble periods in panel A, it has no power in panel B. The book-to-market ratio is a measurement of the valuation of a firm's assets. A higher book-to-market ratio signals undervaluation. During non-bubble periods an undervaluation should lead to a future increase in prices and higher returns, which is suggested by panel A. The book-to-market ratio could lose

its predictive power for returns during bubbles, because on average it could be more difficult for investors to determine the fundamental value of stocks. Just like in panel A, the short term reversal still seems to play a role during bubbles. Investors continue to buy stocks that performed well in the last month, causing them to be overvalued and earn low returns. Although the control for the momentum effect remains to have no significant influence on returns, it seems to be somewhat more robust. These results make intuitive sense for the market conditions during bubbles, where investors are searching for the best stocks to profit from, but it also describes the buying behaviour during non-bubble periods. It is important to notice that the predictive power of all control variables is weaker during bubble periods. Given this fact, the predictive power of TK seems rather high. Maybe the mispricing during bubbles makes it harder for any variables to predict returns consistently. For TK it seems that both standard errors and coefficients are larger on average. Taking the results from the portfolio sort and Figure 7 into account, the relationship between TK and returns might be stronger during bubbles, but only for a smaller set of stocks or particular time and not on average. This is entirely speculative though. In summary, I do not find support for the prediction that high prospect theory value stocks earn lower returns during bubbles, but can confirm the original prediction of its general effect. In the next section, I try to answer part of the question which stocks are favoured by prospect theory, namely if the limits to arbitrage have an impact on TK's relationship with returns. This could also shed some light on whether the limits to arbitrage of a stock have more influence on prices during bubbles.

#### 6.5. Prospect theory value, limits to arbitrage and individual investors

Generally, stocks with high limits to arbitrage tend to be owned by less sophisticated investors. Since it is more costly for arbitrageurs to employ strategies that correct the mispricing, the prices of these kinds of stocks should be affected by the behaviour of individual investors to a greater extent. The prospect theory value is assumed to be a reflection of individual investor's behaviour. Therefore the relationship between high TK values and low returns should be stronger for stocks exhibiting higher limits to arbitrage. The limits to arbitrage are hard to measure, but in section 4.1.3, I presented several approximations. To test whether the effect of TK depends on the proxies, a Fama-MacBeth regression with interaction terms is employed. Each regression includes TK, the previous control variables and adds an interaction term of TK with the proxy for limits to arbitrage. The first proxy for the limits to arbitrage is Mashruwala et al. (2006) arbrisk variable. It captures the idiosyncratic volatility of a stock's past returns. Arbitrage often requires being able to hedge against contrary outcomes. The basic idea for the arbrisk variable is that an arbitrageur can only find substitutes for hedging of the risk explained by the standard market model. She is then exposed to the risk that she cannot hedge, the excess

volatility from her model. The variance of residuals of a standard market model regression, or of models with more factors, therefore is a proxy for arbitrage risk. Arbrisk therefore is the residual variance from a standard market model regression of a stock's return on the market return from month  $t-36$  to month  $t-1$ . Although calculating idiosyncratic volatility with daily returns would probably be more accurate, I do not have the resources to execute the required number of regressions in a large dataset. The second proxy for limits to arbitrage is the dollar trading volume, an indicator for transaction costs. Stoll (2000) finds that the ask-bid price spread is negatively related to measures of trading activity, such as the average daily dollar trading volume. Theoretically there should be fewer frictions when implementing arbitrage strategies for stock with high trading volume. Following Mashruwala et al. (2006), the measure of transaction costs  $tvol$  is the average daily dollar trading over the past year. Size is also an approximation of transaction costs and the limits to arbitrage, since stocks with high market capitalization should be easier to loan or short. The last proxy for the limits to arbitrage is the illiquidity measure of Amihud (2002), which is the ratio of absolute daily stock return to trading volume. It measures the marginal impact of one dollar trading volume on prices. More illiquid stock should be more difficult to incorporate into arbitrage. Table 8 presents the results of the regression using the interaction terms of TK with the proxies for limits to arbitrage. For non-bubble periods, the results confirm the prediction that stocks which are more likely to have higher limits to arbitrage also exhibit a stronger predictive power of TK for returns. The interaction terms all show the expected signs on coefficients. The interaction terms with the measures for arbitrage risk and illiquidity both have significant negative effect on returns, because they measure frictions for arbitrage strategies. The terms with the measures for dollar trading volume and size have a significant positive effect, because they make it easier for arbitrageurs to trade against mispricing introduced by individual investors. During bubbles periods there seems to be no reinforcing effect of the limits to arbitrage to the effect of TK. The interaction terms with arbitrage risk, trading volume, size and illiquidity are not significant anymore. These don't seem to be limits to arbitrage influencing the relationship between prospect theory and returns. Are the limits to arbitrage even rising during bubbles? The arbitrage risk is a highly significant during non-bubble periods, but does this limit to arbitrage rise during bubble periods? I test for difference between means for bubble and non-bubble periods. Indeed I find that the arbitrage risk is significantly higher during bubble periods. This suggests that although the limits to arbitrage are rising, they might not affect the relationship between prospect theory and returns. Of course this is just a narrow view of a much broader problem in research. There are many other limits to arbitrage that would need to be considered. Instead of directly measurable limits to arbitrage, there are also models explaining mispricing by investor behaviour. In the model of Abreu and Brunnermeier (2003) arbitrage is delayed, because arbitrageurs try to profit from

**Table 8:** Prospect theory value and proxies for limits to arbitrage (Based on Barberis et al., 2016, p. 3096).

The table reports the result of Fama-MacBeth regressions with the independent variable percentage return. The entire table with all control variables can be found in Appendix E. Arbrisk is the residual variance from a standard market model regression on market returns from month t-36 to t-1. Tvol is the average daily dollar trading value over year t-1. Size is the log equity market value at month t-1; the product of common shares outstanding in December of year t-1 and monthly stock price. Illiq is the illiquidity measure of Amihud (2002), scaled by 105. TK is scaled by 100. Panel A shows the results of the Fama-MacBeth regressions for the months (N=355) not stamped as a bubble using a method by Phillips et al. (2015a), while Panel B shows the results for months (N=65) stamped as a bubble using the same method. The numbers in brackets show the t-statistics.

Panel	A. Non-Bubble				B. Bubble			
	(1) Arbrisk	(2) Tvol	(3) Size	(4) Illiq	(1) Arbrisk	(2) Tvol	(3) Size	(4) Illiq
TK	0.140*** (8.17)	-0.310*** (-5.41)	-0.213*** (-5.77)	-0.068** (-3.16)	0.084 (1.87)	-0.169 (-1.20)	-0.195 (-1.86)	-0.184** (-3.26)
TKarbrisk	-0.022** (-3.07)				-0.032 (-1.52)			
TKtvol	0.017*** (4.34)				-0.001 (-0.11)			
TKsize			0.023*** (5.06)				0.001 (0.08)	
TKilliq			-0.115*** (-4.85)				-0.047 (-1.58)	

t-statistics in parentheses \*p<0.05, \*\*p<0.01, \*\*\*p<0.001

the bubble. Once the majority of arbitrageurs enter in a selling effort, the initial mispricing is corrected as the bubble collapses. In other words, arbitrageurs could knowingly ignore the mispricing and would not necessarily care about the measurable limits to arbitrage during bubbles. While delayed arbitrage is a possible cause for mispricing and bubbles, it is hard to test empirically. Only investor level data shines some light on the issue. For example, Griffin et al. (2011) find that institutional investors i.e. arbitrageurs were more heavily invested in technology stocks during the tech bubble than individual investors. It is also a coordinated selling effort of institutions in mid-March of 2000 that drives the collapse of the bubble and not individual investors. Individual investors should have more influence on prices during bubbles, because arbitrageurs have more interest in riding it than correcting the mispricing. Other proposition to why the tech bubble built up is the increase number of individual investors. Brennan, for example, argues for an increase in individual investors during the 1990s due to changes in pension plan policy, from benefit pension plans to contribution plans. Benefit pension plans were managed by the employer and protected against failure, while contribution plans put the responsibility for the investment in the hands of the employee. Contribution plans also began to lean more towards equity investments. In addition, households in the 1990s followed the same trend, increasing the share of equities in their

portfolio from 13.2% in 1990 to 32.5% in 1999<sup>21</sup>. Although the discussion on limits to arbitrage and individual investors during bubbles shows that overall individual investor's preferences should be able to manifest themselves to a greater degree, this seems to have no impact on the prospect theory variable. How does the mechanism of the prospect theory value work at a more fundamental level? Maybe the mechanism offers some insights on the issue. In the next section, I therefore analyse the effect of the components of prospect theory on returns individually.

#### 6.6. On the relevance of prospect theory components

In order to better understand what drives the results, I test the importance of the three components of prospect theory reflected in the equations of the value function and the probability weighting function. Each of the components reflects certain behavioural patterns individuals exhibited in experiments. By analysing the components it could be possible to infer which of these patterns is especially important for the predictive power of the prospect theory value. Since I focus on the comparison of bubble and non-bubble periods, the following analysis can also show whether the relative importance of the components undergoes significant changes through time. Hence the analysis could give insights into the evolution of behavioural patterns of investors, de-

<sup>21</sup>See Brennan, 2004, p.6.

pending on the situation they are facing in the stock markets. The empirical test is based on the Fama-MacBeth approach used in Table 7. Specifically it is Model 5 of Table 7, which includes all control variables. But instead of using all the parameters provided by Tversky and Kahnemann (1992) to construct the variable TK, parameters are chosen such that only one specific component of prospect theory is active. This means that either probability weighting, loss aversion or diminishing sensitivity is active. In order to study the predictive power of TK constructed with a focus on probability weighting, the parameters for loss aversion and diminishing sensitivity are set to one. So TK in the column labelled PW (probability weighting) is computed with the parameters  $\lambda=1; \alpha=1; \gamma=0.61; \delta=0.69$ . The same rationale applies to the other columns in Table 7. TK in the column labelled LA (loss aversion) corresponds to applying the parameters  $\lambda=2.25; \alpha=1; \gamma=1; \delta=1$  and TK in the column labelled DS (diminishing sensitivity) to the parameters  $\lambda=1; \alpha=0.88; \gamma=1; \delta=1$ . In contrast to Table 5, the independent variables in Table 7 are normalized to have a mean of zero and a standard deviation of one. Barberis et al. (2016) present a very similar analysis, but also look at interactions of the three components. The results I get for this extension do not provide much additional information. Therefore I don't present them here and avoid making the table more complex than needed, especially since there are two sections of results comparing bubble to non-bubble periods. The first three columns in Table 9 show the result for non-bubble periods. It seems that loss aversion by itself explains a great deal of the effect of prospect theory. Probability weighting and diminishing sensitivity on the other hand become insignificant predictors after including all control variables. Especially the control for the long term reversal has a great impact on their t statistics and on the sign of the coefficients. Before including *ltrev*, probability weighting is actually a significant predictor of returns on its own, with a negative sign on the coefficient. Maybe the turning of the sign is due to the high correlation between *ltrev* and TK constructed with probability weighting and diminishing sensitivity. However, I do not find any evidence for collinearity after testing for the VIF measure. Taking all this into account, I cautiously conclude that the positive sign on the coefficients should not be assigned too much weight, since they are negative for any model excluding *ltrev*. In general probability weighting by itself is a more significant predictor of returns than diminishing sensitivity. This suggests that the skewness of a return distribution should also be somewhat attractive to investors with prospect theory preferences. If this is true, then the combination of probability weighting and loss aversion in the TK variable should result in the most predictive power for returns. I indeed find that this TK variable has the most predictive power for returns, even higher than the original specification. The respective results are located in Appendix F. During bubbles the relative importance of the prospect theory components does not seem to change. Loss aversion remains to be the most reliable predictor of returns by itself. Overall, the effects are less significant, which reflects the re-

sult from Table 7 to some extent. If loss aversion is important during both bubble and non-bubble periods, what does this mean for the return distribution investors acting according to prospect theory prefer? Under loss aversion a loss has a greater weight than a gain of similar size. Therefore investors should prefer return distributions with low standard deviation. In summary then, the optimal distribution of returns should first and foremost have low standard deviation and combine this with some skewness. These results are just guesses, as a comparison of t statistics across the columns is no conclusive test of the matter. Since loss aversion seems to play such a big role, it is possible that the effect on returns is not unique to prospect theory. After all, expected utility also has a concave value function, indicating that individuals are loss averse. With this alternative setup, the prediction would be that high expected utility stocks earn low subsequent returns. In the next section I test this conjecture using the expected power utility formula from Barberis et al. (2016).

#### 6.7. Prospect theory versus expected utility

The test for expected utility will consist of the Fama-MacBeth regressions in Table 7, but this time I compare the effect of the TK variable with the effect of the variable EU, the power expected utility:

$$EU = (1/60) \sum_{i=-m}^n (1 + r_i)^{1-\theta} / 1 - \theta, \quad (26)$$

where the parameter theta is initially set at 5, indicating a high degree of loss aversion. EU is constructed using the same return distributions as the TK variable, where each return is weighted with equal probability. Of course it is stated in terms of final wealth; therefore it is  $1 + r_i$  instead of just  $r_i$ . This means that the loss aversion has the most impact in the region near-1. There, the curve for expected utility is especially steep. In addition meaning of subscripts m and n is in accordance with equation (6), m is the lowest ranked return and n is the highest ranked return in the distribution. Table 10 presents the results of this exercise. Similar to the prospect theory value, I find a negative relationship between the expected utility and returns. This would mean that the effect observed in Table 7 is not unique to prospect theory. A possible cause for this has been discussed earlier. Both models for decision making under risk share the incorporation of loss aversion. Maybe loss aversion is the dominating force behind the results and it does not matter whether returns are measured over final wealth or gains and losses. This result also would stand in contra to Barberis et al. (2016) that find no relationship between returns and expected utility. In fact, the regressions leading up to Table 7 show that the relationship between expected utility and returns is spurious. In the model that only tests expected utility and returns, the relationship is positive. This positive relationship holds until the control for the long term reversal is included. Again one could cautiously conclude that the elevated correlation between EU and *ltrev* has a confounding impact on the regression. Therefore the interpretation of the result from Table 10 is speculative. To quickly test what constitutes a high

**Table 9:** Relative importance of prospect theory components (Based on Barberis et al., 2016, p.3101).

The table presents the results of Fama-MacBeth regressions on the monthly percentage return as dependent variable. The independent control variables are the same as in Table 7, but normalized to have mean zero and standard deviation one. In each column, TK is computed with only one active parameter for probability weighting (PW), loss aversion (LA) and diminishing sensitivity. The sample runs from January 1983 until December 2017, but is split into bubble (N=355) and non-bubble (N=65) periods. T-statistics are in parentheses.

	Non-Bubble			Bubble		
	PW	LA	DS	PW	LA	DS
TK	0.071 (1.17)	-0.336*** (-3.34)	0.164 (0.52)	0.179 (1.00)	-0.813** (-2.76)	0.091 (0.11)
Beta	0.202* (2.49)	0.128 (1.89)	0.204* (2.55)	0.337* (2.33)	0.180 (1.76)	0.351* (2.33)
Size	-0.320*** (-5.64)	-0.225*** (-3.82)	-0.341*** (-5.99)	-0.320 (-1.79)	-0.103 (-0.56)	-0.375* (-2.02)
Bm	0.110*** (3.04)	0.135*** (3.83)	0.113** (3.14)	0.031 (0.27)	0.124 (1.33)	0.042 (0.37)
Mom	-0.040 (-0.53)	0.077 (1.28)	-0.089 (-0.66)	-0.012 (-0.07)	0.286 (1.68)	0.033 (0.08)
Illiq	0.209*** (4.23)	0.202*** (4.01)	0.210*** (4.22)	0.189 (1.80)	0.160 (1.45)	0.193 (1.86)
Rev	-0.709*** (-13.36)	-0.688*** (-14.34)	-0.727*** (-12.46)	-0.872** (-7.00)	-0.806** (-6.78)	-0.861** (-4.59)
Ltrev	-0.241*** (-3.48)	-0.006 (-0.09)	-0.342 (-1.27)	-0.509** (-2.86)	-0.106 (0.49)	-0.457 (-0.65)
Max	0.518* (2.13)	0.364 (1.69)	0.531* (2.21)	1.252 (1.78)	0.589 (1.03)	1.360 (1.87)
Min	-0.225*** (-6.11)	-0.275*** (-8.73)	-0.224*** (-6.11)	-0.131 (-1.77)	-0.235** (-4.07)	-0.135 (-1.79)

t-statistics in parentheses \*p<0.05, \*\*p<0.01, \*\*\*p<0.001

expected utility stock, I sort stocks into deciles each month and calculate the expected value for the different control variables. Then I analyse the time series average of these statistics, with a special focus on the return measures. The average returns in the high expected utility decile are lower than those of the first decile. In addition, the long term reversal is decreasing in the expected utility decile. Therefore I would expect a positive relationship between expected utility if anything, as the counterpart to the effect of prospect theory value. I don't present the resulting table here, because the point was to quickly back test which relationship between expected utility and returns could make sense. However, the corresponding table can be found in Appendix G. In any case, the prospect theory has higher predictive power over returns and the estimation process is more robust. In summary, I do not find definitive proof that expected utility has a similar ef-

fect on returns like prospect theory. Generally, the prospect theory value seems to reflect investor behaviour in the U.S. stock markets more accurately. Then, the following question arises from these results: Which return characteristics do investors with prospect theory preferences like the most? Previous results suggest that low standard deviation and skewness should be preferred. In order to get a broader picture regarding the return characteristics, the next section presents the results of a decile sort on prospect theory value, similar to the one for expected utility.

#### 6.8. Decile sorts on prospect theory value

Each month I sort stocks into deciles based on their prospect theory value. For each decile the average value for several of the control variables is computed. The sample periods starts in January 1983 and ends in December

**Table 10:** Prospect theory versus expected utility.

The Table reports the results of Fama-MacBeth regressions with percentage returns as the dependent variable. The control variables are the same as in Table 7. EU is a stock's expected utility based on its past return distribution of 60 months. The sample runs from January 1983 until December 2017, but is split into bubble (N=355) and non-bubble (N=65) periods. T-statistics are in parentheses.

	Non-Bubble		Bubble	
	(1)	(2)	(1)	(2)
TK	-0.092*** (-3.45)		-0.243** (-2.94)	
EU		-0.057** (-3.07)		-0.147** (2.69)
Beta	0.411 (1.58)	0.543* (2.23)	0.573 (1.45)	0.834* (2.04)
Size	-0.116*** (-4.07)	-0.119*** (-4.42)	-0.073 (-0.76)	-0.082 (-1.00)
Bm	0.160*** (3.60)	0.157*** (3.71)	0.169 (1.50)	0.127 (1.09)
Mom	-0.001 (-0.57)	-0.003 (1.31)	0.005 (1.05)	-0.004 (-0.95)
Illiq	0.344*** (3.33)	0.323*** (3.95)	0.264 (1.32)	0.249 (1.38)
Rev	-0.058*** (-12.62)	-0.060*** (-13.72)	-0.073*** (-6.38)	-0.075*** (-7.76)
Ltrev	-0.071 (-0.84)	-0.443*** (-5.04)	0.128 (0.49)	-0.950*** (-4.35)
Max	0.015** (2.90)	0.007 (1.84)	0.016 (1.12)	0.012 (1.21)
Min	-0.053*** (-6.76)	-0.050*** (-8.08)	-0.049*** (-3.74)	-0.039** (-3.31)
Skew	0.117* (2.56)	-0.005 (-0.18)	0.187 (1.56)	-0.027 (0.44)

t-statistics in parentheses \*p<0.05, \*\*p<0.01, \*\*\*p<0.001

2017. The reported values are the averages of the time series consisting of the average value per variable and month. In addition to sourcing from the control variables introduced earlier, the standard deviation is also included. Sd is computed as the standard deviation of monthly returns over the past three years. Including sd makes it possible to test whether loss aversion indeed plays a role when individual investors pick stocks. Table 11 presents the results of this exercise. First, prospect theory values are negative on average, even in the high TK decile. The prospect theory value

of a stock is also increasing in size and decreasing in the book-to-market ratio. This is a first indication that large and less volatile stocks are preferred by individual investors with prospect theory preferences. In accordance with this conjecture, the measure for illiquidity is decreases monotonically from decile one to decile ten. The regressions from Table 7 have shown that past return measures have a large impact on the predictive power of the prospect theory value. Table 11 makes the same point in a different way. Rev and ltrev increase monotonically from decile one through decile

**Table 11:** Prospect theory value portfolios (Based on Barberis et al., 2016, p.3100).

The Table reports the time series average of the monthly mean statistic of variables for deciles formed on TK. The description of variables can be found in Table 2. Sd is the standard deviation of monthly returns over the past five years. The sample starts in January 1983 and ends in December 2017.

Deciles	TK	Beta	Size	Bm	Rev	Illiq	Ltrev	Skew	Sd
1	-0.138	1.200	4.317	-0.285	-0.007	0.567	-0.403	0.166	0.184
2	-0.107	1.159	5.313	-0.445	0.003	0.211	0.126	-0.044	0.153
3	-0.093	1.120	5.789	-0.519	0.006	0.127	0.283	-0.166	0.136
4	-0.084	1.089	5.984	-0.562	0.008	0.093	0.384	-0.183	0.124
5	-0.076	1.055	6.107	-0.581	0.009	0.068	0.437	-0.145	0.114
6	-0.068	1.029	6.152	-0.589	0.010	0.055	0.468	-0.066	0.105
7	-0.062	1.001	6.218	-0.617	0.012	0.042	0.501	0.045	0.097
8	-0.054	0.986	6.317	-0.625	0.013	0.033	0.553	0.178	0.090
9	-0.046	0.961	6.487	-0.656	0.015	0.023	0.628	0.318	0.084
10	-0.032	0.989	6.772	-0.763	0.023	0.019	0.925	0.646	0.087
Total	-0.076	1.060	5.945	-0.564	0.010	0.124	0.390	0.075	0.117

ten. At first sight this could lead to the conclusion that TK really measures the effects of these reversals. Although these measures take significant bites out of the predictive power of TK, its negative relationship with returns remains significant. The results from Table 9 suggest that skewness should be preferred by investors with prospect theory preferences. As expected, skewness generally increases from decile one to decile ten. But the relationship is less straightforward than assumed. Apparently, returns in the first decile are skewed more than returns up to decile seven. This could be connected to the fact that the standard deviation in the first decile is at its highest point over all deciles. Returns exhibiting a high standard deviation are more likely to be skewed as well. The skewness of returns particularly increases from decile seven on, with the sharpest rise from decile nine to ten. The standard deviation also increases from decile nine to ten, while it decrease from decile one to nine. Again, this could be due to the relationship between skewness and standard deviation. However, the standard deviation remains below its average value. Stocks in the high prospect theory value deciles exhibit the characteristics expected earlier. They have a high return, high skewness and low standard deviation. Investors with prospect theory preferences should indeed like high returns combined with a low standard deviation, because they are loss averse. Loss aversion causes these investors to “feel” losses more heavily than gains, so they tilt towards low standard deviation. On the other hand investors with prospect theory preferences like skewness because of probability weighting. They overweight returns that are less likely to occur and tend to buy “stock lotteries”; high reward and low probability. Together with the results from previous tests, this leads to the conclusion that investors evaluating stock using prospect theory look for a certain mix of both worlds. High prospect theory value stocks are those which optimally combine skewness, volatility and high returns.

## 7. Conclusion and summary of results

There are several questions related to recent research I try to address with the analysis in this paper. The first goal was to test the model by Barberis et al. (2016). The fundamental assumption in this framework is that individual investors evaluate stocks based on the prospect theory value of their past return distributions. Individual investors add stocks with a high prospect theory value to their portfolio and sell those with low prospect theory value. Therefore stocks with high (low) prospect theory value become overvalued (undervalued) and earn low (high) subsequent returns. I find support for this prediction in the cross section of U.S. stock markets during the period of January 1983 to December 2017. This extends the evidence of the impact the preferences of individual investors have on prices in the stock market. Prospect theory seems to capture the behaviour and choices of individual investors to a certain degree. Based on the formal model of Barberis et al. (2016), I made the additional prediction that the relationship between prospect theory value and returns should be more pronounced during bubbles. This prediction was based on a two-fold argument: that the limits to arbitrage rise during bubbles and the number of individual investors increases. The limits to arbitrage are assumed to rise, because rational investors delay arbitrage and arbitrage becomes more costly. The increase in individual investors on the other hand comes from the attention grabbing nature of stocks during bubbles. Testing the prediction during bubbles is a way to find out whether the predictive power of the prospect theory value varies through time. It is also related to the question how investor behaviour changes depending on the market conditions. In order to test the prediction, I had to define what constitutes a bubble and apply a method to time stamp the start and collapse of the bubbles. The BSADF test developed by Phillips et al. (2015a) is tool for the detection of bubbles, which has been proven to work in many different markets. It focuses on the nature of the stochastic processes underlying the price development. The

application of this method has led to the conclusion that two bubbles existed in the price dividend ratio of the S&P 500 during the time period of January 1983 to December 2017. The first corresponds to the lead up to Black Monday in October and spans from February 1987 to August 1987. The second coincides with the tech bubble during the late stages of the past century, ranging from December 1995 to September 2000. The implementation of the BSADF test in the context of the time series of the S&P 500 supports the reliability of the approach. The time stamping of the start and end of the detected bubbles seems to be consistent across different studies, if the underlying time series used is similar. But I also find some issues for the specification of the test, especially concerning the necessary assumption of the minimum bubble duration, where the application of economic intuition becomes necessary. In the main empirical test I find no support for the prediction that high prospect theory value stocks earn lower subsequent returns during bubbles compared to non-bubble periods. In contrast, the returns to a strategy that buys stocks with low prospect theory value and sells stock with high prospect theory value are higher during bubbles. However both tests suggest that the significance could be a problem and the relationship between prospect theory value and returns is, on average, not stronger during bubbles. The changes in significance are not large enough to justify an argument for an increase or decrease in the predictive power of the prospect theory value. The main analysis also sheds some light on the predictive power of other commonly used variables in financial research. Variables drawing their power from a measurement of over- or undervaluation suffer noticeably during bubbles. The same goes for predictors that are based on past performance, although they mostly remain significant. The short term reversal especially has shown a very robust relationship with returns. In light of the findings, the almost constant effect of prospect theory on returns shows that behavioural explanations of asset prices can offer valuable insights. Being able to time stamp bubbles in the stock markets might make it possible to better understand the forces determining prices. This is an interesting avenue for future research in asset pricing. The analysis on the accuracy of the limits to arbitrage argument has also proven insightful. The first and maybe most important question is how to measure the limits to arbitrage. Most of the research is focussed on approximations for costs or risks involved with arbitrage. But they might not be able to explain many risks that arbitrageurs face in the stock market. The delayed arbitrage model by [Abreu and Brunnermeier \(2003\)](#) is just one example for this. By applying several proxies for the limits to arbitrage I was able to show that stocks with higher limits to arbitrage are more prone to the effect of the prospect theory value. Future research in the measurement of limits to arbitrage could be helpful, although it is already quite extensive. Looking into each of the individual components of prospect theory has shown that loss aversion and probability weighting together predict returns with the highest significance. Because of loss aversion, investors with prospect theory preferences like stocks that have a low standard deviation. They

also tend to prefer stocks with high skewness, because low probability outcomes receive more weight in the return distribution due to probability weighting. In general, investors seem to prefer a combination of low standard deviation and high skewness under prospect theory. There are several open questions with regards to prospect theory that future research could address. One important aspect is the definition of gains and losses. The modelling of gains and losses based on recent expectations is a concept that, with further testing, could prove valuable to applications of prospect theory. Empirically testing the reference points of investors could prove difficult though, since the necessary data has to be available on the investor level. Further research on the information sources could provide equally helpful guidelines in this matter. In the short review under section 2.3.2 several interesting questions arise. Where do investors get information on stock prices and returns? How does the presentation form influence the perceived attractiveness of a stock? An overview that connects the findings of the related research would be very helpful for the empirical application of prospect theory. It is also possible to consider other models of individual's choices under risk, as the section on expected utility has shown. The assumption that investors evaluate stocks based on a stock's past return distribution could maybe be lifted in such models. It is possible, that these models have even greater explanatory power for asset prices. After all, one of the most interesting findings of my analysis is that behavioural models like prospect theory might be more consistent predictors of stock prices over time than measures of valuation or risk-return relationships.



## References

- Abdelloui, M. Parameter-free elicitation of utility and probability weighting functions. *Management Science*, 46(11):1497–1512, 2000.
- Abreu, D. and Brunnermeier, M. K. Bubbles and crashes. *Econometrica*, 71(1):173–204, 2003.
- Agnew, J. R. and Szykman, L. R. Asset allocation and information overload: The influence of information display, asset choice, and investor experience. *The Journal of Behavioral Finance*, 6(2):57–70, 2005.
- Alexander, G. J., Jones, J. D., and Nigro, P. J. Mutual fund shareholders: Characteristics, investor knowledge, and sources of information. *Financial Services Review*, 7(4):301–316, 1998.
- Ali, A., Hwang, L. S., and Trombley, M. A. Arbitrage risk and the book-to-market anomaly. *Journal of Financial Economics*, 69(2):355–373, 2003.
- Amihud, Y. Illiquidity and stock returns cross-section and time-series effects. *Journal of Financial Markets*, 5(1):31–56, 2002.
- Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299, 2006.
- Asquith, P., Pathak, P. A., and Ritter, J. R. Short interest, institutional ownership, and stock returns. *Journal of Financial Economics*, 78(2):243–276, 2005.
- Baillon, A., Bleichrodt, H., and Spinu, V. Searching for the reference point. Working paper, 2015.
- Barber, B. M. and Odean, T. All that glitters: The effect of attention and news on the buying behaviour of individual and institutional investors. *The Review of Financial Studies*, 2008.
- Barberis, N. and Huang, M. Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review*, 98(5):2066–2100, 2008.
- Barberis, N. and Thaler, R. A survey of behavioral finance. *Handbook of the Economics of Finance*, 1:1053–1128, 2003.
- Barberis, N., Mukherjee, A., and Wang, B. Prospect theory and stock returns: An empirical test. *The Review of Financial Studies*, 29(11):3068–3107, 2016.
- Benartzi, S. and Thaler, R. H. Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, 110(1):73–92, 1995.
- Berk, J. B. Sorting out sorts. *The Journal of Finance*, 55(1):407–427, 2000.
- Bhargava, A. On the theory of testing for unit roots in observed time series. *The Review of Economic Studies*, 53(3):369–384, 1986.
- Brav, A., Heaton, J. B., and Li, S. The limits of the limits of arbitrage. *Review of Finance*, 14(1):157–187, 2009.
- Brennan, M. J. How did it happen. *Economic Notes*, 33(1):3–22, 2004.
- Bruhlin, A., Fehr-Duda, H., and Epper, T. Risk and rationality. Uncovering heterogeneity in probability distortion. *Econometrica*, 78(4):1375–1412, 2010.
- Busetti, R. and Taylor, A. R. Tests of stationarity against a change in persistence. *Journal of Econometrics*, 123(1):33–66, 2004.
- Caspi, I., Katzke, N., and Gupta, R. Date stamping historical periods of oil price explosivity. 1876–2014. *Energy Economics*, 70:582–587, 2018.
- Chen, J., Hong, H., and Stein, J. C. Breadth of ownership and stock returns. *Journal of Financial Economics*, 66(2-3):171–205, 2002.
- Cheung, A., Roca, E., and Su, J. J. Crypto-currency bubbles: an application of the Phillips–Shi–Yu (2013) methodology on Mt. Gox bitcoin prices. *Applied Economics*, 47(23):2348–2358, 2015.
- Chu, Y., Hirshleifer, D., and Ma, L. The causal effect of limits to arbitrage on asset pricing anomalies (No. w24144). National Bureau of Economic Research, 2017.
- Daniel, K. and Titman, S. Market reactions to tangible and intangible information. *The Journal of Finance*, 61(4):1605–1643, 2006.
- D’Avolio, G. The market for borrowing stock. *Journal of Financial Economics*, 66(2-3):271–306, 2002.
- De Bondt, W. F. and Thaler, R. Does the stock market overreact? *The Journal of Finance*, 40(3):793–805, 1985.
- De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. Noise trader risk in financial markets. *Journal of Political Economy*, 98(4):703–738, 1990.
- Diba, B. T. and Grossman, H. I. Explosive rational bubbles in stock prices. *American Economic Review*, 78(3):520–530, 1988.
- Duffie, D., Garleanu, N., and Pedersen, L. H. Securities lending, shorting, and pricing. *Journal of Financial Economics*, 66(2-3):307–339, 2002.
- Escobari, D. and Jafarnejad, M. Date stamping bubbles in real estate investment trusts. *The Quarterly Review of Economics and Finance*, 60:224–230, 2016.
- Evans, G. W. Pitfalls in testing for explosive bubbles in asset prices. *The American Economic Review*, 81(4):922–930, 1991.
- Fama, E. F. and French, K. R. The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465, 1992.
- Fama, E. F. and French, K. R. Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1):55–84, 1996.
- Fama, E. F. and MacBeth, J. D. Risk, return, and equilibrium empirical tests. *Journal of Political Economy*, 81(3):607–636, 1973.
- Garber, P. M. Famous first bubbles. *Journal of Economic Perspectives*, 4(2):35–54, 1990.
- Glaser, M., Langer, T., Reynders, J., and Weber, M. Framing effects in stock market forecasts. The difference between asking for prices and asking for returns. *Review of Finance*, 11(2):325–357, 2007.
- Glaser, M., Ilieva, Z., and Weber, M. Thinking about prices versus thinking about returns in financial markets. *The Journal of Finance*, 74(6):2997–3039, 2019.
- Griffin, J. M., Harris, J. H., Shu, T., and Topaloglu, S. Who drove and burst the tech bubble. *The Journal of Finance*, 66(4):1251–1290, 2011.
- Gromb, D. and Vayanos, D. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics*, 66(2-3):361–407, 2002.
- Gromb, D. and Vayanos, D. Limits of arbitrage. *Annual Review of Financial Economics*, 2(1):251–275, 2010.
- Gürkaynak, R. S. Econometric tests of asset price bubbles taking stock. *Journal of Economic Surveys*, 22(1):166–186, 2008.
- Hirshleifer, D., Teoh, S. H., and Yu, J. J. Short arbitrage, return asymmetry, and the accrual anomaly. *The Review of Financial Studies*, 24(7):2429–2461, 2011.
- Hommel, U. and Breitung, J. Testing for speculative bubbles in stock markets a comparison of alternative methods. *Journal of Financial Econometrics*, 10(1):198–231, 2012.
- Hu, Y. and Oxley, L. Are there bubbles in exchange rates? Some new evidence from G10 and emerging market economies. *Economic Modelling*, 64:419–442, 2017.
- Jegadeesh, N. and Titman, S. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1):65–91, 1993.
- Kahneman, D. and Tversky, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–292, 1979.
- Keller, K. L. and Staelin, R. Effects of quality and quantity of information on decision effectiveness. *Journal of Consumer Research*, 14(2):200–213, 1987.
- Kim, J. Y. Detection of change in persistence of a linear time series. *Journal of Econometrics*, 95(1):97–116, 2000.
- Kőszegi, B. and Rabin, M. A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165, 2006.
- Kumar, A. Who gambles in the stock market? *The Journal of Finance*, 64(4):1889–1933, 2009.
- Kumar, A. and Lim, S. S. How do decision frames influence the stock investment choices of individual investors. *Management Science*, 54(6):1052–1064, 2008.
- Lehmann, B. N. Fads, martingales, and market efficiency. *The Quarterly Journal of Economics*, 105(1):1–28, 1990.
- Lin, W. and Meng, J. The Status Quo or Expectation. Estimating Investors’ Reference Points on Stock Returns. Working Paper, 2015.
- Lurie, N. H. and Swaminathan, J. M. Is timely information always better. The effect of feedback frequency on decision making. *Organizational Behavior and Human Decision Processes*, 108(2):315–329, 2009.
- Markowitz, H. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.
- Marsh, T. A. and Merton, R. C. Dividend variability and variance bounds tests for the rationality of stock market prices. *American Economic Review*, 76(3):483–498, 1986.
- Mashruwala, C., Rajgopal, S., and Shevlin, T. Why is the accrual anomaly not arbitrated away? The role of idiosyncratic risk and transaction costs. *Journal of Accounting and Economics*, 42(1-2):3–33, 2006.
- Menkhoff, L. and Schmeling, M. A prospect-theoretical interpretation of momentum returns. *Economic Letters*, 93(3):360–366, 2006.
- Miller, E. M. Risk, uncertainty, and divergence of opinion. *The Journal of Finance*, 32(4):1151–1168, 1977.

- Nagel, S. Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics*, 78(2):277–309, 2005.
- Neumann, M., J.v. annd Oskar. *Theory of Games and Economic Behavior*. Princeton, NJ, Princeton University Press, 1944.
- Nolte, S. and Schneider, J. C. How price path characteristics shape investment behavior. *Journal of Economic Behavior & Organization*, 154(33–59), 2018.
- Novy-Marx, R. The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28, 2013.
- Ofek, E. and Richardson, M. The valuation and market rationality of internet stock prices. *Oxford Review of Economic Policy*, 18(3):265–287, 2002.
- Phillips, P. C. and Shi, S. Real time monitoring of asset markets bubbles and crises. *Cowles Foundation Discussion Paper (no.2152)*, 2018.
- Phillips, P. C., Wu, Y., and Yu, J. Explosive behavior in the 1990s Nasdaq. When did exuberance escalate asset values. *International Economic Review*, 52(1):201–226, 2011.
- Phillips, P. C., Shi, S., and Yu, J. Specification sensitivity in right-tailed unit root testing for explosive behaviour. *Oxford Bulletin of Economics and Statistics*, 76(3):315–333, 2014.
- Phillips, P. C., Shi, S., and Yu, J. Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review*, 56(4):1043–1078, 2015a.
- Phillips, P. C., Shi, S., and Yu, J. Testing for multiple bubbles: Limit theory of real-time detectors. *International Economic Review*, 56(4):1079–1134, 2015b.
- Quiggin, J. A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4):323–343, 1982.
- Rabin, M. and Weizsäcker, G. Narrow bracketing and dominated choices. *American Economic Review*, 99(4):1508–43, 2009.
- Scheinkman, J. A. and Xiong, W. Overconfidence and speculative bubbles. *Journal of political Economy*, 111(6):1183–1220, 2003.
- Shiller, R. C. Irrational exuberance. *Philosophy & Public Policy Quarterly*, 20(1):18–23, 2000.
- Shiller, R. J. The use of volatility measures in assessing market efficiency. *The Journal of Finance*, 36(2):291–304, 1981.
- Shleifer, A. and Vishny, R. W. The limits of arbitrage. *The Journal of Finance*, 52(1):35–55, 1997.
- Stoll, H. R. Presidential address: friction. *The Journal of Finance*, 55(4):1479–1514, 2000.
- Tom, S. M., Fox, C. R., Trepel, C., and Poldrack, R. A. The neural basis of loss aversion in decision-making under risk. *Science*, 315(5811):515–518, 2007.
- Tversky, A. and Kahnemann, D. The framing of decisions and the psychology of choice. *Science*, 211(4481):453–458, 1981.
- Tversky, A. and Kahnemann, D. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323, 1992.
- West, K. D. A specification test for speculative bubbles. *The Quarterly Journal of Economics*, 102(3):553–580, 1987.
- White, E. N. The stock market boom and crash of 1929 revisited. *Journal of Economic Perspectives*, 4(2):67–83, 1990.



# Impact of Weather on the Stock Market Returns of Different Industries in Germany

Astrid Schulte-Huermann

WHU – Otto Beisheim School of Management

## Abstract

Weather affects people's mood, according to psychological studies. For example, low temperature can cause aggression, whereas high temperature can induce apathy. Therefore, it may be possible that weather-influenced mood, driven by mood's impact on decision-making, exerts an influence on investment decisions and risk-taking behaviour. Thereby, it might also impact stock market returns. I examine market returns of nine industries in Germany. Empirical results illustrate two findings. First, a statistically significant, negative correlation between market returns and temperature, second, a different effect of weather on industrial sectors is identified. A significant correlation can be found for six out of nine sectors.

**Keywords:** Stock markets; investor behaviour; weather effect; market returns; decision-making.

## 1. Introduction

The impact of climate change has become noticeable on a global level, and Germany has experienced more extreme weather phenomena during the last years. According to the German weather service (Deutscher Wetterdienst (DWD)), the summer of 2018 marks the second hottest summer on record since 1881 (Imbery et al., 2018). Further, weather and climate implications on companies' earnings are increasingly discussed on executive earnings calls, and rating agencies have started to focus on weather implications on companies' economic performance (Williams et al., 2018).

Investigating weather implications on the economy is a relevant field of study that is already supported by extensive research. For example, weather has been linked to stock market performance. The underlying reason for connecting these fields is that research has found evidence that weather conditions affect mood and spawn mood misattribution. A change in emotional state can trigger an alternated risk assessment and in turn a revised investment behaviour. This can ultimately lead to changes in stock market returns. Based on this finding, my study analyses if a relationship between weather and stock market returns exists.

Several researchers have carried out studies in the field of weather implications on psychology. Howarth and Hoffman (1984) observed that human performance is positively correlated with sunshine, and aggressiveness increases with low temperature. On the opposite, high temperatures have been associated with apathy (Cao and Wei, 2005). On sunny days,

people are inclined to rate their life satisfaction higher than on cloudy days (Schwarz, 1990). People's mood is linked to their decision-making (Schwarz, 1990). Thus, it is also a component of their investment decisions.

Turning to empirical research, Saunders (1993) was among the first who linked weather to investors' behaviour. He reported a negative correlation between NYSE index returns and the degree of cloudiness. His results support the conjecture that investors are overoptimistic on sunny days and more pessimistic on cloudy days. Further evidence for the negative relation between cloud cover and stock market returns were given by Hirshleifer and Shumway (2003). In 18 out of 26 cities, a negative relationship could be reported. Other research produced evidence for a negative correlation between stock returns and temperature (Cao and Wei, 2002; Chang et al., 2006; Kang et al., 2010).

Moreover, the seasonal affective disorder (SAD) has an impact on stock returns. Affected people show depressive symptoms during winter, and consequently, stock market returns are lower during this time of the year. Kamstra et al. (2003) call this effect winter blues.

### 1.1. Relevance

My paper expands on the existing research by examining the probable linkage between stock returns and weather variables temperature and precipitation with a focus on Germany and differentiating between sectors. The former is advocated to be among the most essential meteorological variables to alter people's mood. The latter is only included in a minor

part of literature yet and therefore will be further explored. In particular, my analysis examines the impact of temperature and precipitation on the stock market return of different industries in Germany. Hence, I will contribute to a yet unexplored field by differentiating stock market return behaviour between sectors. Hitherto, only stock market indices on a global scale or across industries have been analysed. While most research focuses on adjusting the independent variable set containing weather data, refinements of the dependent variable in the form of separation of industries are somewhat limited.

Weather does not only impact mood and decision-making but also it has economic implications on companies' economic efficiency. Therefore, my study expands upon the relationship between weather and stock market returns by also incorporating economic implications. It is evident that the climate and weather implications on the economy might vary across sectors. This is because integral parts of the supply chain or the use of the finished product depend to ranging degrees on environmental circumstances as the weather. Temperature, for example, might influence the demand for some products. Especially in the Retail and Utilities industry, demand may increase during periods of lower temperatures, whereas the economic performance of other companies is harmed when temperatures are high. These effects might also be present in the opposite direction because high and low temperatures can both be beneficial and unfavourable, depending on the industry. I also conjecture that precipitation might carry weight in the determination of stock market returns.

Germany as the country of research has only been investigated once on a national level by Krämer and Runde (1997). Their study concluded that no systematic relationship exists for the German stock market returns and weather variables. The rejection of the null hypothesis is rather determined by the manner the hypothesis is phrased. My study will elucidate whether those outcomes can be confined and can be replicated for a later period of time.

As psychological literature has outlined, weather influences people's mood and behaviour. Consequently, the risk assessment and investment decisions of traders could be altered. If this holds, trader's evaluation of companies might be over- or under-optimistic depending on the weather. Likewise, hot temperatures might induce a change in the evaluation of some industries. Another approach could be that changes in weather influence the mood of traders only for some particular industries. Certain sector-returns might not be affected by the weather at all. The inference seems appropriate, if one considers the impact of changing weather conditions through global warming on politics and industries today. By dividing into industry groups, it is finally possible to assess whether all company returns are affected by the weather or if this phenomenon is limited to specific sectors.

From a psychological point of view, potential findings could be used for weather-based trading strategies. From an economic standpoint, a sector might be attractive for stock trading, if its market returns are influenced by weather. In-

dustries that are not correlated to weather variables might be more valuable for a buy and hold strategy. In general, the findings could not only be interesting for traders, but also be interesting for investors, and financial managers.

## 1.2. Structure

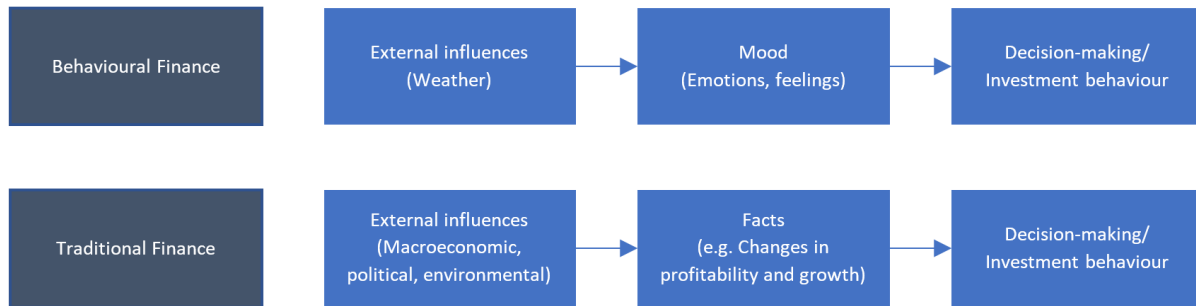
The remaining part of this paper is structured as follows. Previous qualitative and quantitative work is discussed in section two. Section three outlines the methodology and section four describes the data set which was collected for both weather and stock market returns, including descriptive statistics. Section five outlines the underlying hypotheses and leads to the central part of this paper. A detailed analysis of the relationship of weather with stock market returns is given in section six by using both a bin-test and an EGARCH model. Section seven will discuss my findings. Section eight will provide a further outlook and highlights the potential limitation of my work; afterwards, section nine concludes.

## 2. Literature review

Thorough research on the topic has made it evident that behavioural finance uses a systematic approach for the analysis of the potential impact of weather on an investors' psychology. The first part, thus, will elaborate on the effects of weather on human beings' mood. The second part will discuss the impact of mood on decision-making, and the third part presents research that links weather conditions with investment behaviour and consequently stock markets. The underlying process is illustrated in Figure 1.

### 2.1. Weather and mood

One-third of humans respond sensitively to weather, which is characterised by changes in mood and human well-being (Kals, 1982). More precisely, weather can provoke compound physical and psychological responses, including changes in performance, aggressive behaviour, and interpersonal interactions (Lu and Chou, 2012; Shu and Hung, 2009). The phenomenon has been discussed in a large body of literature, where the weather has been linked with human behaviour (Hirshleifer and Shumway, 2003; Howarth and Hoffman, 1984; Parsons, 2001; Pilcher et al., 2002; Rind, 1996; Watson, 2000). In current research, temperature is the most prominent influencing factor for people's mood. Especially, low temperature can fortify aggressive behaviour and discomfort. This effect also holds for high temperatures where apathy and hysteria are caused in addition to aggression (Cao and Wei, 2005; Pilcher et al., 2002). Extreme temperatures can also induce a decline in task performance (Pilcher et al., 2002), as well as a greater willingness to offer help to other people (Cunningham, 1979). Sunshine, on the opposite, can lead to higher satisfaction on sunny days in comparison to cloudy days, and consequently can enhance people's optimism (Howarth and Hoffman, 1984). Moreover, sunshine has been proven to influence the amount of tipping, as confirmed by Rind (1996). On the contrary, lack



**Figure 1:** Self-created process model of the impact of external influences on mood and decision-making

of sunshine was noted to cause depression (Eagles, 1994) and suicide (Tietjen and Kripke, 1994). Sariannidis et al. (2016) claim that wind is directly linked with investors' feeling of safety. With higher wind speed, air pollution is diluted, and therefore particles of dirt accumulate less. Humidity is an indicator that can lower human comfort and simultaneously lead to more aggressive behaviour (Cao and Wei, 2005). Since ancient times, the impact of the moon, especially the full moon, has been associated with changes in human feelings. Research has confirmed a shift in the behaviour of people: higher occurrence of somnambulism, crime, and mental disorders (Zheng et al., 2001).

## 2.2. Mood and decision-making

As weather can affect mood in various ways, this part will review psychological evidence that a change in mood, feelings, or emotions can influence decision-making. Among scholars, no consensus on the definition of the three terms exists, and the words are frequently used interchangeably.

Schwarz was among the first to identify that mood influences decision-making (Schwarz, 1990). He constitutes that mood can have an impact on the judgement of the information capabilities of a person. Further, a good mood can induce a more positive appraisal of life satisfaction (Wright and Bower, 1992) and prospects (Hirshleifer and Shumway, 2003). Also, a more intense utilisation of simplifying heuristics by individuals was identified in the respective research (Bless et al., 1996). Consequently, individuals base their judgement on stereotypes and are more likely to identify a deviation of available information from pre-existing knowledge. Other research has found that positive mood can lead to a higher receptiveness to weak arguments (Mackie and Worth, 1991), while it can enhance problem-solving abilities and creativity (Isen, 2002). According to Schwarz (1990), not only mood but also feelings are linked to decision-making. He further claims that through misattribution, emotions are connected to a false origin and consequently, induce incorrect judgements. Thus, the capability to process information is hindered and especially investors falsely respond to irrelevant information in a state of happiness. Also, visceral factors exert an influence on human beings' preferences (Hirshleifer and Shumway, 2003). Summarising, mood affects judgement, choices, and decision-making.

## 2.3. Stock markets and weather

Behavioural Finance is a field of research that has as its central premise the investment decision of individuals. In contrast to Traditional Finance, it is based on the theory that decision-making by individuals is not always rational but based on the impact of emotions and mood. Kahneman and Tversky are considered essential contributors to Behavioural Finance, promoting the non-existence of a "homo economicus". Instead, cognitive biases and heuristics guide an individual's behaviour and lead to irrational decisions. Consequently, arguments of the field of Behavioural Finance support the view that investment decisions are linked to people's mood. As summarised in section 2.1, weather conditions can influence mood and can cause for instance misattribution. They might also lead to changes in risk assessment and consequently altered tendencies in investment activities. Pleasant weather can lead to more communication, which expands information transfer, thereby encouraging investment activity (Symeonidis et al., 2010). Favourable weather conditions can also induce good mood. Investors transfer this positive state and evaluate economic prospects more optimistic (Shu, 2010) as they underestimate risk and overestimate the probability of success (Wright and Bower, 1992). As a result, stock market returns might be influenced by weather-induced effects. Existing research has not only linked mood to weather but also has investigated potential relationships between stock market returns and various meteorological variables. There is both literature supporting and rejecting the relationship between weather and stock returns.

Temperature (Cao and Wei, 2005; Chang et al., 2008; Floros, 2008; Keef and Roush, 2002), number of sunshine hours (Pardo and Valor, 2003; Shu and Hung, 2009), cloud cover (Hirshleifer and Shumway, 2003; Saunders, 1993) and air humidity levels (Chang et al., 2006; Pardo and Valor, 2003; Sariannidis et al., 2016) are, among others, the standard variables to measure weather's influence on stock market returns. Also, other nature-related variables such as the daylight-savings time change (Kamstra et al., 2000), the length of the night (Kamstra et al., 2003), and the lunar phases of the moon (Dichev and Janes, 2001; Zheng et al., 2001), have been used as weather proxies that might influence an investor's mood.

### 2.3.1. Evidence of the weather and stock market relationship

Saunders (1993) was among the first researchers to connect weather variables with stock market returns. By using daily returns of the Dow Jones Industrial Average from 1927 to 1989, as well as equal- and value-weighted daily percentage changes in the NYSE/AMEX index from 1962 to 1989, he identified a significant negative correlation between the percentage of cloud cover and stock returns. This effect is robust with respect to various market anomalies, including the Monday and January effect. Consequently, Saunders conjectures that the mood of investors is overoptimistic on sunny days.

Further evidence for this sunshine effect has been given by Hirshleifer and Shumway (2003) who assess the level of cloudiness in 26 countries from 1982 to 1997. An OLS regression model and a logit model both indicate that cloudiness exhibits a strong negative effect on returns in the majority of countries. The results remain stable after controlling for seasonality. This finding emphasises the conjecture that overcast sky reinforces a downbeat mood. Therefore, the researchers raise the question of whether a trading strategy based on a sunshine effect exists.

The research of Dowling and Lucey (2005) emphasises the relationship between stock market returns and rain and concludes that both variables are negatively correlated. This outcome is in line with the study of Saunders (1993), as “85% of rain occurs on days with 100% cloud cover” (Dowling and Lucey, 2005, p. 18).

A vast base of stock-market-focused literature uses temperature as a weather proxy. For most of the studies, a negative relationship has been found between temperature and stock market returns: The higher the temperature, the smaller the return.

For example, in an investigation of eight stock exchanges, Cao and Wei (2002, 2005) have found a negative correlation between temperature and market returns.

Negative correlation of temperature, sunshine, and humidity on the returns of the Shanghai stock market has been documented (Kang et al., 2010). Similar results were identified by Yoon and Kang (2009) for Korean stock returns. Here, for the whole period of 1990 to 2006, high temperature and humidity negatively influenced stock returns, while high cloudiness and low humidity led to positive stock returns. For the Taiwan stock exchange, temperature and cloud cover have a significant effect while humidity is not significantly correlated (Chang et al., 2006).

Keef and Roush (2007, 2002) have studied both the Australian and the New Zealand Stock Exchange, deriving the conclusion that temperature negatively affects returns. Whereas wind speed and cloud cover do not have a significant impact on the Australian stock returns, wind is indeed significant for returns in New Zealand. A further study of the researchers from Keef and Roush (2005) does indicate a negative correlation between New Zealand stock prices and wind, as well as a positive effect of sunshine on bank bills.

Kamstra et al. (2000) draw attention to the effect of sleep desynchronosis, an interruption in sleep patterns, induced

through daylight saving time changes on stock returns. Further, the seasonal affective disorder, caused by fewer daily sunshine hours, has a significant impact on stock returns in several countries (Hirshleifer and Shumway, 2003). The researchers provide evidence that longer nights correlate with lower stock returns.

Europe has also served as a ground for studies in the field of weather impact on stock exchanges (Floros, 2011, 2008; Sariannidis et al., 2016). Hereby, negative relationships between weather variables and stock returns were again identified (Shu and Hung, 2009).

### 2.3.2. Evidence against the weather and stock market relationship

Contradicting with previously discussed findings is the research by Trombley (1997) who replicated Saunders results without finding any significant correlations. Thus, he states that “[Saunders’s] choice is the only comparison during this period that would produce a statistically significant test statistic and does not consider that the returns on the zero percent days are inconsistent with the existence of a weather effect (p.13).”

Replication of Saunders result with German data has also supported Trombley’s view (Krämer and Runde, 1997). Further, the researchers Krämer and Runde claim, that the significance of results, and a systematic relationship between weather and stock returns, does mostly depend on the phrasing of the null hypothesis and the variables. Loughran and Schultz (2004) receive little evidence in their analysis of the effect of cloudiness in the location of the company’s headquarter on the returns of NASDAQ stocks.

Moreover, wind speed is no indicator for returns, as proven by Shu and Hung (2009) who analysed 18 European countries from 1994 to 2004.

Extant work by Zhu and Goetzmann (2003), who focused on data of individual traders during the period 1992 to 1996, further proved that not the weather, but the bid-ask spread led to investors’ propensity to buy or sell. These findings are consistent with Pardo and Valor (2003) who provide evidence that humidity is not correlated with the Madrid Stock Exchange. This serves as an example of rational human behaviour. Lu and Chou (2012) also derive that the relationship between weather and stock market returns is insignificant. However, they identified that humidity and wind exert a positive effect on the volatility of stock indices, as shown with the Shanghai Stock Exchange.

In contrast to the research of Dichev and Janes (2001), who identified a strong effect of the lunar cycle on stock returns that is consistent over the last 100 years and for the majority of the 24 countries analysed, Rotton and Rosenberg (1984) have gathered evidence that shows no significant effect of the lunar phases on the Dow Jones Index prices.

Taken together, the relationship between weather and stock returns is controversial (Floros, 2011; Sariannidis et al., 2016) and research results differ. Locations, data, and the time horizon used are distinct among extant research and cannot lead to a generalised conclusion. However, it can be

said that there is a significant body of literature that has provided evidence on an existing relationship between weather-related variables and stock market returns. If that holds, the marginal investor's behaviour is dependent on the weather, the season of the year, the phases of the moon, the daylight-savings time, and the length of the night. According to [Zhu and Goetzmann \(2003\)](#), the market is therefore not efficient, and this leaves opportunities to exploit its inefficiency. Investors can benefit from the awareness of the impact of their mood, as [Hirshleifer and Shumway \(2003\)](#) suggest. Nevertheless, they also claim that mood is influenced not only by weather but also by other factors and biases.

The existence of the weather's influence on stock returns cannot be denied and raises the question on whether this effect is limited to countries and the industry as a whole or whether sectors are affected differently by weather variables. Having a profound body of literature as a base, the material will be used as a guide to analyse a yet underexplored field of research.

### 3. Methodology

The purpose of this paper is to gain valuable insight into the impact of weather on stock market returns in different industries in Germany. Literature has made use of a variety of analysis tools to investigate the relationship between market returns and weather. For this purpose, weather variable data and industry return data have been collected. The following paragraphs will describe the methodology of my analysis which will be grouped into two types of tests. Afterwards, the data collection shall be discussed.

#### 3.1. Bin test

The first analysis, following [Saunders \(1993\)](#), is called Bin test based on [Cao and Wei \(2005\)](#). For this, I will use both industry index returns and the meteorological variable temperature. It will enable the identification of relationships between temperature and returns by matching temperature and return data, grouping it into bins, and calculating z-scores to evaluate the statistical difference between the bins. However, this test cannot identify specific correlations and lacks the ability to control for known stock return anomalies. The underlying methodology and detailed steps will be explained in the following.

First, the daily return and respective temperature data are sorted in descending order and separated into four bins. For each group, the mean return and the frequency of positive return are calculated, and the mean returns of the first and the fourth bin, covering the lowest temperature and the highest temperature respectively, will be compared. This will allow for the evaluation of the potential significance in the difference in mean return. Another comparison will be made between the percentage of positive returns in the four bins to verify whether the return difference is caused by outliers.

In detail, first, the four bins are identified by taking the range between maximum and minimum temperature and dividing by the total number of  $k$  bins with ( $k = 1, \dots, 4$ ).

$$\Delta = (\text{Temperature}_{\max} - \text{Temperature}_{\min}) / k \quad (1)$$

The bin ranges are then calculated as followed:  $[\text{Temp}_{\min} + \Delta]$ ,  $[\text{Temp}_{\min} + 2\Delta]$  and so forth. Based on the procedure by [Saunders \(1993\)](#), a z-test is performed to verify whether mean return differences of bin one and four are statistically significant.

$$z_{-} \text{score}_{4,1}^{\text{mean}} = (\mu_4 - \mu_1) / \sqrt{\frac{\sigma_4^2}{n_4} + \frac{\sigma_1^2}{n_1}} \quad (2)$$

The variable  $\mu_i$  describes mean return,  $\sigma_i^2$  is variance of return and  $n_i$  the number of observations per bin  $i$  with ( $i = 1, \dots, k$ ). A z-statistic is also performed for the percentage of positive returns per bin, to examine whether there is a significant difference between frequency of positive returns in the two extreme bins.

$$z_{-} \text{score}_{4,1}^{\text{frequency}} = (p_4 - p_1) / \sqrt{\frac{p_4(1-p_4)}{n_4} + \frac{p_1(1-p_1)}{n_1}} \quad (3)$$

In this case,  $p_i$  is the percentage of positive return for bin  $i$  ( $i = 1, \dots, k$ ).

#### 3.2. EGARCH model

In financial research, the generalised autoregressive conditional heteroscedasticity (GARCH) model has been used extensively to investigate the relation between weather variables and stock market returns. The Glosten, Jagannathan, and Runkle (GJR) GARCH model has been utilised by [Chang et al. \(2006\)](#) and [Symeonidis et al. \(2010\)](#), other researchers employed the autoregressive threshold GARCH (AR-TGARCH) (1,1) model ([Floros, 2011](#)) or a simple GARCH(1,1) analysis ([Kang et al., 2010](#)). My study will make use of an exponential GARCH (EGARCH) model by [Nelson \(1991\)](#), and it resembles the core of my thesis.

The EGARCH model is applied because it is frequently argued that stock return shocks are asymmetric. Positive return shocks have a lower impact on volatility than negative return shocks of the same magnitude. This phenomenon is called leverage effect. Further, the model captures volatility clustering, which is typical for financial time series. The omitted parameter restriction can result in more reliable optimisations. It also makes this model robust against changes in volatility and against long-run memory effects which are a result of time-varying clusters. Besides, the EGARCH model uses the natural logarithm  $\ln(\cdot)$  which guarantees the positivity of the parameters. As a result, the EGARCH model has been often seen to be superior to other models.

The regression equation of the EGARCH model including return  $R$ , prices  $P$ , and a stochastic error term, is expressed as follows:

$$R_{t,j} = b_0 + b_1 \sigma_t^2 R_{t-1,j} + b_2 \sigma_t^2 \text{Temp}_t + b_3 \sigma_t^2 \text{Prec}_t + \xi_t \quad (4)$$

$R_{t,j}$  is the daily industry return gain or loss for day  $t$ , calculated:

$$R_{t,j} = \ln(P_{t,j}) - \ln(P_{t-1,j}) \tag{5}$$

For the industry  $j$ , the subscript  $j \in$  Financials, Telecommunication, Technology, Consumer Non-Cyclicals, Industrials, Healthcare, Basic Materials, Consumer Cyclicals, Utilities.  $R_{t-1}$ , the lagged return variable, is incorporated to account for potential nonsynchronous trading effects.  $Temp_t$  is the daily average temperature at day  $t$ ,  $Prec_t$  the precipitation height in mm at day  $t$ , and  $\varepsilon_t$  is the error term at day  $t$ . The formula is consistent across all industries. The same set of variables is used for all industries.

The variable  $\sigma_{t,j}^2$  is the volatility of industry  $j$  at day  $t$  of the EGARCH (p,q) model. It is an estimate of the one period ahead variance including past information that is considered relevant. For this reason, it is called conditional variance. The formula for each single industry is given by:

$$\ln(\sigma_{t,j}^2) = \alpha_0 + \sum_{l=1}^p \gamma_l \ln \sigma_{t-l,j}^2 + \sum_{l=1}^q \alpha_l [ |z_{t-l,j}| - E |z_{t-l,j}| ] + \sum_{l=1}^q \xi_l z_{t-l,j} \tag{6}$$

The conditional variance consists of several parts: First,  $Z_{t-l} = \frac{\varepsilon_{t-l}}{\sigma_{t-l}}$  is the standardized residual at time  $t-l$  and  $l$  describes the number of lags.  $\gamma_l$  measures the persistence in conditional volatility irrespective of any activity in the market. Second,  $\alpha_l$  describes the GARCH effect, which represents the magnitude effect or the symmetry effect. It gives insights into the degree to which the size of shock influences the stock market response. Third, the parameter  $\xi_l$  measures the leverage effect, which takes into consideration the asymmetry, which is an advantage of the EGARCH model. For  $\xi_l < 0$ , positive shocks generate less volatility than negative shocks, which is evidence for the leverage effect. For asymmetric shocks, the variable  $\xi_l$  must be non-zero ( $\xi_l \neq 0$ ).

Returns are not normal distributions, as the JB test in section 4.3 will prove, and standardized residuals are leptokurtic. For heavier tails, the generalized error distribution (GED) is a more appropriate distribution than the t-distribution and will offer more reliable test results. The GED has the density function:

$$f(\mu_t, \sigma_t^2, \nu) = \frac{\nu}{2} \left( \Gamma\left(\frac{3}{\nu}\right) \right)^{\frac{1}{2}} \left( \Gamma\left(\frac{1}{\nu}\right) \right)^{-\frac{3}{2}} \left( \frac{1}{\sigma_t} \right) \exp \left\{ - \left[ \Gamma\left(\frac{3}{\nu}\right) / \Gamma\left(\frac{1}{\nu}\right) \right]^{\frac{\nu}{2}} |z_t|^\nu \right\} \tag{7}$$

with  $\mu_t$  as the conditional mean,  $\sigma_t^2$  as conditional variance, and  $\Gamma(\cdot)$  as the gamma function. The shape parameter  $\nu$  indicated the degrees of freedom under the GED distribution. The GED distribution can be identical with the normal distribution for  $\nu = 2$  and to the Laplace distribution for  $\nu = 1$ .

For  $1 = 2, \nu = 2$ , and  $\nu \rightarrow \infty$  degrees of freedom, the GED distribution converges to uniform density.

An estimation for the maximum of the log-likelihood function, which can also be considered a quasi-maximum likelihood function, is given as follows:

$$L(\theta) = \sum_{t=1}^T \log f(\mu_{t-1}, \sigma_{t-1}^2, \nu). \tag{8}$$

#### 4. Data set

The dataset used will be described in the following paragraphs. For my analysis, both stock market data and weather data were collected, and the rationale for the specific data selection will be explained.

##### 4.1. Stock market data

My investigation is based on companies of the Deutsche Aktienindex (DAX) 30, Mid-cap DAX (MDAX), and Small-cap DAX (SDAX). The TechDAX is not included, because recently, TechDAX companies were granted access to the MDAX and SDAX and are, thus, already incorporated in the given indices. The constitution of all three indices follows rules that are checked every quarter. Hence, for my analysis I use the companies which are contained in the first quarter of 2019. This is especially relevant for the MDAX and SDAX, which significantly change their composition regularly. A total of 121 companies is included, which is selected based on available stock market data. Data is retrieved from Datastream for a time horizon from 01 January 2009 to 28 February 2019 on a daily basis. This time frame was selected by reason of two arguments. First, it complies with time horizons used in extant research (Hirshleifer and Shumway, 2003; Shu, 2010) and second, it still comprises a relatively large set of companies, that is representative for the German business landscape. Moreover, a significant body of research has proven that the utilisation of the major index of a country serves as a representative dataset (Floros, 2008; Symeonidis et al., 2010). Therefore, the DAX, MDAX, and SDAX are suitable for an analysis of the German market.

Based on Datastream results, companies were sorted into nine industries, being Basic Materials, Consumer Cyclicals, Consumer Non-Cyclicals, Financials, Healthcare, Industrials, Telecommunication, Technology, and Utilities. The number of companies per sector is disparate which is explained through the historically shaped industrial landscape of Germany. It contains more companies in the sectors Industrials and Consumer Cyclicals than Telecommunication and Utilities as seen in Table 1. The appendix includes a list of companies in each sector for further reference.

To calculate the industry return indices, two option exists: First, the value-weighted average gives a better representation of the economy by putting more weight on companies with a larger number of outstanding shares and a higher market price per share than small companies. However, this



**Table 1:** Overview of the number of companies included per industry

Industry	Number of companies included
Basic Materials	14
Consumer Cyclical	19
Consumer Non-Cyclical	4
Financials	19
Healthcare	12
Industrials	30
Telecommunication	4
Technology	16
Utilities	3

method does not incorporate the quick growth of small enterprises and is exposed to the risk of overweighting a few incumbent companies. Second, an equally-valued approach, in turn, distributes market returns equally along with all included companies. This method benefits from higher exposure to small and medium-sized companies, which hold significant growth potential. A downturn is the missing realistic depiction of the stock-economy relation. In a considerable body of literature, both equally- and value-weighting methods have been used, but results for both cases were very similar (Cao and Wei, 2005; Kamstra et al., 2003; Saunders, 1993; Zheng et al., 2001). Zhu and Goetzmann (2003) used the equal weighted average only, because “average of the spread better reflects the average market maker’s behavior” (p. 12). There are both arguments for and against the equally- and value-weighted strategy. For the calculation of the industry return indices, the equally-weighted method has been used.

The compounded returns of the industry indices have been used to receive a stationary series of stock market returns. Returns for each industry have been calculated as the change in the natural logarithm of the company price for the two successive days:

$$R_{t,j} = \ln(P_{t,j}) - \ln(P_{t-1,j}) \quad (9)$$

$P_{t,j}$  is the adjusted price at day  $t$  of industry  $j$ , incorporating stock splits and other corporate actions, apart from dividends.  $P_{t-1,j}$  is the price of the previous day of industry  $j$ . This calculation is also in line with extant research (Floros, 2011; Sariannidis et al., 2016).

The data plots in Figure 2 to Figure 10 in the appendix depict the fluctuation of market returns for each of the nine industries. It can be noticed that returns are highly volatile and show evidence for volatility clustering.

#### 4.2. Weather data

Weather data was obtained from the Climate Data Center (CDC) of the DWD. The platform offers a comprehensive set of historical data for several weather variables. For my analysis, historical weather data of Frankfurt am Main is used. The weather station with the station ID 1420 is located close to Frankfurt Airport at an altitude of 100m (8.5213, 50.059).

Weather data has been collected on a daily basis for the variables: the average temperature at an altitude of 2m in °C and precipitation height in mm per square meter.

It is also necessary to address the topic of location choice. Frankfurt is one of the most important financial centres in Europe and the most important one in Germany. With Xetra and Börse Frankfurt, it inherits the largest of the seven German stock exchange. The dispersion of market participants in Germany is, compared with the US, much smaller. Consequently, Frankfurt is a good proxy for the behaviour of traders located in Munich, Dusseldorf, and Stuttgart as well (Krämer and Runde, 1997). As the weather conditions are relatively moderate all over Germany, the weather of Frankfurt is utilised as a representative for the weather in Germany. Research has moreover proven that only the largest exchange in a country determines prices, as stated by Krämer and Runde. As a result, changes in market prices in Frankfurt are representative of other exchanges. This fact also holds through the effect of arbitrage. Saunders (1993) states that traders on Wall Street might have a severe impact on security prices because they are concerned about market prices. As they work at the same location every day, a mood variable might influence traders more than geographically spread market participants. This opinion is shared by Zhu and Goetzmann (2003) who include, next to market makers, also news providers and other agents who are physically located in the exchange hosting city. The behaviour of those market participants is conjectured to be responsible for a relation between weather and return. As traders are based in Frankfurt, it appears plausible to utilise this location. For my analysis, I will not further investigate e-trading, as this process is done automatically and is probably not influenced by traders’ or investors’ mood.

#### 4.3. Descriptive statistics

Table 2 below presents Pearson correlation coefficients for the nine industry indices. One can observe a strong correlation between all industry returns, which are all significant at a 1% level. Correlation between industry returns implies an interconnectedness of industries. Industry indices seem to move in lockstep with each other. The correlation is highest for Basic Materials and Industrials with 0.8792\*\*\* and lowest for Consumer Non-Cyclicals and Telecommunication

**Table 2:** Correlation matrix for stock market returns

	Financials	Telecommu- nication	Technology	Consumer Non- Cyclicals	Industrials	Healthcare	Basic Materials	Consumer Cyclicals	Utilities
Financials	1								
Telecommunication	.6556***	1							
Technology	.7268***	.6188***	1						
Consumer Non-Cyclicals	.5985***	.5089***	.5459***	1					
Industrials	.8424***	.6822***	.8162***	.6380***	1				
Healthcare	.6666***	.6230***	.7060***	.5883***	.7429***	1			
Basic Materials	.7966***	.6309***	.7435***	.6002***	.8792***	.6768***	1		
Consumer Cyclicals	.7925***	.6551***	.7592***	.6053***	.8703***	.7054***	.8358***	1	
Utilities	.8312***	.6977***	.8090***	.6561***	.9181***	.8534***	.9374***	.9332***	1

Note: Asterisks indicate significance as follows: (\*)10%, (\*\*) 5%, (\*\*\*) 1%. In parentheses: Standard errors.

with 0.5089\*\*\*. Therefore, knowing that industry return's behaviour is quite similar, it will be interesting to investigate, whether the correlation with weather is identical for all industries.

Simple summary statistics of the market return and weather data are included in my analysis. Table 3 reports market returns. It provides insight into the nature of the distribution of returns that should be analysed before proceeding with further analysis. The first column shows the mean return, which varies across industries but is always slightly positive. Nevertheless, it is not possible to reject the null hypothesis that mean returns are not statistically different from zero. Standard deviation is relatively similar across all sectors. Basic Materials returns have the highest standard deviation and consequently the highest variance in returns. The return series of all industries are strongly skewed to negative returns and leptokurtic to the normal distribution. To evaluate the goodness-of-fit, a Jarque-Bera (JB) test is conducted, to analyse whether the skewness and kurtosis of the sample data match a normal distribution. Normally distributed data would have an expected kurtosis and skewness of zero. Deviation of a normal distribution will increase the JB value. In the selected sample, JB values range from 2,240.7704\*\*\* to 13,227.4538\*\*\*. The Augmented Dickey-Fuller (ADF) test confirms that the market return variables are stationary, and therefore the hypothesis that a unit root exists is rejected.

Table 4 displays statistics of the weather data. Temperature is recorded in °C and precipitation in mm per square meter. Therefore, they do not have high explanatory value when compared with each other. However, skewness is valuable for interpretation. Temperature has a negative skew, just like returns. Consequently, it might be interesting to test for correlation between those variables and market returns. Further, precipitation has a substantial skewness of 4.3147 which indicates a strongly asymmetrical distribution. Its kurtosis is also extreme with 30.5830. Therefore, precipitation exhibits tail data that exceeds the tails of a normal distribution. This is a signal that Frankfurt is sometimes exposed to intense rainfall which could potentially influence investor's

behaviour.

## 5. Hypothesis

Market returns are the result of a complex construct of influencing factors. Generally, the price is determined by supply and demand. However, this is just a theoretical approach, and also technical factors impact stock returns. Inflation, the economic strength of the market, and competitors influence investors' sentiments, attitude, and expectations. Further political and economic decisions can induce stock market activity. Moreover, the weather has often been argued to influence stock market returns. Therefore, keeping in mind that the trader's attitude can impact investment decisions, leads me to the hypotheses I want to test with the dataset and the methods discussed:

H0: Weather variables are not correlated with stock market returns. The investor is not significantly influenced by the weather.

H1: Weather variables are correlated with stock market returns. Investors are significantly influenced by weather.

The analysis of existing research implies a possible relationship between weather and stock market returns that would implicate to reject the null hypothesis. This will be discussed in the following section six "Empirical Results" in greater detail. Additionally, the purpose of this paper is not only to identify a potential relationship between stock market returns and weather, but also to investigate whether industries are affected differently by weather. This could result either in a difference in positive or a negative correlation with respect to each industry or in equally existent correlation.

H2: Industry market returns are equally influenced by weather.

H3: Industry market returns respond significantly different to weather.

## 6. Empirical results

The following paragraphs will investigate the implied hypotheses after the data and methodology used for my empir-

**Table 3:** Descriptive statistics of stock market returns

Statistics	Financials	Telecommuni- -cation	Technology	Consumer Non- Cyclicals	Industrials	Healthcare	Basic Materials	Consumer Cyclicals	Utilities
Mean	0.0005	0.0006	0.0008	0.0002	0.0005	0.0006	0.0002	0.0004	0.0004
Maximum	0.0728	0.0829	0.0725	0.0389	0.0631	0.0459	0.0803	0.0659	0.0520
Minimum	-0.0560	-0.1584	-0.0741	-0.0693	-0.0704	-0.0509	-0.0777	-0.0635	-0.0640
SD	0.0111	0.0142	0.0129	0.0100	0.0124	0.0105	0.0145	0.0118	0.0112
Skewness	-0.1965	-0.5350	-0.3526	-0.2723	-0.4000	-0.2827	-0.1475	-0.2639	-0.3041
Kurtosis	7.4156	10.8906	4.9359	5.2427	5.3707	4.4684	5.7647	5.8631	5.1275
Jarque-Bera	6,091.2391***	13,227.4538***	2,746.0117***	3,068.8088***	3,256.7562***	2,240.7704***	3,680.3264***	3,827.8807***	2,945.0084***
ADF	-47.9430***	-49.2850***	-48.6610***	-51.1070***	-46.9800***	-50.0000***	-48.8410***	-46.1710***	-47.9750***
Observations	2,651	2,651	2,651	2,651	2,651	2,651	2,651	2,651	2,651

Note: This table reports bin-test results for the equal-size sample. The sample covers the period 01.02.2009 to 28.02.2019 with a sample size of 2651 observations. All samples are in logarithmic first difference form and expressed as a percentage on a daily basis. Jarque-Bera test matches the skewness and kurtosis to the normal distribution. ADF is the Augmented Dicked-Fuller test for a unit root. Asterisks indicate significance as follows: (\*) 10%, (\*\*) 5%, (\*\*\*) 1%.

**Table 4:** Descriptive statistics of weather data

Statistics	Temperature	Precipitation
Mean	11.1436	1.5558
Maximum	29.7000	50.2000
Minimum	-11.0000	0.0000
SD	7.5970	3.6679
Skewness	-0.0587	4.3147
Kurtosis	2.2562	30.5830
Observations	2,651	2,651

ical tests has been described, and it will summarise the outcomes of my analysis. The bin-test is discussed first. Afterwards, the results of the EGARCH model are described with the aim to find a relationship between weather and industry stock market returns. This will be done by testing two sets of variables. My study ends with a discussion of the outcomes and a further outlook.

### 6.1. Bin-test

The bin-test was performed for each industry separately, and results are depicted in Table 5. Generally, one can observe a negative correlation between return and temperature for seven of the nine industries. Thus, low temperatures are typically associated with greater returns than high temperatures. Financials and Telecommunication, however, show a positive correlation between returns and temperature. The percentage of positive returns is always higher for low temperatures when there is a negative correlation between returns and temperature. The z-score for differences in return means is only significant at the 5% level for Consumer Non-Cyclicals. The percentage of positive returns does not have any significant z-scores.

The results in Table 5 conform with the literature, as a negative correlation is reported in the majority of the sectors. High temperatures seem to be associated with lower returns and low temperatures with higher returns. This finding can be explained with psychological research, indicating that high temperature can cause apathy leading to lower risk-taking. Consequently, apathy must strongly dominate aggres-

sion. On the opposite, aggression is dominant on days with low temperature leading to higher risk-taking which in turn leads to greater market returns. These conjectures are in line with the findings of [Cao and Wei \(2005\)](#).

The return behaviour in the Financials and Telecommunication sector raises the question for the underlying reason for this anomaly and could be explained with existing literature. Both high and low temperature can induce aggressive behaviour. Therefore, contradicting with the findings derived from the majority of industries, aggression might dominate apathy in certain occasions.

### 6.2. EGARCH

Table 6 and 7 report estimated parameters for equation (4) and (6) as outlined in section 3.2. Table 6 depicts estimates for the variance equations, having the underlying mean equation which is displayed in Table 7 with respective mean equations' estimates for each industry. The analysis has been conducted by using an EGARCH (1,1) model with GED. The reason for choosing this method is that the variance of stock-market returns varies over time and is dependent on past performance which has been further explained in section 3.2.

First, I will analyse the variance equation that offers insights into the distribution and behaviour of stock market returns. The persistence in conditional volatility, measured by  $\gamma_1$ , is highly significant and mostly has positive coefficient values close to one. Therefore, it is evident that GARCH effects are apparent in the German stock market and stock market

**Table 5:** Bin-test results

		Bin 1	Bin 2	Bin 3	Bin 4	z-score(4,1)
Financials	Return mean	0.00040	0.00053	0.00035	0.00062	0.20866
	% of pos. returns	0.53623	0.53800	0.52330	0.54912	0.26165
Telecommunication	Return mean	0.00034	0.00090	0.00040	0.00053	0.13921
	% of pos. returns	0.50000	0.52000	0.52061	0.50882	0.17844
Technology	Return mean	0.00124	0.00102	0.00076	0.00038	-0.68063
	% of pos. returns	0.52899	0.53100	0.53405	0.53652	0.15287
Consumer Non-Cyclicals	Return mean	0.00062	0.00018	0.00036	-0.00062	-1.26603
	% of pos. returns	0.55072	0.49700	0.50538	0.46851	-1.67111**
Industrials	Return mean	0.00071	0.00065	0.00035	0.00031	-0.34299
	% of pos. returns	0.57246	0.55100	0.52867	0.53401	-0.78498
Healthcare	Return mean	0.00073	0.00058	0.00073	0.00007	-0.71451
	% of pos. returns	0.55072	0.52100	0.54839	0.50630	-0.90264
Basic Materials	Return mean	0.00042	0.00027	0.00006	0.00023	-0.13717
	% of pos. returns	0.51449	0.48900	0.49552	0.51134	-0.06393
Consumer Cyclicals	Return mean	0.00021	0.00041	0.00064	-0.00014	-0.31193
	% of pos. returns	0.55797	0.51200	0.52419	0.51889	-0.79505
Utilities	Return mean	0.00045	0.00042	0.00048	0.00005	-0.38838
	% of pos. returns	0.56522	0.52700	0.53226	0.52393	-0.84120

Note: This table reports bin-test results for the equal-size sample. The sample covers the period 01.02.2009 to 28.02.2019 with a total of 2651 observations. The mean return and the percentage of positive returns are reported for each of the four bins. Z-scores are computed as follows:  $z\_score_{4,1}^{mean} = (\mu_4 - \mu_1) / \sqrt{\frac{\sigma_4^2}{n_4} + \frac{\sigma_1^2}{n_1}}$  and  $z\_score_{4,1}^{frequency} = (p_4 - p_1) / \sqrt{\frac{p_4(1-p_4)}{n_4} + \frac{p_1(1-p_1)}{n_1}}$ . Asterisks indicate significance as follows: (\*)10%, (\*\*) 5%, (\*\*\*) 1%.

**Table 6:** Estimates for the conditional variance

Coefficient	Financials	Telecommunication	Technology	Consumer			Basic Materials	Consumer Cyclicals	Utilities
				Non-Cyclicals	Industrials	Healthcare			
$\alpha_0$	-0.1907*** (0.0395)	-3.2453*** (0.6492)	-0.4332*** (0.0854)	-0.2907*** (0.0812)	-0.3863*** (0.0681)	-0.6388*** (0.1178)	-0.0926*** (0.0235)	-0.2059*** (0.0476)	-0.2836*** (0.0530)
$\alpha_1$	0.1541*** (0.0199)	0.3920*** (0.0728)	0.1612*** (0.0242)	0.1192*** (0.0228)	0.1781*** (0.0231)	0.1610*** (0.0276)	0.0969*** (0.0151)	0.1646*** (0.0214)	0.1580*** (0.0215)
$\xi_0$	-0.0937*** (0.0121)	-0.1055*** (0.0447)	-0.1004*** (0.0164)	-0.0511*** (0.0136)	-0.1343*** (0.0152)	-0.1291*** (0.0188)	-0.0898*** (0.0106)	-0.0801*** (0.0141)	-0.1204*** (0.0149)
$\gamma_1$	0.9792*** (0.0043)	0.6148*** (0.0775)	0.9506*** (0.0097)	0.9683*** (0.0088)	0.9570*** (0.0075)	0.9306*** (0.0127)	0.9892*** (0.0027)	0.9768*** (0.0052)	0.9690*** (0.0058)
$\nu$	1.4203*** (0.0551)	0.8780*** (0.0301)	1.4870*** (0.0583)	1.3938*** (0.0430)	1.4283*** (0.0591)	1.4531*** (0.0593)	1.5798*** (0.0674)	1.3506*** (0.0405)	1.4867*** (0.0622)
Log-likelihood	8629.7920	7778.1860	7986.2680	8602.1610	8250.7640	8528.4020	7836.6410	8370.9450	8491.2720
Adj. R <sup>2</sup>	0.0039	0.0011	0.0025	-0.0002	0.0077	0.0003	0.0019	0.0107	0.0042

Note: This table presents parameter coefficients for the EGARCH(1,1) specification in order to estimate the conditional variance. The model is based on the conditional variance equation (6):  $\ln(\sigma_t^2) = \alpha_0 + \sum_{l=1}^p \gamma_l \ln \sigma_{t-1}^2 + \sum_{l=1}^q \alpha_l [ |z_{t-1}| - E|z_{t-1}| ] + \sum_{l=1}^q \xi_l z_{t-1}$ . Asterisks indicate significance as follows: (\*)10%, (\*\*) 5%, (\*\*\*) 1%. In parentheses: Standard errors.

volatility is highly dependent on the previous period. Also, the magnitude effect  $\alpha_l$  is highly significant, implicating that the size of shock influences the stock market response. However, the values are relatively small, which translates into a short memory of variance.  $\xi_l$ , measures the leverage effect. Its values are negative and highly significant for all industries, implicating that an asymmetrical response to past residuals

is observed. Thus, higher volatility is perceived for negative shocks than for positive shocks of the same magnitude. This observation indicates that investors are more influenced by negative compared to positive news. This fact could be translated to weather news, and consequently, especially high and low temperatures should be tested.

Further, the endogenously estimated shape parameter  $\nu$

**Table 7:** EGARCH parameter estimates

Coefficient	Financials	Telecommu- -nication	Technology	Consumer Non- Cyclicals	Industrials	Healthcare	Basic Materials	Consumer Cyclicals	Utilities
b0	0.0010*** (0.0003)	-0.0004 (0.0003)	0.0011*** (0.0004)	0.0005 * (0.0003)	0.0012*** (0.0003)	0.0008*** (0.0003)	0.0007 ** (0.0003)	0.0008*** (0.0003)	0.0009*** (0.0003)
b1	0.0679*** (0.0191)	0.0377*** (0.0146)	0.0514*** (0.0195)	-0.0111 (0.0186)	0.0761*** (0.0196)	0.0046 (0.0194)	0.0517*** (0.0197)	0.0941*** (0.0190)	0.0555*** (0.0196)
b2	-0.0000** (0.0000)	0.0001*** (0.0000)	-0.0000 (0.0000)	-0.0000* (0.0000)	-0.0001** (0.0000)	-0.0000 (0.0000)	-0.0001** (0.0000)	-0.0000 (0.0000)	-0.0000* (0.0000)
b3	0.0000 (0.0000)	0.0001 (0.0000)	0.0000 (0.0001)	0.0000 (0.0000)	0.0000 (0.0001)	0.0000 (0.0000)	-0.0000 (0.0001)	-0.0000 (0.0000)	-0.0000 (0.0000)

Note: The estimated parameters rely on the model, eq.(4):  $R_{t,j} = b_0 + b_1\sigma_t^2 R_{t-1,j} + b_2\sigma_t^2 \text{Temp}_t + b_3\sigma_t^2 \text{Prec}_t + \xi_{t,j}$ . Parameters are estimated simultaneously with eq. (6):  $\ln(\sigma_{t,j}^2) = \alpha_0 + \sum_{i=1}^p \gamma_i \ln \sigma_{t-1,j}^2 + \sum_{i=1}^q \alpha_i [|z_{t-1,j}| - E|z_{t-1,j}|] + \sum_{i=1}^q \xi_i z_{t-1,j}$ . Asterisks indicate significance as follows: (\*) 10%, (\*\*) 5%, (\*\*\*) 1%. In parentheses: Standard errors.

shows that we can reject a normal or Laplace distribution for all industries except the Telecommunication sector. As  $\nu < 2$ , the tails are heavier than in a normal distribution. The distribution is leptokurtic and has positive excess kurtosis. Values are all significant at a 1% level, and range between 0.8780\*\*\* and 1.3506\*\*\*. This implies a distribution that is more similar to Laplace than to a normal distribution. The distribution of Telecommunication is smaller than zero, while all other industry indices return values are greater than zero. The rationale behind this is that the sample size forming the industry index is very small and object to flaws.

The adjusted  $R^2$  differs across industries. On average, the previous day's market return and weather data explain 0.36% of the market return. Weather explains only a minimal amount of market returns. In terms of magnitude and pattern, the adjusted  $R^2$  is in line with research by [Saunders \(1993\)](#), [Cao and Wei \(2005\)](#) and [Kamstra et al. \(2003\)](#). Explanatory power is greatest for Consumer Cyclicals with 1.07% and smallest for Consumer Non-Cyclicals with 0.00% displaying that this model has no predictive value for this particular industry.

In the following, I will present results for the mean equation and analyse whether market returns are correlated with the weather variables, temperature and precipitation and if the relationships are significant.

Overall, the model affirms the existence of a relationship between weather and stock market returns. The lagged market return value is highly significant and positive for all industries. This finding goes hand in hand with the EGARCH model choice, as it proves the conjecture that past returns influence today's returns.

Temperature is also significantly correlated with industry index returns when considered together with precipitation. The coefficient parameter b2 is mostly negative and highly significant. Values are all close to zero and range from 0.00005\*\*\* for Telecommunication to -0.00006\*\* for Basic Materials. The coefficients for Telecommunication and Financials are significant at the 1% level, the values for In-

dustrials and Basic materials at the 5% level, and the coefficient for Consumer Non-Cyclicals and Utilities at the 10% level. Consequently, traders are influenced by temperature to a different extent depending on the industry. The small parameter coefficients come as no surprise and are in line with existing research, which frequently reports very small coefficient values for weather variables ([Saunders, 1993](#)). The fact that temperature mostly has a negative effect on market returns is in agreement with previous findings ([Cao and Wei, 2005](#); [Floros, 2011, 2008](#); [Keef and Roush, 2002](#)). To illustrate the economic significance inherited in the temperature effect, I will include an example that shows how daily returns react to a temperature shock of one standard deviation. The standard deviation for temperature measured in Frankfurt is 7.5970°C; the coefficient of temperature for Industrials is -0.00005. Consequently, the return impact of one standard deviation of temperature is 0.0004.

The one standard deviation impact, thus, is three-quarters of the average daily return for the Industrials index.

Apart from temperature, precipitation is included in the model, which has a slightly positive, but no significant, effect on market returns. [Hirshleifer and Shumway \(2003\)](#) have also brought up evidence for the insignificance of rain on stock market returns.

In conclusion, according to this model, the null hypothesis, that weather variables and stock market returns are not correlated, can be rejected. My analysis provides evidence that traders are influenced by temperature and the past performance of stock prices. Further, the hypothesis that industry market returns are equally affected by weather is rejected as well. My outcomes reveal that significance in return depends on the industry.

### 6.3. Robustness test

My analysis has made use of an EGARCH model. As part of a thorough review, the robustness of the EGARCH model and regression must be examined. I will perform a robustness test for the EGARCH model, and for the regression analysis.

First, a likelihood ratio (LR) test was conducted to verify the relative explanatory power of the EGARCH model. The underlying equation for the LR test is as follows:

$$LR = -2 * (L_0 - L_1) \quad (10)$$

LR is based on a chi-squared distribution, with  $L_0$  being the likelihood function used under the null hypothesis and  $L_1$  corresponding to the alternative hypothesis. The test, which is also called deviance, measures the difference in the LR of the nested model to more sophisticated models. The GARCH model is considered the nested model,  $L_0$  and the alternative GARCH specifications are treated as  $L_1$ . The smaller the absolute difference, the better the model fit. Table 8 illustrates the log likelihood values of each model and Table 9 the LR ratio test. LR is smallest for the EGARCH model, supporting my model choice. Additionally, for all industries, apart from Telecommunications, the EGARCH model yields the highest log-likelihood values. Therefore, this model has the best fit and verifies the robustness of the model choice.

Secondly, I consider an alternative regression function for the EGARCH model which incorporates a calendar related anomaly to achieve a sufficient robustness test. I have chosen the Monday effect, also called the weekend effect, because it has been employed in a variety of research (Floros, 2008; Hirshleifer and Shumway, 2003). Studies have shown that the Friday closing price has an influence on the Monday opening price (Rogalski, 1984). Therefore, this effect has been added to the mean equation (4). The Monday effect is incorporated as a dummy variable  $D_{Mon}$ , being one for the day Monday and zero otherwise. It yields the regression equation:

$$R_{t,j} = b_0 + b_1 R_{t-1,j} + b_2 Temp_t + b_3 Prec_t + b_4 D_{Mon} + \xi_t \quad (11)$$

After controlling for this effect, the regression results should remain similar in order to achieve robustness in the utilised model. Indeed, temperature remains a significant variable after the dummy variable is added. Only Telecommunication does not report a significant correlation with weather anymore. Precipitation is still insignificant for all industries. Also, the day Monday, as the beginning of the trading week, does not influence stock market returns. The insignificance of the Monday effect contradicts the majority of existing research. However, it is in line with the findings of Krämer and Runde (1997), who analysed the German DAX returns and did not find any significant effect for specific days or the week.

In summary, the regression robustness test, my model is robust against the Monday effect. The temperature correlation coefficient remains negative for all industries and significant for the majority of sectors. Consequently, the EGARCH model has high informative value.

## 7. Discussion

From the empirical results, which have been discussed in detail, it can be concluded that temperature impacts stock

market returns of particular industries. Observations from the underlying dataset identified Financials, Consumer Non-Cyclicals, Industrials, Basic Materials, and Utilities as industries that have a significant negative correlation between temperature and stock market performance. This chapter will elaborate on possible reasons for the observed relationship in the industries mentioned above and focuses on the analysis of economic impacts. It has been extensively discussed in section 2 that weather impacts mood, decision-making, and consequently investment decisions and risk-taking behaviour of traders. In addition to this effect, weather influences companies' economic performance. This effect is partially factored into the market price, but the extent to which investors are subconsciously influenced by weather effects is unknown. Therefore, I conjecture that both effects are subliminally integrated into the behaviour of traders and are therefore relevant aspects in the discussion.

Hot and cold temperature can influence several industrial areas. This paragraph will outline economic reasons, why low temperature leads to higher market returns and vice versa. The Industrials industry consists of companies in the logistics area, machinery, aviation, construction, and many more. Focussing on aviation, companies as Lufthansa are affected in their daily business by high temperatures. The take-off performance is negatively affected, because the air density decreases, the aeroplane will climb slower, leading to delays and requiring more fuel. The same problem exists for many logistics companies. In addition, extremely hot weather induces physical activity to be especially strenuous and is harmful to workers' health constitution and performance.

The Basic Material sector is very energy intensive. Especially in summer, machines need to be cooled down, which increases energy consumption and consequent energy costs.

In the Utilities industry, the primary energy sources are predominantly used for heating. Thus, low temperatures can be favourable because demand rises which in turn increases revenue.

For Consumer Non-Cyclicals, the interpretation is not as intuitive as for the other sectors. My sample includes industries active in the food producing sector (Suedzucker), growers and seed producers (KWS), personal care (Beiersdorf), and the Baywa group, which is active in agriculture, energy, and building materials. This sample, thus, still includes companies engaged in different fields. Therefore, my interpretation concentrates on the agricultural companies Baywa and KWS. The former benefits from the advantages of low temperatures for the same reasons as Utilities companies. This could outweigh the effect that increasing fruit and vegetable consumption is expected for higher temperatures. The latter might be influenced by hot spells implicating declining seed sales volume. The economic performance of Suedzucker might depend on the growth of the sugar beet. This argumentation is not exhaustive, and no reasonable explanations can be found for temperature implications on the economics of Beiersdorf.

Turning to the Financial industry, my conjectures are

**Table 8:** Loglikelihood values

Statistics	Financials	Telecommu- nication	Technology	Consumer Non- Cyclicals	Industrials	Healthcare	Basic Materials	Consumer Cyclicals	Utilities
GJR	8463.963***	7764.910***	7904.216***	8562.111***	8092.710***	8457.401***	7647.490***	8211.179***	8326.612***
PARCH	8456.612***	7759.040***	7888.662***	8560.148***	8080.146***	8448.126***	7642.939***	8208.130***	8318.698***
GARCH	8607.276***	7836.726***	7972.032***	8588.664***	8216.340***	8506.574***	7805.412***	8355.436***	8458.939***
NARCH	8467.850***	7764.164***	7905.959***	8561.830***	8096.731***	8456.661***	7650.093***	8213.454***	8328.849***
EGARCH	8629.792***	7778.186***	7986.268***	8602.161***	8250.764***	8528.402***	7836.641***	8370.945***	8491.272***

Note: This table reports the log-likelihood values with respect to the selected GARCH model: GJR GARCH, Power ARCH (PARCH), Nonparametric ARCH (NARCH), EGARCH. In highlighted numbers indicate the greatest log-likelihood value per industry. Asterisks indicate significance as follows (\*) 10%, (\*\*) 5%, (\*\*\*) 1%.

**Table 9:** Likelihood ratio statistics

Statistics	Financials	Telecommu- nication	Technology	Consumer Non- Cyclicals	Industrials	Healthcare	Basic Materials	Consumer Cyclicals	Utilities
GJR	-286.626***	-143.632***	-135.632***	-53.106***	-247.260***	-98.346***	-315.844***	-288.514***	-264.654***
PARCH	-301.328***	-155.372***	-166.740***	-57.032***	-272.388***	-116.896***	-324.946***	-294.612***	-280.482***
NARCH	-278.852***	-145.124***	-132.146***	-53.668***	-239.218***	-99.826***	-310.638***	-283.964***	-260.180***
EGARCH	45.032***	-117.080***	28.472***	26.994***	68.848***	43.656***	62.458***	31.018***	64.666***

Note: This table illustrates the LR statistic in order to test the significance of the models (GJR, PARCH, NARCH, and EGARCH) to the nested model GARCH. The LR statistic is calculated as follows:  $LR = -2(L_0 - L_1)$ . Asterisks indicate significance as follows: (\*) 10%, (\*\*) 5%, (\*\*\*) 1%.

**Table 10:** EGARCH robustness test including seasonality related anomaly variable

Coefficient	Financials	Telecommu- nication	Technology	Consumer Non- Cyclicals	Industrials	Healthcare	Basic Materials	Consumer Cyclicals	Utilities
b0	0.0010 *** (0.0003)	0.0005 *** (0.0003)	0.0010 *** (0.0004)	0.0004 (0.0003)	0.0012 *** (0.0003)	0.0040 ** (0.0194)	0.0006 * (0.0003)	0.0008 *** (0.0003)	0.0002 *** (0.0003)
b1	0.0679 *** (0.0191)	0.0520 (0.0176)	0.0512 *** (0.0195)	-0.0125 (0.0186)	0.0762 *** (0.0195)	-0.0000 (0.0000)	0.0517 *** (0.0197)	0.0941 *** (0.0190)	0.0550 *** (0.0196)
b2	-0.0000** (0.0000)	-0.0000 (0.0000)	-0.0000 (0.0000)	-0.0000* (0.0000)	-0.0001** (0.0000)	0.0000 (0.0000)	-0.0001** (0.0000)	-0.0000 (0.0000)	-0.0000* (0.0000)
b3	0.0000 (0.0000)	0.0000 (0.0001)	0.0000 (0.0001)	0.0000 (0.0000)	0.0000 (0.0001)	0.0006 (0.0004)	-0.0000 (0.0001)	-0.0000 (0.0000)	-0.0000 (0.0000)
b4	0.0001 (0.0004)	0.0007 (0.0005)	0.0004 (0.0005)	0.0006 (0.0004)	0.0001 (0.0004)	0.0006 (0.0003)	0.0002 (0.0005)	0.0000 (0.0004)	0.0002 (0.0004)

Note: The estimated parameters rely on the model, eq. (11):  $R_{t,j} = b_0 + b_1 R_{t-1,j} + b_2 Temp_t + b_3 Prec_t + b_4 D_{Mon} + \xi_{t,j}$ . Parameters are estimated simultaneously with eq. (6):  $\ln(\sigma_{t,j}^2) = \alpha_0 + \sum_{l=1}^p \gamma_l \ln \sigma_{t-l,j}^2 + \sum_{l=1}^q \alpha_l [|z_{t-l,j}| - E|z_{t-l,j}|] + \sum_{l=1}^q \xi_l z_{t-l,j}$ . Asterisks indicate significance as follows: (\*) 10%, (\*\*) 5%, (\*\*\*) 1%. In parentheses: Standard errors.

mainly based on the findings of Subak et al. (2000). By analysing the implications of the anomalously hot weather during summer in 1995 in the UK, they find that the recorded insurance claims for property damage through subsidence were much higher than for bursting pipes. Soil subsidence is a natural cause of lasting drought and heat. Furthermore, outdoor fires are sometimes caused by heatwaves, increasing the number of insurance cases. Consequently, high-temperature periods constitute a less economically profitable

scenario for companies in this sector, than more moderate weather circumstances.

It has been previously reported in the literature that climate changes, in particular temperature, can impact the economic dimension in the Telecommunications sector (Adams and Steeves, 2014). Specifically, high temperatures lead to more need for I equipment cooling in base stations or exchanges, which can result in an escalating failure rate. Also, malfunction and reduced life span of network-and telecom-

munication equipment are possible. Energy demand increases as a result of intensified cooling activities and might lead to a rise in energy prices. These are, among others, reasons explaining why high temperatures are harmful to companies in the telecommunication sector.

Taken together, industry performance is impacted by climate conditions. However, the argumentation outlining why low temperatures are favourable for business performance and high temperatures harm the industrial segment is one-sided, and it must be noted that the effects of climate can also be opposing. In other words, there are numerous unfavourable implications of cold weather.

Another relevant aspect of this topic is the psychological implication for investor behaviour. As a result of the economic impact I conjecture that these complex correlations between temperature and stock markets are triggered intuitively in market makers. Thus, on days with hotter temperatures traders subconsciously incorporate the environmental circumstances into their evaluation of the respective industries. In addition, they also consider these factors in their investment decisions.

## 8. Conclusion

My findings are essential to understand the implications of weather on stock market returns. The gained knowledge could be insightful for developing weather-based trading strategies. To derive a benefit from this strategy, transaction costs need to be very low because the correlation coefficient of stock market returns is minimal and consequently, economic gains are marginal (Hirshleifer and Shumway, 2003). Another finding of this paper is that it raises the general awareness that people are likely to be influenced by climate factors. If traders, or investors, know that their investment decisions might be impacted by the outside temperature, ill-conceived investments could be prevented.

### 8.1. Outlook and limitations

Additionally, seven caveats must be addressed that concern potential flaws in methodology and data choice. First, the investor location must be addressed. By choosing Frankfurt, I can depict the behaviour of traders. This does not include independent market makers or individual investors who are widely distributed in Germany and over the whole world (Sariannidis et al., 2016). If order-submitting investors are price setters, and investors' mood is affected by weather, Frankfurt is not a good proxy and can potentially deteriorate results. This limitation could be addressed by calculating a population-weighted temperature based on the greatest German cities (Cao and Wei, 2002), but as investors are not limited to Germany, this is not a sound solution.

Second, traders usually work in office buildings which are furnished with air cooling as well as heating systems. Indoor temperature levels are retained at a medium temperature of approximately 21°C. If the temperature does correlate with

human behaviour and market returns, this occurrence is confined to the extent to which traders are exposed to the outside temperature, e.g., during their commute. Accordingly, it is "the psychological imprint of the extreme temperature that mediates people's mood and influences their behaviour" (Cao and Wei, 2002, p. 1562). Nevertheless, it remains questionable whether the exposure to the outside environment is significant enough to utilise temperature as a regression variable.

Third, other economy-related variables could be included in the regression model because it is evident, that weather, if at all, only explains a vanishing small portion of market return behaviour. Especially, variables related to global environmental emissions might affect stock market returns (Sariannidis et al., 2016).

Fourth, weather and the economy are dynamic, and climate anomalies are never alike. Therefore, it is not reasonable to draw implications out of a climate anomaly to a future period.

Fifth, the researchers Krämer and Runde (1997) claim that findings of Saunders could also be a result of a type one error due to statistical inference. This flaw could be limited by collecting data from a longer period of time than ten years. Further, they claim that the significance of a correlation between the variables weather and market returns is subject to data mining. Depending on the definition of good and bad weather, relationships can either be identified or dismissed. Zheng et al. (2001) also share this view. They claim that by studying historical stock market returns one can find significant relationships just by chance.

Sixth, existing research has mostly focused on global or national stock market indices that do not differentiate between industries. As my analysis has proven, correlations do only exist for specific industries. Thus, by combining all sectors in an index, the weather effect is diluted because significant and insignificant industries are combined. This fact could explain why researchers obtain diverging results and would challenge the outcome correctness of extant work.

Seventh, no psychological literature exists that directly analyses the impact of weather on investment behaviour. This lack of research has already been pointed out by Cao and Wei (2002). The conjunction is generally made by financial researchers that extrapolate from feelings and mood to investment activities. Consequently, it is questionable whether a direct link between weather and market returns exists. One needs to strictly differentiate between correlation and causation. My paper could only prove correlation which does not guarantee any causal relations. This limitation is based on potential endogeneity problems within the model implicating that an independent variable is potentially correlated with the error term. Three sources of endogeneity will be briefly discussed.

It is difficult to avoid omitted variables. Uncontrolled variables can be correlated with both independent and dependent variables or independent variables and the error term. The confounding variables lead to inconsistent estimates and biases. I cannot rule out the existence of omitted



variables from my model. Another source of endogeneity is the simultaneity bias which means that the explanatory and dependent variable are jointly determined. This bias is unlikely as weather is not caused by stock market but by meteorological dynamics. Additionally, measurement errors of the dependent variable can result in inconsistent parameter estimates. Market returns can be calculated based on various methods. The prices and returns chosen in this model might be prone to errors. Instrumental variables, among other techniques, can be utilized to address the problem of endogeneity. However, completely exogenous instruments are very difficult to find.

All limitations suggest that more research is needed in this field, in particular, with respect to the analysis of weather implications on industry returns. Especially, this field should be evaluated on a global level as has been done in research that investigated the relationship of a variety of global indices with weather. Besides, further studies should incorporate a finer depiction of weather data and incorporate individual and institutional investors into the model. Hitherto, only weather and seasonal variables have been included in research. The literature is outdated to a great extent. On the one hand, one could argue, that volatility in weather does not change. On the other hand, global warming is increasingly getting public attention. Hence, it would be interesting to connect market returns not only with weather but include a climate change variable that depicts profound variations in weather trends and influences mood to a yet unknown extent. This procedure is also valuable, considering that climate change has already had major implications on industry standards and policies, as the automotive industry has illustrated (Williams et al., 2018).

## 8.2. Final remarks

In the psychological literature, it is well established that weather can influence people's mood and feelings. Mood, in particular, impacts people's decision-making behaviour. From a financial point of view, mood also impacts investment-decisions and risk assessment. The empirical literature has identified temperature as one of the most critical weather variables that affect human behaviour and mood: Low temperature can cause aggression, high temperature might induce apathy. These mood alternations can lead to either higher or lower risk attitude. A body of empirical literature has revealed a relationship between weather and market returns. Examples are sunshine (Saunders, 1993), temperature (Cao and Wei, 2005), length of the night (Kamstra et al., 2003), and many more.

It is examined if weather alternates mood and investment behaviour. This is done by analysing the relationship between market returns and the weather variables temperature and precipitation. For this, a Bin-test and an EGARCH model are applied. First, I hypothesise that weather variables are correlated with stock market returns. Market returns include the German indices DAX 30, MDAX, and SDAX and are separated by industries. I consider the second hypothesis: Industry returns respond significantly differently to the weather.

After the examination of the industry indices, I can provide evidence that traders are influenced by temperature and the past performance of stock prices. The correlation between temperature and stock market returns is negative and significant for six out of nine industries. Lagged returns are all significant and mostly positively correlated with stock market returns. Precipitation, on the other hand, is not significantly correlated with stock market returns. The significance in return depends on the industry which supports the hypothesis that industry market returns do not respond equally to weather effects. The industries Technology, Healthcare, and Consumer Cyclical do not show a significant correlation. However, for the other six industries, a significant negative correlation between temperature and stock market returns was identified. This relationship is robust to changes in the underlying model and seasonal anomalies as the Monday effect.

All in all, temperature has a negative relationship with some industries. Future research should elaborate on the underlying reasons, why particular industries' market returns are affected by weather variables. Furthermore, studies on an industry-wide level should be extended to a global scale.

## References

- Adams, P. and Steeves, J. Climate risks study for telecommunications and data center services Risks Study for Telecommunications and Data Center Service: Report prepared for the general service administration, 2014.
- Bless, H., Schwarz, N., and Kimmelmeier, M. Mood and Stereotyping: Affective States and the Use of General Knowledge Structures. *European Review of Social Psychology*, 7(1):63–93, 1996. ISSN 1046-3283. doi: 10.1080/14792779443000102.
- Cao, M. and Wei, J. Stock market returns: A temperature anomaly: Working Paper. 2002.
- Cao, M. and Wei, J. Stock market returns: A note on temperature anomaly. *Journal of Banking & Finance*, 29(6):1559–1573, 2005. ISSN 03784266. doi: 10.1016/j.jbankfin.2004.06.028.
- Chang, S., Chen, S., Chou, R. K., and Lin, Y. Weather and intraday patterns in stock returns and trading activity. *Journal of Banking & Finance*, 32(9):1754–1766, 2008. ISSN 03784266. doi: 10.1016/j.jbankfin.2007.12.007.
- Chang, T., Nieh, C.-C., Yang, M. J., and Yang, T.-Y. Are stock market returns related to the weather effects? Empirical evidence from Taiwan. *Physica A: Statistical Mechanics and its Applications*, 364:343–354, 2006. ISSN 03784371. doi: 10.1016/j.physa.2005.09.040.
- Cunningham, M. R. Weather, mood, and helping behavior: Quasi experiments with the sunshine samaritan. *Journal of Personality and Social Psychology*, 37(11):1947–1956, 1979. ISSN 0022-3514. doi: 10.1037/0022-3514.37.11.1947.
- Dichev, I. D. and Janes, T. D. Lunar cycle effects in stock returns. *The Journal of Private Equity*, 6(4), 2001. URL [SSRN281665](https://ssrn.com/abstract=281665).
- Dowling, M. and Lucey, B. M. Weather, biorhythms, beliefs and stock returns—Some preliminary Irish evidence. *International Review of Financial Analysis*, 14(3):337–355, 2005. ISSN 10575219. doi: 10.1016/j.irfa.2004.10.003.
- Eagles, J. M. The relationship between mood and daily hours of sunlight in rapid cycling bipolar illness. *Biological Psychiatry*, 36(6):422–424, 1994.
- Floros, C. Stock market returns and the temperature effect: new evidence from Europe. *Applied Financial Economics Letters*, 4(6):461–467, 2008. ISSN 1744-6546. doi: 10.1080/17446540801998585.
- Floros, C. On the relationship between weather and stock market returns. *Studies in Economics and Finance*, 28(1):5–13, 2011. ISSN 1086-7376. doi: 10.1108/108673711111110525.
- Hirshleifer, D. and Shumway, T. Good Day Sunshine: Stock Returns and the Weather. *The Journal of Finance*, 58(3):1009–1032, 2003.
- Howarth, E. and Hoffman, M. S. A multidimensional approach to the relationship between mood and weather. *British Journal of Psychology*, 75(1):15–23, 1984. ISSN 00071269. doi: 10.1111/j.2044-8295.1984.tb02785.x.
- Imbery, F., Friedrich, K., Koppe, C., Janssen, W., Pfeifroth, U., Daßler, J., and Bissolli, P. 2018 wärmster Sommer im Norden und Osten Deutschlands, 2018.
- Isen, A. M. Missing in Action in the AIM: Positive Affect's Facilitation of Cognitive Flexibility, Innovation, and Problem Solving. *Psychological Inquiry*, 13(1):57–65, 2002.
- Kals, W. S. *Your health, your moods, and the weather*. New York, N.Y., Doubleday Books, New York, 1 edition, 1982.
- Kamstra, M. J., Kramer, L. A., and Levi, M. D. Losing Sleep at the Market: The Daylight Saving Anomaly. *American Economic Review*, 90(4):1105–1111, 2000.
- Kamstra, M. J., Kramer, L. A., and Levi, M. D. Winter Blues: A SAD Stock Market Cycle. *American Economic Review*, 93(1):324–343, 2003.
- Kang, S. H., Jiang, Z., Lee, Y., and Yoon, S.-M. Weather effects on the returns and volatility of the Shanghai stock market. *Physica A: Statistical Mechanics and its Applications*, 389(1):91–99, 2010. ISSN 03784371. doi: 10.1016/j.physa.2009.09.010.
- Keef, S. P. and Roush, M. L. The Weather and Stock Returns in New Zealand. *Quarterly Journal of Business and Economics*, 41(1/2):61–79, 2002.
- Keef, S. P. and Roush, M. L. Influence of weather on New Zealand financial securities. *Accounting and Finance*, 45(3):415–437, 2005. ISSN 0810-5391. doi: 10.1111/j.1467-629x.2004.00135.x.
- Keef, S. P. and Roush, M. L. Daily weather effects on the returns of Australian stock indices. *Applied Financial Economics*, 17(3):173–184, 2007. ISSN 0960-3107. doi: 10.1080/09603100600592745.
- Krämer, W. and Runde, R. Stocks and the weather: An exercise in data mining or yet another capital market anomaly? *Empirical Economics*, 22(4):637–641, 1997.
- Loughran, T. and Schultz, P. Weather, Stock Returns, and the Impact of Localized Trading Behavior. *The Journal of Financial and Quantitative Analysis*, 39(2):343–364, 2004.
- Lu, J. and Chou, R. K. Does the weather have impacts on returns and trading activities in order-driven stock markets? Evidence from China. *Journal of Empirical Finance*, 19(1):79–93, 2012. doi: 10.1016/j.jempfin.2011.10.001.
- Mackie, D. M. and Worth, L. T., editors. *Feeling good, but not thinking straight: The impact of positive mood on persuasion*. Emotion and social judgment. Amsterdam, Pergamon/Elsevier, Oxford, 1991.
- Nelson, D. B. Conditional Heteroskedasticity in Asset Returns: a new approach. *Econometrica - The Econometric Society*, 59(2):347–370, 1991.
- Pardo, A. and Valor, E. Spanish Stock Returns: Where is the Weather Effect? *European Financial Management*, 9(1):117–126, 2003.
- Parsons, A. G. The Association Between Daily Weather and Daily Shopping Patterns. *Australasian Marketing Journal (AMJ)*, 9(2):78–84, 2001. ISSN 14413582. doi: 10.1016/S1441-3582(01)70177-2.
- Pilcher, J. J., Nadler, E., and Busch, C. Effects of hot and cold temperature exposure on performance: a meta-analytic review. *Ergonomics*, 45(10):682–698, 2002. ISSN 0014-0139. doi: 10.1080/00140130210158419.
- Rind, B. Effect of Beliefs About Weather Conditions on Tipping. *Journal of Applied Social Psychology*, 26(2):137–147, 1996. ISSN 0021-9029. doi: 10.1111/j.1559-1816.1996.tb01842.x.
- Rogalski, R. New findings regarding day of the week returns over trading and non-trading periods. *Journal of Finance*, 39:1603–1614, 1984.
- Rotton, J. and Rosenberg, M. Lunar cycles and the stock market: Time-series analysis for environmental psychologists. 1984.
- Sariannidis, N., Giannarakis, G., and Partalidou, X. The effect of weather on the European stock market. *International Journal of Social Economics*, 43(9):943–958, 2016. doi: 10.1108/IJSE-04-2015-0079.
- Saunders, E. D. Stock Prices and Wall Street Weather. *The American Economic Review*, 83(5):1337–1345, 1993.
- Schwarz, N. Feelings as information: Informational and motivational functions of affective states. *Handbook of motivation and cognition: Foundations of social behavior*, 2, 1990.
- Shu, H.-C. Investor mood and financial markets. *Journal of Economic Behavior & Organization*, 76(2):267–282, 2010. ISSN 01672681. doi: 10.1016/j.jebo.2010.06.004.
- Shu, H.-C. and Hung, M.-W. Effect of wind on stock market returns: evidence from European markets. *Applied Financial Economics*, 19(11):893–904, 2009. ISSN 0960-3107. doi: 10.1080/09603100802243766.
- Subak, S., Palutikof, J. P., d. Agnew, M., Watson, S. J., Bentham, C. G., Cannel, M. G. R., Hulme, M., McNally, S., Thornes, J. E., Waughray, and Woods, J. C. The Impact of the Anomalous Weather of 1995 on the U.K. Economy. *Climatic Change*, 44(1/2):1–26, 2000. ISSN 01650009. doi: 10.1023/A:1005566710285.
- Symeonidis, L., Daskalakis, G., and Markellos, R. N. Does the weather affect stock market volatility? *Finance Research Letters*, 7(4):214–223, 2010. ISSN 15446123. doi: 10.1016/j.frl.2010.05.004.
- Tietjen, G. H. and Kripke, D. F. Suicides in California (1968–1977): Absence of seasonality in Los Angeles and Sacramento counties. *Psychiatry Research*, 53(2):161–172, 1994. ISSN 01651781. doi: 10.1016/0165-1781(94)90107-4.
- Trombley, M. A. Stock Prices and Wall Street Weather: Additional Evidence. *Quarterly Journal of Business and Economics*, 36(3):11–21, 1997. URL <https://www.jstor.org/stable/pdf/40473319.pdf?refreqid=excelsior%3A8adc329de54eefe6f55c56792d31c163>.
- Watson, D. *Emotions and social behavior. Mood and temperament*. New York, N.Y., Guilford Press, New York, 2000.
- Williams, J., Ferguson, M., Petkov, M., and Wilkins, M. The effect of weather events on corporate warnings are gathering force, 2018.
- Wright, W. F. and Bower, G. H. Mood effects on subjective probability assessment. *Organizational behavior and human decision processes*, 52(2):276–291, 1992.
- Yoon, S.-M. and Kang, S. H. Weather effects on returns: Evidence from the Korean stock market. *Physica A: Statistical Mechanics and its Applications*, 388(5):682–690, 2009. ISSN 03784371. doi: 10.1016/j.physa.2008.11.017.
- Zheng, L., Yuan, K., and Zhu, Q. Are Investors Moonstruck? - Lunar Phases

and Stock Returns. *Journal of Empirical Finance*, 13(1):1–23, 2001. doi:  
10.2139/ssrn.283156.

Zhu, N. and Goetzmann, W. N. Rain or Shine: Where is the Weather Effect?  
*European Financial Management*, 11(5):559–578, 2003.



## Culture as a Determinant of Individuals' Risk-Taking Behavior – Empirical Evidence from the PGA TOUR

### Kultur als Determinante des Risikoverhaltens von Individuen – Empirische Evidenz von der PGA TOUR

Andreas Johannes Dambaur

Universität Zürich

#### Abstract

This work examines the relationship between individuals' cultural background and their risk-taking behavior in the context of the exercise of their profession. To this end, the risk-taking behavior of professional golfers playing on the PGA TOUR is analyzed using hybrid regression models. The risk-taking behavior is, on the one hand, measured by the decision of golfers to go for the green and, on the other hand, by the standard deviation of players' scores relative to par on the 18 holes of a round. The players' cultural background is quantified using the national cultural dimensions according to Hofstede. In addition, it is controlled for various socio-demographic risk determinants, as well as for risk determinants that play a role in tournament situations. The results show that golfers' cultural background has no robust, significant effect on their risk-taking behavior. At the same time, however, the results confirm the relevance of other risk determinants known in the literature.

#### Zusammenfassung

Diese Arbeit untersucht den Zusammenhang zwischen dem kulturellen Hintergrund eines Individuums und dessen Risikoverhalten im Rahmen der Ausübung seines Berufs. Hierfür wird das Risikoverhalten von professionellen Golfspielern, die auf der PGA TOUR spielen, mithilfe von hybriden Regressionsmodellen analysiert. Das Risikoverhalten wird dabei einerseits anhand der Entscheidung eines Golfers, das Grün anzugreifen, sowie andererseits anhand der Standardabweichung der Scores eines Spielers relativ zu Par auf den 18 Löchern einer Runde gemessen. Der kulturelle Hintergrund eines Spielers wird mithilfe der nationalen Kulturdimensionen nach Hofstede quantifiziert. Darüber hinaus wird für verschiedene sozio-demographische Risikodeterminanten sowie für Risikodeterminanten, die in Turniersituationen eine Rolle spielen, kontrolliert. Die Ergebnisse zeigen, dass der kulturelle Hintergrund eines Golfers keinen robusten, signifikanten Effekt auf dessen Risikoverhalten hat. Gleichzeitig bestätigen die Resultate jedoch die Relevanz von anderen in der Literatur bekannten Risikodeterminanten.

**Keywords:** risk-taking; cultural background; sports economics; behavioral economics.

#### 1. Einleitung

Studien, die den Einfluss kultureller Unterschiede auf ökonomische Entscheidungen von Individuen untersuchen, erfreuen sich in letzter Zeit zunehmender Beliebtheit. So untersuchen [Bedendo et al. \(2019\)](#) beispielsweise den Einfluss des kulturellen Hintergrundes eines Managers auf die Finanzierungsstruktur von kleinen Firmen. [Lee \(2019\)](#) analysiert

die Rolle von Kultur bei Werbung. [Rieger et al. \(2015\)](#) und [Wang et al. \(2017\)](#) beobachten kulturelle Unterschiede in Entscheidungen unter Risiko. Auch [Oosterbeek et al. \(2004\)](#), [Guiso et al. \(2006\)](#) und [Herrmann et al. \(2008\)](#) analysieren den Einfluss von Kultur auf verschiedene ökonomische Entscheidungen.

Dabei legen diverse Studien nahe, dass das Risikoverhalten eines Individuums von dessen kulturellem Hintergrund

abhängt (Brumagim und Xianhua, 2005; Cheung et al., 2013; Eriksson und Simpson, 2010; Hsee und Weber, 1999; Lau und Ranyard, 2005; Marshall et al., 2011; Nobre et al., 2016; Rieger et al., 2015; Schneider et al., 2017; Statman, 2008; Statman und Weng, 2010; Statman, 2010, Statman, 2015; Tavor und Garyn-Tal, 2016; Vieider et al., 2015a; Wang und Fischbeck, 2008; Wang et al., 2017; Weber und Hsee, 1998; Weber, 2014). Die bisherige Forschung hierüber basiert weitestgehend auf Experimenten bzw. Umfragen und beschränkt sich zudem auf finanzielle Risiken (Brumagim und Xianhua, 2005; Cheung et al., 2013; Eriksson und Simpson, 2010; Hsee und Weber, 1999; Lau und Ranyard, 2005; Marshall et al., 2011; Nobre et al., 2016; Rieger et al., 2015; Schneider et al., 2017; Statman, 2008; Statman und Weng, 2010; Statman, 2010, Statman, 2015; Tavor und Garyn-Tal, 2016; Vieider et al., 2015a; Wang und Fischbeck, 2008; Wang et al., 2017; Weber und Hsee, 1998; Weber, 2014).

Ein Teilbereich der Literatur, die sich mit Entscheidungen unter Risiko beschäftigt, zeigt jedoch, dass das Risikoverhalten von Individuen kontext- bzw. domänenspezifisch ist (Barseghyan et al., 2011; Breuer et al., 2016; Chetty und Szeidl, 2007; Einav et al., 2012; Hanoch et al., 2006; Rolison et al., 2014; Slovic, 1972; Tavor und Garyn-Tal, 2016; Tyagi et al., 2017; Weber et al., 2002; Wolf und Pohlman, 1983). So ist es möglich, dass Individuen, die eine moderate Risikobereitschaft bei finanziellen Risiken zeigen, eine hohe Risikobereitschaft bei der Wahl ihrer Freizeitaktivitäten haben und beispielsweise Bungee springen (Hanoch et al., 2006).

Im Allgemeinen ist das Risikoverhalten von Individuen eines der Schlüsselemente ökonomischer Entscheidungen. Das Risikoverhalten im Kontext ihrer Berufsausübung hat zudem ökonomisch besonders weitreichende Konsequenzen. In dieser Arbeit wird deshalb untersucht, ob der kulturelle Hintergrund eines Individuums dessen Risikoverhalten bei der Ausübung seines Berufs beeinflusst. Eine Betrachtung von professionellen Golfspielern, die auf der PGA TOUR spielen, ist hierfür ideal. Diese Vorgehensweise erlaubt es nicht nur, das Risikoverhalten in einer anderen Domäne als der in anderer Literatur zumeist untersuchten Finanzdomäne zu analysieren. Im Gegensatz zu oftmals hypothetischen Umfragen oder Experimenten werden auf diese Weise nämlich außerdem reale Entscheidungen, bei denen sehr hohe Beträge<sup>1</sup> auf dem Spiel stehen, von Experten untersucht. Die typischen Probleme von Umfragen und Experimenten können somit umgangen werden.

Diese Arbeit trägt zu drei verschiedenen Bereichen der Literatur bei. Erstens setzt sie an den anfangs erwähnten Studien, die sich mit dem Einfluss von Kultur auf ökonomische Fragestellungen beschäftigen, an. Des Weiteren stellt diese Arbeit einen Beitrag zu jener Literatur dar, die sich mit den verschiedenen Determinanten des Risikoverhaltens von Individuen befasst. Ein Literaturüberblick über diese Studien gibt Kapitel 2.1. Außerdem lässt sich diese Arbeit in jene

Literatur einreihen, die zur Analyse (ökonomischer) Fragestellungen auf Golf als Datengrundlage zurückgreift. Als eine der ersten Autoren verwenden Ehrenberg und Bognanno (1990a), Ehrenberg und Bognanno (1990b) Daten aus dem Profigolf sport, um die Anreizeffekte von Turnieren empirisch zu untersuchen. Auch Orszag (1994) analysiert Anreizeffekte in Golfturnieren. Guryan et al. (2009) betrachten anhand von Golfturnieren Peer-Effekte am Arbeitsplatz. Während Hickman und Metz (2015) anhand der PGA TOUR den Einfluss von Druck auf die Leistung eines Individuums untersuchen, befassen sich Hickman et al. (2019) mit dem Effekt der Platzierung im Klassement auf die Leistung. Broadie (2008), Broadie (2012) beschäftigt sich damit, wie sich die Leistung eines Golfers messen lässt. Connolly und Rendleman (2012) untersuchen, wie man Turniere auf der PGA TOUR gewinnt. Pope und Schweitzer (2011) analysieren, ob professionelle Golfer Verlustaversion zeigen. Brown (2011) betrachtet die negativen Anreizeffekte, die durch die Turnierteilnahme eines Superstars entstehen können. Ozbeklik und Smith (2017), Adams und Waddell (2018) sowie McFall und Rothhoff (2018) beschäftigen sich mit unterschiedlichen Aspekten des Risikoverhaltens in (Golf-)Turnieren.

Der Rest dieser Arbeit ist wie folgt aufgebaut. In Kapitel 2 werden zunächst die theoretischen Grundlagen vorgestellt, die zur Beantwortung der Frage, welchen Einfluss der kulturelle Hintergrund eines Individuums auf dessen Risikoverhalten hat, benötigt werden. Hierzu werden in Kapitel 2.1 die wichtigsten Begrifflichkeiten aus dem Bereich des Risikoverhaltens geklärt und anschließend ein Literaturüberblick über wichtige Determinanten des Risikoverhaltens gegeben. Die wichtigsten Grundlagen aus dem (professionellen) Golf sport werden in Kapitel 2.2 dargestellt. In Kapitel 2.3 wird der Begriff Kultur definiert sowie eine Möglichkeit, Kultur quantitativ zu messen, vorgestellt. Kapitel 2.4 fasst auf den vorherigen Kapiteln aufbauend den zu erwartenden Zusammenhang zwischen Kultur und dem Risikoverhalten zusammen. Kapitel 3 beschreibt die Daten, die in der empirischen Analyse in Kapitel 4 verwendet werden. In Kapitel 5 wird ein Fazit gezogen und Kapitel 6 schließt mit einem Ausblick ab.

## 2. Theoretische Grundlagen

### 2.1. Determinanten von Risikoverhalten

#### 2.1.1. Begriffsklärung

Bevor im Folgenden auf diverse Determinanten des Risikoverhaltens von Individuen eingegangen wird, ist es hilfreich, zunächst kurz zu definieren, was im Rahmen dieser Arbeit unter dem Begriff Risikoverhalten verstanden wird. Hierfür ist es wichtig, die in der Literatur verwendeten Begriffe Risikoverhalten, Risikobereitschaft, Risikowahrnehmung, Risikoeinschätzung, Risikoeinstellung sowie Risikopräferenz voneinander abzugrenzen.

Sitkin und Pablo (1992, S. 11) schreiben: „Risk behavior may be characterized by the degree of risk associated with the decisions made“. Auch Breuer et al. (2016) merken an, dass das Risikoverhalten unmittelbar durch Aktionen, die Individuen in Entscheidungssituationen wählen, gemessen werden

<sup>1</sup>Der Sieger der U.S. Open 2017 und 2018, Brooks Koepka, erhielt ein Preisgeld von je 2.160.000 US-Dollar.

kann. Unter Risikoverhalten werden im Rahmen dieser Arbeit daher tatsächlich beobachtete Entscheidungen eines Individuums unter Risiko verstanden. Risikobereitschaft hingegen hat hypothetische Entscheidungen unter Risiko zum Gegenstand. Sowohl das Risikoverhalten als auch die Risikobereitschaft eines Individuums resultieren aus dessen Risikoeinstellung (auch Risikopräferenz genannt) sowie dessen Risikowahrnehmung (auch Risikoeinschätzung genannt) (Breuer et al., 2016; van Winsen et al., 2016; Weber, 2001).

Unter der Risikoeinstellung eines Individuums lässt sich die Wahrscheinlichkeit verstehen, mit der dieses Individuum riskante Alternativen wählt (Cheung et al., 2013). Während risikofreudige Individuen bei zwei Alternativen, die sich lediglich im Grad des Risikos unterscheiden, die riskante Option der risikolosen vorziehen, bevorzugen Individuen mit risikoaverser Einstellung in solchen Fällen die risikolose Alternative gegenüber der riskanten. Individuen mit risikoneutraler Risikoeinstellung wiederum sind in diesem Fall indifferent zwischen der riskanten und der risikolosen Alternative. Die Risikowahrnehmung eines Individuums lässt sich nach Sitkin und Pablo (1992, S. 12) definieren als „a decision maker's assessment of the risk inherent in a situation“.

Unterschiede im Risikoverhalten von Individuen resultieren also entweder aus unterschiedlicher Risikowahrnehmung oder aus unterschiedlichen Risikoeinstellungen. Folgendes Beispiel von Weber (2001) soll dies anhand von zwei Szenarien veranschaulichen: Während Person A und Person B beide eine risikoaverse Einstellung haben, empfindet Person A Pokerspielen riskanter als Fallschirmspringen und entscheidet sich als risikoaverses Individuum deshalb für Fallschirmspringern. Person B hingegen schätzt Pokerspielen als die risikoärmere Alternative ein und entscheidet sich als ebenfalls risikoaverse Person deshalb dafür, Poker zu spielen. Person C und D hingegen stimmen beide darüber ein, dass Fallschirmspringen die riskantere Alternative ist, unterscheiden sich aber in ihrer Risikoeinstellung. Person C als risikofreudigere Person zieht Nutzen aus dem Risiko und entscheidet sich deshalb für das Fallschirmspringen, während die risikoaversere Person D mehr Nutzen aus der risikoärmeren Alternative Pokerspielen zieht und sich deshalb hierfür entscheidet. Während sich im ersten Szenario (Person A vs. B) der Unterschied im Risikoverhalten der beiden Personen durch Unterschiede in der Risikowahrnehmung erklären lässt, sind im Fall des zweiten Szenarios (Person C vs. D) Unterschiede in der Risikoeinstellung dafür verantwortlich, dass sich die betrachteten Personen jeweils unterschiedlich verhalten.

Ein Bereich der Risikoliteratur befasst sich damit, herauszufinden, ob unterschiedliches Risikoverhalten empirisch betrachtet aus unterschiedlichen Risikoeinstellungen oder aber aus Unterschieden in der Risikowahrnehmung resultiert. So finden beispielsweise Weber und Milliman (1997) in zwei Experimenten Belege dafür, dass die Risikoeinstellung eines Individuums ein stabiler Persönlichkeitszug ist und situative Unterschiede im Risikoverhalten aus unterschiedlicher Risikoeinschätzung resultieren. Im Rahmen dieser Arbeit liegt der Fokus jedoch lediglich darauf, die Frage zu beantworten, ob der kulturelle Hintergrund eines Individuums einen Ein-

fluss auf dessen Risikoverhalten hat. Die Beantwortung der Frage, über welchen der beiden Kanäle der kulturelle Hintergrund auf das Risikoverhalten gegebenenfalls wirkt, bliebe dann offen für weitere Forschung.

### 2.1.2. Erwartungsnutzentheorie und Prospect Theorie

Folgt man der Erwartungsnutzentheorie nach von von Neumann und Morgenstern (1944), der Standardtheorie zu Entscheidungen unter Risiko (Starmer, 2000), so hängt das Risikoverhalten eines Individuums (bei gegebener Risikowahrnehmung) lediglich von der Krümmung seiner Nutzenfunktion ab. Somit wäre das (End-)Vermögen eines Individuums die einzig endogene Determinante seines Risikoverhaltens.

Weil jedoch einige Phänomene, die empirisch robust zu beobachten sind, nicht mit den Axiomen der Erwartungsnutzentheorie vereinbar sind, kritisieren Kahneman und Tversky (1979) diese als deskriptiv ungeeignet und entwickeln stattdessen mit der Prospect Theorie ein alternatives Modell für Entscheidungen unter Risiko. Zu den wichtigsten Merkmalen der Prospect Theorie gehört, dass nicht länger das Endvermögen eines Individuums relevant ist, sondern Veränderungen gegenüber einem Referenzpunkt ausschlaggebend für die Bewertung von Alternativen sind. Typischerweise zeigen Individuen nämlich risikofreudiges Verhalten, wenn sie sich in der Verlustdomäne befinden (d. h. schlechter gestellt sind als der Referenzpunkt) und risikoaverses Verhalten in der Gewinn-domäne. Deshalb ist die Wertefunktion, welche an die Stelle der Nutzenfunktion aus der Erwartungsnutzentheorie rückt, konkav für Gewinne und konvex für Verluste. Die beobachtete Verlustaversion von Individuen wird implementiert, indem die Wertefunktion für Verluste steiler ist als für Gewinne. Außerdem werden die Werte eines riskanten Ergebnisses nicht wie in der Erwartungsnutzentheorie mit deren Eintrittswahrscheinlichkeit gewichtet, sondern mit Entscheidungsgewichten, welche kleine Wahrscheinlichkeiten übergewichten und größere Wahrscheinlichkeiten untergewichten. Später verfeinern Tversky und Kahneman (1992) ihre Prospect Theorie als kumulative Prospect Theorie.

Folglich spielt es für das Risikoverhalten eines Individuums eine Rolle, ob ein Ergebnis als Verlust oder als Gewinn gesehen wird. Pope und Schweitzer (2011) belegen die deskriptive Gültigkeit der Prospect Theorie, indem sie herausfinden, dass Golfer „Par“ als salienten Referenzpunkt benutzen und entsprechend abhängig davon, ob sie in der Verlust- oder Gewinn-domäne sind, unterschiedlich riskantes Verhalten zeigen. Darüber hinaus finden sie Belege dafür, dass professionelle Golfspieler Verlustaversion zeigen.

Auch andere Forscher (u. a. Allais (1953) und Rabin (2000)) üben Kritik an der Erwartungsnutzentheorie als deskriptive Theorie für Entscheidungen unter Risiko. Deshalb werden in den folgenden Kapiteln Determinanten des Risikoverhaltens von Individuen vorgestellt, welche in (empirischen) Forschungsarbeiten festgehalten sind.

### 2.1.3. Vorherige Ergebnisse

Ein Faktor, der das Risikoverhalten eines Individuums bei wiederholten Entscheidungen unter Risiko beeinflusst, sind die realisierten Ergebnisse vorangegangener Risiken derselben Art. Ein möglicher Grund hierfür ist das sehr weit verbreitete Phänomen von übermäßigem Selbstvertrauen von Individuen, welches in der meist englischsprachigen Literatur unter dem Begriff „Overconfidence“ bekannt ist. So merken [de Bondt und Thaler \(1995, S. 389\)](#) an: „Perhaps the most robust finding in the psychology of judgment is that people are overconfident“.

Übermäßiges Selbstbewusstsein drückt sich darin aus, dass Individuen die Reliabilität ihrer Einschätzung von Wahrscheinlichkeiten überschätzen, zu enge Konfidenzintervalle wählen und ihre eigenen Fähigkeiten überschätzen ([de de Bondt und Thaler, 1995](#)). Auch [Tversky und Kahneman \(1971\)](#), [Tversky und Kahneman \(1974\)](#) beobachten derartige Fehleinschätzungen. So stellen [Tversky und Kahneman \(1971\)](#) fest, dass Individuen an das „Gesetz der kleinen Zahlen“ glauben, demzufolge das Gesetz der großen Zahlen auch für kleine Stichproben gilt. [Tversky und Kahneman \(1974\)](#) beobachten außerdem, dass Individuen die Repräsentativität von kurzen Sequenzen, die durch einen (Zufalls-)Prozess generiert werden, überschätzen, indem sie davon ausgehen, dass auch diese kurze Abfolge von Ereignissen die wesentlichen Eigenschaften dieses Prozesses verkörpern. Beobachten Individuen also mehrere Erfolge in vorherigen Versuchen, folgern sie möglicherweise fälschlicherweise, dass die Erfolgswahrscheinlichkeit höher liegt, als ursprünglich angenommen.

Ist Overconfidence gepaart mit diesem Glauben an das „Gesetz der kleinen Zahlen“, bzw. ist dieser Glaube die Ursache von übermäßigem Selbstvertrauen, so sorgt ein vorheriger Erfolg dafür, dass die eigenen Fähigkeiten bzw. die Erfolgswahrscheinlichkeit einer riskanten Strategie (noch) stärker überschätzt werden. Entsprechend wird das zugrundeliegende Risiko somit als (noch) geringer betrachtet, was in risikofreudigerem Risikoverhalten mündet. Übereinstimmend mit dieser Argumentation finden [Pan und Statman \(2012\)](#) in ihrer Onlineumfrage mit 2.512 Teilnehmern heraus, dass übertrieben selbstbewusste Individuen eine größere Risikobereitschaft haben als weniger übertrieben selbstbewusste. Die Ergebnisse von [Heath und Tversky \(1991\)](#) legen nahe, dass Individuen sich bei ihren Vorhersagen in einem Bereich, in dem sie sich ihrer eigenen Meinung nach auskennen, sicherer fühlen, während ihre tatsächliche Vorhersagefähigkeit unverändert bleibt. Auch die Ergebnisse von [Glaser et al. \(2013\)](#) belegen, dass Fachleute stärkere Overconfidence zeigen als Laien. Es ist daher davon auszugehen, dass professionelle Golfer, die Gegenstand der empirischen Untersuchung dieser Arbeit sind, übermäßig selbstbewusst sind und deshalb in ihrem Risikoverhalten von vorherigen Ergebnissen beeinflusst werden.

[Gilovich et al. \(1985\)](#) führen eine Umfrage mit Basketballfans, die selbst Basketball spielen, durch und beobachten dabei, dass die meisten Fans glauben, dass die Chan-

ce, dass ein Spieler einen Wurf verwandelt, höher ist, wenn er die vorangegangenen Würfe verwandelt hat, als wenn er diese nicht verwandelt hat. Die Idee, dass ein Sportler im nächsten Versuch eine höhere Erfolgswahrscheinlichkeit hat, wenn der vorherige Versuch erfolgreich war, ist auch bekannt als „Hot-Hand“-Phänomen oder Momentum. Etliche Studien überprüfen die empirische Validität der Hot-Hand. [Bar-Eli et al. \(2006\)](#) geben einen Literaturüberblick über solche Studien und kommen zu dem Ergebnis, dass die Frage, ob die Hot-Hand wirklich existiert oder nicht, nicht abschließend geklärt sei. Während beispielsweise [Gilovich et al. \(1985\)](#) bei der Auswertung von Basketballdaten zu der Folgerung kommen, dass keine positive Korrelation zwischen aufeinanderfolgenden erfolgreichen Würfen besteht, finden unter anderem [Green und Zwiebel \(2018\)](#) in ihrer Analyse von Daten aus der Major League Baseball Belege für die Existenz der Hot-Hand. Auch wenn die Mehrzahl der empirischen Studien nahelegt, dass die umstrittene Hot-Hand nicht existiert ([Bar-Eli et al. \(2006\)](#)), führt bereits der Glaube des Athleten daran zu einem veränderten Risikoverhalten. Ist ein Sportler davon überzeugt, dass die Hot-Hand tatsächlich existiert, so schätzt er das Risiko eines Fehlwurfs nach einem vorangegangenen erfolgreichen Versuch als geringer ein und verhält sich deshalb im nächsten Versuch bei gegebener Risikoeinstellung risikofreudiger. Da [Gilovich et al. \(1985\)](#) in Interviews mit professionellen Basketballspielern der Philadelphia 76er herausfinden, dass auch diese an die Existenz der Hot-Hand glauben, ist nicht auszuschließen, dass auch professionelle Golfer daran glauben und deshalb in ihrem Risikoverhalten von vorherigen Ergebnissen beeinflusst werden.

Losgelöst von den Überlegungen zu übermäßigem Selbstvertrauen sowie der Hot-Hand finden [Wik et al. \(2004\)](#) in einem Experiment, bei dem die Teilnehmer zwischen unterschiedlichen Lotterien auswählen müssen, heraus, dass Individuen weniger risikoavers sind, wenn sie in vorherigen Lotterien Glück hatten und gewonnen haben. Auch [Thaler und Johnson \(1990\)](#) untersuchen mithilfe eines Lotterieexperiments mit Wirtschaftsstudenten, wie deren Risikoverhalten durch vorangegangene Verluste und Gewinne beeinflusst wird. Dabei beobachten sie einerseits, dass sich Individuen unter Umständen risikofreudiger verhalten, wenn sie in der Runde zuvor gewonnen haben und risikoaverser, wenn sie in der vorherigen Runde verloren haben. Als Erklärung für das risikofreudigere Verhalten nach Gewinnen führen sie den „House Money Effect“ an, hinter welchem die Idee steckt, dass einem Gewinn folgende Verluste – solange bis der Gewinn wieder komplett verspielt ist – lediglich als Reduktion des Gewinns betrachtet werden, was weniger schmerzt, als das eigene Bargeld zu verspielen. Gleichzeitig beobachten die Autoren, dass sich Individuen nach vorangegangenen Verlusten risikofreudig verhalten, wenn sich die Chance bietet, diese Verluste durch einen Erfolg in der aktuellen Runde wieder auszugleichen.

Zusammenfassend ist festzuhalten, dass positive Realisationen vorheriger Risiken typischerweise zu risikofreudigerem Verhalten führen.

#### 2.1.4. Einkommen und Vermögen

Nachdem im vorherigen Kapitel gezeigt wurde, dass die realisierten Ergebnisse vorheriger Risiken aus unterschiedlichen Gründen eine Determinante des Risikoverhaltens von Individuen darstellen, werden in diesem Kapitel die Ergebnisse empirischer Studien bezüglich des Einflusses von Einkommen und Vermögen eines Individuums dargelegt. Wie in Kapitel 2.1.2 beschrieben, ist das Vermögen in der Erwartungsnutzentheorie, abgesehen von der Krümmung der Nutzenfunktion, die einzige Determinante des Risikoverhaltens.

Diverse Studien stellen fest, dass sich Individuen risikofreudiger verhalten, je größer ihre finanziellen Ressourcen sind. So analysieren [Rosenzweig und Binswanger \(1993\)](#) die landwirtschaftlichen Portfolios (Rinder, Geflügel, Boden, Maschinen, etc.) von Bauern aus Indien und finden heraus, dass vermögendere Bauern riskanter investieren als weniger vermögende. [Liu \(2013\)](#) beobachtet für chinesische Baumwollfarmer, dass ärmere Individuen risikoaverser sind. Auch [Miyata \(2003\)](#) beobachtet in ihrem Experiment mit indonesischen Dorfbewohnern, dass Haushalte, die mehr Vermögen haben, risikofreudigeres Verhalten in einem Investmentspiel zeigen. Die Experimentalergebnisse von [Wik et al. \(2004\)](#) deuten ebenfalls darauf hin, dass sich vermögendere Individuen risikofreudiger verhalten. [Schneider et al. \(2017\)](#) analysieren die Bereitschaft von Individuen, finanzielle Risiken einzugehen und beobachten, dass diese Bereitschaft mit dem Vermögen zunimmt. [Calvet und Sodini \(2014\)](#) untersuchen für den Zeitraum 1999 bis 2002 das Anlageverhalten von schwedischen Zwillingen und stellen dabei fest, dass der Anteil des Portfolios, der in riskante Anlagen investiert ist, eine zunehmende Funktion des eigenen Vermögens ist. Außerdem entdecken die beiden Autoren, dass das Vermögen auch die zugrunde liegende Entscheidung, überhaupt in riskante Anlagen zu investieren, positiv beeinflusst.

In einem Rahmen, in dem sehr hohe Beträge auf dem Spiel stehen, finden [Hopland et al. \(2016\)](#) heraus, dass Teilnehmer der norwegischen Spielshow „Millionsjansen“ risikofreudigere Entscheidungen treffen, je mehr Einkommen sie haben. [Donkers et al. \(2001\)](#) stellen anhand einer Umfrage ebenfalls fest, dass Individuen mit höherem Einkommen risikofreudiger sind. Zudem beobachten sie bei der Schätzung eines parametrischen Modells nach der kumulativen Prospect Theorie, dass sowohl das Vermögen eines Individuums als auch dessen Einkommen die Wertefunktion und die Wahrscheinlichkeitsgewichtungsfunktion beeinflussen. In einem Experiment mit 504 Haushalten in Äthiopien zeigen [Viider et al. \(2018\)](#), dass Haushalte mit höherem Einkommen eine größere Risikobereitschaft haben.

[Dohmen et al. \(2011\)](#) sowie [Weber \(2014\)](#) kommen bei ihren Analysen von Daten des deutschen Sozio-oekonomischen Panels ebenfalls zu dem Ergebnis, dass vermögendere Individuen sowie solche mit höherem Einkommen eine größere Risikobereitschaft haben. [Barsky et al. \(1997\)](#) beobachten in ihrer Analyse einen nicht-linearen Zusammenhang zwischen den finanziellen Ressourcen eines Individuums und dessen Risikobereitschaft. So beobachten sie,

dass die Risikobereitschaft bis zur Mitte der Einkommens-/Vermögensverteilung hin mit dem Vermögen und Einkommen abnimmt und anschließend wieder zunimmt. Übereinstimmend damit finden [Halek und Eisenhauer \(2001\)](#) bei ihrer Untersuchung von Daten über den Abschluss von Lebensversicherungen heraus, dass die (relative) Risikoaversion von Individuen zunächst mit dem Vermögen zunimmt und erst ab einem gewissen Niveau wieder abnimmt.

[Tanaka et al. \(2010\)](#) entdecken mithilfe eines Experiments in vietnamesischen Dörfern, dass zwar kein Zusammenhang zwischen dem Einkommen und dem Risikoverhalten besteht, Individuen aus wohlhabenderen Dörfern aber weniger risikoavers sind und eine weniger stark ausgeprägte Verlustaversion zeigen. Im Gegensatz dazu stellen [Viider et al. \(2015b\)](#) mithilfe von Laborexperimenten mit über 2.900 Studenten in 30 Ländern fest, dass Studienteilnehmer aus wohlhabenderen Ländern risikoaverser sind. Auch [Rieger et al. \(2015\)](#) beobachten diesen Zusammenhang in ihrer Studie. Bezugnehmend auf die Ergebnisse der beiden letztgenannten Studien sprechen [Viider et al. \(2018\)](#) in ihrer Arbeit von einem Risiko-Einkommensparadoxon, da innerhalb eines Landes die Risikoaversion typischerweise mit dem Einkommen abnimmt, zwischen Ländern ein höheres Pro-Kopf-Einkommen jedoch zu risikoaverserem Verhalten führt. Auch [Statman \(2008\)](#), [Statman \(2010\)](#), [Statman \(2015\)](#) beobachten in Umfragen in mehr als 20 Ländern, dass Individuen aus Ländern mit geringem Pro-Kopf-Einkommen eher bereit sind, finanzielle Risiken einzugehen.

Während etliche Studien also zu dem Ergebnis kommen, dass sich Individuen mit größeren finanziellen Ressourcen risikofreudiger verhalten, können manche Arbeiten keinen signifikanten Zusammenhang beobachten ([Binswanger, 1980](#), [Binswanger, 1981](#); [Buccioli und Miniaci, 2018](#); [Cardenas und Carpenter, 2008](#); [Henrich und McElreath, 2002](#); [Holt und Laury, 2002](#); [Mosley und Verschoor, 2005](#)).

#### 2.1.5. Alter

Eine weitere Determinante des Risikoverhaltens, die Untersuchungsgegenstand diverser empirischer Studien ist, stellt das Alter des Entscheidungsträgers dar. Der Großteil der Literatur beobachtet dabei, dass ältere Menschen ceteris paribus risikoaverseres Verhalten bzw. eine geringere Risikobereitschaft zeigen.

[Botwinick, 1966](#) beobachtet, dass ältere Studienteilnehmer weniger riskante Entscheidungen bezüglich 24 hypothetischen Lebenssituationen, welche jeweils zwei Entscheidungsoptionen – eine riskante, ansonsten aber superiore Option sowie eine sichere Option – beinhalten. Auch die Ergebnisse der Onlineumfrage von [Pan und Statman \(2012\)](#) zeigen, dass ältere Studienteilnehmer gemessen anhand von hypothetischen Fragen, weniger bereit sind, Risiko einzugehen als jüngere. [Buccioli und Miniaci \(2018\)](#) finden in ihrer Analyse von Daten der DNB Household Survey für den Zeitraum 1995-2015 ebenfalls heraus, dass das Alter eines Individuums einen negativen Effekt auf dessen Risikobereitschaft hat.

Auch [Tavor und Garyn-Tal \(2016\)](#) sowie [Schneider et al. \(2017\)](#) beobachten, dass ältere Menschen risikoaverser sind.



Bellante und Green (2004) untersuchen Portfolioallokationen von älteren Personen anhand des Asset and Health Dynamics Among the Oldest Old Datensatzes, der von der University of Michigan erhoben wird, und beobachten dabei, dass die (relative) Risikoaversion eines Individuums mit dem Alter unabhängig von dessen Vermögensniveau zunimmt.

Auch Jianakoplos und Bernasek (2006) untersuchen die Bereitschaft von Individuen, finanzielle Risiken einzugehen, anhand von beobachteten Portfolioallokationen sowie auf Grundlage einer Umfrage bezüglich der Bereitschaft, finanzielle Risiken einzugehen. Als Datengrundlage dienen dabei die Surveys of Consumer Finances der Jahre 1989, 1995 und 2001. Im Unterschied zu den meisten vorherigen Studien zerlegen die beiden Autoren den Effekt des Alters in einen echten Alterseffekt, einen Kohorteneffekt sowie einen Kalenderjahreffekt, was es erlaubt, für unterschiedliche makroökonomische Bedingungen in den Erhebungsjahren zu kontrollieren. Die Ergebnisse zeigen einerseits, dass jüngere Kohorten risikoaverser als ihre älteren Pendants sind. Andererseits nimmt die Risikobereitschaft, gemessen an den Umfrageantworten, mit dem Alter ab und auch die beobachteten Portfolioallokationen zeugen von risikoaverserem Verhalten. Der Zusammenhang von Alter ist dabei nicht-linear, sodass die maximale Risikobereitschaft bei einem Alter von 22,5 Jahren liegt.

Dohmen et al. (2017) kontrollieren in ihrer Untersuchung von Paneldaten aus der DNB Haushaltsumfrage sowie dem deutschen Sozio-oekonomischen Panel ebenfalls für Kohorten- und Kalenderjahreffekte. Sie betonen dabei insbesondere ihre neuartige Vorgehensweise<sup>2</sup>, um zwischen Kohorten-, Alters- und Kalenderjahreffekten unterscheiden zu können. Während der Fokus der DNB Umfrage auf finanziellen Risiken liegt, wird mit den Fragen des deutschen Sozio-oekonomischen Panels die allgemeine Risikobereitschaft abgefragt. Auch hier kommen die Autoren zu dem Ergebnis, dass die Risikobereitschaft eines Individuums im Verlaufe seines Lebens zunächst linear abnimmt, ehe die Steigung ab einem Alter von 65 Jahren abflacht. In einer früheren Arbeit untersuchen Dohmen et al. (2011) ebenfalls Daten des deutschen Sozio-oekonomischen Panels. Zusätzlich führen sie ein Feldexperiment durch, um zu untersuchen, ob das Risikoverhalten der Teilnehmer auch mit der von ihnen in der Umfrage angegebenen Risikobereitschaft übereinstimmt. Dabei stellen die Autoren fest, dass die Risikobereitschaft mit dem Alter abnimmt und dass die Antworten der Studienteilnehmer auf die Frage zur generellen Risikobereitschaft des Sozio-oekonomischen Panels konsistent mit ihrem Risikoverhalten im Feldexperiment sind. Auch Weber (2014) stellt fest, dass die generelle sowie finanzielle Risikobereitschaft, gemessen anhand der Selbsteinschätzungsfragen des deut-

schen Sozio-oekonomischen Panels aus dem Jahr 2009 mit dem Alter abnimmt.

Donkers et al. (2001) finden basierend auf einer Analyse von acht hypothetischen Fragen zu Lotterien, die im Rahmen der niederländischen CentER Savings Survey gestellt wurden, ebenfalls heraus, dass ältere Studienteilnehmer eine geringere Risikobereitschaft haben. Darüber hinaus stellen die Autoren beim Schätzen eines parametrischen Modells nach der (kumulativen) Prospect Theorie fest, dass das Alter eines Individuums sowohl auf die Wahrscheinlichkeitsgewichtungsfunktion als auch auf die Wertefunktion einen signifikanten Einfluss hat. Bonsang und Dohmen (2015) finden mithilfe von Daten aus der Survey of Health, Ageing, and Retirement in Europe heraus, dass ein nicht unerheblicher Anteil der altersbedingten Unterschiede in der Risikobereitschaft, finanzielle Risiken einzugehen, auf einen Rückgang der kognitiven Fähigkeiten zurückzuführen ist. Auch sie beobachten dabei, dass ältere Menschen eine geringere Risikobereitschaft haben.

Nicht alle Studien kommen jedoch zu der Schlussfolgerung, dass ältere Individuen risikoaverser sind. So beobachten Halek und Eisenhauer (2001) in ihrer Analyse von Daten über den Abschluss von Lebensversicherungen zwar, dass die Risikoaversion von Individuen ab einem Alter von 65 Jahren signifikant ansteigt. Davor nimmt die Risikoaversion jedoch mit dem Alter ab. Rieger et al. (2015) beobachten in einer Umfrage, dass ältere Individuen in der Gewinn-domäne weniger risikoavers sind als jüngere, in der Verlustdomäne hingegen risikofreudiger. Barsky et al. (1997) analysieren die Umfrageergebnisse der Health and Retirement Study aus 1992, welche anhand von hypothetischen Situationen die Risikobereitschaft der Umfrageteilnehmer, mit dem Lebens-einkommen zu spielen, misst. Dabei stellen sie fest, dass die jüngste und die älteste Altersgruppe die größte Risikobereitschaft haben.

Rolison et al. (2014) untersuchen die Risikobereitschaft von Individuen in verschiedenen Domänen und kommen dabei zu folgenden Ergebnissen. Während die Risikobereitschaft in der Finanzdomäne stark mit dem Alter abnimmt, nimmt die Risikobereitschaft in der sozialen Domäne zunächst leicht zu von jungem zu mittlerem Alter, ehe sie später wieder stark abnimmt. In der Freizeitdomäne nimmt die Risikobereitschaft von jungem zu mittlerem Alter stärker ab als später. Zu guter Letzt nimmt die Risikobereitschaft in der Gesundheits- sowie der Ethikdomäne relativ sanft mit dem Alter ab.

Wik et al. (2004) finden im Rahmen ihres Experiments in Nordsambia keinen signifikanten Zusammenhang zwischen dem Alter eines Individuums und dessen Risikoverhalten. Auch Holt und Laury (2002) können in ihrem Laborexperiment an drei Universitäten in den USA keinen Einfluss des Alters auf das Risikoverhalten der 175 Studienteilnehmer beobachten. Bei ihrer Untersuchung von Daten über die norwegische Spielshow „Millionsjansen“, bei der der Teilnehmer zwischen einem sicheren Preis oder einer Lotterie wählt und bei der hohe Einsätze auf dem Spiel stehen, finden Hopland et al. (2016) ebenfalls keinen Alterseffekt auf das Risikover-

<sup>2</sup>Die Autoren nutzen anstelle von Periodendummies Determinanten von Risikopräferenzen, die zwar von der Kalenderzeit abhängen, sich jedoch nicht linear mit der Kalenderzeit ändern (konkret: das BIP-pro-Kopfwachstum), als Proxy für Kalenderzeiteffekte, da sie davon ausgehen, dass die makroökonomischen Rahmenbedingungen die treibende Kraft hinter Kalenderzeiteffekten sind.

halten der Kandidaten. Und auch [Nobre et al. \(2016\)](#) finden keinen signifikanten Zusammenhang zwischen dem Alter eines Umfrageteilnehmers und dessen Bereitschaft, finanzielle Risiken einzugehen.

#### 2.1.6. Höhe des Einsatzes

Eine weitere wichtige Determinante des Risikoverhaltens stellt die Höhe des Betrages, der auf dem Spiel steht, dar. So beobachtet bereits [Binswanger \(1980\)](#), [Binswanger \(1981\)](#) in einem Experiment im ländlichen Indien, bei dem die Teilnehmer zwischen verschiedenen Lotterien wählen müssen, dass Individuen risikoaverseres Verhalten an den Tag legen, je höher die Beträge in den zur Wahl stehenden Lotterien sind. Auch [Grisley und Kellog \(1987\)](#) führen ein Lotteriexperiment mit Bauern aus zwei Dörfern in Nordthailand durch und kommen zu dem Ergebnis, dass sich diese risikoaverser verhalten, je höher die Beträge sind, die auf dem Spiel stehen. Ebenfalls Dorfbewohner als Untersuchungsgruppe verwendet [Miyata \(2003\)](#). In ihrem Experiment spielen 400 Haushalte aus indonesischen Dörfern ein Investmentspiel. Dabei ist zu beobachten, dass sich die Teilnehmer risikoaverser verhalten, wenn die Auszahlungsniveaus der unterschiedlichen Investmentalternativen steigen.

[Smith und Walker \(1993\)](#) führen ein Laborexperiment zu Erstpreisauktionen mit Studierenden der Indiana University sowie der University of Arizona durch und beobachten risikoaversere Gebote der Teilnehmer, je höher die Auszahlungsniveaus sind. [Holt und Laury \(2002\)](#) untersuchen ebenfalls in einem Laborexperiment, bei dem die Teilnehmer zwischen zwei Lotterien wählen müssen, den Einfluss der Höhe der Beträge in diesen Lotterien auf das beobachtete Risikoverhalten und stellen fest, dass höhere Beträge zu risikoaverserem Verhalten führen. Auch die Ergebnisse des Experiments von [Wik et al. \(2004\)](#) deuten darauf hin, dass sich Individuen weniger risikofreudig verhalten, wenn höhere Beträge auf dem Spiel stehen.

Mithilfe von Experimenten an der Beijing University in China, einer mittelgroßen kanadischen Universität und an einer großen Universität in den USA, bei denen die Studienteilnehmer Sicherheitsäquivalente für 25 verschiedene Lotterien angeben, stellen auch [Kachelmeier und Shehata \(1992\)](#) fest, dass sich die Teilnehmer risikoaverser verhalten, wenn die real ausgezahlten Beträge der Lotterien verzehnfacht werden. Diesen Einfluss der Höhe der Beträge finden die Autoren jedoch nicht, wenn die Sicherheitsäquivalente, wie im Falle der kanadischen Studienteilnehmer, von rein hypothetischen Lotterien untersucht werden.

[Lefebvre et al. \(2010\)](#) beobachten in ihrem Experiment, dass sich Individuen grundsätzlich risikofreudig bei Lotterien mit geringen Gewinnwahrscheinlichkeiten verhalten. Eine Erhöhung der Beträge, die gewonnen werden können, führt jedoch dazu, dass diese Risikofreude abnimmt. Auch [Hsee und Weber \(1999\)](#) führen ein Experiment mit Studenten durch, wobei die Teilnehmer zwischen einer sicheren Auszahlung und einer Lotterie wählen. Während das Verhalten der Teilnehmer in der Gewinn-domäne weniger risikofreudig ist, wenn die auf dem Spiel stehenden Beträge erhöht

werden, verhalten sich die Studenten in der Verlustdomäne risikofreudiger, wenn die Beträge erhöht werden.

[Becker und Huselid \(1992\)](#) analysieren Daten aus dem Motorsport und stellen dabei fest, dass Rennfahrer, gemessen an der Anzahl von Unfällen, mehr Risiko eingehen, je ungleicher Preisgelder auf die unterschiedlichen Platzierungen verteilt werden, jedoch nur, wenn ein sehr großer Unterschied zwischen den Preisgeldern für die unterschiedlichen Platzierungen besteht. In Rennen, in denen hingegen nur geringe Preisgeldunterschiede bestehen, verhalten sich die Fahrer risikoaverser, wenn die Preisgeldunterschiede zunehmen. Die (absolute) Höhe der Preisgelder, die ausgeschüttet werden, hat also auch hier einen signifikanten Einfluss auf das Risikoverhalten.

Zusammenfassend lässt sich festhalten, dass sich Individuen typischerweise risikoaverser verhalten, je mehr auf dem Spiel steht.

#### 2.1.7. Zwischenplatzierung

Ein Faktor, der das Risikoverhalten von Individuen in Wettbewerbssituationen beeinflusst, ist der Zwischenrang eines Individuums bzw. die Information über die bisherige relative Leistung des Individuums in diesem Wettbewerb. Empirische Literatur hierüber untersucht diesen Zusammenhang meistens anhand von Daten über Investmentfonds oder aber anhand von Sportwettbewerben.

Eine der ersten Analysen des Zusammenhangs zwischen Risikoverhalten und der Zwischenplatzierung anhand von Investmentfonds ist die Arbeit von [Brown et al. \(1996\)](#). Die Autoren untersuchen 334 wachstumsorientierte Fonds zwischen 1980 und 1991 und finden dabei heraus, dass Fondsmanager, deren Vergütung an die relative Performance ihres Fonds gekoppelt ist, für den Rest des Jahres mehr Risiko – verglichen mit den Investitionsentscheidungen der Halbjahres-Gewinner – eingehen, wenn der von ihnen gemanagte Fonds zur Jahresmitte schlechter abschneidet als die Fonds der Konkurrenz. Auch [Chevalier und Ellison \(1997\)](#) finden heraus, dass Manager von Fonds mit unterdurchschnittlicher Rendite im Zeitraum von Jahresbeginn bis September das Risiko des Fondsportfolios zum Ende des Jahres hin erhöhen, um zu den Renditen konkurrierender Fonds aufzuschließen und so Kapitalzuflüsse in den Fonds zu sichern. [Koski und Pontiff \(1999\)](#) bestätigen dieses Ergebnis, dass Fonds mit unterdurchschnittlicher Performance im Laufe des Jahres das Risiko gegen Ende des Jahres erhöhen.

[Busse \(2001\)](#) beobachtet bei einer Betrachtung von monatlichen Renditedaten ebenfalls, dass Fonds mit unterdurchschnittlicher Performance das Risiko verglichen mit Fonds mit besserer Performance erhöhen. Seine Untersuchung von täglichen Renditen führt jedoch zu dem Ergebnis, dass Fondsmanager mit einer besseren Performance als der Median-Performance das eingegangene Risiko stärker erhöhen als Manager von Fonds, deren Performance schlechter ist als die des Median-Fonds, und stellt in Frage, ob Fondsmanager wirklich aktiv das Risiko ihrer Portfolios an die vorherige Performance anpassen.

Kempf et al. (2009) betrachten Manager von US-Aktienfonds über den Zeitraum 1980-2003 und stellen dabei fest, dass Fondsmanager konkurrierende Anreize haben und die Antwort auf die Frage, welcher der beiden Anreize dominiert, dafür ausschlaggebend ist, welchen Effekt die relative Halbjahres-Performance auf das Risikoverhalten hat. So führen Vergütungsanreize dazu, dass Manager von Fonds mit schlechter Halbjahres-Performance im Vergleich zu denen mit guter Halbjahres-Performance mehr Risiko eingehen, um zu Halbjahres-Gewinnern aufzuschließen. Arbeitsplatzanreize (Anreize, den eigenen Job zu behalten) hingegen führen dazu, dass Manager von schwachen Fonds das Risiko des Fondsportfolios verringern. Während die Autoren zu dem Ergebnis kommen, dass in Bullenmärkten die Vergütungsanreize dominieren, dominieren die Arbeitsplatzanreize in Bärenmärkten.

Zu der Literatur, die den Zusammenhang zwischen der relativen Leistung und dem Risikoverhalten anhand von Sportwettbewerben untersucht, gehört die Arbeit von Grund und Gürtler (2005). Die beiden Forscher untersuchen Fußballspiele aus der 1. Bundesliga in Deutschland, die in der Saison 2003/2004 ausgetragen wurden. Dabei finden sie heraus, dass Trainer der führenden Mannschaft in einem Spiel zu einer risikoärmeren Strategie wechseln, indem sie die Formation ihrer Mannschaft durch Spielerwechsel (Spieler mit anderer taktischer Position) defensiver gestalten, während Trainer des zurückliegenden Teams durch die Einwechslung eines offensiveren Spielers zu einer offensiveren Formation wechseln. Der Wechsel zu einer riskanteren Formation durch die Einwechslung eines offensiveren Spielers zahlt sich jedoch weder in Form von einer Spielergebnisverbesserung noch in Form von einem zusätzlichen Punktgewinn aus. Bei der Analyse von Spielen aus der NBA-Saison 2007/2008 beobachten Grund et al. (2013), dass auch im Basketball die zurückliegende Mannschaft während der letzten Minuten eines Spiels, gemessen an der Anzahl von 3-Punkte-Würfen vermehrt Risiko eingeht. Je größer der Rückstand ist, desto mehr Risiko gehen Mannschaften dabei ein. Auch hier zahlt sich das zusätzlich eingegangene Risiko in den meisten Situationen jedoch nicht aus.

Genakos und Pagliero (2012) betrachten professionelle Wettbewerbe im Gewichtheben, die im Zeitraum von 1990 bis 2006 im Rahmen von Europa- und Weltmeisterschaften sowie von Olympischen Spielen ausgetragen wurden. Dabei finden sie heraus, dass das Risikoverhalten eines Athleten eine umgekehrte U-Funktion seiner Zwischenplatzierung ist. Die Autoren erklären dies damit, dass zurückliegende Athleten mehr Risiko eingehen, um besserplatzierte Sportler zu überholen. Ist der platzierungsmäßige Rückstand jedoch sehr groß, wählen Gewichtheber eine weniger riskante Strategie, da kaum noch die (realistische) Chance besteht, die Führenden einzuholen. Außerdem stellen die beiden Forscher fest, dass die Bereitschaft, mehr Risiko einzugehen, zunimmt, je umkämpfter und prestigeträchtiger ein Wettbewerb ist.

Auch Bothner et al. (2007) stellen bei ihrer Analyse der NASCAR Winston Cup Serie über den Zeitraum 1990-2003 fest, dass Sportler, gemessen an der Häufigkeit von Unfällen,

mehr Risiko eingehen, je größer die Gefahr ist, von einem schlechter platzierten Konkurrenten überholt zu werden. Dabei ist die Gefahr, überholt zu werden, umso größer, je mehr schlechter platzierte Fahrer in unmittelbarer Reichweite sind. Übereinstimmend mit der Idee von Verlustaversion beobachten die Autoren außerdem, dass Rennfahrer die Gefahr, von schlechter platzierten Fahrern überholt zu werden, bei der Wahl ihres Risikoverhaltens stärker gewichten als die Chance, besserplatzierte Fahrer zu überholen.

Mit Ozbeklik und Smith (2017) sowie Adams und Waddell (2018) nutzen zwei Forschergruppen Daten aus dem Profigolf, um den Zusammenhang zwischen der Zwischenplatzierung eines Sportlers und dessen Risikoverhalten zu untersuchen. Ozbeklik und Smith (2017) betrachten dafür 579 Partien Lochspielgolf, die von 2003 bis 2013 im Rahmen der WGC-Accenture Match Play Championship auf der PGA TOUR ausgetragen wurden. Dabei finden sie heraus, dass schwächere und zurückliegende Spieler risikofreudiger spielen. Je größer der Rückstand und je fortgeschrittener ein Duell ist, desto mehr Risiko wird dabei eingegangen, da die Chancen auf einen Sieg zunehmend geringer werden. Im Gegensatz dazu wählen in Führung liegende Golfer generell weniger riskante Strategien und gehen umso weniger Risiko ein, je größer ihr Vorsprung ist. Adams und Waddell (2018) dagegen analysieren für den Zeitraum 2002 bis 2016 Zählspieltourniere auf der PGA TOUR. Sie kommen dennoch ebenfalls zu dem Ergebnis, dass sich Golfer risikofreudiger verhalten, je schlechter ihr Zwischenrang ist, insbesondere wenn sich ein Spieler unterhalb des Cuts befindet und somit die ernsthafte Gefahr besteht, aus dem Turnier auszuschneiden.

Die meisten Studien sagen also riskanteres Verhalten voraus, je schlechter platziert ein Wettbewerbsteilnehmer ist.

#### 2.1.8. Nationalität und Ethnie

Eine Vielzahl an Studien, die das Risikoverhalten beziehungsweise die Risikobereitschaft von Individuen untersuchen, stellt fest, dass es nationale und ethnische Unterschiede gibt. Im Gegensatz zu dieser Arbeit werden dort jedoch üblicherweise nur qualitative Vergleiche zwischen zumeist zwei Ländern angestellt.

So führen Viieder et al. (2015a) das gleiche Experiment in zwei chinesischen Städten sowie auf zwei Campussen in Äthiopien durch und zeigen dabei, dass die Unterschiede in der Risikoeinstellung innerhalb eines Landes zu vernachlässigen sind, Äthiopier aber signifikant risikoaverser sind als Chinesen. Kleinhesselink und Rosa (1991) beobachten im Rahmen einer Umfrage mit japanischen und amerikanischen Studenten signifikante Unterschiede in der Risikowahrnehmung der beiden Nationalitäten. Eriksson und Simpson (2010) stellen in einem Experiment fest, dass Inder wesentlich eher bereit sind, an einer (riskanten) Lotterie teilzunehmen als Amerikaner. Nobre et al. (2016) führen eine Umfrage mit Brasilianern und Amerikanern bezüglich ihrer Risikoeinstellung in der Finanzdomäne durch. Dabei decken sie auf, dass Brasilianer risikofreudiger sind als Amerikaner und eher eine aggressivere Investmentstrategie wählen. Amerikaner hingegen sind eher bereit, an einer riskanten Lotterie teilzuneh-

men und halten mit geringerer Wahrscheinlichkeit Anleihen als sichere Geldanlage. [Cheung et al. \(2013\)](#) vergleichen die Risikoeinstellung sowie die Risikowahrnehmung von Studenten aus Hong Kong und Festlandchina und stellen dabei fest, dass Studenten aus Hong Kong in allen getesteten Domänen Risiken als vergleichsweise größer einstufen. Außerdem beobachten sie, dass Studenten vom chinesischen Festland in der sozialen Domäne sowie in der Freizeit-, Finanz- und Gesundheits-/Sicherheitsdomäne eine höhere Risikobereitschaft haben als Studenten aus Hong Kong. [Lau und Ranyard \(2005\)](#) beobachten anhand einer hypothetischen Wette auf Pferderennen, dass Chinesen risikofreudiger sind als Engländer. [Breuer et al. \(2016\)](#) vergleichen das Risikoverhalten von Chinesen und Deutschen in verschiedenen Domänen.

Sehr viele Studien vergleichen das Risikoverhalten von Amerikanern mit dem von Chinesen. So zeigen [Brumagim und Xianhua \(2005\)](#), dass sich Chinesen sowohl in der Verlust- als auch in der Gewinn-domäne risikofreudig verhalten, während sich Amerikaner gemäß den Vorhersagen der Prospect Theorie verhalten. [Wang und Fischbeck \(2004\)](#) fragen in einer Umfrage die hypothetische Zahlungsbereitschaft von Amerikanern und Chinesen für eine Krankenversicherung ab. Dabei beobachten sie, dass amerikanische Teilnehmer risikoavers in der Gewinn-domäne und risikoneutral bis risikoavers in der Verlust-domäne sind, während Chinesen risikoneutral für Gewinne und risikofreudig bei Verlusten sind. In einer weiteren Studie, die zum Ziel hat, systematische Unterschiede in der Risikowahrnehmung und Risikoeinstellung zwischen Amerikanern und Chinesen zu untersuchen, stellen [Wang und Fischbeck \(2008\)](#) fest, dass chinesische Studienteilnehmer weniger risikoavers sind als amerikanische. Auch [Statman und Weng \(2010\)](#) vergleichen Chinesen mit Amerikanern. Hierzu führen sie in der San Francisco Bay Region eine Umfrage mit 326 Amerikanern, 82 in China Geborenen sowie 89 in den USA Geborenen, deren Eltern aber aus China kommen, durch. Dabei stellen sie fest, dass Sino-Amerikaner, d. h. Amerikaner, die in China, Hong Kong oder Taiwan geboren wurden sowie deren Kinder, sich in ihrer Risikobereitschaft bei Geldanlagen nicht signifikant von Amerikanern unterscheiden, deren Eltern keine Immigranten sind. Die Autoren sehen hierfür die Assimilation von Sino-Amerikanern mit der amerikanischen Kultur als mögliche Ursache. Mit [Hsee und Weber \(1999\)](#), [Fan und Xiao \(2006\)](#) sowie [Pyles et al. \(2016\)](#) betrachten weitere Autoren Unterschiede zwischen Chinesen und Amerikanern.

Zu der Literatur, die mehrere Nationen vergleicht, gehört die Arbeit von [Statman \(2008\)](#), [Statman \(2010\)](#), [Statman \(2015\)](#), in welcher er die finanzielle Risikobereitschaft von Studenten aus mehr als 20 Ländern analysiert. [Cummings et al. \(1971\)](#) stellen in ihrer Umfrage mit Managern aus Griechenland, Spanien, Mitteleuropa, Skandinavien sowie den USA fest, dass deren selbstangegebene Einstellung gegenüber Risiko regionale Unterschiede aufweist. [Sivak et al. \(1989\)](#) analysieren in einem Experiment mit je 80 amerikanischen, spanischen, westdeutschen und brasilianischen Autofahrern Unterschiede in der Risikowahrnehmung von Verkehrssituationen. Dabei fällt auf, dass Spanier gegebene

Verkehrssituationen als am riskantesten bewerten, während die amerikanischen Teilnehmer das Risiko als am geringsten bewerten. Auch [Goszczyńska et al. \(1991\)](#) vergleichen die Risikowahrnehmung bezüglich verschiedener Risiken über mehrere Länder (Polen, Ungarn, Norwegen und USA) hinweg. US-Amerikaner schätzen dabei die Risiken durchschnittlich am höchsten ein, gefolgt von Polen, Norwegen und Ungarn. Die Autoren folgern daraus, dass die Größe des Landes und nicht etwa der soziale, wirtschaftliche oder kulturelle Hintergrund eines Landes die Risikowahrnehmung beeinflusst.

[Weber \(2014\)](#) untersucht, ob die Nationalität eines Individuums dessen generelle sowie finanzielle selbst eingeschätzte Risikoeinstellung beeinflusst. Anhand von den Umfrageergebnissen des deutschen Sozio-oekonomischen Panels von 2009 findet er heraus, dass Türken, Italiener, Griechen und Ex-Jugoslawen risikoaverser sind als andere Umfrageteilnehmer. Außerdem nutzt er die Religion und Religiosität eines Individuums, um dessen kulturellen Hintergrund zu messen. [Salacuse \(1998\)](#) stellt mit einer Umfrage mit Teilnehmern aus zwölf Ländern über deren Verhandlungsstil fest, dass zwischen Teilnehmern aus unterschiedlichen Nationen signifikante Unterschiede in der Risikobereitschaft beim Verhandeln bestehen. Während Japaner am risikoaverssten sind, zeigen Franzosen die größte Risikobereitschaft. [Marshall et al. \(2011\)](#) finden mit einer Umfrage heraus, dass Singapur und Chinesen bei finanziellen Risiken sowohl in der Verlust- als auch in der Gewinn-domäne eine größere Risikobereitschaft haben als Niederländer und Neuseeländer. In einer weiteren Umfrage stellen sie außerdem fest, dass amerikanische und neuseeländische Konsumenten sowohl in der Gewinn- als auch in der Verlust-domäne risikoaverser sind als japanische und singapurische. [Rieger et al. \(2015\)](#) sowie [Wang et al. \(2017\)](#) untersuchen mithilfe eines Experiments das Risikoverhalten von Studienteilnehmern aus 53 Ländern. [Bontempo et al. \(1997\)](#) finden in ihrem Vergleich von Studienteilnehmern aus den USA, Niederlanden, Taiwan und Hong Kong nationale Unterschiede in der Wahrnehmung finanzieller Risiken und erklären diese mit Unsicherheitsvermeidung. [Weber und Hsee \(1998\)](#) vergleichen mit einer Umfrage die finanzielle Risikobereitschaft von Polen, Chinesen, Amerikanern und Deutschen.

Andere Studien beobachten das Verhalten verschiedener ethnischer Gruppen bezüglich ihrer Entscheidungen unter Risiko. So stellen [Bellante und Green \(2004\)](#) in ihrer Arbeit fest, dass Nicht-Weiße 3,22% weniger in riskante Anlagen investieren als Weiße. [Holt und Laury \(2002\)](#) finden in ihrem Experiment heraus, dass sich hispanische Teilnehmer signifikant risikofreudiger verhalten als nicht-hispanische. Auch [Halek und Eisenhauer \(2001\)](#) zeigen, dass Menschen hispanischer Herkunft sowie Schwarze weniger risikoavers sind als Weiße und andere Ethnien. [Henrich und McElreath \(2002\)](#) beobachten in einem Experiment, dass Mapuche risikofreudiger sind als Huinca. [Rosenboim et al. \(2010\)](#) sowie [Tavor und Garyn-Tal \(2016\)](#) vergleichen das Risikoverhalten verschiedener Bevölkerungsgruppen in Israel.

## 2.2. Grundlagen und Begrifflichkeiten aus dem (Profi-) Golfsport

Da für die empirische Analyse in Kapitel 4 Daten aus dem Golfsport verwendet werden, werden in diesem Kapitel zunächst die wichtigsten Grundlagen des (professionellen) Golfsports erläutert.

Ziel im Golf ist es, mithilfe eines von insgesamt bis zu 14 Golfschlägern<sup>3</sup> den Golfball mit möglichst wenigen Schlägen vom Abschlag, auch „Tee“ genannt, einer Spielbahn ins Loch dieser Spielbahn zu befördern. Die einzelnen Spielbahnen werden daher auch Löcher genannt. Die 18 Löcher eines Golfplatzes unterscheiden sich nicht nur in ihrer Länge, sondern auch in der Platzierung von Hindernissen aus Sand, sogenannte „Bunker“, und Wasser sowie in den natürlichen Gegebenheiten (z.B. Baumbestand oder Hanglagen), welche den Weg vom Tee zum Loch für die Spieler erschweren. Darüber hinaus wird das Gras auf der Spielbahn auf unterschiedliche Höhen gemäht, was ebenfalls die Schwierigkeit des nächsten Schläges beeinflusst: Der Bereich um das Loch einer Spielbahn wird „Grün“ genannt. Hier ist das Gras am kürzesten gemäht. Der zentrale Bereich einer Bahn zwischen Abschlag und Grün ist das „Fairway“. Hier ist das Gras kurz gemäht, sodass Spieler, die einen guten Schlag gemacht haben, auch bestmögliche Voraussetzungen für ihren nächsten Schlag vorfinden. Direkt an das Fairway grenzen das „Semi-Rough“ (etwas höheres Gras) bzw. „Rough“ (hohes Gras), welche weniger kurz gemäht werden, um schlechte Schläge durch eine schlechtere Ausgangslage für den nächsten Schlag zu bestrafen. Abbildung 1 verdeutlicht den Aufbau einer Spielbahn grafisch anhand des ersten Loches des Golfplatzes TPC Sawgrass Stadium Course.

Grundsätzlich gibt es zwei verschiedene Formate, in denen professionelle Turniere auf der PGA TOUR, jener Turnierserie, auf der die besten Golfspieler der Welt antreten, ausgetragen werden. Die meisten Turniere werden als sogenanntes Zählspiel über üblicherweise vier Runden an vier aufeinanderfolgenden Tagen ausgetragen. Eine Runde besteht dabei aus den 18 Löchern des Golfplatzes. Sieger von Turnieren im Zählspielmodus ist jener Spieler, der die wenigste kumulierte Schlaganzahl benötigt, um alle Löcher aller Runden zu bewältigen. Zusätzlich gibt es bei den meisten Turnieren auf der PGA TOUR einen „Cut“ nach zwei gespielten Runden. Üblicherweise wird der Cut dabei so festgesetzt, dass jene Spieler, die nach den ersten beiden Runden schlechter platziert sind als die 70 bestplatzierten (inklusive aller Spieler, die sich den 70. Platz teilen)<sup>4</sup>, aus dem Turnier ausscheiden. Während alle Golfer, die den Cut überstehen, einen Teil des Gesamtpreisdeldes erhalten, gehen jene, die am Cut scheitern, leer aus. Die Höhe des insgesamt ausgeschütteten Preisgeldes variiert dabei zwar von Turnier zu Turnier, die Verteilung des Preisgeldes auf die einzelnen Endplatzierungen folgt jedoch

einer festen Formel. Diese Verteilungsformel resultiert in einer nicht-linearen Verteilung des Preisgeldes. So erhält der Sieger eines Turniers auf der PGA TOUR 18% des gesamten Preisgeldes und der kumulierte Anteil der Top 10-Platzierten beträgt 60,05% (PGA TOUR, 2017).

Das andere auf der PGA TOUR verwendete Format ist das sogenannte Lochspiel, bei dem jeder Golfspieler in jeder Runde im direkten Duell gegen einen anderen Athleten antritt. Im Lochspiel werden die benötigten Schläge im Gegensatz zum Zählspiel nicht über Löcher und Runden hinweg kumuliert, sondern es gewinnt derjenige Sportler das Duell, der mehr Löcher für sich entscheiden kann, indem er am entsprechenden Loch weniger Schläge benötigt als sein direkter Konkurrent. Nach einer Gruppenphase, in der die Spieler einer Gruppe jeweils ein Duell gegen die anderen Gruppenmitglieder absolvieren, kommen die jeweiligen Gruppensieger in die nächste Runde. Die Verlierer dieses Duells scheiden aus, während die Gewinner in die nächste Runde einziehen, bis letztlich der Sieger des Finalduells als Turniersieger feststeht.

Die aus dem Lochspielformat resultierende spieltheoretische Dynamik erschwert die Analyse des Risikoverhaltens von Spielern in Lochspieltournieren. Deshalb liegt der Fokus dieser Arbeit in der Analyse von Turnieren, die im Zählspielmodus ausgetragen werden.

Die Distanz zwischen Abschlag und Loch (gegebenenfalls angepasst um Faktoren wie die Neigung des Geländes) bestimmt das sogenannte „Par“ dieses Loches. Unter Par wird die Schlagzahl verstanden, die ein sehr guter Golfspieler durchschnittlich für dieses Loch benötigen sollte. Üblicherweise gibt es Par 3, Par 4 und Par 5 Löcher. Dabei sind Par 3 Löcher so angelegt, dass sehr gute Golfspieler das Grün mit dem Abschlag erreichen sollten und dann zwei „Putts“ benötigen, um den Ball auf dem Grün einzulochen. Entsprechend sind für Par 4 (Par 5) Löcher zwei (drei) Schläge zum Erreichen des Grüns sowie zwei Putts zum Einlochen vorgesehen. Benötigt ein Golfer für ein Loch einen Schlag bzw. zwei Schläge weniger als es das Par dieses Loches vorsieht, so hat er ein „Birdie“ bzw. ein „Eagle“ erzielt. Benötigt ein Spieler hingegen einen Schlag oder zwei Schläge mehr als Par, so hat er an diesem Loch ein „Bogey“ bzw. ein „Doppel-Bogey“ zu verzeichnen.

Da professionelle Golfspieler das Golfspiel besonders hervorragend beherrschen, ist es nicht unüblich, dass sie in der Lage sind, den Golfball derart weit zu schlagen, dass sie auf manchen Par 4 (Par 5) Löchern bereits mit dem ersten (zweiten) Schlag das Grün erreichen können und somit die Aussicht auf ein Birdie an diesem Loch haben. Solche Par 4 und Par 5 Löcher, an denen Golfspieler sich entscheiden, das Grün anzugreifen (d.h. versuchen, das Grün mit einem Schlag weniger als bei der Bestimmung von Par vorgesehen, zu erreichen), werden als „Going-for-it“-Löcher bezeichnet. Auch professionelle Golfer müssen jedoch bei der Schläger- und Schlagwahl zwischen Schlagweite und Schlaggenauigkeit abwägen, da weitere Schläge nur auf Kosten von Genauigkeit zu erreichen sind. Entscheidet sich ein Spieler, das Grün anzugreifen, ist somit zwar die Chance, einen Schlag gegenüber der Konkurrenz zu gewinnen, erhöht, gleichzeitig besteht je-

<sup>3</sup>Die Wahl des Golfschlägers hat einen Einfluss auf die Flugbahn des Balles sowie die zurückgelegte Schlagweite, die ein Golfer bei einem Schlag mit dem jeweiligen Schläger erreichen kann.

<sup>4</sup>Für weitere Details siehe PGA TOUR (2017).



**Abbildung 1:** Bereiche einer Spielbahn (Quelle: In Anlehnung an [Golf Digest \(2018\)](#))

doch bei einem nicht perfekt ausgeführten Schlag auch die erhöhte Gefahr, dass der Golfball in einem Wasserhindernis, Bunker oder einer anderweitig schwierigen Lage landet, was möglicherweise zu einer höheren Anzahl an benötigten Schlägen, auch Score genannt, führt, als wenn der Golfer sich nicht dazu entschieden hätte, das Grün anzugreifen. Aufgrund der nicht-linearen Verteilung der Preisgelder bei PGA TOUR-Turnieren kann jeder zusätzliche Schlag bzw. jeder Schlag weniger, den ein Spieler im Vergleich zu den anderen Teilnehmern benötigt, zu großen Unterschieden beim verdienten Preisgeld führen. Deshalb gilt es bei der Entscheidung, das Grün anzugreifen, die Chancen und Risiken sehr gut abzuwägen.

Abbildung 2 verdeutlicht anhand von zwei Par 5 Löchern des TPC Sawgrass Stadium Course die Abwägung, die Golfer bei ihrer Entscheidung, das Grün anzugreifen, machen müssen. So ist es auf diesen beiden Spielbahnen zwar möglich, bereits mit dem zweiten Schlag das Grün zu erreichen, allerdings ist hierfür ein langer zweiter Schlag nötig, welcher bei Misslingen mit großer Wahrscheinlichkeit in einem Bunker, hohem Gras oder gar in einem Wasserhindernis, welche sich unmittelbar um das Grün herum befinden, landet.

Abbildung 3 zeigt anhand des Loch 11 des TPC Sawgrass Stadium Course exemplarisch, dass sich Tiger Woods und Rickie Fowler in ihrer ersten Runde im Rahmen der Player's Championship 2018, an jenem Loch für die sicherere Variante, den Ball mit dem zweiten Schlag vorzulegen, entschieden. Phil Mickelson hingegen entschied sich für die riskantere Variante, das Grün mit dem zweiten Schlag anzugreifen. Weil Mickelsons riskanter Schlag gelang, konnte er ein Birdie auf dieser Bahn verzeichnen, während Woods nur Par spielte und Fowler aufgrund eines missratenen dritten Schlages gar ein Bogey zu verzeichnen hatte.

Löcher, an denen sich Sportler entscheiden können, das

Grün anzugreifen oder nicht, erlauben es somit, das Risikoverhalten von Golfspielern zu untersuchen. Beispielsweise [Adams und Waddell \(2018\)](#), [McFall und Rotthoff \(2018\)](#) sowie [Ozbeklik und Smith \(2017\)](#) betrachten dieses Risikomaß.

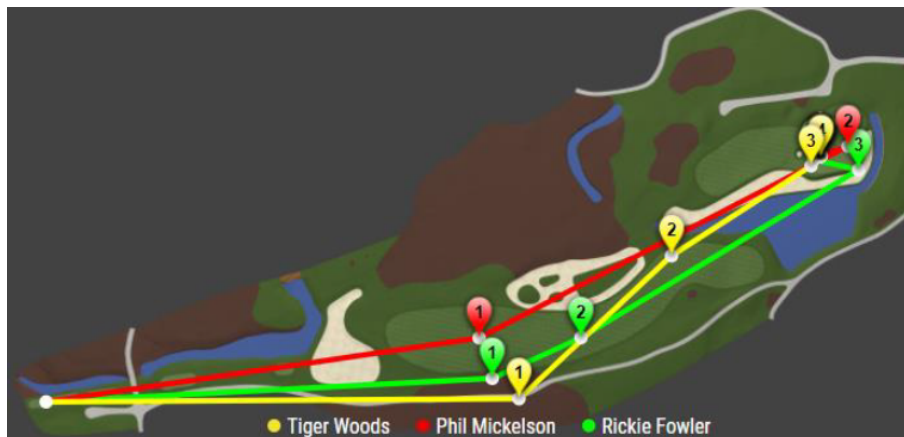
Die Standardabweichung der auf einer Runde erzielten Scores relativ zu Par zu betrachten, ist eine alternative Möglichkeit, das Risikoverhalten im Golf zu untersuchen. Missrät ein Schlag, der länger und somit gleichzeitig eine größere Streuung aufweist, hat der Spieler möglicherweise ein Bogey (oder schlechter) zu verzeichnen, während ein gelungener Schlag ein Birdie (oder besser) ermöglicht. Riskante Schläge führen also dazu, dass auch die Anzahl an benötigten Schlägen relativ zu Par über die 18 Löcher einer Runde hinweg stärker schwankt. Neben [Brown \(2011\)](#) greifen auch [Ozbeklik und Smith \(2017\)](#) auf diese Streuung als Risikomaß zurück.

Für die spätere empirische Analyse ist es wichtig, zu berücksichtigen, dass während eines Turniers ein Teil der Spieler manche Runden auf Loch 1 beginnt und somit auf Loch 18 beendet, während ein anderer Teil manche Runden auf Loch 10 startet und entsprechend auf Loch 9 beendet. Darüber hinaus ist anzumerken, dass manche Turniere nicht nur auf einem einzigen Golfplatz ausgetragen werden, sondern im Verlauf des Turniers Runden auf mehreren verschiedenen Plätzen gespielt werden. Somit besteht beispielsweise die Möglichkeit, dass ein Spieler seine erste Runde auf Golfplatz A, die zweite Runde auf Golfplatz B und die dritte sowie vierte Runde dann wieder auf Golfplatz A spielt, während ein anderer Spieler Runde eins auf Platz B, Runde zwei auf Platz A sowie die Runden drei und vier ebenfalls auf Platz A absolviert. Darüber hinaus variieren die genauen Positionen des Tees sowie des Lochs üblicherweise von Runde zu Runde, sodass beispielsweise Spielbahn 1 am ersten Tag nicht identisch

A) Loch 11 (Par 5) TPC Sawgrass Stadium Course    B) Loch 16 (Par 5) TPC Sawgrass Stadium Course



**Abbildung 2:** Going-for-it-Löcher (Quelle: In Anlehnung an Golf Digest (2018))



**Abbildung 3:** Loch 11, Runde 1 des The Players 2018 (Quelle: PGA TOUR (2018))

Anmerkung: Die Abbildung zeigt die 11. Spielbahn aus der ersten Runde des Turniers „The Players“, aus dem Jahr 2018. Die farbigen Linien in Kombination mit den entsprechenden Zahlen veranschaulichen die einzelnen Schläge von Tiger Woods, Phil Mickelson und Rickie Fowler auf diesem Going-for-it-Loch.

ist mit Spielbahn 1 der zweiten Runde.

### 2.3. Definition und Messbarkeit von Kultur

Kultur ist ein abstraktes Konstrukt, weshalb dafür viele verschiedene Definitionen existieren. Im Folgenden werden daher vier Definitionen des Begriffs Kultur vorgestellt und anschließend miteinander verglichen. Hofstede (2001, S. 9) definiert Kultur als: „the collective programming of the mind that distinguishes the members of one group or category of people from another“. Des Weiteren schreibt er, dass Kultur über mehrere Generationen hinweg stabil ist. Statman und Weng (2010, S. 38) merken an: „Culture consists of beliefs and values that ethnic, religious, and social groups transmit

from generation to generation“. Guiso et al. (2006, S. 23) definieren Kultur „as those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation“ (kursiv im Original).

Und für das GLOBE-Projekt ist Kultur definiert als (kursiv im Original):

shared motives, values, beliefs, identities, and interpretations or meanings of significant events that result from common experiences of members of collectives that are transmitted across generations (House und Javidan (2004), S. 15)

Was diese Definitionen eint, ist, dass sie Kultur als ein

Wertekonstrukt beschreiben, welches von einem Kollektiv geteilt wird und über Generationen hinweg überliefert wird.

Nachdem im Rahmen dieser Arbeit der Einfluss des kulturellen Hintergrundes eines Individuums auf dessen Risikoverhalten auf eine Art und Weise untersucht werden soll, die über das einfache Vergleichen von Ländern hinaus geht, ist es nötig, eine Möglichkeit zu finden, Kultur zu messen. Die wohl weitverbreitetste Methode, Kultur zu messen, stellen die in der Arbeit von Hofstede (2001) entwickelten nationalen Kulturdimensionen Machtdistanz, Unsicherheitsvermeidung, Individualismus, Maskulinität sowie Langzeitorientierung dar.

Hofstede (2001, S. 98) versteht unter Machtdistanz dabei (fett im Original):

The extent to which the less powerful members of institutions and organizations within a country expect and accept that power is distributed unequally.

Machtdistanz fasst also zusammen, wie mit Ungleichheit in einer Gesellschaft umgegangen wird.

Als Unsicherheitsvermeidung definiert Hofstede (2001, S. 161) (fett im Original):

The extent to which the members of a culture feel threatened by uncertain or unknown situations.

Hofstede (2001) hält zwar fest, dass Individuen aus Gesellschaften mit hoher Unsicherheitsvermeidung Aufgaben mit sicherem Ergebnis und ohne Risiko bevorzugen, weist jedoch auch explizit darauf hin, dass Unsicherheitsvermeidung nicht mit Risikovermeidung zu verwechseln ist. Außerdem beobachtet er, dass, während in Gesellschaften mit hoher Unsicherheitsvermeidung die Angst, zu versagen, im Fokus steht, es in solchen mit geringerer Unsicherheitsvermeidung die Hoffnung auf Erfolg ist.

Individualismus versteht Hofstede (2001, S. 225) folgendermaßen (fett im Original):

Individualism stands for a society in which the ties between individuals are loose: Everyone is expected to look after him/herself and her/his immediate family only. Collectivism [as its opposite] stands for a society in which people from birth onwards are integrated into strong, cohesive in-groups, which throughout people's lifetime continue to protect them in exchange for unquestioning loyalty.

Über die Dimension Maskulinität schreibt Hofstede (2001, S. 297) (fett im Original):

Masculinity stands for a society in which social gender roles are clearly distinct: Men are supposed to be assertive, tough, and focused on material success; women are supposed to be more modest, tender, and concerned with the quality of life. Femininity [as its opposite] stands for a society in which social gender roles overlap: Both men and women are supposed to be modest, tender, and concerned with the quality of life.

In Kulturen mit hoher Maskulinität sind das Ich, Geld, (Durchsetzungs-)Stärke und Leistung wichtiger, während in feminineren Gesellschaften soziale Beziehungen sowie Bescheidenheit wichtiger sind (Hofstede (2001)). Das Streben, der Beste zu sein, steht in maskulinen Gesellschaften im Vordergrund, wohingegen in femininen der Fokus darauf liegt, das zu tun, was man gerne tut (Hofstede et al., 2010).

Langzeitorientierung beschreibt Hofstede (2001, S. 359) wie folgt (fett im Original):

Long Term Orientation stands for the fostering of virtues oriented towards future rewards, in particular, perseverance and thrift. Its opposite pole, Short Term Orientation, stands for the fostering of virtues related to the past and present, in particular, respect for tradition, preservation of 'face' and fulfilling social obligations.

Zu einem späteren Zeitpunkt wurden die Kulturdimensionen mit Genuss um eine weitere Dimension ergänzt (Hofstede et al. (2010, S. 281)), welche wie folgt definiert ist (teilweise kursiv im Original):

Indulgence stands for a tendency to allow relatively free gratification of basic and natural human desires related to enjoying the life and having fun. Its opposite pole, restraint, reflects a conviction that such gratification needs to be curbed and regulated by strict social norms.

Menschen aus Kulturen mit hohen Werten für Genuss erinnern sich eher an positive Emotionen, haben eine geringere moralische Disziplin und sind optimistischer, ein niedriger Wert für diese Kulturdimension geht hingegen einher mit mehr Pessimismus sowie moralischer Disziplin (Hofstede et al. (2010)).

Ramírez und Tadesse (2009) kritisieren die Vorgehensweise von Hofstede dafür, dass unternehmensinterne Umfragen mit IBM-Angestellten als Datengrundlage dienen und nicht etwa ein Fragebogen, der auf einer Theorie basiert und von einer zufälligen Stichprobe beantwortet wurde. Des Weiteren kritisieren sie, dass die Arbeit von Hofstede, die in den 1970er Jahren entwickelt wurde, heutzutage möglicherweise nicht mehr gültig ist, da sich nationale Kulturen im Laufe der Zeit verändert haben könnten. Als eine alternative Möglichkeit, Kultur zu messen, ohne mit den oben geschilderten Problemen konfrontiert zu sein, schlagen sie deshalb die Kulturdimensionen des GLOBE-Projekts (House et al. (2004)) vor. Aus diesem Grund wird im empirischen Teil dieser Arbeit ein Robustheitstest durchgeführt, der anstatt auf die Dimensionen nach Hofstede auf die Kulturdimensionen nach dem GLOBE-Projekt zurückgreift.

## 2.4. Erwarteter Zusammenhang zwischen Kultur und Risikoverhalten

Basierend auf den Definitionen der Kulturdimensionen sowie auf bestehender Literatur werden im Folgenden kurz



die zu erwartenden Zusammenhänge zwischen den einzelnen Kulturdimensionen und dem Risikoverhalten eines Individuums beschrieben. Dabei ist es jedoch wichtig, in Erinnerung zu behalten, dass das Risikoverhalten von Individuen domänenspezifisch ist (Barseghyan et al., 2011; Breuer et al., 2016; Chetty und Szeidl, 2007; Einav et al., 2012; Hanoch et al., 2006; Rolison et al., 2014; Slovic, 1972; Tavor und Garyn-Tal, 2016; Tyagi et al., 2017; Weber et al., 2002; Wolf und Pohlman, 1983). Ergebnisse aus bisherigen Studien können deshalb nicht zwangsläufig auf das Risikoverhalten von professionellen Golfspielern im Rahmen ihrer Berufsausübung übertragen werden. Nachdem jedoch keine Theorie existiert, die erklärt, wie sich Unterschiede im Risikoverhalten zwischen verschiedenen Domänen verhalten, werden die empirischen Ergebnisse bisheriger Literatur gegebenenfalls dennoch dafür verwendet, eine Erwartung über den Zusammenhang zwischen der entsprechenden Kulturdimension und dem Risikoverhalten zu bilden.

#### 2.4.1. Machtdistanz

Wang et al. (2017) beobachten, dass ein höherer Wert für die Machtdistanz korreliert ist mit höherer Verlustaversion von Individuen. Geht man davon aus, dass Verlustaversion dazu führt, dass Individuen in Situationen, in denen sie bei Eingehen des Risikos zwar die Chance auf einen Gewinn haben, jedoch gleichzeitig die Gefahr eines Verlustes besteht<sup>5</sup>, dazu neigen, die weniger riskante Alternative zu wählen, ist anzunehmen, dass sich Individuen aus Ländern mit einem hohen Wert für Machtdistanz risikoaverser verhalten als jene aus Ländern mit geringer Machtdistanz. In Übereinstimmung damit stellen Kreiser et al. (2010) die Hypothese auf, dass sich Unternehmen aus Ländern mit hoher Machtdistanz risikoaverser verhalten, da in diesen Ländern die Betonung darauf liegt, den aktuellen Status in der sozialen Rangordnung beizubehalten. Ihre anschließende Analyse von über 1.000 Unternehmen bestätigt diese Vermutung. Auch Mihet (2013) findet heraus, dass sich Unternehmen aus Ländern mit höherer Machtdistanz risikoaverser verhalten. Sie argumentiert dabei, dass in Ländern mit hoher Machtdistanz ein latenter Konflikt zwischen den Menschen mit Macht und den Machtlosen besteht, bei dem den Machtlosen nicht vertraut werden kann. In Kombination mit dem Ergebnis von Das und Teng (2004), dass größeres Vertrauen zu geringerer Risikoaversion führt, folgert sie, dass sich Unternehmen in Ländern mit geringer Machtdistanz risikofreudiger verhalten, da mehr Vertrauen untereinander herrscht.

Auch wenn sich diese Arbeit auf das Risikoverhalten von Individuen konzentriert, sind Beobachtungen für Unternehmen nicht gänzlich irrelevant. Schließlich sind auch in Unternehmen Individuen die Entscheidungsträger. Es ist daher zu erwarten, dass ein höherer Wert für die Machtdistanz zu risikoaverserem Verhalten führt.

<sup>5</sup>Die Situation, in der ein Golfspieler entscheiden muss, ob er das Grün auf einem Going-for-it-Loch angreift, stellt eine solche Situation dar.

#### 2.4.2. Unsicherheitsvermeidung

Bontempo et al. (1997) beobachten interkulturelle Unterschiede bei der Wahrnehmung finanzieller Risiken zwischen Umfrageteilnehmern aus Hong Kong, Taiwan, USA und den Niederlanden. Die Unterschiede bestehen dabei insbesondere bei einem Vergleich der beiden westlichen Länder mit den beiden asiatischen Ländern. Die Autoren folgern daraus, dass die Kulturdimension der Unsicherheitsvermeidung einen Einfluss auf die Risikowahrnehmung und somit auch das Risikoverhalten von Individuen hat.

Wie in Kapitel 2.3 geschildert, bevorzugen Individuen aus Gesellschaften mit hoher Unsicherheitsvermeidung Aufgaben ohne Risiko und die Angst, zu versagen, steht im Fokus. Zudem beobachten Hofstede et al. (2010), dass Menschen aus Ländern mit hoher Unsicherheitsvermeidung konservativere Investitionen tätigen, während Individuen aus Ländern mit geringer Unsicherheitsvermeidung riskanter investieren. Darüber hinaus zeigen Rieger et al. (2015) in einem Experiment, dass eine stärkere Unsicherheitsvermeidung mit stärkerer Risikoaversion in der Gewinn-domäne und stärkerer Risikofreude in der Verlust-domäne einhergeht. Auch die Ergebnisse der Umfrage von Statman (2010) legen eine geringere Risikobereitschaft von Individuen aus Ländern mit einem hohen Wert für die Unsicherheitsvermeidung nahe.

Ramírez und Tadesse (2009) untersuchen das Risikoverhalten von Firmen und beobachten, dass Firmen in Ländern mit hoher Unsicherheitsvermeidung höhere Barbestände halten, was als Zeichen für eine stärkere Risikoaversion interpretiert wird. Kreiser et al. (2010) argumentieren, dass das Eingehen von Risiko ein hohes Level an Ergebnisunsicherheit generiert und erwarten deshalb risikoaverseres Verhalten in Ländern mit hoher Unsicherheitsvermeidung. Übereinstimmend damit beobachten sie risikoaverseres Verhalten von Unternehmen aus Ländern mit hoher Unsicherheitsvermeidung. Mihet (2013) findet ebenfalls heraus, dass Unternehmen aus Ländern mit geringer Unsicherheitsvermeidung risikofreudiger sind.

Basierend auf den Ergebnissen wird daher erwartet, dass sich Individuen aus Ländern mit hoher Unsicherheitsvermeidung risikoaverser verhalten.

#### 2.4.3. Individualismus

Für den Einfluss der Kulturdimension Individualismus auf das Risikoverhalten existieren zwei gegensätzliche Hypothesen. Einerseits wird in individualistischen Kulturen das Ich als Individuum betont und dementsprechend ein stärkerer Wert auf persönliche Leistungen gelegt (Hofstede (2001)). Außerdem finden Chui et al. (2010) einen positiven Zusammenhang zwischen Individualismus und übermäßigem Selbstbewusstsein. Auch Yamaguchi et al. (2005) beobachten, dass Amerikaner, Mitglieder einer individualistischen Kultur, ihre Fähigkeiten, die Kontrolle über den Ausgang eines Ereignisses zu haben, eher überschätzen als Japaner, die Teil einer kollektivistischen Kultur sind. Wie bereits in Kapitel 2.1.3 geschildert, führt übermäßiges Selbstbewusstsein dazu, dass sich Individuen risikofreudiger verhalten. Deshalb ist davon

auszugehen, dass sich Individuen aus individualistischeren Ländern risikofreudiger verhalten.

Übereinstimmend mit dieser Hypothese finden Breuer et al. (2014) mithilfe einer Umfrage heraus, dass Studienteilnehmer aus individualistischeren Ländern eine größere Risikobereitschaft bezüglich finanzieller Risiken haben und führen dies auf übermäßiges Selbstbewusstsein zurück. Des Weiteren zeigen Li et al. (2013), dass sich Unternehmen aus individualistischeren Ländern risikofreudiger verhalten. Mihet (2013) beobachtet ebenfalls risikofreudigeres Verhalten von Unternehmen, die aus individualistischeren Nationen kommen.

Andererseits finden etliche Studien heraus, dass sich Individuen aus kollektivistischeren Kulturen in der Finanzdomäne risikofreudiger verhalten (Breuer et al., 2016; Fan und Xiao, 2006; Hsee und Weber, 1999; Pyles et al., 2016; Statman, 2008, Statman, 2010, Statman, 2015; Wang et al., 2017; Weber und Hsee, 1998; Weber et al., 1998). Eine Erklärung für diese Beobachtung liefert die von Hsee und Weber (1999) entwickelte „Cushion“-Hypothese. Diese Hypothese beruht auf der Idee, dass Individuen aufgrund des engen Zusammenhalts der Mitglieder einer sozialen Gruppe, der für kollektivistische Kulturen typisch ist, in finanziellen Notsituationen auf finanzielle Unterstützung durch die anderen Gruppenmitglieder vertrauen können. Dies führt dazu, dass finanzielle Risiken von Individuen aus kollektivistischeren Gesellschaften als geringer wahrgenommen werden, was wiederum in risikofreudigerem Verhalten mündet. Da das soziale Netz eines Individuums zwar finanzielle Unterstützung leisten kann, jedoch abgesehen von seelischer Unterstützung keinerlei Hilfe bei Krankheit oder akademischen Themen leisten kann, gilt die Cushion-Hypothese jedoch nur in der Finanzdomäne. Dementsprechend betrachten Hsee und Weber (1999) das Ergebnis, dass sich in einem Experiment mit Chinesen und Amerikanern Chinesen, als Mitglieder einer kollektivistischen Kultur, nur in der Investmentdomäne risikofreudiger verhalten als Amerikaner, als Mitglieder einer individualistischen Kultur, als Beleg für ihre Cushion-Hypothese.

In Einklang mit der Cushion-Hypothese beobachtet Miyata (2003) im Rahmen eines Investmentspiels, dass Individuen, die gemeinsam mit ihren Eltern in einem Haushalt leben, risikofreudiger sind. Als Erklärung liefert sie der Idee von Hsee und Weber (1999) folgend, dass die Eltern als eine Art Sicherheitsnetz dienen. Auch Schneider et al. (2017) untersuchen den Einfluss von (privaten und staatlichen) Sicherheitsnetzen und stellen dabei fest, dass Individuen mit größeren sozialen Netzwerken eine größere Risikobereitschaft bezüglich finanzieller Risiken haben. Ein starkes staatliches soziales Sicherheitsnetz senkt hierbei die Bedeutung des sozialen Netzwerkes für die Risikobereitschaft.

Im Widerspruch zur Cushion-Hypothese stehen die Resultate von Rosenboim et al. (2010). In ihrer Arbeit vergleichen sie israelische Kibbuzmitglieder mit Stadtbewohnern und finden dabei heraus, dass Kibbuzmitglieder, die verglichen mit Stadtbewohnern in einer kollektivistischeren Gemeinschaft leben, eine geringere Risikobereitschaft bei finanziellen Ris-

ken haben. Die Autoren vermuten, dass das kollektive Management der Finanzen bei den Kibbuzmitgliedern das Finanzwissen des Einzelnen sowie dessen Erfahrung in finanziellen Angelegenheiten verringert, was zu risikoaverseren Entscheidungen führt, die das Sicherheitsnetz einer kollektivistischen Gesellschaft überwiegen. Auch Tavor und Garyn-Tal (2016) vergleichen israelische Kibbuzmitglieder als Mitglieder einer kollektivistischen Kultur mit verschiedenen anderen Bewohnergruppen in Israel. Dabei kommen sie ebenfalls zu dem Ergebnis, dass individualistischere Bewohner Israels eine größere Risikobereitschaft haben als jene einer kollektivistischen Gemeinschaft.

Außerdem beobachten Tavor und Garyn-Tal (2016), dass Kibbuzmitglieder weniger risikoavers bei Versicherung und Glücksspiel jedoch risikoaverser beim Autofahren und Sport sind. Auch Breuer et al. (2016) stellen in einer Umfrage mit chinesischen sowie deutschen Teilnehmern fest, dass Teilnehmer aus China, einer kollektivistischen Nation, zwar übereinstimmend mit der Cushion-Hypothese risikofreudiger sind als Teilnehmer aus Deutschland, einer moderat kollektivistischen Nation. In allen anderen untersuchten Domänen jedoch sind Chinesen tendenziell risikoaverser als Deutsche. Rieger et al. (2015) finden in ihrem Experiment keinen signifikanten Zusammenhang zwischen dem Risikoverhalten und der Kulturdimension Individualismus, wenn sie für Unsicherheitsvermeidung kontrollieren.

Welchen Einfluss die Kulturdimension Individualismus auf das Risikoverhalten von professionellen Golfspielern bei der Ausübung ihres Berufs hat, lässt sich daher nicht eindeutig vorhersagen und ist somit eine Frage, die es empirisch zu beantworten gilt.

#### 2.4.4. Maskulinität, Langzeitorientierung und Genuss

Wie in Kapitel 2.3 geschildert, liegt in Ländern mit hoher Maskulinität die Betonung auf Leistung und Wettbewerbsfähigkeit. Außerdem herrschen in solchen Kulturen eine starke Geld- und Vermögensorientierung sowie Sympathie für den Erfolgreichen und Starken. Deshalb ist aus theoretischer Perspektive davon auszugehen, dass ein höherer Wert für die Maskulinität mit risikofreudigerem Verhalten einhergeht.

Allerdings finden weder Mihet (2013) noch Kreiser et al. (2010) in ihren Analysen des Risikoverhaltens von Unternehmen einen signifikanten Einfluss von Maskulinität. Wang et al. (2017) hingegen beobachten, dass Individuen aus maskulineren Ländern eine stärker ausgeprägte Verlustaversion haben. Der Argumentation in Kapitel 2.4.1 folgend wäre daher zu erwarten, dass sich Individuen aus maskulineren Nationen dementsprechend risikoaverser verhalten. Letztlich gilt es also, empirisch herauszufinden, welchen Effekt die Kulturdimension Maskulinität auf das Risikoverhalten von Individuen hat.

Auch für die Kulturdimension der Langzeitorientierung ist unklar, welcher Zusammenhang mit dem Risikoverhalten eines Individuums besteht. Es ist aus theoretischer Sicht davon auszugehen, dass Individuen aus Nationen mit stark ausgeprägter Langzeitorientierung aufgrund ihres Ausdauervermögens sowie ihrer Weitsicht wohlüberlegt handeln und des-

halb keine Risiken eingehen, um nicht das zukünftige Ergebnis negativ zu beeinflussen. Hierfür spricht auch die Beobachtung von Hofstede et al. (2010), dass in Kulturen mit Kurzfristorientierung schnelle Ergebnisse erwartet werden, wohingegen in Kulturen mit Langzeitorientierung nachhaltiger Aufwand zu langsamen Ergebnissen führen soll. Der Effekt von Langzeitorientierung auf das Risikoverhalten ist jedoch bisher ein Aspekt, dem in der empirischen Literatur offensichtlich kaum Beachtung geschenkt wurde. Aus diesem Grund ist auch der Zusammenhang zwischen der Langzeitorientierung und dem Risikoverhalten eine offene Frage, die es im empirischen Teil dieser Arbeit zu beantworten gilt.

Wie der Beschreibung zur Kulturdimension Genuss in Kapitel 2.3 zu entnehmen ist, erinnern sich Menschen aus Kulturen mit hohen Werten für Genuss eher an positive Emotionen, haben eine geringere moralische Disziplin und sind optimistischer. Zudem wurde in den Kapiteln 2.1.3 sowie 2.4.3 bereits darauf hingewiesen, dass sich optimistischere Individuen risikofreudiger verhalten. Es wird daher erwartet, dass sich Individuen aus Nationen, die höhere Werte für die Genussdimension erzielen, risikofreudiger verhalten.

### 3. Daten

#### 3.1. Datensatz und Variablendefinitionen

Um den Zusammenhang von Kultur und Risikoverhalten zu analysieren, werden professionelle Golfturniere, die im Rahmen der PGA TOUR ausgetragen wurden, betrachtet. Professionelles Golf ist ein idealer Bereich für derartige Untersuchungen, da es eine Art Laborexperiment mit professionellen Teilnehmern, die reale Entscheidungen bei gleichzeitig hohen Einsätzen treffen, darstellt. Darüber hinaus sind Daten leicht verfügbar. So wurde ein Paneldatensatz von der PGA TOUR zur Verfügung gestellt, der für über 2.000 Spieler aus 53 Nationen Daten, aufgeschlüsselt bis zur Lochebene, über 481 gespielte Turniere für den Zeitraum 2008 bis 2018 beinhaltet.<sup>6</sup>

Tabelle 1 fasst für jedes Jahr dieses Zeitraums die Anzahl an unterschiedlichen Spielern, unterschiedlichen Turnieren sowie die Anzahl an insgesamt gespielten Löchern zusammen. Da für diese Arbeit vor allem jene Löcher von Interesse sind, an denen Golfer die Möglichkeit haben, das Grün anzugreifen, sind in Klammern die entsprechenden Anzahlen angegeben, wenn lediglich Going-for-it-Löcher betrachtet werden. Auch wenn lediglich knapp 16% aller beobachteten gespielten Löcher auf Beobachtungen an Going-for-it-Löchern zurückgehen, beinhaltet der Datensatz mehr als eine halbe Million Beobachtungen, die für die Analyse des Risikoverhaltens genutzt werden können.

Um den Zusammenhang von Kultur und Risikoverhalten untersuchen zu können, wird der Datensatz der PGA TOUR

mit Werten für die Kulturdimensionen nach Hofstede zusammengeführt. Diese Daten stammen von Hofstede Insights<sup>7</sup>. In einem Robustheitstest (siehe Kapitel 4.3) werden außerdem Daten für die Kulturdimensionen nach dem GLOBE-Projekt, abgerufen von der Website des GLOBE-Projekts<sup>8</sup>, verwendet. Um für die in Kapitel 2.1 beschriebenen Risikodeterminanten kontrollieren zu können, wird der Datensatz um Daten über das kaufkraftbereinigte jährliche Bruttoinlandsprodukt, welche von der World Bank Group<sup>9</sup> stammen, sowie um Werte für den Human Development Index, welche auf der Website des United Nations Development Programme<sup>10</sup> verfügbar sind, ergänzt.

Wie bereits in Kapitel 2.2 erwähnt, bieten sich zwei verschiedene Variablen an, um das Risikoverhalten von Golfern zu analysieren. Während eine Untersuchung der Entscheidung, das Grün anzugreifen, definitionsgemäß auf Going-for-it-Löcher beschränkt ist, erlaubt es eine Betrachtung der Standardabweichung der Scores eines Spielers relativ zu Par auf den 18 Löchern einer Runde, sämtliche Beobachtungen zu nutzen. Allerdings geht hierbei im Vergleich zur Untersuchung der Entscheidung, ob ein Spieler auf einem Going-for-it-Loch das Grün angreift, aufgrund der Aggregation über eine Runde hinweg Detailgenauigkeit verloren.

Als Going-for-it-Loch gelten Par 4 Löcher im Rahmen dieser Arbeit, wenn mindestens 10% der ersten Schläge aller Spieler über das gesamte Turnier hinweg, sowie mindestens 10% der ersten Schläge aller Spieler an diesem Loch in der entsprechenden Runde auf dem Grün bzw. innerhalb eines Umkreises von 30 Yards um das Grün herum landen. Par 5 Löcher hingegen werden bereits als Going-for-it-Loch gewertet, wenn mindestens ein Golfer in der entsprechenden Runde das Grün angreift, d.h. in der entsprechenden Runde ein zweiter Schlag von mindestens einem Spieler auf dem Grün, in einem Umkreis von 30 Yards um das Grün oder im Wasserhindernis landet.<sup>11</sup> Going-for-it-Löcher werden deshalb auf Rundenbasis definiert, weil die Position des Abschlags und des Loches einer Spielbahn typischerweise zwischen den einzelnen Runden variiert<sup>12</sup> und somit ein Loch, welches in einer Runde ein Going-for-it-Loch ist, nicht zwangsläufig auch in den anderen Runden ein Going-for-it-Loch sein muss. Da es für die Betrachtung der Entscheidung, das Grün anzugreifen oder nicht, leider nicht möglich ist, die tatsächliche Absicht eines Spielers zu beobachten, wird die binäre Variable angegriffen so definiert, dass sie den Wert eins annimmt, wenn der erste Schlag eines Golfers auf einem Par 4 Going-for-it-Loch oder dessen zweiter Schlag auf einem Par 5 Going-for-it-Loch auf dem Grün oder in einem Umkreis von 30 Yards um

<sup>7</sup>Siehe <https://www.hofstede-insights.com/product/compare-countries/>

<sup>8</sup>Siehe <https://globeproject.com/data/GLOBE-Phase-2-Aggregated-Societal-Culture-Data.xls>

<sup>9</sup>Siehe <https://data.worldbank.org/indicator/NY.GDP.PCAR.PPKD>

<sup>10</sup>Siehe <http://hdr.undp.org/en/data>

<sup>11</sup>Par 3 Löcher können definitionsgemäß kein Going-for-it-Loch sein, da Par an diesen Löchern bereits vorsieht, dass der erste Schlag auf dem Grün landet

<sup>12</sup>Vgl. hierzu Kapitel 2.2.

<sup>6</sup>Nach Bereinigung der Daten. Für die Beschreibung der Vorgehensweise zur Bereinigung der Daten siehe Anhang A.

**Tabelle 1:** Überblick über den Datensatz

Jahr		Spieler	Turniere	gespielte Löcher
2008	Alle Beobachtungen	553	45	331. 344
	(nur Going-for-it-Löcher)	(453)	(37)	(52.103)
2009	Alle Beobachtungen	567	43	314. 568
	(nur Going-for-it-Löcher)	(461)	(36)	(50.832)
2010	Alle Beobachtungen	604	44	319. 788
	(nur Going-for-it-Löcher)	(465)	(38)	(51.373)
2011	Alle Beobachtungen	572	42	301. 752
	(nur Going-for-it-Löcher)	(454)	(36)	(48.865)
2012	Alle Beobachtungen	562	41	297.036
	(nur Going-for-it-Löcher)	(463)	(36)	(48.865)
2013	Alle Beobachtungen	518	36	262.836
	(nur Going-for-it-Löcher)	(429)	(32)	(42.979)
2014	Alle Beobachtungen	589	43	311. 148
	(nur Going-for-it-Löcher)	(462)	(36)	(46.948)
2015	Alle Beobachtungen	629	45	327. 060
	(nur Going-for-it-Löcher)	(501)	(38)	(52.038)
2016	Alle Beobachtungen	597	44	317. 430
	(nur Going-for-it-Löcher)	(457)	(37)	(49.219)
2017	Alle Beobachtungen	613	44	319. 770
	(nur Going-for-it-Löcher)	(532)	(38)	(52.641)
2018	Alle Beobachtungen	641	45	326. 052
	(nur Going-for-it-Löcher)	(547)	(38)	(52.647)
Gesamt	Alle Beobachtungen	2.023	472	3.428.784
	(nur Going-for-it-Löcher)	(1.604)	(402)	(547.941)

Anmerkungen: Die Tabelle fasst für jedes Jahr die Anzahl an Beobachtungen von unterschiedlichen Spielern und gespielten Turnieren sowie die entsprechende Anzahl an insgesamt gespielten Löchern zusammen. Die Zahlen in Klammern geben die entsprechenden Werte an, wenn nicht der gesamte Datensatz berücksichtigt wird, sondern lediglich jene Beobachtungen, die an Going-for-it-Löchern gemacht wurden.

das Grün landet bzw. im Falle von Par 5 Going-for-it-Löchern auch dann, wenn der zweite Schlag im Wasser landet.

Im Folgenden werden die wichtigsten Variablen, die für die Untersuchungen in dieser Arbeit verwendet werden, kurz definiert. Die Variablen Machtdistanz, Individualismus, Maskulinität, Unsicherheitsvermeidung, Langzeitorientierung und Genuss spiegeln, basierend auf der Nationalität eines Golfers, die entsprechenden Werte für die Kulturdimensionen nach Hofstede wider. Der Wertebereich jedes Indexes liegt zwischen 0 und 100. Je höher der Wert des entsprechenden Indexes ist, desto stärker ausgeprägt ist die entsprechende Kulturdimension in einem Land. Analog hierzu sind Unsicherheitsvermeidung (Praxis), Zukunftsorientierung (Praxis), Machtdistanz (Praxis), Kollektivismus I (Praxis), Humanorientierung (Praxis), Leistungsorientierung (Praxis), Kollektivismus II (Praxis), Geschlechteregalitarismus (Praxis), Durchsetzungsvermögen (Praxis), Unsicherheitsvermeidung (Wert), Zukunftsorientierung (Wert), Machtdistanz (Wert), Kollektivismus I (Wert), Humanorientierung (Wert), Leistungsorientierung (Wert), Kollektivismus II (Wert), Geschlechteregalitarismus (Wert) sowie Durchsetzungsvermögen (Wert) die der Nationalität eines Spielers entsprechenden Werte für die gesellschaftlichen Kulturdi-

mensionen nach dem GLOBE-Projekt<sup>13</sup>. Der Wertebereich liegt hier zwischen 1 und 7. Höhere Werte bedeuten dabei eine stärkere Ausprägung.

Die Variable Alter gibt das Alter eines Spielers in Jahren an. ln(Vermögen) ist definiert als die logarithmierte Summe aller offiziellen Gelder in USD, die ein Spieler seit Beginn seiner Karriere bis zum Ende der jeweils vorherigen Saison auf der PGA TOUR eingenommen hat. ln(Einkommen) ist der natürliche Logarithmus aller Gelder in USD, die ein Golfer in der Vorsaison auf der PGA TOUR, der PGA TOUR Champions, der (ehemaligen) Web.com Tour<sup>14</sup> sowie der European Tour erspielt hat. Das logarithmierte, kaufkraftbereinigte Bruttoinlandsprodukt-pro-Kopf des Herkunftslandes eines Golfers, gemessen in konstanten 2011 internationalen Dollar, definiert ln(BIP-pro-Kopf). Zwischenrang gibt die Platzierung eines Spielers am Ende der vorherigen Runde im Turnier an. Geteilt ist eine binäre Variable, die den Wert eins annimmt, wenn aufgrund von gleicher kumulierter

<sup>13</sup>Für das GLOBE-Projekt werden für jede Kulturdimension Werte dafür erhoben, wie die Kultur in den Augen von in dieser Kultur lebenden Individuen ist („Praxis“) und wie die Kultur in den Augen dieser Individuen sein sollte („Werte“) (Javidan et al., 2004).

<sup>14</sup>Im Jahr 2019 wurde die Web.com Tour umbenannt in Korn Ferry Tour.

Schlaganzahlen mehrere Golfer den gleichen Zwischenrang haben und dient somit als Proxy für die Intensität des Wettbewerbs. Rückstand gibt an, wie viele Schläge Rückstand ein Golfer nach dem Ende der vorangegangenen Runde auf den Führenden hat.

Preisgeld gibt für jedes Turnier den Betrag, gemessen in 100.000 USD, an, den der Sieger des Turniers erhält. Körpergröße misst die Größe eines Sportlers in Fuß. Verlustdomäne ist eine binäre Variable, die den Wert eins annimmt, wenn der Gesamtscore eines Spielers, gemessen zu Beginn des jeweiligen Loches, für die entsprechende Runde schlechter ist als Par, der Golfer also verglichen mit dem salienten Referenzpunkt „Par“ einen „Verlust“ erlebt. Äquivalent hierzu nimmt die Dummyvariable Gewinn-domäne den Wert eins an, wenn der bis zum betrachteten Loch kumulierte Score für die Runde zu Beginn des jeweiligen Loches unter Par liegt. Die Variable Superstar ist eine Dummyvariable, die den Wert eins annimmt, wenn Tiger Woods am Turnier teilnimmt. Major und WGC sind Indikatorvariablen, die den Wert eins annehmen, wenn ein Turnier zu den sogenannten Major-Turnieren (Masters Tournament, PGA Championship, U.S. Open und The Open Championship) respektive zu den World Golf Championship Turnieren zählt. Die Variablen Abschlag, Annäherung, Kurzspiel sowie Putting sind Variablen, welche, basierend auf der von Broadie (2008), Broadie (2012) entwickelten Statistik „Strokes Gained“<sup>15</sup>, die Leistung eines Golfers im jeweiligen Teilbereich des Golfspiels relativ zur durchschnittlichen Leistung aller Spieler auf der PGA TOUR quantifizieren und somit als Proxy für die (relativen) Fähigkeiten eines Spielers dienen. Runde 1, Runde 2, Runde 3, Runde 4 sowie Runde 5 sind Indikatorvariablen, die jeweils den Wert eins annehmen, wenn die vorliegende Beobachtung in der entsprechenden Runde eines Turniers gemacht wurde.

Während die oben definierten Variablen für beide Varianten, das Risikoverhalten von Golfern zu analysieren, von Bedeutung sind, erlaubt es eine Analyse der Entscheidung, das Grün anzugreifen, zusätzlich zu diesen Variablen, weitere potenzielle Determinanten des Risikoverhaltens zu betrachten. Gleichzeitig erfordert diese Untersuchung zusätzliche Kontrollvariablen. Deshalb werden zusätzlich folgende Variablen definiert. Die Variable Erfolgsquote ist definiert als die Anzahl an erfolgreichen Grünangriffsversuchen in der vorangegangenen Runde des Turniers, geteilt durch die Anzahl aller Versuche eines Spielers, das Grün in ebendieser Runde anzugreifen. Dabei gilt ein Golfer als Spieler, der das Grün angreift, wenn sein erster Schlag auf einem Par 4 Loch oder sein zweiter Schlag auf einem Par 5 Loch auf oder in einem Umkreis von 30 Yards um das Grün landet bzw. im Falle von Par 5 Löchern auch dann, wenn der zweite Schlag im Wasser landet. Entfernung zum Loch gibt an, welche Distanz, gemessen in Yards, ein Spieler mit seinem Schlag zurückzulegen hätte, falls er sich für die riskante Option, das Grün anzugreifen, entscheiden sollte. Die Variablen Tee, Fairway, Semi-Rough,

Rough, Wasserhindernis, Bunker sowie Sonstige<sup>16</sup> sind Indikatorvariablen, die jeweils den Wert eins annehmen, wenn der Ball vor Ausführung des Schläges, mit dem das Grün angegriffen werden könnte, im entsprechenden Bereich der Spielbahn liegt.

### 3.2. Deskriptive Statistiken

Tabelle 2 fasst die in Kapitel 3.1 definierten Variablen unter Angabe der Anzahl an Beobachtungen, Mittelwert, Standardabweichung sowie Minimum und Maximum für jene Beobachtungen, welche an Going-for-it-Löchern gemacht wurden, zusammen. Einen entsprechenden Überblick über die statistischen Merkmale jener Variablen, die für die Analyse der Standardabweichung, bei welcher die gesamte Stichprobe in aggregierter Form verwendet werden kann, relevant sind, gibt Abbildung A2 im Appendix. Sowohl bei Betrachtung der gesamten Stichprobe als auch bei Betrachtung der auf Going-for-it-Löcher begrenzten Stichprobe fällt auf, dass die Werte für die Hofstede- und GLOBE-Kulturdimensionen eine relativ geringe Variation aufweisen. Dies gilt es, bei der weitergehenden empirischen Analyse sowie der Interpretation der Ergebnisse zu berücksichtigen. Eine mögliche Ursache für die geringe Variation der Kulturdimensionen wird durch Tabelle 4 aufgezeigt.

Wie in Kapitel 2.2 beschrieben, ist zu erwarten, dass Golfer, die auf Going-for-it-Löchern die riskante Strategie – das Grün anzugreifen – wählen, im Durchschnitt einen niedrigeren Score zu verzeichnen haben, als Golfer, die sich für die sichere Strategie – den Ball vorzulegen – entscheiden. Gleichzeitig ist jedoch auch zu erwarten, dass aufgrund des größeren Risikos, die Standardabweichung für den Score in den Fällen, in denen Spieler das Grün angreifen, im Durchschnitt größer ist, als wenn sie dies nicht tun. Übereinstimmend damit zeigt Tabelle 3, dass der durchschnittliche Score eines Golfers sowohl auf Par 4 Löchern als auch auf Par 5 Löchern, die als Going-for-it-Löcher gelten, niedriger ist, wenn er das Grün angreift, als wenn er dies nicht tut. Wilcoxon Rangsummen-Tests (p-Wert jeweils 0,000) zeigen, dass diese Unterschiede statistisch signifikant sind. Außerdem ist im Fall von Par 5 Going-for-it-Löchern die beobachtete Standardabweichung für die riskantere Strategie größer, wenn auch nur marginal (0,6803 vs. 0,6797). Entgegen der Erwartungen ist jedoch im Fall von Par 4 Going-for-it-Löchern die beobachtete durchschnittliche Standardabweichung für die weniger riskante Strategie größer. Eine mögliche Erklärung hierfür stellen besonders misslungene Versuche, das Grün anzugreifen, dar. Endet ein Versuch, das Grün anzugreifen, nämlich weiter als 30 Yards vom Grün entfernt oder, im Fall von Par 4 Löchern zusätzlich, in einem beliebigen Wasserhindernis, so wird dieser Versuch aufgrund der Definition der Variable angegriffen nicht als solcher gewertet.<sup>17</sup> Gleichzeitig ist

<sup>15</sup>Ein positiver Wert für Strokes Gained bedeutet, dass die Leistung des Golfers besser war als die durchschnittliche Leistung eines Spielers auf der PGA TOUR.

<sup>16</sup>Zum Bereich Sonstige zählen beispielsweise Wege und Brücken, die die Spielbahn kreuzen.

<sup>17</sup>Auch McFall und Rotthoff (2018) merken in ihrer Arbeit an, dass eine derartige Definition der Variable angegriffen möglicherweise dazu führt, dass übermäßig viele Beobachtungen als Entscheidung für die weniger riskante Strategie, das Grün nicht anzugreifen, gewertet werden.

**Tabelle 2:** Variablenausprägungen für Going-for-it-Löcher

	N	Mittelwert	Standardabweichung	Min	Max
<b>Maß für Risikoverhalten</b>					
angegriffen	547.941	0,501	0,5	0	1
<b>Hofstede Kulturdimensionen</b>					
Machtdistanz	543.934	41,203	8,402	11	100
Individualismus	543.934	84,115	17,251	12	91
Unsicherheitsvermeidung	543.934	48,421	11,152	8	100
Maskulinität	543.934	59,97	11,274	5	100
Langzeitorientierung	541.787	31,396	15,923	13	100
Genuss	541.768	66,44	9,218	24	100
<b>GLOBE Kulturdimensionen</b>					
Unsicherheitsvermeidung (Praxis)	532.917	4,205	0,28	3,392	5,371
Zukunftsorientierung (Praxis)	532.917	4,123	0,189	3,08	5,068
Machtdistanz (Praxis)	532.917	4,938	0,201	3,894	5,667
Kollektivismus I (Praxis)	532.917	4,274	0,276	3,246	5,224
Humanorientierung (Praxis)	532.917	4,131	0,207	3,23	5,12
Leistungsorientierung (Praxis)	532.917	4,408	0,202	3,204	4,935
Kollektivismus II (Praxis)	532.917	4,329	0,371	3,529	6,362
Geschlechteregalitarismus (Praxis)	532.917	3,349	0,191	2,497	3,926
Durchsetzungsvermögen (Praxis)	532.917	4,433	0,262	3,376	4,624
Unsicherheitsvermeidung (Wert)	532.917	4,06	0,254	3,158	5,61
Zukunftsorientierung (Wert)	532.917	5,319	0,174	4,331	6,203
Machtdistanz (Wert)	532.917	2,803	0,135	2,041	3,529
Kollektivismus I (Wert)	532.917	4,231	0,238	3,9	5,62
Humanorientierung (Wert)	532.917	5,526	0,107	4,485	5,807
Leistungsorientierung (Wert)	532.917	6,07	0,198	5,172	6,452
Kollektivismus II (Wert)	532.917	5,764	0,141	4,943	6,253
Geschlechteregalitarismus (Wert)	532.917	5,004	0,199	3,675	5,172
Durchsetzungsvermögen (Wert)	532.917	4,179	0,326	2,812	5,564
<b>Determinanten von Risikoverhalten</b>					
Alter	546.573	34,281	6,976	14,508	67,369
(Alter) <sup>2</sup>	546.573	1.223,834	500,152	210,488	4.538,575
ln(Vermögen)	503.148	15,25	1,699	7,867	18,517
ln(Einkommen)	473.813	13,602	1,233	6,706	16,383
ln(BIP-pro-Kopf)	547.135	10,693	0,409	7,398	11,161
Erfolgsquote	318.293	0,218	0,319	0	1
Zwischenrang	378.684	46,386	34,902	1	180
(Zwischenrang) <sup>2</sup>	378.684	3.369,76	4.621,367	1	32.400
Geteilt	378.684	0,969	0,172	0	1
Rückstand	378.684	7,855	3,77	0	38
Preisgeld	547.941	11,511	2,746	5,4	21,6
Körpergröße	508.210	5,983	0,19	5,083	6,583
Verlustdomäne	547.941	0,35	0,477	0	1
Gewinndomäne	547.941	0,369	0,482	0	1
<b>Golfspezifische Kontrollvariablen</b>					
Superstar	547.941	0,185	0,389	0	1
Major	547.941	0,021	0,142	0	1
WGC	547.941	0,037	0,19	0	1

(Continued)

Table 2—continued

Abschlag	547.920	0,002	0,747	-10,112	4,87
Annäherung	547.920	0,003	1,054	-9,658	5,346
Kurzspiel	547.920	0,001	0,669	-6,729	3,823
Putting	547.920	0,358	3,209	-14,709	13,8
Runde 1	547.941	0,309	0,462	0	1
Runde 2	547.941	0,31	0,463	0	1
Runde 3	547.941	0,191	0,393	0	1
Runde 4	547.941	0,187	0,39	0	1
Runde 5	547.941	0,003	0,053	0	1
Entfernung zum Loch	547.941	275,341	42,625	4	625
(Entfernung zum Loch) <sup>2</sup>	547.941	77.629,56	23.998,67	16	390.625
Tee	547.941	0,147	0,354	0	1
Fairway	547.941	0,508	0,5	0	1
Semi-Rough	547.941	0,048	0,215	0	1
Rough	547.941	0,198	0,398	0	1
Wasserhindernis	547.941	0,007	0,081	0	1
Bunker	547.941	0,058	0,234	0	1
Sonstige	547.941	0,034	0,182	0	1

Anmerkungen: Die Tabelle gibt für alle Variablen die Anzahl an Beobachtungen (N), den Mittelwert, die Standardabweichung sowie Minimum und Maximum für jene Beobachtungen an, welche an Going-for-it-Löchern gemacht wurden.

die Wahrscheinlichkeit eines besonders hohen Scores jedoch gerade nach einem besonders misslungenen Schlag erhöht. Dies führt einerseits dazu, dass die Standardabweichung des Scores – insbesondere im Fall von Par 4 Löchern, auf welchen es ein Wasserhindernis gibt – für die riskante Strategie im Vergleich zur sichereren Strategie unterschätzt wird. Andererseits würde dies auch erklären, wieso die höchsten beobachteten Schlagzahlen (10 für Par 4 Going-for-it-Löcher und 13 für Par 5 Going-for-it-Löcher) jeweils der vermeintlich sicheren Strategie zuzuordnen sind. Diese Problematik gilt es, bei der ökonomischen Analyse in Kapitel 4 zu berücksichtigen.

Da zur Beantwortung der Frage, welche Rolle der kulturelle Hintergrund eines Individuums auf dessen Risikoverhalten hat, als Zwischenschritt auf die Nationalität eines Golfers zurückgegriffen wird, wird mithilfe von Tabelle 4 zunächst untersucht, ob Sportler mit unterschiedlichen Nationalitäten unterschiedlich häufig die riskante Strategie, das Grün anzugreifen, wählen. Hierfür wird für jede Nationalität neben der absoluten Häufigkeit und des relativen Anteils dieser Nationalität in der Stichprobe der Going-for-it-Löcher auch der Mittelwert der Variable angegriffen dargestellt. Während beispielsweise venezolanische Golfer durchschnittlich in circa 68% der Fälle das Grün angreifen, wenn sich die Gelegenheit dazu bietet, entscheiden sich ihre dänischen Kollegen in lediglich etwas mehr als 42% der Fälle dazu. Dies ist ein Indiz dafür, dass die Nationalität und damit auch der kulturelle Hintergrund eines Spielers einen Einfluss auf dessen Risikoverhalten haben könnten.

Außerdem fällt bei Betrachtung der Anteile der Nationen an der auf Going-for-it-Löcher beschränkten Stichprobe auf, dass fast 70% aller Beobachtungen von Spielern mit US-

amerikanischer Nationalität stammen. Dies erklärt die oben festgestellte geringe Variation der Werte für die Kulturdimensionen. Da diese geringe Variation der Kulturdimension-Variablen zu einer unpräzisen Schätzung des Zusammenhangs von Kultur und Risikoverhalten führen kann, werden in einem Robustheitstest in Abschnitt 4.3 nur jene Golfer berücksichtigt, die nicht aus den USA stammen.

In einem nächsten Schritt werden kulturell ähnliche Länder basierend auf Gupta und Hanges (2004) zu Kulturclustern zusammengefasst. Tabelle 5 gibt für jedes dieser Kulturcluster die Anzahl an Beobachtungen, den relativen Anteil an der auf Going-for-it-Löcher begrenzten Stichprobe sowie den Mittelwert der Variable angegriffen an. Ein Vergleich der Mittelwerte für die verschiedenen Cluster legt nahe, dass Golfspieler mit unterschiedlichem kulturellem Hintergrund unterschiedlich häufig die riskante Strategie, das Grün anzugreifen, wählen. So zeigen Spieler, die basierend auf ihrer Nationalität, dem Cluster „Lateinamerika“ zugeordnet werden, das risikofreudigste Verhalten, während jene, die dem osteuropäischen Cluster zugeordnet werden, wesentlich risikoaverseres Verhalten an den Tag legen. Ein Kruskal-Wallis-Tests ( $p$ -Wert = 0,0001) zeigt, dass die Unterschiede zwischen den Clustern auch statistisch signifikant sind.

Bevor in Kapitel 4 der Einfluss von Kultur und Risikoverhalten mithilfe von Regressionen tiefergehend untersucht wird, wird als letzter Schritt dieses Kapitels der Zusammenhang zwischen den unterschiedlichen Kulturdimensionen nach Hofstede und der Entscheidung eines Spielers, das Grün anzugreifen, analysiert. Abbildung 4 gibt daher für jede der 6 Kulturdimensionen nach Hofstede den entsprechend beobachteten Mittelwert für angegriffen als Funktion des Wertes der jeweiligen Kulturdimension an. Zusätzlich gibt die rote

**Tabelle 3:** Zusammenhang zwischen Score und eingegangenem Risiko

	N	Mittelwert	Standardabweichung	Min	Max
Par 4					
angegriffen	40.732	3,603	0,67	2	8
nicht angegriffen	39.621	3,9353	0,6734	2	10
Par 5					
angegriffen	233.598	4,4395	0,6803	2	11
nicht angegriffen	233.990	4,9324	0,6797	3	13

Anmerkungen: Die Tabelle fasst den Score auf Going-for-it Löchern zusammen. Dafür werden, gruppiert nach dem Par des Loches, dem die jeweilige Beobachtung zuzuordnen ist, und unterteilt danach, ob ein Golfer das Grün angegriffen hat oder nicht, die Anzahl an Beobachtungen (N), der Mittelwert, die Standardabweichung sowie Minimum und Maximum für den auf dem entsprechenden Loch erzielten Score angegeben.

Linie jeweils den linearen Trend für diese Funktionen an. Eine Betrachtung der Trends für die einzelnen Kulturdimensionen legt nahe, dass die Dimensionen Unsicherheitsvermeidung und Machtdistanz einen leicht positiven Einfluss auf die Entscheidung, das Grün auf Going-for-it-Löchern anzugreifen, hat. Im Gegensatz hierzu scheinen Golfer aus Ländern mit hohem Individualismus weniger risikofreudig zu sein und das Grün deshalb weniger oft anzugreifen. Die Kulturdimensionen Maskulinität, Genuss und Langzeitorientierung scheinen gemessen an der durchschnittlichen Häufigkeit, das Grün anzugreifen, keinen Einfluss auf das Risikoverhalten zu haben.

#### 4. Ökonometrische Analyse

##### 4.1. Spezifikation und Methode

Die deskriptiven Ergebnisse aus dem vorherigen Kapitel deuten darauf hin, dass der kulturelle Hintergrund eines Spielers, gemessen an den der Nationalität eines Golfers entsprechenden Werten der Kulturdimensionen nach Hofstede, einen Einfluss auf dessen Risikoverhalten hat. Die folgenden beiden Regressionsmodelle dienen als Basis für eine tiefergehende Untersuchung dieses Zusammenhangs in den folgenden Kapiteln. Gleichung (1) beschreibt hierfür das Basisregressionsmodell für die Analyse des Risikoverhaltens, gemessen anhand der Beobachtung, ob ein Golfer das Grün an einem Going-for-it-Loch angegriffen hat oder nicht.

$$\begin{aligned}
 \text{angegriffen}_{i,jtrl} = & \alpha + \beta \text{Kultur}_i \\
 & + \gamma \text{Risikodeterminanten}_{i,jtrl} \\
 & + \delta G_{i,jtrl} + \zeta \text{Leistung}_{i,jt} \\
 & + \eta_1 \text{Entfernung zum Loch}_{i,jtrl} \\
 & + \eta_2 (\text{Entfernung zum Loch})^2_{i,jtrl} \\
 & + \theta \text{Lage}_{i,jtrl} + \vartheta \text{Runde}_{i,jtrl} \\
 & + \phi \text{Jahr}_{jtrl} + c_i + \epsilon_{i,jtrl}
 \end{aligned} \quad (1)$$

Die binäre Variable  $\text{angegriffen}_{i,jtrl}$  nimmt dabei, wie in Kapitel 3.1 definiert, den Wert eins an, wenn Golfer  $i$  im Jahr  $j$  im Turnier  $t$  in Runde  $r$  auf Loch  $l$  das Grün angegriffen

hat.  $\text{Kultur}_i$  ist ein Vektor, der für Spieler  $i$  die Werte für die Kulturdimensionen nach Hofstede (Machtdistanz, Individualismus, Maskulinität, Unsicherheitsvermeidung, Langzeitorientierung und Genuss) beinhaltet. Somit liegt das primäre Interesse dieser Arbeit darin, Schätzer für den Vektor  $\beta$  zu erhalten.

Die Literaturanalyse in Kapitel 2.1 zeigt, dass das Alter eines Individuums (möglicherweise in nicht-linearer Weise), dessen Vermögen und Einkommen sowie das BIP-pro-Kopf für das Land, in dem ein Individuum lebt, Determinanten des Risikoverhaltens sind. Deshalb werden die Variablen  $\text{Alter}$ ,  $\text{Alter}^2$ ,  $\ln(\text{Vermögen})$ ,  $\ln(\text{Einkommen})$ ,  $\ln(\text{BIP-pro-Kopf})$  als Kontrollvariablen aufgenommen. Außerdem deuten die Ergebnisse der Literaturanalyse darauf hin, dass der Zwischenrang eines Spielers (möglicherweise in nicht-linearer Weise), die Intensität des Wettbewerbs, der Rückstand auf den Führenden, das Preisgeld, welches auf dem Spiel steht, sowie etwaige Erfahrungen über Erfolge/Misserfolge, die bei vorherigen Entscheidungen gemacht wurden, das Risikoverhalten beeinflussen könnten. Aus diesem Grund dienen Zwischenrang,  $\text{Zwischenrang}^2$ ,  $\text{Geteilt}$ , Rückstand, Preisgeld und Erfolgsquote als weitere Kontrollvariablen. Da [Dohmen et al. \(2011\)](#) herausfinden, dass größere Individuen eine größere Risikobereitschaft haben und die Körpergröße eines Golfspielers gleichzeitig auch Auswirkungen auf dessen physische Grundvoraussetzungen für das Golfspiel hat, wird Körpergröße ebenfalls als Variable mit aufgenommen. Nachdem die Prospect-Theorie zudem vorhersagt, dass das Risikoverhalten davon abhängt, ob sich ein Individuum in der Gewinn- oder Verlustdomäne gegenüber einem salienten Referenzpunkt befindet, werden auch die Variablen  $\text{Verlustdomäne}$  und  $\text{Gewinndomäne}$  als Kontrollvariablen mit aufgenommen. Der Vektor  $\text{Risikodeterminanten}$  fasst diese Kontrollvariablen für die in Kapitel 2.1 beschriebenen Determinanten des Risikoverhaltens von Individuen zusammen.

[Hvide und Kristiansen \(2003\)](#) sagen in einem theoretischen Modell über das Risikoverhalten von Individuen in Wettbewerben voraus, dass eine höhere Qualität des Teilnehmerfeldes zu riskanterem Verhalten führt. Außerdem beobachten [Genakos und Pagliero \(2012\)](#), dass sich Athleten risikofreudiger verhalten, je prestigeträchtiger und intensiver ein Wettbewerb ist. Da die vier Major-Turniere und insbeson-



**Tabelle 4:** Risikoverhalten nach Nationalität

Nationalität	N	Anteil in %	Mittelwert angegriffen
Argentinien	6.604	1,21	0,571
Australien	39.468	7,21	0,513
Belgien	465	0,08	0,619
Brasilien	489	0,09	0,464
Chile	257	0,05	0,533
China	628	0,11	0,546
Costa Rica	9	0,00	0,222
Dänemark	974	0,18	0,423
Deutschland	3.280	0,60	0,512
England	15.680	2,86	0,494
Fidschi	2.138	0,39	0,572
Finnland	85	0,02	0,412
Frankreich	415	0,08	0,506
Französisch-Polynesien	4	0,00	0,75
Griechenland	12	0,00	0,333
Indien	2.103	0,38	0,512
Irland	2.676	0,49	0,549
Island	4	0,00	0,5
Italien	1.425	0,26	0,51
Japan	5.538	1,01	0,49
Kanada	12.325	2,25	0,477
Kolumbien	2.563	0,47	0,522
Malaysia	37	0,01	0,324
Mexiko	1.408	0,26	0,551
Namibia	19	0,00	0,526
Neuseeland	3.639	0,66	0,442
Niederlande	233	0,04	0,451
Nordirland	2.865	0,52	0,582
Norwegen	223	0,04	0,583
Österreich	293	0,05	0,444
Paraguay	502	0,09	0,436
Philippinen	41	0,01	0,585
Portugal	8	0,00	0,5
Puerto Rico	60	0,01	0,35
Schottland	4.627	0,84	0,529
Schweden	13.934	2,54	0,502
Schweiz	10	0,00	0,1
Simbabwe	2.699	0,49	0,504
Singapur	8	0,00	0,375
Slowakei	2.448	0,45	0,584
Spanien	4.114	0,75	0,593
Südafrika	13.778	2,52	0,498
Südkorea	13.723	2,51	0,479
Taipei	579	0,11	0,546
Thailand	574	0,10	0,502
United States	382.321	69,80	0,496
Venezuela	1804	0,33	0,68
Wales	629	0,11	0,534

Anmerkungen: Die Tabelle gibt basierend auf der auf Going-for-it-Löcher beschränkten Stichprobe für jede Nation die absoluten sowie relativen Häufigkeiten dieser Nation in der Stichprobe an. Außerdem wird für jede Nation der durchschnittliche Wert für die Variable angegriffen, den Golfer aus diesem Land erzielt haben, angegeben.

**Tabelle 5:** Risikoverhalten nach Kulturcluster

Cluster	N	Anteil in %	Mittelwert angegriffen
Anglo	469.887	88,17	0,497
Germanisches Europa	3.816	0,72	0,502
Konfuzianisches Asien	19.897	3,73	0,484
Lateinamerika	12.877	2,42	0,57
Lateineuropa	5.962	1,12	0,567
Nordisches Europa	14.993	2,81	0,496
Osteuropa	12	0,00	0,333
Sub-Sahara Afrika	2.718	0,51	0,504
Südasien	2.755	0,52	0,508

Anmerkungen: Die Tabelle gibt basierend auf der auf Going-for-it-Löcher beschränkten Stichprobe für jede Nation die absoluten sowie relativen Häufigkeiten dieser Nation in der Stichprobe an. Außerdem wird für jede Nation der durchschnittliche Wert für die Variable angegriffen, den Golfer aus diesem Land erzielt haben, angegeben.

dere auch die World Golf Championship Turniere aufgrund ihrer Qualifikationsvoraussetzungen jene Turniere im professionellen Golfsport sind, an denen nur die allerbesten Spieler teilnehmen, und diese Turniere auch zu den prestigeträchtigsten Turnieren zählen, werden deshalb die beiden Dummyvariablen WGC und Major als weitere Kontrollvariablen aufgenommen. [Brown \(2011\)](#) findet heraus, dass Golfer eine schlechtere Leistung in Turnieren zeigen, wenn der Superstar Tiger Woods im Teilnehmerfeld dieses Turniers ist. Um zu untersuchen, ob die Ursache dieser Leistungsverschlechterung darin liegt, dass Spieler in der Anwesenheit von Tiger Woods eine riskantere Strategie wählen, wird Superstar als weitere Variable mitaufgenommen. Diese (golf-)turnierspezifischen Kontrollvariablen ergeben Vektor G.

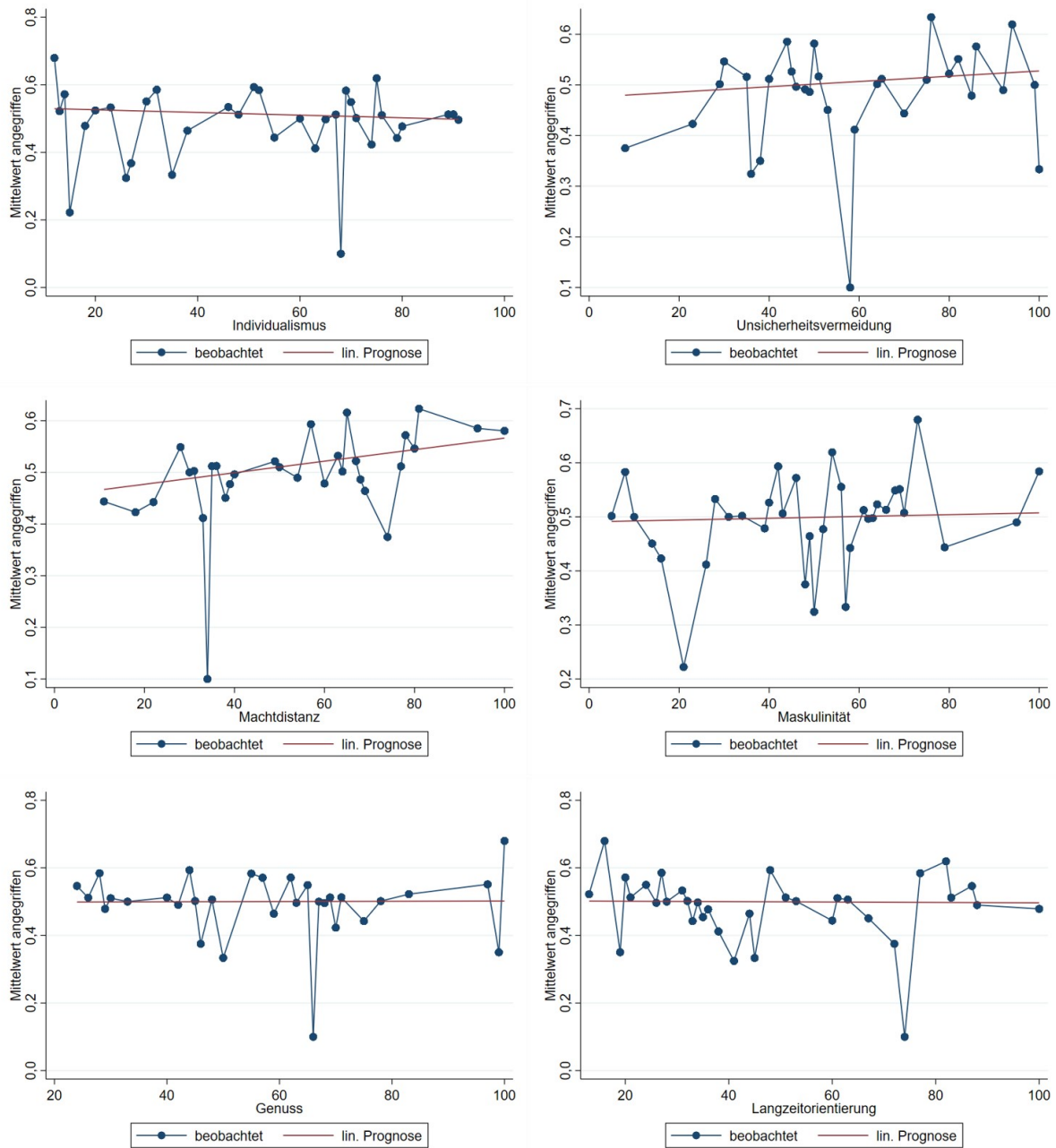
Der Vektor Leistung kontrolliert für die Fähigkeiten eines Spielers, indem er die Variablen Abschlag, Annäherung, Kurzspiel und Putting umfasst. Dabei wird die Leistung jeweils auf Turnierebene betrachtet, da die Fähigkeiten eines Spielers gemessen an dessen Leistung im Golfsport von Turnier zu Turnier sehr stark schwanken können. So gewann Tiger Woods, als Weltranglistenrster, im Jahr 2009 beispielsweise das AT&T National Turnier, scheiterte jedoch zwei Wochen darauf bereits am Cut der Open Championship und gewann wiederum weitere zwei Wochen später die Buick Open. Golfer-Fixed-Effects allein reichen also nicht aus, um für die offensichtlich zeitvarianten Fähigkeiten eines Spielers zu kontrollieren.

Wie bereits in Kapitel 2.2 geschildert, müssen Golfer bei jedem Schlag zwischen Genauigkeit und Weite abwägen. Der Grad an Risiko, den ein Spieler eingeht, hängt also primär von der Entfernung ab, die der Golfball zurücklegen müsste, um das Grün mit dem ersten Schlag auf Par 4 Löchern bzw. dem zweiten auf Par 5 Löchern zu erreichen. Dabei gilt, dass ein Schlag umso riskanter ist, je größer die Entfernung bis zum Loch ist. Gleichzeitig gibt es aus physikalisch-technischen Gründen eine Grenze dafür, wie weit ein Golfer einen Golfball maximal schlagen kann. Um für einen möglicherweise nicht-linearen Einfluss der Entfernung zum Loch zu kontrollieren, werden die Variablen Entfernung zum Loch sowie (Entfernung zum Loch)<sup>2</sup> aufgenommen.

Ein weiterer Faktor, der neben der Entfernung zum Loch bestimmt, wie riskant es ist, das Grün anzugreifen, ist die Lage des Balles. Die Lage des Balles hat nämlich einen erheblichen Einfluss darauf, wie präzise ein Golfschlag ausgeführt werden kann und welche Schlagweite erzielt werden kann. Deshalb kontrolliert der Vektor Lage, der aus den Indikatorvariablen Fairway, Semi-Rough, Rough, Wasserhindernis, Bunker und Sonstige zusammengesetzt ist, dafür, von welchem Bereich der Spielbahn aus der potenzielle Angriffsschlag aufs Grün ausgeführt werden muss. Als Referenzkategorie in den Basisregressionsmodellen gelten somit Schläge, die vom Tee aus ausgeführt werden.

Aufgrund der Gestaltung der meisten Turniere auf der PGA TOUR mit einem Cut nach Ende der zweiten Runde sowie Bestimmung des Siegers mit Ende der vierten Runde, ist es möglich, dass für Golfer, was die Wahl des Risikos betrifft, unterschiedliche Anreize je nach Runde bestehen. Auch die Ergebnisse von [McFall und Rotthoff \(2018\)](#) deuten darauf hin, dass das Risikoverhalten von Golfern abhängig von der Runde ist. Daher werden Rundendummies, welche im Vektor Runde zusammengefasst werden, verwendet. Die Dummies werden dabei so gewählt, dass Runde 3 als Referenzkategorie dient. Abschließend werden mit  $\text{Jahr}_{jtrl}$  Jahrdummies aufgenommen, um für etwaige Regeländerungen sowie technologische Entwicklungen beim Equipment, welche beide typischerweise zwischen zwei Saisons stattfinden, zu kontrollieren.  $c_i$  beschreibt einen individuen-spezifischen zeitvarianten unbeobachteten Effekt (unbeobachtete Heterogenität),  $\epsilon_{i,jtrl}$  ist der idiosynkratische Fehlerterm.

Aufgrund der in Kapitel 3.2 geschilderten Problematik bei der Analyse von der Entscheidung eines Spielers, das Grün anzugreifen, wird die Standardabweichung der Scores eines Spielers relativ zu Par auf den 18 Löchern einer Runde als alternatives Risikomaß betrachtet. Gleichung (2) beschreibt das entsprechende Basisregressionsmodell, wenn das Risikoverhalten eines Spielers anhand der Standardabweichung gemessen wird.



**Abbildung 4:** Zusammenhang zwischen Risikoverhalten und Kultur

Anmerkungen: Die blauen Punkte geben für jede der 6 Kulturdimensionen nach Hofstede (Individualismus, Unsicherheitsvermeidung, Machtdistanz, Maskulinität, Genuss und Langzeitorientierung) den beobachteten Mittelwert für angegriffen als Funktion des Wertes der jeweiligen Kulturdimension an. Zusätzlich gibt die rote Linie jeweils den linearen Trend für diese Funktionen an.

$$\begin{aligned}
 \text{Standardabweichung}_{i,jtr} = & \alpha + \beta \text{Kultur}_i \\
 & + \gamma \text{Risikodeterminanten}_{i,jtr} \\
 & + \delta G_{i,jtr} + \zeta \text{Leistung}_{i,jt} \quad (2) \\
 & + \vartheta \text{Runde}_{i,jtr} + \phi \text{Jahr}_{jtr} \\
 & + c_i + \epsilon_{i,jtr}
 \end{aligned}$$

Analog zu Gleichung (1) wird mit Standardabweichung<sub>i,jtr</sub>

die Standardabweichung des Scores für die Runde relativ zu Par von Golfer i im Jahr j im Turnier t in Runde r bezeichnet. Grundsätzlich werden dieselben Kontrollvariablen wie bei der Analyse der Entscheidung, das Grün anzugreifen, verwendet. Da die Standardabweichung jedoch nicht auf Loebene, sondern lediglich auf der nächsthöheren, aggregierten Rundenebene gemessen wird, können die Kontrollvariablen Erfolgsquote, Verlustdomäne und Gewinn domäne nicht verwendet werden. Die Kontrollvariablen Entfernung zum

Loch, (Entfernung zum Loch)<sup>2</sup> sowie der Lage-Vektor werden hingegen nicht als Kontrollvariablen aufgenommen, da diese Variablen spezifisch für die Entscheidung, das Grün anzugreifen, sind bzw. lediglich für Entscheidungen auf Schlagenebene relevant sind.

Um für unbeobachtete Heterogenität  $c_i$  zu kontrollieren, sollten für die beschriebenen Spezifikationen Fixed-Effects-Modelle geschätzt werden. Fixed-Effects-Modelle erlauben es jedoch nicht, die Parameter von individuen-spezifischen, zeitinvarianten Variablen zu schätzen. Das Hauptinteresse dieser Arbeit besteht aber gerade darin, die Parameter für den Vektor Kultur <sub>$i$</sub> , welcher per Definition zeitinvariant ist<sup>18</sup>, zu schätzen. Random-Effects-Modelle hingegen erlauben es, auch Parameter für zeitinvariante Variablen zu schätzen. Ist die grundlegende Annahme, dass die unbeobachtete Heterogenität  $c_i$  nicht mit den erklärenden Variablen korreliert ist, jedoch verletzt, sind Random-Effects-Schätzer im Gegensatz zu Fixed-Effects-Modellen inkonsistent.

Aus diesem Grund werden hybride Modelle nach Allison (2009), welche eine Mischung aus Random-Effects und Fixed-Effects darstellen, geschätzt. Bei diesen Modellen werden die Effekte der individuen-spezifischen und zeitvarianten Regressoren in sogenannte Within-Cluster-Effekte,  $\beta_W$ , und Between-Cluster Effekte,  $\beta_B$ , aufgesplittet, während die Koeffizienten für zeitinvariante Regressoren,  $\gamma$ , mithilfe von Random-Effects geschätzt werden.<sup>19</sup> Die Within-Cluster-Effekte werden dabei geschätzt, indem nur die Variation innerhalb eines Clusters verwendet wird. Analog dazu werden die Between-Cluster Effekte geschätzt, indem lediglich die Variation zwischen den unterschiedlichen Clustern genutzt wird.

#### 4.2. Ergebnisse

Die Regressionsergebnisse für hybride Modelle für Gleichungen (1) und (2) werden in Tabelle 6 zusammengefasst. Während in der ersten Spalte angegriffen die abhängige Variable ist, wird in der zweiten Spalte das Ergebnis für Gleichung (2), d. h. für die Standardabweichung als abhängige Variable, angegeben. Ein Golfer kann sich definitionsgemäß nur an Going-for-it-Löchern dazu entscheiden, das Grün anzugreifen. Entsprechend umfasst die Stichprobe für die erste Spalte alle Going-for-it-Löcher, sofern die zurückzulegende Distanz zum Loch für den Schlag, mit dem das Grün angegriffen werden müsste, maximal 355 Yards beträgt. Die zusätzliche Einschränkung nach der Entfernung zum Loch wird gemacht, da physikalische und physische Grenzen dafür bestehen, wie weit ein Spieler einen Golfball maximal schlagen kann. Ab einer gewissen Distanz stellt eine vermeintliche

Gelegenheit, das Grün anzugreifen, auch wenn es sich um ein Going-for-it-Loch handelt, deshalb keine realistische Gelegenheit mehr dar. Da 99% aller Schläge, mit denen Golfer tatsächlich versucht haben, das Grün anzugreifen, maximal aus einer Entfernung von 355 Yards zum Loch gemacht wurden, wird diese Distanz als Grenze betrachtet. Die Stichprobe für Spalte 2 hingegen umfasst alle gespielten Runden.

Während für die Analyse der Standardabweichung ein lineares hybrides Modell geschätzt wird, wird für die Untersuchung der Entscheidung, das Grün anzugreifen, ein hybrides Logit-Modell geschätzt, um der binären Struktur der Variable angegriffen Rechnung zu tragen. In beiden Fällen wird angenommen, dass die unbeobachtete Heterogenität  $c_i$  linear von den Mittelwerten der individuen-spezifischen und zeitvarianten Regressoren abhängt. Indem Spieler als Cluster betrachtet werden, werden Spieler-Fixed-Effects geschätzt. Da unter anderem Schunck und Perales (2017) anmerken, dass bei der Untersuchung von Paneldaten, wie es in dieser Arbeit der Fall ist, das Interesse primär in der Interpretation der Effekte innerhalb eines Clusters liegt, wird bei der folgenden Interpretation der Ergebnisse auch nur auf diese sowie die Random-Effects-Komponenten genauer eingegangen. Darüber hinaus merkt Schunck (2013) an, dass die Berechnung von marginalen Effekte in hybriden Modellen problematisch ist. Deshalb wird bei der Interpretation der Fokus auf die Richtung des Effekts gelegt.

Entgegen der deskriptiven Ergebnisse aus Kapitel 3.2, legen statistisch insignifikante Koeffizienten für Machtdistanz, Individualismus, Maskulinität, Unsicherheitsvermeidung, Langzeitorientierung und Genuss in beiden Spalten von Tabelle 6 nahe, dass der kulturelle Hintergrund eines Spielers, gemessen an den Werten für die Kulturdimensionen nach Hofstede seines Heimatlandes keinen Einfluss auf dessen Risikoverhalten hat und zwar unabhängig davon, ob das Risikoverhalten anhand der Standardabweichung oder anhand der Entscheidung, das Grün auf Going-for-it-Löchern anzugreifen, gemessen wird. Möglicherweise sind diese insignifikanten Schätzer von einer zu geringen Variation der erklärenden Variablen, d.h. den Kulturdimensionsvariablen, getrieben, welche zu einer unpräzisen Schätzung des Zusammenhangs von Kultur und Risikoverhalten führen kann. Um dies ausschließen zu können, werden in einem Robustheitstest in Abschnitt 4.3 nur jene Golfer berücksichtigt, die nicht aus den USA kommen, da Tabelle 4 nahelegt, dass die geringe Variation daraus resultiert, dass knapp 70% der Beobachtungen in der Stichprobe von US-amerikanischen Spielern stammen.

Betrachtet man die in Kapitel 2.1 beschriebenen Determinanten des Risikoverhaltens von Individuen, so fällt Folgendes auf: Statistisch signifikant positive Parameter für die Körpergröße eines Spielers in beiden Modellen verdeutlichen, dass größere Spieler risikofreudigeres Verhalten zeigen. Diese Beobachtung ist konsistent mit den Ergebnissen von Dohmen et al. (2011). Im Widerspruch zu den Ergebnissen von unter anderem Vielder et al. (2015b) hat das Bruttoinlandsprodukt-pro-Kopf, welches als Proxy für den Wohlstand des Heimatlandes eines Spielers dient, keinen

<sup>18</sup>Die Nationalität eines Spielers und somit auch die entsprechenden Ausprägungen für die unterschiedlichen Kulturdimensionen ändern sich nicht mit der Zeit.

<sup>19</sup>Bezeichnet  $x_{i,t}$  eine Variable, die innerhalb und zwischen Clustern variiert,  $\bar{x}_i$  den clusterspezifischen Mittelwert dieser Variable,  $d_i$  eine Variable, die nur zwischen Clustern variiert,  $a_i$  einen individuen-spezifischen zeitinvarianten unbeobachteten Effekt und  $u_{i,t}$  den idiosynkratischen Fehler, so lässt sich ein hybrides Modell für die abhängige Variable  $y_{i,t}$  durch folgende Gleichung beschreiben:  $y_{i,t} = \beta_W(x_{i,t} - \bar{x}_i) + \beta_B\bar{x}_i + \gamma d_i + a_i + u_{i,t}$

Effekt auf dessen Risikoverhalten. Dies signalisieren insignifikante Koeffizienten in beiden Modellen. Übereinstimmend mit den Ergebnissen der meisten Studien über den Zusammenhang von Alter und Risikoverhalten<sup>20</sup>, verhalten sich Golfer *ceteris paribus* risikoaverser, wenn sie älter sind verglichen mit den Zeitpunkten, zu denen sie jünger waren. Negative und statistisch signifikante Koeffizienten für die Within-Cluster-Komponente des Alters zeigen dies. Dieser Alterseffekt ist nicht-linear, wenn das Risikoverhalten anhand der Entscheidung, das Grün anzugreifen, gemessen wird, was der signifikant negative Parameter für (Alter)<sup>2</sup> in Spalte 1 zeigt. Während das Vermögen eines Spielers dessen Risikoverhalten nicht zu beeinflussen scheint, hat dessen Einkommen zumindest auf die Entscheidung, das Grün anzugreifen, einen Effekt. Denn der auf dem 1%-Niveau signifikante, positive Within-Cluster-Koeffizient für  $\ln(\text{Einkommen})$  in Spalte 1 legt nahe, dass Zeiten, in denen ein Golfer *ceteris paribus* ein höheres Einkommen hat, einhergehen mit risikofreudigerem Verhalten. Dieses Ergebnis stimmt mit dem Großteil der in Kapitel 2.1.4 vorgestellten Literatur überein.

Der statistisch signifikante, positive Within-Cluster-Koeffizient für Verlustdomäne in Modell 1 zeigt, dass sich Golfer *ceteris paribus* risikofreudiger an jenen Going-for-it-Löchern verhalten, an denen sie sich gegenüber einem salienten Referenzpunkt in der Verlustdomäne befinden als an Going-for-it-Löchern, an denen sie sich genau an diesem Referenzpunkt befinden. Befindet sich ein Spieler an einem Going-for-it-Loch hingegen in der Gewinn-domäne gegenüber diesem Referenzpunkt, so deutet der signifikante, negative Within-Cluster-Koeffizient für Gewinn-domäne darauf hin, dass sich der Sportler an jenen Löchern *ceteris paribus* risikoaverser verhält, verglichen mit Going-for-it-Löchern, an denen er sich genau am Referenzpunkt befindet. Dieses Verhalten ist konsistent mit der Prospect Theorie sowie mit den Ergebnissen von Pope und Schweitzer (2011), welche zeigen, dass Golfer Par als salienten Referenzpunkt benutzen und entsprechend abhängig davon, ob sie in der Verlust- oder Gewinn-domäne sind, unterschiedlich riskantes Verhalten zeigen.

Der statistisch signifikante, negative Within-Cluster-Koeffizient für Erfolgsquote deutet darauf hin, dass Golfprofis *ceteris paribus* das Grün seltener angreifen, wenn sie in der vorherigen Runde einen größeren Anteil an erfolgreichen Versuchen, das Grün anzugreifen, hatten. Dieses Ergebnis steht im Kontrast zu den Ergebnissen des Literaturüberblicks aus Kapitel 2.1.3.

Wie die positiven, signifikanten Within-Cluster-Koeffizienten für Zwischenrang und (Zwischenrang)<sup>2</sup> in Spalte 1 zeigen, greifen Spieler das Grün auf Going-for-it-Löchern mit größerer Wahrscheinlichkeit an, wenn sie *ceteris paribus* schlechter platziert sind, wobei diese Wahrscheinlichkeit, risikofreudigeres Verhalten zu beobachten, für schlechter platzierte Golfer stärker zunimmt. Dieses Ergebnis, dass zurückliegende Individuen mehr Risiko eingehen,

stimmt mit den Ergebnissen der meisten vorherigen Studien<sup>21</sup> überein. Eine Erklärung hierfür liefert die in Kapitel 2.2 beschriebene nicht-lineare Gestaltung der Preisgelder auf der PGA TOUR. Gelingt der Angriff auf das Grün und gewinnt der Golfer dadurch einen Schlag auf die Konkurrenz, so bedeutet dies möglicherweise eine bessere Platzierung und damit ein höheres Preisgeld. Allerdings besteht gleichzeitig die Gefahr, durch einen missratenen Versuch, das Grün anzugreifen, einen oder gar mehrere Schläge auf die Mitstreiter zu verlieren und somit auf eine schlechtere Platzierung zurückzufallen. Die nicht-lineare Preisgeldgestaltung sorgt dafür, dass die Chance im Vergleich zur Gefahr umso attraktiver wird, je schlechter platziert ein Spieler ist. Wird das Risikoverhalten allerdings mithilfe der Variable Standardabweichung gemessen, so hat die Zwischenplatzierung eines Spielers keinen signifikanten Einfluss auf dessen Risikoverhalten.

Eine Betrachtung der geschätzten Within-Cluster-Koeffizienten für Geteilt zeigt, dass höhere Intensität des Wettbewerbs, übereinstimmend mit den Ergebnissen von Genakos und Pagliero (2012), mit einem leicht signifikant risikofreudigeren Verhalten einhergeht. Allerdings ist es im Golf ohnehin die Regel, dass sich mindestens zwei Spieler aufgrund von Schlaggleichheit eine Platzierung teilen. So zeigt Tabelle 2, dass sich in ca. 97 % aller Beobachtungen an Going-for-it-Löchern mindestens zwei Spieler einen Zwischenplatz teilen. Im Widerspruch zu den Ergebnissen von Grund et al. (2013) sowie von Ozbeklik und Smith (2017) stehend, zeigen die Koeffizienten für Rückstand, dass Golfer nicht mehr Risiko eingehen, je größer ihr Rückstand ist. Wird das Risikoverhalten anhand der Standardabweichung gemessen (Spalte 2), so scheinen Spieler sogar weniger riskante Strategien zu wählen, je größer ihr Rückstand ist.

Der statistisch signifikante, negative Within-Cluster-Koeffizient für Preisgeld in Spalte 1 verdeutlicht, dass Golfer mit geringerer Wahrscheinlichkeit das Grün auf einem Going-for-it-Loch angreifen, je höher das Preisgeld ist, welches auf dem Spiel steht. Dieses Ergebnis stimmt mit der in Kapitel 2.1.6 analysierten Literatur überein, welche feststellt, dass Individuen risikoaverser sind, je höher die Beträge sind, die auf dem Spiel stehen. Im Gegensatz hierzu weist ein positiver, signifikanter Koeffizient in Spalte 2 darauf hin, dass Golfspieler in den Turnieren, in denen das Preisgeld höher ist, riskantere Strategien wählen. Möglicherweise sind die Turniere, in denen besonders hohe Preisgelder ausgeschüttet werden, zugleich jene Turniere, in denen auf besonders schweren Golfplätzen gespielt wird. Somit ist eine höhere Standardabweichung nicht auf risikofreudigeres Verhalten der Spieler, sondern auf die Schwierigkeit der Löcher zurückzuführen.

Während der geschätzte Within-Cluster-Effekt für Superstar signifikant negativ ist für Modell 1, ist er signifikant positiv in Modell 2. Somit greift ein Spieler *ceteris paribus* mit geringer Wahrscheinlichkeit das Grün auf einem Going-for-it-Loch an, wenn Tiger Woods im Teilnehmerfeld eines Tur-

<sup>20</sup>Vergleiche hierzu Kapitel 2.1.5

<sup>21</sup>Vergleiche hierzu Kapitel 2.1.7.

niers ist. Dies suggeriert, dass, übereinstimmend mit den Ergebnissen von McFall und Rothhoff (2018), die Anwesenheit des Superstars dazu führt, dass sich andere Golfer risikoaverser verhalten. Der geschätzte Effekt für das Modell der Standardabweichung des Scores relativ zu Par legt dagegen nahe, dass sich andere Turnierteilnehmer bei der Wahl ihrer Strategie risikofreudiger verhalten, wenn der Superstar im Teilnehmerfeld ist. Es bleibt festzuhalten, dass die Anwesenheit von Tiger Woods in einem Turnier mit einer Änderung der Strategie anderer Spieler einhergeht.

Statistisch signifikante, positive Koeffizienten für Major sowie WGC in Spalte 1, zeigen, dass Golfer in Turnieren, die prestigeträchtiger sind und in denen die Qualität des Teilnehmerfeldes höher ist, das Grün auf Going-for-it-Löchern mit größerer Wahrscheinlichkeit angreifen. Somit stimmen die Ergebnisse mit Genakos und Pagliero (2012) überein und belegen gleichzeitig die Vorhersage des theoretischen Modells von Hvide und Kristiansen (2003) empirisch. Die entsprechenden Koeffizienten für die Standardabweichung als abhängige Variable stehen jedoch im Widerspruch dazu. So sind die geschätzten Within-Cluster-Koeffizienten in Spalte 2 negativ, wobei lediglich der Koeffizient für World Golf Championship Turniere statistisch signifikant ist.

Die Within-Cluster-Parameterschätzer für Abschlag, Annäherung, Kurzspiel und Putting haben im Modell für die Entscheidung, das Grün anzugreifen, die erwarteten Vorzeichen. Je besser die Abschläge und die Annäherungen eines Golfers in einem Turnier sind, desto mehr Ballkontrolle hat er bei der Ausübung seiner Schläge. Hat ein Spieler mehr Kontrolle über einen Schlag, so sinkt – bei gegebenen Rahmenbedingungen – auch das Risiko dieses Schlages. Ein Sportler greift in diesem Fall das Grün also nicht deshalb mit größerer Wahrscheinlichkeit an, weil er risikofreudiger ist, sondern lediglich deshalb, weil das subjektive Risiko des Schlages bei gegebener Risikoeinstellung geringer ist. Diese Argumentation wird dadurch bekräftigt, dass mit Abschlägen und Annäherungen nur die beiden Teilleistungen einen signifikanten Effekt haben, die bei Schlägen, mit denen auf Going-for-it-Löchern das Grün angegriffen werden kann, von besonderer Bedeutung sind. Auf Par 4 Going-for-it-Löchern wird das Grün nämlich mit dem Abschlag angegriffen und auf Par 5 Löchern mit einem Annäherungsschlag. Bei der Interpretation der Koeffizienten für die vier Fähigkeitsproxys im Modell mit Standardabweichung als abhängige Variable gilt es zu berücksichtigen, dass die Standardabweichung nicht nur ein Risikomaß, sondern auch ein Maß für die Leistung eines Spielers ist. Spielt ein Golfer besser, so sollte seine Leistung konstanter und damit auch die Standardabweichung seines Scores relativ zu Par geringer sein. Insofern ist es nicht verwunderlich, dass die Vorzeichen der Fähigkeitsproxys in Spalte 2 signifikant negativ sind.

Die Vorzeichen der Koeffizienten für jene Variablen, die spezifisch für die Entscheidung, das Grün anzugreifen, sind, stimmen mit den theoretischen Überlegungen aus Kapitel 4.1 überein. So zeigen statistisch signifikant negative Schätzer für Entfernung zum Loch und  $(\text{Entfernung zum Loch})^2$ , dass, je größer die Entfernung zum Loch ist, die ein Golfer

mit seinem Grünangriffsschlag zurückzulegen hat, es umso unwahrscheinlicher ist, dass dieser Golfer sich letztlich dazu entscheidet, das Grün anzugreifen. Außerdem zeigen die geschätzten Koeffizienten für die Indikatorvariablen Fairway, Semi-Rough, Rough, Wasserhindernis, Bunker und Sonstige, welche beschreiben, von welchem Bereich der Spielbahn aus der Angriffsschlag aufs Grün auszuführen ist, dass die Wahrscheinlichkeit, dass ein Golfer das Grün angreift, größer ist, je besser die Lage des Balles für diesen Schlag ist. Wie bereits in Kapitel 2.2 geschildert wurde, nimmt die Höhe des Grasses von Fairway über Semi-Rough hin zum Rough immer weiter zu, was einen sauberen Kontakt zwischen Schläger und Ball und somit einen kontrollierten Schlag zunehmend erschwert. In Bunkern erschwert Sand, der den Schlägerkopf bei der Ballberührung sehr stark abbremst, die Kontrolle und in Wasserhindernissen ist es häufig gar unmöglich, den Ball zu spielen. Beachtet man, dass die Indikatorvariablen so gewählt wurden, dass das Tee, von wo aus Golfer den Schlag am besten kontrollieren können, da hier der sauberste Ballkontakt möglich ist, als Referenzkategorie dient, lässt sich folgende Reihenfolge bezüglich der Wahrscheinlichkeit, dass ein Spieler das Grün auf einem Going-for-it-Loch angreift, festhalten. Am wahrscheinlichsten greift ein Spieler das Grün auf einem Going-for-it-Loch an, wenn der Ball dafür vom Tee gespielt werden kann, gefolgt vom Fairway. Auf Platz drei folgen Schläge, die vom Semi-Rough aus ausgeführt werden müssen, auf Platz vier jene aus dem Rough, gefolgt von Schlägen aus sonstigen Bereichen sowie aus Bunkern. Am unwahrscheinlichsten greift ein Golfer das Grün an, wenn sein Ball für diesen Schlag in einem Wasserhindernis liegt. Diese Reihenfolge stimmt mit der oben erläuterten Reihenfolge der Kontrollierbarkeit eines Schlages überein.

Darüber hinaus weisen die in beiden Modellen signifikanten Koeffizienten für die Dummies Runde 2 und Runde 4 sowie ein signifikant positiver Parameter für Runde 5 in Spalte 2 darauf hin, dass Golfspieler nach überstandenen Cut, welcher nach Ende der zweiten Runde gemacht wird, risiko-freudigeres Verhalten zeigen. Dabei nimmt die Bereitschaft, Risiko einzugehen, zu, je weiter fortgeschritten das Turnier ist<sup>22</sup>. Eine mögliche Erklärung für dieses Muster ist, dass Golfer in ihrer zweiten Runde nicht riskieren möchten, den Cut durch riskante Entscheidungen, zu verpassen, da nur jene Spieler, die den Cut überstehen, auch ein Preisgeld erhalten. Hat ein Sportler den Cut erst einmal überstanden und somit auch ein Preisgeld sicher, so führt möglicherweise die nicht-lineare Verteilung der Preisgelder dazu, dass er, je näher das Ende des Turniers rückt, mehr Risiko eingeht, um Sprünge nach weiter vorne im Klassement zu schaffen und somit ein höheres Preisgeld zu gewinnen.

Abschließende regressionsbasierte Hausman Tests (p-Wert jeweils 0,0000) deuten darauf hin, dass für beide Risikomaße ein hybrides Modell einem einfachen Random-Effects-Modell vorzuziehen ist.

<sup>22</sup>Anmerkung: Runde 3 dient als Referenzkategorie. Beobachtungen aus Runde 1 fließen nicht mit in die Analyse ein, da die Variable Erfolgsquote so definiert ist, dass diese in Runde 1 fehlende Werte aufweist.

**Tabelle 6:** Determinanten des Risikoverhaltens von Golfspielern

Basismodell		
	(1)	(2)
Abhängige Variable	angegriffen	Standardabweichung
Random-Effects		
Machtdistanz	0,00063 (0,00310)	0,00008 (0,00022)
Individualismus	-0,00013 (0,00240)	-0,00013 (0,00012)
Maskulinität	-0,00156 (0,00171)	-0,00005 (0,00010)
Unsicherheitsvermeidung	0,00231 (0,00235)	-0,00002 (0,00012)
Langzeitorientierung	0,00115 (0,00177)	-0,00005 (0,00009)
Genuss	0,00153 (0,00248)	-0,00008 (0,00017)
Körpergröße	0,20051*** (0,07559)	0,02224*** (0,00438)
ln(BIP-pro-Kopf)	-0,01704 (0,06330)	0,00261 (0,00331)
Within-Cluster Effekte		
Alter	-0,20610*** (0,04054)	-0,01013*** (0,00289)
(Alter) <sup>2</sup>	-0,00111*** (0,00040)	0,00002 (0,00003)
ln(Vermögen)	-0,01715 (0,01392)	0,00119 (0,00093)
ln(Einkommen)	0,03840*** (0,00965)	0,00103 (0,00075)
Verlustdomäne	0,15488*** (0,01574)	
Gewinndomäne	-0,03893*** (0,01412)	
Erfolgsquote	-0,05808*** (0,01922)	
Zwischenrang	0,00141** (0,00068)	-0,00006 (0,00006)
(Zwischenrang) <sup>2</sup>	0,00002*** (0,00000)	-0,00000 (0,00000)
Geteilt	0,06663* (0,03700)	0,00257 (0,00309)
Rückstand	0,00370 (0,00294)	-0,00270*** (0,00026)
Preisgeld	-0,04477*** (0,00338)	0,00287*** (0,00028)
Superstar	-0,15957*** (0,01976)	0,00478*** (0,00154)
Major	0,42703*** (0,04655)	-0,00669 (0,00467)

(Continued)

Table 6—continued

WGC	0,28243*** (0,03300)	-0,02689*** (0,00275)
Abschlag	0,22114*** (0,01207)	-0,03826*** (0,00105)
Annäherung	0,08486*** (0,00786)	-0,02331*** (0,00070)
Kurzspiel	0,01192 (0,01093)	-0,03625*** (0,00096)
Putting	0,00364 (0,00241)	-0,00415*** (0,00021)
Entfernung zum Loch	-0,03250*** (0,00277)	
(Entfernung zum Loch) <sup>2</sup>	-0,00003*** (0,00001)	
Fairway	-1,81780*** (0,03104)	
Semi-Rough	-2,59501*** (0,04106)	
Rough	-4,09621*** (0,03983)	
Wasserhindernis	-9,98317*** (0,42518)	
Bunker	-5,75381*** (0,05018)	
Sonstige	-5,57920*** (0,05643)	
Runde 2	-0,03950** (0,01612)	-0,01939*** (0,00144)
Runde 4	0,06861*** (0,01463)	0,01254*** (0,00147)
Runde 5	0,28596 (0,18196)	0,03061*** (0,01069)
Konstante	Ja	Ja
Between-Cluster Effekte	Ja	Ja
Jahresdummies	Ja	Ja
Spieler-Fixed-Effects	Ja	Ja
Beobachtungen	236.020	85.011

Anmerkungen: Die Tabelle gibt die Ergebnisse von hybriden Regressionen wieder. Für Spalte 1 wird dafür ein Logit-Modell, für Spalte 2 ein lineares Modell verwendet. Für das Modell mit angegriffen als abhängiger Variable wird die Stichprobe auf Going-for-it-Löcher beschränkt und Beobachtungen sind auf Loch-Ebene. Außerdem werden nur jene Beobachtungen berücksichtigt, bei denen die Entfernung, die mit dem Grünangriffsschlag zurückzulegen wäre (ist), maximal 355 Yards beträgt. Für das Modell mit Standardabweichung als abhängige Variable wird die gesamte Stichprobe, aggregiert auf Runden-Ebene verwendet. Robuste, nach Golfern geclusterte Standardfehler sind in Klammern angegeben. \*\*\*, \*\* und \* kennzeichnen Signifikanz auf dem 1%-, 5%- bzw. 10%-Signifikanzniveau.

#### 4.3. Robustheit

In diesem Kapitel werden die Ergebnisse aus dem vorhergehenden Kapitel auf ihre Robustheit untersucht. In einem ersten Robustheitstest werden hierbei, wie von Schunck und Perales (2017) empfohlen, für jene Variablen, deren geschätzte Within-Cluster-Effekte sich im Basismodell nicht auf dem 10%-Signifikanzniveau von den geschätz-

ten Between-Cluster-Effekten unterscheiden<sup>23</sup>, aus Effizienzgründen Random-Effects geschätzt. Für die Variablen, deren geschätzte Within-Cluster-Effekte sich auf dem 10%-Signifikanzniveau von den geschätzten Between-Cluster-Effekten unterscheiden, werden dabei weiterhin Within- und Between-Cluster-Effekte geschätzt. Wie die erste und zweite Spalte aus Abbildung A3 im Appendix zeigen, hat dies je-

<sup>23</sup>Hierfür wird ein Wald-Test benutzt, welcher überprüft, ob  $\beta_W = \beta_B$  gilt.



doch keinerlei Einfluss auf die geschätzten Koeffizienten für die Kulturdimensionen nach Hofstede. Unabhängig davon, ob angegriffen oder Standardabweichung als Maß für das Risikoverhalten verwendet wird, spielt der kulturelle Hintergrund eines Golfers keine statistisch signifikante Rolle.

Eine Alternative zu hybriden Modellen, welche es ebenfalls erlaubt, für unbeobachtete Heterogenität zu kontrollieren und gleichzeitig Schätzer für zeitinvariante Variablen zu erhalten, stellt ein Modell nach Hausman und Taylor (1981) dar. Bei diesem Modell werden die Variablen in zeitinvariante und zeitvariante Variablen unterteilt, welche wiederum in exogene und endogene Variablen unterteilt werden. Spalten 3 und 4 in Abbildung A3 im Appendix fassen die Ergebnisse zusammen, wenn Gleichung (1) und (2) mithilfe eines Hausman-Taylor Modells geschätzt werden. Als endogene Variablen werden für Spalte 3 dabei die Variablen  $\ln(\text{Vermögen})$ ,  $\ln(\text{Einkommen})$ , Erfolgsquote, Zwischenrang,  $(\text{Zwischenrang})^2$ , Entfernung zum Loch,  $(\text{Entfernung zum Loch})^2$ , Abschlag, Annäherung, Kurzspiel, Putting, Verlustdomäne, Gewinn-domäne sowie die Lageindikatorvariablen Fairway, Semi-Rough, Rough, Wasserhindernis, Bunker und Sonstige betrachtet. Für Spalte 4 werden die Variablen  $\ln(\text{Vermögen})$ ,  $\ln(\text{Einkommen})$ , Zwischenrang,  $(\text{Zwischenrang})^2$  sowie die Fähigkeitsproxys Abschlag, Annäherung, Kurzspiel und Putting als endogen angesehen. Auch in diesem Robustheitstest deuten statistisch insignifikante Koeffizienten für die Kulturdimensionen nach Hofstede darauf hin, dass kein Zusammenhang zwischen dem kulturellen Hintergrund eines Golfers und dessen Risikoverhalten besteht.

Wie Tabelle 4 zeigt, stammen fast 70% aller Beobachtungen für Going-for-it-Löcher von US-amerikanischen Spielern, was eine geringe Variation<sup>24</sup> der Werte für die Kulturdimensionen zur Folge hat und eine präzise Schätzung des Zusammenhangs von Kultur und Risikoverhalten erschwert. Aus diesem Grund werden in einem weiteren Robustheitstest dieselben Spezifikationen wie in Kapitel 4.2 geschätzt, allerdings wird die Stichprobe dabei auf jene Beobachtungen begrenzt, die sich nicht US-amerikanischen Golfern zuordnen lassen. Die Ergebnisse fassen die erste und zweite Spalte in Abbildung A4 im Appendix zusammen. Die geschätzten Koeffizienten zeigen, dass entgegen der Erwartungen aus Kapitel 2.4.1 ein höherer Wert für die Kulturdimension Machtdistanz für das Heimatland eines Golfers einhergeht mit risikofreudigerem Verhalten, sowohl wenn das Risikoverhalten anhand der beobachteten Entscheidung, das Grün anzugreifen, als auch wenn es anhand der Standardabweichung gemessen wird.<sup>25</sup> Während die übrigen Kulturdimensionen für die Entscheidung, das Grün anzugreifen, keine statistisch signifikante Rolle spielen, suggeriert ein auf dem 5%-Signifikanzniveau signifikanter, negativer Parameterschätzer für Langzeitorien-

tierung in Spalte 2, dass sich Spieler aus Ländern, in denen die Zukunft eine wichtige Rolle spielt, gemessen an der Standardabweichung ihres Scores relativ zu Par den Erwartungen aus Kapitel 2.4.4 entsprechend risikoaverser verhalten. Die Ergebnisse stehen somit weitestgehend im Einklang mit den Ergebnissen des Basismodells. Dass die Insignifikanz der Kulturdeterminanten im Basismodell aus einer zu geringen Variation, die wiederum auf einen hohen Anteil US-amerikanischer Golfer in der Stichprobe zurückzuführen ist, resultiert, ist daher unwahrscheinlich.

Als weiterer Robustheitstest wird in Spalte 3 und 4 von Abbildung A4 im Appendix wie in Rieger et al. (2015) anstelle des logarithmierten BIP-pro-Kopf der Human Development Index als Proxy für die wirtschaftliche und soziale Entwicklung des Heimatlandes eines Golfers verwendet. Für das Modell mit angegriffen als abhängiger Variable hat dies keinerlei Konsequenz was die Insignifikanz der Kulturdimensionen sowie des Proxys für die Entwicklung des Heimatlandes eines Sportlers anbelangt. Wie Spalte 4 nahelegt, wird ein höherer Entwicklungsstand der Heimatnation eines Golfers begleitet von risikofreudigerem Verhalten, gemessen an der Standardabweichung des Scores relativ zu Par. Außerdem geht in diesem Modell ein höherer Wert für Individualismus einher mit leicht statistisch signifikant risikoaverserem Verhalten. Dieses Ergebnis steht im Einklang mit der in Kapitel 2.4.3 behandelten Cushion-Hypothese.

Spalte 5 in Abbildung A4 im Appendix zeigt, dass die Ergebnisse aus Kapitel 4.2, was die Entscheidung, das Grün anzugreifen, betrifft, auch dann robust sind, wenn für Gleichung (1) anstelle eines hybriden Logit Modells ein hybrides lineares Modell geschätzt wird.

In einem weiteren Robustheitstest werden der Vorgehensweise von Adams und Waddell (2018) folgend Spieler-Jahr-Turnier-Fixed-Effects anstelle von Spieler-Fixed-Effects geschätzt, um für unbeobachtete Heterogenität zu kontrollieren. Abbildung A5 im Appendix zeigt, dass diese veränderte Spezifikation dazu führt, dass im Modell für die Standardabweichung (Spalte 2) der positive Koeffizient für Machtdistanz signifikant auf dem 5%-Signifikanzniveau ist. Golfer aus Ländern, deren Kultur von stärkerer Akzeptanz von Unterschieden im Machtgefüge geprägt ist, zeigen also risikofreudigeres Verhalten in Form von einer höheren Standardabweichung. Im Modell für die Entscheidung, das Grün anzugreifen (Spalte 1), hat die Anpassung der Spezifikation gegenüber dem Basismodell zur Folge, dass der negative Effekt von Maskulinität sowie der positive Effekt von Unsicherheitsvermeidung Signifikanz auf dem 10%- bzw. 1%-Signifikanzniveau erlangen. Diese Ergebnisse widersprechen den Erwartungen aus Kapitel 2.4.4 und 2.4.2. Der positive Koeffizient für Unsicherheitsvermeidung ist jedoch ein Beleg dafür, dass Unsicherheitsvermeidung, wie von Hofstede (2001) betont, nicht gleichzusetzen ist mit risikoaverserem Verhalten.

Wie bereits in Kapitel 3.2 geschildert und auch von McFall und Rotthoff (2018) angemerkt, führt die Definition der Variable angegriffen möglicherweise dazu, dass übermäßig viele Beobachtungen als Entscheidung für die weniger riskan-

<sup>24</sup>Vgl. hierzu Tabelle 2 und Abbildung A2 im Appendix

<sup>25</sup>Während der Zusammenhang für Spalte 2 (Standardabweichung als abhängige Variable) auf dem 5%-Signifikanzniveau signifikant ist, ist er dies in Spalte 1 (angegriffen als abhängige Variable) lediglich auf dem 10%-Signifikanzniveau.

te Strategie, das Grün nicht anzugreifen, gewertet werden. Dies liegt daran, dass besonders misslungene Versuche, das Grün anzugreifen, definitionsgemäß nicht als solche gewertet werden. Da davon auszugehen ist, dass diese Problematik insbesondere an Par 4 Löchern auftritt<sup>26</sup>, wird die Analyse der Entscheidung, das Grün anzugreifen, in einem weiteren Robustheitstest auf Par 5 Going-for-it-Löcher beschränkt. Die Ergebnisse in Spalte 3 in Abbildung A5 im Appendix zeigen, dass eine Beschränkung der Untersuchung auf Par 5 Going-for-it-Löcher die Insignifikanz der für die im Basismodell geschätzten Koeffizienten der Kulturdimensionen nach Hofstede nicht beeinflusst. Berücksichtigt man, dass lediglich 80.353 Beobachtungen der insgesamt 547.941 Beobachtungen für Going-for-it-Löcher im Datensatz von Par 4 Löchern stammen<sup>27</sup>, überrascht dieses Ergebnis nicht.

In einem weiteren Robustheitstest wird die Beschränkung der Stichprobe auf jene (potenziellen) Grünangriffsschläge, die aus maximal 355 Yards Entfernung zum Loch auszuführen sind (wären), aufgehoben, weil es durchaus Golfer gibt, die auch bei Schlägen über mehr als 355 Yards die Kontrolle über den Ball behalten. So zeigten beispielsweise Bubba Watson, Ryan Palmer, Boo Weekley und Sergio Garcia bereits Schläge, bei denen der Ball mehr als 400 Yards zurücklegte und die trotzdem ausreichend präzise waren, um das Fairway zu treffen. Im Jahr 2018 gelang Dustin Johnson an Loch 7 der Finalrunde der Deutsche Bank Championship gar ein Schlag mit einer Länge von 463 Yards<sup>28</sup>. Weil die Entfernung zum Loch lediglich für die Entscheidung, das Grün anzugreifen, relevant ist, wird dieser Robustheitstest nur für das Modell mit angegriffen als abhängiger Variable durchgeführt. Spalte 4 in Abbildung A5 im Appendix gibt die entsprechenden geschätzten Koeffizienten wieder. Auch dieser Robustheitstest bestätigt das Ergebnis aus den Basismodellen, dass kein statistisch signifikanter Zusammenhang zwischen dem kulturellen Hintergrund eines Golfers und dessen Risikoverhalten auf dem Golfplatz besteht.

Da es Kritik an der Vorgehensweise der Bestimmung der Werte für die Kulturdimensionen nach Hofstede (2001) gibt<sup>29</sup>, wird die Robustheit der Ergebnisse mit zwei finalen Tests überprüft, in welchen die Kulturdimensionen des GLOBE-Projekts (House et al., 2004) anstelle der Dimensionen nach Hofstede als erklärende Variablen verwendet werden<sup>30</sup>. Abbildung A6 im Appendix fasst die Ergebnisse zusammen, wenn für die Analyse dieselben Stichproben verwendet werden wie in den Basismodellen. Weil jedoch auch die Werte für die GLOBE-Kulturdimensionen aufgrund des hohen Anteils US-amerikanischer Golfer in der Stichprobe, eine relativ geringe Variation aufweisen, wird die Untersuchung für Abbildung A7 im Appendix mit auf Beobachtungen von nicht US-amerikanischen Spielern beschränkten Stichproben wiederholt. Vergleicht man die Koeffizienten aus

Spalte 1 in Abbildung A6 im Appendix mit Spalte 2 in Abbildung A6, so fällt auf, dass im Modell mit Standardabweichung als erklärender Variable lediglich die Kulturdimension Leistungsorientierung (Wert) einen auf dem 10%-Niveau signifikanten, negativen Zusammenhang mit dem Risikoverhalten eines Spielers hat. Stammt ein Golfer also aus einem Land, in welchem die Einwohner eher der Ansicht sind, dass übermäßige Leistungen gefördert und gefordert werden sollte, verhält sich dieser ceteris paribus risikoaverser. Wird das Risikoverhalten hingegen mit der Variablen angegriffen gemessen, so haben neben der Dimension Leistungsorientierung (Wert) auch die Dimensionen Leistungsorientierung (Praxis), Machtdistanz (Wert) sowie Kollektivismus II (Wert) einen signifikanten Einfluss. Während der Effekt von Leistungsorientierung, unabhängig davon, ob die Praxis oder die Werte im Heimatland des Golfers betrachtet werden, negativ ist, haben die Vorstellungen darüber, wie loyal und stolz Individuen ihren Familien gegenüber sein sollten (gemessen durch Kollektivismus II (Wert)), sowie darüber, wie gleichmäßig Macht innerhalb einer Gemeinschaft verteilt sein sollte (Machtdistanz (Wert)), einen Risikofreude fördernden Effekt.

Werden hingegen die Stichproben auf Beobachtungen von nicht US-amerikanischen Golfern beschränkt, so hat im Fall von der Standardabweichung (Spalte 2 in Abbildung A7 im Appendix) als abhängiger Variable nicht länger Leistungsorientierung (Wert) einen signifikanten, negativen Einfluss auf das Risikoverhalten, sondern mit Leistungsorientierung (Praxis) der Grad, zu dem eine Gesellschaft Leistung tatsächlich fördert und fordert. Alle anderen Kulturdimensionen haben wie in der unbeschränkten Stichprobe keinen signifikanten Effekt auf das mithilfe der Standardabweichung gemessene Risikoverhalten eines Golfers. Vergleicht man die Schätzer für das Modell, bei dem angegriffen als Maß für das Risikoverhalten dient, bei Verwendung der vollständigen Stichprobe (Spalte 1 in Abbildung A6 im Appendix) mit jenen für die beschränkte Stichprobe (Spalte 1 in Abbildung A7 im Appendix), so fällt auf, dass Vorzeichen und Signifikanz der Kulturdimensionen Leistungsorientierung (Praxis) sowie Kollektivismus II (Wert) unabhängig davon sind, ob US-amerikanische Golfer in der Stichprobe inkludiert werden oder nicht. Im Gegensatz hierzu werden der negative Koeffizient für Leistungsorientierung (Wert) sowie der positive Koeffizient für Machtdistanz (Wert) insignifikant, wenn die Stichprobe zur Analyse von angegriffen beschränkt wird. Gleichzeitig werden durch die Einschränkung der Stichprobe die negativen Koeffizienten für Kollektivismus I (Praxis), Geschlechteregalitarismus (Praxis) sowie Zukunftsorientierung (Wert) signifikant auf dem 5%-, 10%- bzw. 10%-Niveau. Golfer verhalten sich, gemessen an ihren Entscheidungen, das Grün anzugreifen, also ceteris paribus risikoaverser, wenn sie aus einem Land kommen, welches kollektivistischer ist, in welchem Geschlechterunterschiede in hohem Maße minimiert werden oder in welchem die Einwohner der Ansicht sind, dass bei Entscheidungen die Zukunft eine wichtige Rolle spielen sollte. Alle anderen GLOBE-Kulturdimensionen haben in beiden Stichproben keinen signifikanten Effekt auf

<sup>26</sup>Vgl. hierzu Kapitel 3.2.

<sup>27</sup>Vgl. hierzu Tabelle 3.

<sup>28</sup>Dieser Abschlag landete allerdings auch im Rough.

<sup>29</sup>Vgl. hierzu Kapitel 2.3.

<sup>30</sup>Für eine Definition der GLOBE-Kulturdimensionen siehe Abbildung A1 im Appendix.

das Risikoverhalten. Sowohl der Vergleich der Ergebnisse zwischen den unterschiedlichen Stichproben aus Abbildung A6 und Abbildung A7 im Appendix, als auch insbesondere der Vergleich der beiden Modelle aus Abbildung A7 im Appendix legt dabei nahe, dass die Ergebnisse zumindest nicht ausschließlich durch eine zu geringe Variation der Kulturdimensionen getrieben sind.

## 5. Fazit

Die deskriptiven Ergebnisse aus Kapitel 3.2 legen nahe, dass bei einer Betrachtung der Daten in auf Landesebene aggregierter Form ein schwacher Zusammenhang zwischen dem kulturellen Hintergrund eines Individuums und dessen Risikoverhalten besteht. So zeigen sie, dass höhere Werte für die Individualismus-Dimension nach Hofstede mit risikoaverserem Verhalten einhergehen. Des Weiteren deuten die deskriptiven Ergebnisse darauf hin, dass Golfer aus Ländern mit höheren Werten für die Hofstede-Kulturdimensionen Unsicherheitsvermeidung sowie Machtdistanz risikofreudigeres Verhalten zeigen.

Im Gegensatz dazu stehen die Ergebnisse aus Kapitel 4. Die ökonometrische Analyse zeigt, dass der kulturelle Hintergrund eines Spielers keinen signifikanten Einfluss auf dessen Risikoverhalten hat. Dieses Resultat ist dabei ziemlich robust. Einzig in den Spezifikationen, in denen die Stichprobe auf Beobachtungen, die von nicht US-amerikanischen Spielern stammen, beschränkt wird, sowie in den Spezifikationen, in denen Spieler-Jahr-Turnier-Fixed-Effects anstelle von Spieler-Fixed-Effects geschätzt werden, sind für beide Maße des Risikoverhaltens signifikante Koeffizienten für einzelne Kulturdimensionen zu beobachten. Dabei lässt sich in Übereinstimmung mit den deskriptiven Ergebnissen feststellen, dass sich Golfer aus Nationen mit einem höheren Wert für Machtdistanz sowie für Unsicherheitsvermeidung signifikant risikofreudiger verhalten. Diese Ergebnisse stehen jedoch im Widerspruch zu den Erwartungen aus Kapitel 2.4. Ebenfalls im Widerspruch zu den Erwartungen legen die Ergebnisse dieser Robustheitstests nahe, dass sich Individuen aus maskulineren Ländern risikoaverser verhalten. Dieses Resultat für Maskulinität widerspricht auch den deskriptiven Ergebnissen. Übereinstimmend mit den Erwartungen aus Kapitel 2.4.4 hingegen weisen die Ergebnisse der auf nicht US-amerikanische Spieler beschränkten Stichprobe darauf hin, dass sich Golfer aus Nationen mit einem höheren Wert für die Langzeitorientierung nach Hofstede risikoaverser verhalten, sofern das Risiko anhand der Standardabweichung gemessen wird. Wird anstelle des BIP-pro-Kopf der Human Development Index als Proxy für die Entwicklung des Heimatlandes eines Golfers verwendet, so legen die Ergebnisse übereinstimmend mit den deskriptiven Ergebnissen sowie der Cushion-Hypothese nahe, dass sich Spieler aus individualistischeren Ländern leicht signifikant risikoaverser verhalten. Allerdings ist dieser Effekt nur signifikant, wenn das Risikoverhalten anhand der Standardabweichung gemessen wird. Auch eine Betrachtung der GLOBE-Kulturdimensionen

anstelle der Kulturdimensionen nach Hofstede lässt kein klares Muster für den Einfluss des kulturellen Hintergrundes eines Spielers auf dessen Risikoverhalten erkennen.

## 6. Diskussion und Ausblick

Es gibt verschiedene Erklärungen dafür, wieso in der ökonometrischen Analyse dieser Arbeit im Gegensatz zu großen Teilen der bestehenden Literatur kein robuster, signifikanter Zusammenhang zwischen dem kulturellen Hintergrund eines Individuums und dessen Risikoverhalten gefunden werden kann.

Eine mögliche Erklärung dafür, dass in dieser Arbeit kein robuster Zusammenhang gefunden werden kann, besteht darin, dass fast 70% der Beobachtungen in der verwendeten Stichprobe von US-amerikanischen Spielern stammt, was eine geringe Variation der Kulturvariablen zur Folge hat und eine präzise Schätzung des Zusammenhangs von Kultur und Risikoverhalten erschwert. Robustheitstests legen jedoch nahe, dass diese Tatsache zumindest nicht die alleinige Erklärung ist.

Ein potenzielles Problem in der Vorgehensweise dieser Arbeit liegt darin, dass die Stichprobe lediglich aus professionellen Golfspielern besteht. Golfspieler auf der PGA TOUR stellen möglicherweise jedoch eine derart homogene Gruppe dar, dass nicht länger die Kultur des Heimatlandes eines Golfers dessen Verhalten prägt, sondern die Kultur „Golfspieler“. Übereinstimmend mit dieser Idee stellen [Schmidt und Saller \(2013\)](#) fest, dass Spitzensportler im Vergleich zum Rest der Bevölkerung überdurchschnittlich karriere- und wettbewerbsorientiert sind. Außerdem besitzen sie überdurchschnittlich viel Gelassenheit, Disziplin, Selbstbewusstsein sowie Stress- und Frustrationstoleranz und haben einen hohen Leistungsanspruch. Es ist davon auszugehen, dass diese Persönlichkeitsmerkmale einen erheblichen Einfluss auf das Risikoverhalten von Individuen in Turnieren haben. Insofern sind Kulturdimensionen, die auf Mittelwerten national aggregierter Werte basieren, möglicherweise ungeeignet, um tatsächliche kulturelle Unterschiede zu erfassen. Verschiedene Autoren weisen darauf hin, dass eine Verwendung der national aggregierten Hofstede- sowie GLOBE-Kulturdimensionen für Analysen auf Individuenebene möglicherweise zu Problemen führt, den Effekt von Kultur richtig zu messen ([Brewer und Venaik, 2012](#); [de Mooij, 2013](#); [McSweeney, 2013](#); [Venaik und Brewer, 2013](#)).

Möglicherweise sind die Definitionen der einzelnen Kulturdimensionen auch zu weit gefasst, sodass die Elemente einer Kultur, die das Risikoverhalten eines Individuums beeinflussen, von verschiedenen Kulturdimensionen erfasst werden. Vermischt eine zu weit gefasste Kulturdimension verschiedene Merkmale einer Kultur, kann der Einfluss jener Merkmale, die das Risikoverhalten beeinflussen, nicht länger sauber geschätzt werden.

Eine weitere Erklärung wäre, dass die verwendeten Maße für das Risikoverhalten eines Golfers ungeeignet sind. Allerdings stimmen im Basismodell die geschätzten Effekte – insbesondere für angegriffen als Maß des Risikoverhaltens –

für die übrigen Risikodeterminanten weitestgehend mit denen aus anderen Studien bzw. den theoretischen Erwartungen überein.

Weil im Rahmen dieser Arbeit das Risikoverhalten in einer anderen Domäne als in den meisten bisherigen Studien untersucht wird, besteht aber auch die Möglichkeit, dass Kultur für das Risikoverhalten in dieser Domäne kaum von Bedeutung ist. Alternativ ist es ebenfalls möglich, dass die Bedeutung des Effekts des kulturellen Hintergrundes eines Individuums abnimmt, sobald hohe Beträge auf dem Spiel stehen. Im Gegensatz zu den Experimentalstudien anderer Autoren geht es für Spieler auf der PGA TOUR nämlich um sehr hohe Beträge.

Da das Risikoverhalten von Individuen im Kontext ihrer Berufsausübung ökonomisch weitreichende Konsequenzen hat, ist es unerlässlich, dieses genauer zu verstehen. Insofern ist es wichtig, dass der Einfluss des kulturellen Hintergrundes auf das Risikoverhalten eines Individuums noch tiefergehender untersucht wird. Dabei wäre es hilfreich, wenn zukünftige Forschung eine Möglichkeit findet, Kultur zu messen, ohne mit den oben geschilderten Problemen konfrontiert zu sein.

## Literatur

- Adams, N. R. und Waddell, G. R. Performance and Risk Taking Under Threat of Elimination. *Journal of Economic Behavior & Organization*, 156:41–54, 2018. ISSN 01672681. doi: 10.1016/j.jebo.2018.09.014.
- Allais, M. Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica*, 21(4): 503–546, 1953. ISSN 00129682. doi: 10.2307/1907921.
- Allison, P. *Fixed Effects Regression Models*. California, Thousand Oaks, Sage Publications, 2009. ISBN 9780761924975. doi: 10.4135/9781412993869.
- Bar-Eli, M., Avugos, S., und Raab, M. Twenty Years of "Hot Hand" Research: Review and Critique. *Psychology of Sport and Exercise*, 7(6):525–553, 2006. ISSN 14690292. doi: 10.1016/j.psychsport.2006.03.001.
- Barseghyan, L., Prince, J., und Teitelbaum, J. C. Are Risk Preferences Stable Across Contexts? Evidence from Insurance Data. *The American Economic Review*, 101(2):591–631, 2011. ISSN 0002-8282. doi: 10.1257/aer.101.2.591.
- Barsky, R. B., Juster, F. T., Kimball, M. S., und Shapiro, M. D. Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. *The Quarterly Journal of Economics*, 112(2):537–579, 1997. ISSN 0033-5533. doi: 10.1162/00335539755280.
- Becker, B. E. und Huselid, M. A. The Incentive Effects of Tournament Compensation Systems. *Administrative Science Quarterly*, 37(2):336–350, 1992. ISSN 0001-8392. doi: 10.2307/2393228.
- Bedendo, M., Garcia-Appendini, E., und Siming, L. Cultural Preferences and Firm Financing Choices. *The Journal of Financial and Quantitative Analysis*, pages 1–75, 2019. ISSN 00221090. doi: 10.1017/S0022109019000103.
- Bellante, D. und Green, C. A. Relative Risk Aversion among the Elderly. *Review of Financial Economics*, 13(3):269–281, 2004. ISSN 10583300. doi: 10.1016/j.rfe.2003.09.010.
- Binswanger, H. P. Attitudes Toward Risk: Experimental Measurement in Rural India. *American Journal of Agricultural Economics*, 62(3):395–407, 1980. ISSN 00029092. doi: 10.2307/1240194.
- Binswanger, H. P. Attitudes Toward Risk: Theoretical Implications of an Experiment in Rural India. *The Economic Journal*, 91(364):867–890, 1981. ISSN 00130133. doi: 10.2307/2232497.
- Bonsang, E. und Dohmen, T. Risk Attitude and Cognitive Aging. *Journal of Economic Behavior & Organization*, 112:112–126, 2015. ISSN 01672681. doi: 10.1016/j.jebo.2015.01.004.
- Bontempo, R. N., Bottom, W. P., und Weber, E. U. Cross-Cultural Differences in Risk Perception: A Model-Based Approach. *Risk Analysis*, 17(4):479–488, 1997. ISSN 0272-4332. doi: 10.1111/j.1539-6924.1997.tb00888.x.
- Bothner, M. S., Kang, J.-h., und Stuart, T. E. Competitive Crowding and Risk Taking in a Tournament: Evidence from NASCAR Racing. *Administrative Science Quarterly*, 52(2):208–247, 2007. ISSN 0001-8392. doi: 10.2189/asqu.52.2.208.
- Botwinick, J. Cautiousness in Advanced Age. *Journal of Gerontology*, 21(3): 347–353, 1966. ISSN 0022-1422. doi: 10.1093/geronj/21.3.347.
- Breuer, W., Riesener, M., und Salzmann, A. J. Risk Aversion vs. Individualism: What Drives Risk Taking in Household Finance? *The European Journal of Finance*, 20(5):446–462, 2014. ISSN 1351-847X. doi: 10.1080/1351847X.2012.714792.
- Breuer, W., Kollath, J. A., Salzmann, A. J., und von Nitzsch, R. The assessment of risk behavior: A cross-cultural analysis using the dospert scale. In Venezia, I., editor, *Behavioral Finance*, pages 159–180. World Scientific Publishing, Singapur, 2016. ISBN 978-981-310-008-4. doi: 10.1142/9789813100091{\textunderscore}0005.
- Brewer, P. und Venaik, S. On the Misuse of National Culture Dimensions. *International Marketing Review*, 29(6):673–683, 2012. ISSN 0265-1335. doi: 10.1108/02651331211277991.
- Broadie, M. Assessing golfer performance using golfmetrics. In Crews, D. und Lutz, R., editors, *Science and Golf V: Proceedings of the World Scientific Congress of Golf*, pages 253–262. Energy in Motion, Inc., Mesa, Arizona, 2008. ISBN 978-0978873417.
- Broadie, M. Assessing Golfer Performance on the PGA TOUR. *Interfaces*, 42(2):146–165, 2012. ISSN 0092-2102. doi: 10.1287/inte.1120.0626.
- Brown, J. Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars. *Journal of Political Economy*, 119(5):982–1013, 2011. ISSN 0022-3808. doi: 10.1086/663306.
- Brown, K. C., Harlow, W. V., und Starks, L. T. Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry. *The Journal of Finance*, 51(1):85–110, 1996. ISSN 00221082. doi: 10.2307/2329303.
- Brumagim, A. L. und Xianhua, W. An Examination of Cross-cultural Differences in Attitudes Towards Risk: Testing Prospect Theory in the People's Republic of China. *Multinational Business Review*, 13(3):67–86, 2005. ISSN 1525-383X. doi: 10.1108/1525383X200500015.
- Buccioli, A. und Miniaci, R. Financial Risk Propensity, Business Cycles and Perceived Risk Exposure. *Oxford Bulletin of Economics and Statistics*, 80(1):160–183, 2018. ISSN 03059049. doi: 10.1111/obes.12193.
- Busse, J. A. Another Look at Mutual Fund Tournaments. *The Journal of Financial and Quantitative Analysis*, 36(1):53–73, 2001. ISSN 00221090. doi: 10.2307/2676197.
- Calvet, L. E. und Sodini, P. Twin Picks: Disentangling the Determinants of Risk-Taking in Household Portfolios. *The Journal of Finance*, 69(2):867–906, 2014. ISSN 00221082. doi: 10.1111/jofi.12125.
- Cardenas, J. C. und Carpenter, J. Behavioural Development Economics: Lessons from Field Labs in the Developing World. *The Journal of Development Studies*, 44(3):311–338, 2008. ISSN 0022-0388. doi: 10.1080/00220380701848327.
- Chetty, R. und Szeidl, A. Consumption Commitments and Risk Preferences. *The Quarterly Journal of Economics*, 122(2):831–877, 2007. ISSN 0033-5533. doi: 10.1162/qjec.122.2.831.
- Cheung, H. Y., Wu, J., und Tao, J. Risk Perception and Risk-Taking Attitude: A Comparison Between Hong Kong and Mainland Chinese Undergraduate Students. *The Asia-Pacific Education Researcher*, 22(4):497–506, 2013. ISSN 0119-5646. doi: 10.1007/s40299-012-0048-7.
- Chevalier, J. und Ellison, G. Risk Taking by Mutual Funds as a Response to Incentives. *Journal of Political Economy*, 105(6):1167–1200, 1997. ISSN 0022-3808. doi: 10.1086/516389.
- Chui, A. C. W., Titman, S., und Wei, K. C. J. Individualism and Momentum around the World. *The Journal of Finance*, 65(1):361–392, 2010. ISSN 00221082. doi: 10.1111/j.1540-6261.2009.01532.x.
- Connolly, R. A. und Rendleman, R. J. What It Takes to Win on the PGA TOUR (If Your Name Is "Tiger" or If It Isn't). *Interfaces*, 42(6):554–576, 2012. ISSN 0092-2102. doi: 10.1287/inte.1110.0615.
- Cummings, L. L., Harnett, D. L., und Stevens, O. J. Risk, Fate, Conciliation and Trust: An International Study of Attitudinal Differences Among Executives. *Academy of Management Journal*, 14(3):285–304, 1971. ISSN 0001-4273. doi: 10.2307/255074.
- Das, T. K. und Teng, B.-S. The Risk-Based View of Trust: A Conceptual Framework. *Journal of Business and Psychology*, 19(1):85–116, 2004. ISSN 0889-3268. doi: 10.1023/B:JOBU.0000040274.23551.1b.
- de Bondt, W. F. M. und Thaler, R. H. Financial decision-making in markets and firms: A behavioral perspective. In Jarrow, R. A., Maksimovic, V., und Ziemba, W. T., editors, *Finance*, volume 9 of *Handbooks in Operations Research and Management Science*, pages 385–410. Elsevier, Oxford, 1995. ISBN 9780444890849. doi: 10.1016/S0927-0507(05)80057-X.
- de Mooij, M. On the Misuse and Misinterpretation of Dimensions of National Culture. *International Marketing Review*, 30(3):253–261, 2013. ISSN 0265-1335. doi: 10.1108/02651331311321990.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., und Wagner, G. G. Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association*, 9(3):522–550, 2011. ISSN 15424766. doi: 10.1111/j.1542-4774.2011.01015.x.
- Dohmen, T., Falk, A., Golsteyn, B. H. H., Huffman, D., und Sunde, U. Risk Attitudes Across the Life Course. *The Economic Journal*, 127(605):F95–F116, 2017. ISSN 00130133. doi: 10.1111/eoj.12322.
- Donkers, B., Melenberg, B., und van Soest, A. Estimating Risk Attitudes Using Lotteries: A Large Sample Approach. *Journal of Risk and Uncertainty*, 22(2):165–195, 2001. ISSN 08955646. doi: 10.1023/A:1011109625844.
- Ehrenberg, R. G. und Bognanno, M. L. Do Tournaments Have Incentive Effects? *Journal of Political Economy*, 98(6):1307–1324, 1990a. ISSN 0022-3808. doi: 10.1086/261736.
- Ehrenberg, R. G. und Bognanno, M. L. The Incentive Effects of Tournaments Revisited: Evidence from the European PGA Tour. *Industrial and Labor Relations Review*, 43(3):74S–88S, 1990b. ISSN 00197939. doi: 10.2307/2523572.
- Einav, L., Finkelstein, A., Pascu, I., und Cullen, M. R. How General are Risk Preferences? Choices Under Uncertainty in Different Domains. *The Ame-*

- rican Economic Review*, 102(6):2606–2038, 2012. ISSN 0002-8282. doi: 10.1257/aer.102.6.2606.
- Eriksson, K. und Simpson, B. Emotional Reactions to Losing Explain Gender Differences in Entering a Risky Lottery. *Judgment and Decision Making*, 5(3):159–163, 2010.
- Fan, J. X. und Xiao, J. J. Cross-Cultural Differences in Risk Tolerance: A Comparison between Chinese and Americans. *Journal of Personal Finance*, 5(3):54–75, 2006.
- Genakos, C. und Pagliero, M. Interim Rank, Risk Taking, and Performance in Dynamic Tournaments. *Journal of Political Economy*, 120(4):782–813, 2012. ISSN 0022-3808. doi: 10.1086/668502.
- Gilovich, T., Vallone, R., und Tversky, A. The Hot Hand in Basketball: On the Misperception of Random Sequences. *Cognitive Psychology*, 17(3): 295–314, 1985. ISSN 00100285. doi: 10.1016/0010-0285(85)90010-6.
- Glaser, M., Langer, T., und Weber, M. True Overconfidence in Interval Estimates: Evidence Based on a New Measure of Miscalibration. *Journal of Behavioral Decision Making*, 26(5):405–417, 2013. ISSN 08943257. doi: 10.1002/bdm.1773.
- Golf Digest. Tpc sawgrass stadium course: A hole-by-hole look at tpc sawgrass stadium course, 2018. URL [https://www.golfdigest.com/gallery/photos\\_tpccourse](https://www.golfdigest.com/gallery/photos_tpccourse).
- Goszczyńska, M., Tyszka, T., und Slovic, P. Risk Perception in Poland: A Comparison With Three Other Countries. *Journal of Behavioral Decision Making*, 4(3):179–193, 1991. ISSN 08943257. doi: 10.1002/bdm.3960040305.
- Green, B. und Zwiebel, J. The Hot-Hand Fallacy: Cognitive Mistakes or Equilibrium Adjustments? Evidence from Major League Baseball. *Management Science*, 64(11):5315–5348, 2018. ISSN 0025-1909. doi: 10.1287/mnsc.2017.2804.
- Grisley, W. und Kellog, E. Risk-Taking Preferences of Farmers in Northern Thailand: Measurements and Implications. *Agricultural Economics*, 1(2): 127–142, 1987. ISSN 01695150. doi: 10.1016/0169-5150(87)90009-0.
- Grund, C. und Gürtler, O. An Empirical Study on Risk-Taking in Tournaments. *Applied Economics Letters*, 12(8):457–461, 2005. ISSN 1350-4851. doi: 10.1080/13504850500109881.
- Grund, C., Höcker, J., und Zimmermann, S. Incidence and Consequences of Risk-Taking Behavior in Tournaments – Evidence from the NBA. *Economic Inquiry*, 51(2):1489–1501, 2013. ISSN 00952583. doi: 10.1111/j.1465-7295.2012.00499.x.
- Guiso, L., Sapienza, P., und Zingales, L. Does Culture Affect Economic Outcomes? *Journal of Economic Perspectives*, 20(2):23–48, 2006. ISSN 0895-3309. doi: 10.1257/jep.20.2.23.
- Gupta, V. und Hanges, P. J. Regional and climate clustering of societal cultures. In House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., und Gupta, V., editors, *Culture, Leadership, and Organizations*, pages 178–218. Sage Publications, Thousand Oaks, California, 2004. ISBN 9780761924012.
- Guryan, J., Kroft, K., und Notowidigdo, M. J. Peer Effects in the Workplace: Evidence from Random Groupings in Professional Golf Tournaments. *American Economic Journal: Applied Economics*, 1(4):34–68, 2009. doi: 10.1257/app.1.4.34.
- Halek, M. und Eisenhauer, J. G. Demography of Risk Aversion. *The Journal of Risk and Insurance*, 68(1):1–24, 2001. ISSN 00224367. doi: 10.2307/2678130.
- Hanoch, Y., Johnson, J. G., und Wilke, A. Domain Specificity in Experimental Measures and Participant Recruitment: An Application to Risk-Taking Behavior. *Psychological Science*, 17(4):300–304, 2006. ISSN 0956-7976. doi: 10.1111/j.1467-9280.2006.01702.x.
- Hausman, J. A. und Taylor, W. E. Panel Data and Unobservable Individual Effects. *Econometrica*, 49(6):1377–1398, 1981. ISSN 00129682. doi: 10.2307/1911406.
- Heath, C. und Tversky, A. Preference and Belief: Ambiguity and Competence in Choice Under Uncertainty. *Journal of Risk and Uncertainty*, 4(1):5–28, 1991. ISSN 08955646. doi: 10.1007/BF00057884.
- Henrich, J. und McElreath, R. Are Peasants Risk-Averse Decision Makers? *Current Anthropology*, 43(1):172–181, 2002. ISSN 0011-3204. doi: 10.1086/338291.
- Herrmann, B., Thöni, C., und Gächter, S. Antisocial Punishment Across Societies. *Science*, 319(5868):1362–1367, 2008. doi: 10.1126/science.1153808.
- Hickman, D. C. und Metz, N. E. The Impact of Pressure on Performance: Evidence from the PGA TOUR. *Journal of Economic Behavior & Organization*, 116:319–330, 2015. ISSN 01672681. doi: 10.1016/j.jebo.2015.04.007.
- Hickman, D. C., Kerr, C., und Metz, N. Rank and Performance in Dynamic Tournaments: Evidence from the PGA Tour. *Journal of Sports Economics*, 20(4):509–534, 2019. ISSN 1527-0025. doi: 10.1177/1527002518794788.
- Hofstede, G. *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*. California, Thousand Oaks, Sage Publications, 2 edition, 2001. ISBN 0-8039-7323-3.
- Hofstede, G., Hofstede, G. J., und Minkov, M. *Cultures and Organizations: Software of the Mind: Intercultural Cooperation and its Importance for Survival*. New York, New York City, McGraw-Hill, 3 edition, 2010. ISBN 9780071664189.
- Holt, C. A. und Laury, S. K. Risk Aversion and Incentive Effects. *The American Economic Review*, 92(5):1644–1655, 2002. ISSN 0002-8282. doi: 10.1257/000282802762024700.
- Hopland, A. O., Matsen, E., und Strøm, B. Income and Choice Under Risk. *Journal of Behavioral and Experimental Finance*, 12:55–64, 2016. ISSN 22146350. doi: 10.1016/j.jbef.2016.09.001.
- House, R. J. und Javidan, M. Overview of globe. In House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., und Gupta, V., editors, *Culture, Leadership, and Organizations*, pages 9–28. Sage Publications, Thousand Oaks, California, 2004. ISBN 9780761924012.
- House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., und Gupta, V., editors. *Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies*. California, Thousand Oaks, Sage Publications, 2004. ISBN 9780761924012.
- Hsee, C. K. und Weber, E. U. Cross-National Differences in Risk Preference and Lay Predictions. *Journal of Behavioral Decision Making*, 12(2):165–179, 1999. ISSN 08943257.
- Hvide, H. K. und Kristiansen, E. G. Risk Taking in Selection Contests. *Games and Economic Behavior*, 42(1):172–179, 2003. ISSN 08998256. doi: 10.1016/S0899-8256(02)00538-9.
- Javidan, M., House, R. J., und Dorfman, P. W. A nontechnical summary of globe findings. In House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., und Gupta, V., editors, *Culture, Leadership, and Organizations*, pages 29–48. Sage Publications, Thousand Oaks, California, 2004. ISBN 9780761924012.
- Jianakoplos, N. A. und Bernasek, A. Financial Risk Taking by Age and Birth Cohort. *Southern Economic Journal*, 72(4):981–1001, 2006. ISSN 00384038. doi: 10.2307/20111864.
- Kachelmeier, S. J. und Shehata, M. Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence from the People's Republic of China. *The American Economic Review*, 82(5):1120–1141, 1992. ISSN 0002-8282.
- Kahneman, D. und Tversky, A. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica*, 47(2):263–292, 1979. ISSN 00129682. doi: 10.2307/1914185.
- Kempf, A., Ruenzi, S., und Thiele, T. Employment Risk, Compensation Incentives, and Managerial Risk Taking: Evidence from the Mutual Fund Industry. *Journal of Financial Economics*, 92(1):92–108, 2009. ISSN 0304405X. doi: 10.1016/j.jfineco.2008.05.001.
- Kleinhesselink, R. R. und Rosa, E. A. Cognitive Representation of Risk Perceptions: A Comparison of Japan and the United States. *Journal of Cross-Cultural Psychology*, 22(1):11–28, 1991. ISSN 0022-0221. doi: 10.1177/0022022191221004.
- Koski, J. L. und Pontiff, J. How Are Derivatives Used? Evidence from the Mutual Fund Industry. *The Journal of Finance*, 54(2):791–816, 1999. ISSN 00221082. doi: 10.1111/0022-1082.00126.
- Kreiser, P. M., Marino, L. D., Dickson, P., und Weaver, K. M. Cultural Influences on Entrepreneurial Orientation: The Impact of National Culture on Risk Taking and Proactiveness in SMEs. *Entrepreneurship Theory and Practice*, 34(5):959–984, 2010. ISSN 10422587. doi: 10.1111/j.1540-6520.2010.00396.x.
- Lau, L.-Y. und Ranyard, R. Chinese and English Probabilistic Thinking and Risk Taking in Gambling. *Journal of Cross-Cultural Psychology*, 36(5): 621–627, 2005. ISSN 0022-0221. doi: 10.1177/0022022105278545.
- Lee, W.-N. Exploring the Role of Culture in Advertising: Resolving Persistent Issues and Responding to Changes. *Journal of Advertising*, 48(1):115–125, 2019. ISSN 0091-3367. doi: 10.1080/00913367.2019.1579686.
- Lefebvre, M., Vieider, F. M., und Villeval, M. C. Incentive Effects on Risk Attitude in Small Probability Prospects. *Economics Letters*, 109(2):115–

- 120, 2010. ISSN 01651765. doi: 10.1016/j.econlet.2010.09.002.
- Li, K., Griffin, D., Yue, H., und Zhao, L. How Does Culture Influence Corporate Risk-Taking? *Journal of Corporate Finance*, 23:1–22, 2013. ISSN 09291199. doi: 10.1016/j.jcorpfin.2013.07.008.
- Liu, E. M. Time to Change What to Sow: Risk Preferences and Technology Adoption Decisions of Cotton Farmers in China. *Review of Economics and Statistics*, 95(4):1386–1403, 2013. ISSN 0034-6535. doi: 10.1162/REST\_a\_00295.
- Marshall, R., Huan, T.-C., Xu, Y., und Nam, I. Extending Prospect Theory Cross-Culturally by Examining Switching Behavior in Consumer and Business-to-Business Contexts. *Journal of Business Research*, 64(8):871–878, 2011. ISSN 01482963. doi: 10.1016/j.jbusres.2010.09.009.
- McFall, T. und Rothhoff, K. W. Risk taking dynamics in tournaments: Evidence from professional golf, 2018. URL <http://pirate.shu.edu/~rotthok/papers/Par%20Five%20Strategy.pdf>.
- McSweeney, B. Fashion Founded on a Flaw: The Ecological Mono-Deterministic Fallacy of Hofstede, GLOBE, and Followers. *International Marketing Review*, 30(5):483–504, 2013. ISSN 0265-1335. doi: 10.1108/IMR-04-2013-0082.
- Mihet, R. Effects of Culture on Firm Risk-Taking: A Cross-Country and Cross-Industry Analysis. *Journal of Cultural Economics*, 37(1):109–151, 2013. ISSN 0885-2545. doi: 10.1007/s10824-012-9186-2.
- Miyata, S. Household's Risk Attitudes in Indonesian Villages. *Applied Economics*, 35(5):573–583, 2003. ISSN 0003-6846. doi: 10.1080/0003684022000020823.
- Mosley, P. und Verschoor, A. Risk Attitudes and the 'Vicious Circle of Poverty'. *The European Journal of Development Research*, 17(1):59–88, 2005. ISSN 0957-8811. doi: 10.1080/09578810500066548.
- Nobre, L. H. N., Grable, J. E., da Silva, W. V., und da Veiga, C. P. A Cross Cultural Test of Financial Risk Tolerance Attitudes: Brazilian and American Similarities and Differences. *International Journal of Economics and Financial Issues*, 6(1):314–322, 2016.
- Oosterbeek, H., Sloof, R., und van de Kuilen, G. Cultural Differences in Ultimatum Game Experiments: Evidence from a Meta-Analysis. *Experimental Economics*, 7(2):171–188, 2004. ISSN 1386-4157. doi: 10.1023/B:EXEC.0000026978.14316.74.
- Orszag, J. M. A New Look at Incentive Effects and Golf Tournaments. *Economics Letters*, 46(1):77–88, 1994. ISSN 01651765. doi: 10.1016/0165-1765(94)90080-9.
- Ozbeklik, S. und Smith, J. K. Risk Taking in Competition: Evidence from Match Play Golf Tournaments. *Journal of Corporate Finance*, 44:506–523, 2017. ISSN 09291199. doi: 10.1016/j.jcorpfin.2014.05.003.
- Pan, C. H. und Statman, M. Questionnaires of Risk Tolerance, Regret, Overconfidence, and Other Investor Propensities. *Journal of Investment Consulting*, 13(1):54–63, 2012.
- PGA TOUR. 2017-2018 player handbook & tournament regulations, 2017. URL [https://qualifying.pgatourhq.com/static-assets/uploads/2017-18\\_pga\\_tour\\_handbookregs\\_final.pdf](https://qualifying.pgatourhq.com/static-assets/uploads/2017-18_pga_tour_handbookregs_final.pdf).
- PGA TOUR. The players, 2018. URL <https://www.pgatour.com/competition/2018/the-players-championship/leaderboard.html>.
- Pope, D. G. und Schweitzer, M. E. Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes. *The American Economic Review*, 101(1):129–157, 2011. ISSN 0002-8282. doi: 10.1257/aer.101.1.129.
- Pyles, M. K., Li, Y., Wu, S., und Dolvin, S. D. Cultural Influences on Risk Tolerance and Portfolio Creation. *Journal of Behavioral and Experimental Finance*, 9:43–55, 2016. ISSN 22146350. doi: 10.1016/j.jbef.2016.01.001.
- Rabin, M. Risk Aversion and Expected-utility Theory: A Calibration Theorem. *Econometrica*, 68(5):1281–1292, 2000. ISSN 00129682. doi: 10.1111/1468-0262.00158.
- Ramírez, A. und Tadesse, S. Corporate Cash Holdings, Uncertainty Avoidance, and the Multinationality of Firms. *International Business Review*, 18(4):387–403, 2009. ISSN 09695931. doi: 10.1016/j.ibusrev.2009.02.013.
- Rieger, M. O., Wang, M., und Hens, T. Risk Preferences Around the World. *Management Science*, 61(3):637–648, 2015. ISSN 0025-1909. doi: 10.1287/mnsc.2013.1869.
- Rolison, J. J., Hanoch, Y., Wood, S., und Liu, P.-J. Risk-Taking Differences Across the Adult Life Span: A Question of Age and Domain. *The Journals of Gerontology: Series B, Psychological sciences and social sciences*, 69(6):870–880, 2014. doi: 10.1093/geronb/gbt081.
- Rosenboim, M., Shavit, T., und Shoham, A. Financial Decision Making in Collective Society – A Field Test on Israeli Kibbutz Members and City Residents. *The Journal of Socio-Economics*, 39(1):30–36, 2010. ISSN 10535357. doi: 10.1016/j.socec.2009.07.008.
- Rosenzweig, M. R. und Binswanger, H. P. Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments. *The Economic Journal*, 103(416):56–78, 1993. ISSN 00130133. doi: 10.2307/2234337.
- Salacuse, J. W. Ten Ways that Culture Affects Negotiating Style: Some Survey Results. *Negotiation Journal*, 14(3):221–240, 1998. ISSN 0748-4526. doi: 10.1111/j.1571-9979.1998.tb00162.x.
- Schmidt, S. L. und Saller, T. Kollege Spitzensportler: Chancen für Wirtschaft und Athleten, 2013. URL [https://www.sporthilfe.de/fileadmin/pdf/Studien/Kollege\\_Spitzensportler\\_2013.pdf](https://www.sporthilfe.de/fileadmin/pdf/Studien/Kollege_Spitzensportler_2013.pdf).
- Schneider, C. R., Fehrenbacher, D. D., und Weber, E. U. Catch Me if I Fall: Cross-National Differences in Willingness to Take Financial Risks as a Function of Social and State 'Cushioning'. *International Business Review*, 26(6):1023–1033, 2017. ISSN 09695931. doi: 10.1016/j.ibusrev.2017.03.008.
- Schunck, R. Within and Between Estimates in Random-Effects Models: Advantages and Drawbacks of Correlated Random Effects and Hybrid Models. *The Stata Journal: Promoting communications on statistics and Stata*, 13(1):65–76, 2013. ISSN 1536-867X. doi: 10.1177/1536867X1301300105.
- Schunck, R. und Perales, F. Within- and Between-Cluster Effects in Generalized Linear Mixed Models: A Discussion of Approaches and the Xthybrid Command. *The Stata Journal: Promoting communications on statistics and Stata*, 17(1):89–115, 2017. ISSN 1536-867X. doi: 10.1177/1536867X1701700106.
- Sitkin, S. B. und Pablo, A. L. Reconceptualizing the Determinants of Risk Behavior. *The Academy of Management Review*, 17(1):9–38, 1992. ISSN 03637425. doi: 10.2307/258646.
- Sivak, M., Soler, J., Tränkle, U., und Spagnhol, J. M. Cross-Cultural Differences in Driver Risk-Perception. *Accident Analysis & Prevention*, 21(4):355–362, 1989. ISSN 00014575. doi: 10.1016/0001-4575(89)90026-2.
- Slovic, P. Information Processing, Situation Specificity, and the Generality of Risk-Taking Behavior. *Journal of Personality and Social Psychology*, 22(1):128–134, 1972. ISSN 0022-3514. doi: 10.1037/h0032370.
- Smith, V. L. und Walker, J. M. Rewards, Experience and Decision Costs in First Price Auctions. *Economic Inquiry*, 31(2):237–244, 1993. ISSN 00952583. doi: 10.1111/j.1465-7295.1993.tb00880.x.
- Starmer, C. Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk. *Journal of Economic Literature*, 38(2):332–382, 2000. ISSN 0022-0515. doi: 10.1257/jel.38.2.332.
- Statman, M. Countries and Culture in Behavioral Finance. *CFA Institute Conference Proceedings Quarterly*, 25(3):38–44, 2008. ISSN 1930-2703. doi: 10.2469/cp.v25.n3.6.
- Statman, M. The Cultures of Risk Tolerance. *SSRN Electronic Journal*, 2010. doi: 10.2139/ssrn.1647086.
- Statman, M. Culture in Risk, Regret, Maximization, Social Trust, and Life Satisfaction. *Journal of Investment Consulting*, 16(1):20–30, 2015.
- Statman, M. und Weng, J. A. Investments Across Cultures: Financial Attitudes of Chinese-Americans. *Journal of Investment Consulting*, 11(1):37–44, 2010.
- Tanaka, T., Camerer, C. F., und Nguyen, Q. Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam. *The American Economic Review*, 100(1):557–571, 2010. ISSN 0002-8282. doi: 10.1257/aer.100.1.557.
- Tavor, T. und Garyn-Tal, S. Further Examination of the Demographic and Social Factors Affecting Risk Aversion. *Financial Markets and Portfolio Management*, 30(1):95–110, 2016. ISSN 1934-4554. doi: 10.1007/s11408-016-0260-1.
- Thaler, R. H. und Johnson, E. J. Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice. *Management Science*, 36(6):643–660, 1990. ISSN 0025-1909. doi: 10.1287/mnsc.36.6.643.
- Tversky, A. und Kahneman, D. Belief in the Law of Small Numbers. *Psychological Bulletin*, 76(2):105–110, 1971. ISSN 0033-2909. doi: 10.1037/h0031322.
- Tversky, A. und Kahneman, D. Judgment Under Uncertainty: Heuristics and Biases. *Science*, 185(4157):1124–1131, 1974. doi: 10.1126/science.185.

- 4157.1124.
- Tversky, A. und Kahneman, D. Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323, 1992. ISSN 08955646. doi: 10.1007/BF00122574.
- Tyagi, V., Hanoch, Y., Hall, S. D., Runco, M., und Denham, S. L. The Risky Side of Creativity: Domain Specific Risk Taking in Creative Individuals. *Frontiers in psychology*, 8, 2017. ISSN 1664-1078. doi: 10.3389/fpsyg.2017.00145.
- van Winsen, F., de Mey, Y., Lauwers, L., van Passel, S., Vancauteran, M., und Wauters, E. Determinants of Risk Behaviour: Effects of Perceived Risks and Risk Attitude on Farmer's Adoption of Risk Management Strategies. *Journal of Risk Research*, 19(1):56–78, 2016. ISSN 1366-9877. doi: 10.1080/13669877.2014.940597.
- Venaik, S. und Brewer, P. Critical Issues in the Hofstede and GLOBE National Culture Models. *International Marketing Review*, 30(5):469–482, 2013. ISSN 0265-1335. doi: 10.1108/IMR-03-2013-0058.
- Veider, F. M., Chmura, T., Fisher, T., Kusakawa, T., Martinsson, P., Mattison Thompson, F., und Sunday, A. Within- versus Between-Country Differences in Risk Attitudes: Implications for Cultural Comparisons. *Theory and Decision*, 78(2):209–218, 2015a. ISSN 0040-5833. doi: 10.1007/s11238-014-9418-3.
- Veider, F. M., Lefebvre, M., Bouchouicha, R., Chmura, T., Hakimov, R., Krawczyk, M., und Martinsson, P. Common Components of Risk and Uncertainty Attitudes Across Contexts and Domains: Evidence from 30 Countries. *Journal of the European Economic Association*, 13(3):421–452, 2015b. ISSN 15424766. doi: 10.1111/jeea.12102.
- Veider, F. M., Beyene, A., Bluffstone, R., Dissanayake, S., Gebreegziabher, Z., Martinsson, P., und Mekonnen, A. Measuring Risk Preferences in Rural Ethiopia. *Economic Development and Cultural Change*, 66(3):417–446, 2018. ISSN 0013-0079. doi: 10.1086/696106.
- von Neumann, J. und Morgenstern, O. *Theory of Games and Economic Behavior*. New Jersey, Princeton, Princeton University Press, 1944. ISBN 9781400829460.
- Wang, M. und Fischbeck, P. S. Incorporating Framing into Prospect Theory Modeling: A Mixture-Model Approach. *Journal of Risk and Uncertainty*, 29(2):181–197, 2004. ISSN 08955646. doi: 10.1023/B:RISK.0000038943.63610.16.
- Wang, M. und Fischbeck, P. S. Evaluating Lotteries, Risks, and Risk-Mitigation Programs. *Journal of Risk Research*, 11(6):775–795, 2008. ISSN 1366-9877. doi: 10.1080/13669870801967259.
- Wang, M., Rieger, M. O., und Hens, T. The Impact of Culture on Loss Aversion. *Journal of Behavioral Decision Making*, 30(2):270–281, 2017. ISSN 08943257. doi: 10.1002/bdm.1941.
- Weber, C. S. Determinants of risk tolerance. *International Journal of Economics, Finance and Management Sciences*, 2(2):143–152, 2014. ISSN 2326-9553. doi: 10.11648/j.ijefm.20140202.15.
- Weber, E. U. Personality and risk taking. In Smelser, N. J. und Baltes, P. B., editors, *International Encyclopedia of the Social & Behavioral Sciences*, pages 11274–11276. Elsevier, Oxford, 2001. ISBN 9780080430768. doi: 10.1016/B0-08-043076-7/01782-4.
- Weber, E. U. und Hsee, C. K. Cross-Cultural Differences in Risk Perception, but Cross-Cultural Similarities in Attitudes Towards Perceived Risk. *Management Science*, 44(9):1205–1217, 1998. ISSN 0025-1909. doi: 10.1287/mnsc.44.9.1205.
- Weber, E. U. und Milliman, R. A. Perceived Risk Attitudes: Relating Risk Perception to Risky Choice. *Management Science*, 43(2):123–144, 1997. ISSN 0025-1909. doi: 10.1287/mnsc.43.2.123.
- Weber, E. U., Hsee, C. K., und Sokolowska, J. What Folklore Tells Us about Risk and Risk Taking: Cross-Cultural Comparisons of American, German, and Chinese Proverbs. *Organizational Behavior and Human Decision Processes*, 75(2):170–186, 1998. ISSN 07495978. doi: 10.1006/obhd.1998.2788.
- Weber, E. U., Blais, A.-R., und Betz, N. E. A Domain-Specific Risk-Attitude Scale: Measuring Risk Perceptions and Risk Behaviors. *Journal of Behavioral Decision Making*, 15(4):263–290, 2002. ISSN 08943257. doi: 10.1002/bdm.414.
- Wik, M., Aragie Kebede, T., Bergland, O., und Holden, S. T. On the Measurement of Risk Aversion from Experimental Data. *Applied Economics*, 36(21):2443–2451, 2004. ISSN 0003-6846. doi: 10.1080/0003684042000280580.
- Wolf, C. und Pohlman, L. The Recovery of Risk Preferences from Actual Choices. *Econometrica*, 51(3):843–850, 1983. ISSN 00129682. doi: 10.2307/1912161.
- Yamaguchi, S., Gelfand, M., Ohashi, M. M., und Zemba, Y. The Cultural Psychology of Control: Illusions of Personal Versus Collective Control in the United States and Japan. *Journal of Cross-Cultural Psychology*, 36(6):750–761, 2005. ISSN 0022-0221. doi: 10.1177/0022022105280513.





## The Influencing Phenomenon – Impact of Fitness Influencers on Sports Activity

### Das Phänomen Influencing – Beeinflussung von Fitness Influencern auf die Sportaktivität

Laura Schlittchen

*Johannes Gutenberg-Universität Mainz*

#### Abstract

Social media are increasingly determining people's everyday lives. Especially young adults aged 18 to 38, Generation Y in Germany, spend several hours a day dealing with the topic of social media. In the course of the popularity of social platforms, so-called opinion makers, influencers or influencers emerged. Opinion makers from the fitness industry are particularly popular in the social media. This study deals with the research question 'What role do fitness influencers play on Instagram in Generation Y and their sports activity? Using a quantitative online survey, Generation Y people with an Instagram account were able to anonymously express their opinion. Four supporting hypotheses are put forward for the research question. 240 people (w=149; m=91) participated in the study. The results show that fitness influencers can influence physical activity. This opens up other possibilities for studying the influence of influencers.

#### Zusammenfassung

Die sozialen Medien bestimmen immer mehr den Alltag der Menschen. Vor allem junge Erwachsene im Alter von 18 bis 38 Jahre, die Generation Y in Deutschland, beschäftigen sich täglich mehrere Stunden mit dem Thema Social Media. Im Zuge der Beliebtheit von sozialen Plattformen bildeten sich sogenannte Meinungsmacher, Influencer oder Beeinflusser heraus. Vor allem Meinungsmacher aus der Fitnessbranche sind populär in den Sozialen Medien. Diese Studie beschäftigt sich mit der Forschungsfrage „Welche Bedeutung haben Fitness Influencer auf Instagram bei der Generation Y und deren Sportaktivität?“. Anhand einer quantitativen Online Umfrage konnten Personen der Generation Y mit einem Instagram Account ihre Meinung anonym äußern. Für die Forschungsfrage werden vier unterstützende Hypothesen aufgestellt. An der Studie haben 240 Personen (w=149; m=91) teilgenommen. Aus den Ergebnissen geht hervor, dass Fitness Influencer die sportliche Aktivität beeinflussen können. Daraus ergeben sich weitere Möglichkeiten, den Einfluss von Influencern zu untersuchen.

*Keywords:* Fitness Influencer; Sport; Soziale Medien; Instagram und Training.

#### 1. Einleitung

Das Interesse an Sozialen Medien und deren intensiver Nutzung ist in den letzten Jahren gestiegen (vgl. Grabs et al., 2018, S. 27). Bei den populärsten Netzwerken handelt es sich um Facebook, Instagram, Twitter, YouTube etc. (vgl. Deges, 2018, S. 14). Durch die Sozialen Medien können die Menschen soziale Beziehungen knüpfen und pflegen oder gemeinsame Interessen austauschen (vgl. Schmidt und Tad-dicken, 2017, S. 5). Ein beliebtes Thema auf Social Media

ist Fitness und Gesundheit (vgl. Deloitte, 2019a; Reifegerste und Baumann, 2018, S. 48). Mittels Fitness-Apps- oder Arm-bändern werden sportliche Aktivitäten auf Instagram mit der Community geteilt (vgl. Grimmer, 2019, S. 4, 8). Im Fokus der sportlichen Aktivität auf den Sozialen Medien liegt zu-meist der eigene Körper (ebd.). Der Körper ist für die Men-schen das Symbol für eine glückliche Lebensführung (jung, erfolgreich, fit und schön) (ebd.). Vor allem junge Erwach-sene nutzen die Sozialen Netzwerke als Informationsquelle

oder Motivator für einen fitness- und gesundheitsorientierten Lebensstil (vgl. Vaterlaus et al., 2015; Döring, 2015; Carrotte et al., 2015a).

Im Zuge der Beliebtheit der Plattformen bildeten sich sogenannte Meinungsmacher, Influencer oder Beeinflusser<sup>1</sup> heraus (vgl. Jahnke et al., 2018, S. 1). Influencer können die Meinungen oder das Verhalten anderer durch ihre Beliebtheit beeinflussen und kommunizieren Inhalte zu bestimmten Themen wie Mode, Beauty, Fashion, Food etc. (vgl. Jahnke et al., 2018, S. 4; Deges, 2018, S. 22f.). Auf der Social-Media-Plattform Instagram sind zahlreiche Influencer aktiv (vgl. Deges, 2018, S. 55f.). Vor allem Meinungsmacher aus der Fitnessbranche wie Pamela Reif, Sophia Thiel, Daniel Fuchs usw. sind populär in den Sozialen Medien (vgl. Statista, 2019a). Die Mitglieder der Sozialen Netzwerke vertrauen den Fitness- und Gesundheitstipps der Fitness Influencer (vgl. Pilgrim und Bohnet-Joschko, 2019). Die Meinungsmacher geben den Zuschauern das Gefühl, durch einen gesunden und starken Körper glücklich zu sein (vgl. ebd.; Raggatt et al., 2018). Zu den wesentlichen Themen gehören das sportliche Training und die Ernährung (vgl. Pilgrim und Bohnet-Joschko, 2019). Dazu gehören Trainingsmethoden, -übungen, und -geräte sowie Ernährungsstile, Diäten und Kochrezepte (vgl. Döring, 2015). Aufgrund der Popularität der Fitness Influencer beschäftigt sich diese Arbeit damit, inwiefern die sportlichen Aktivitäten der Follower durch die tägliche Kommunikation der Fitness Influencer auf Instagram beeinflusst werden können. Daher stellt sich die Frage, „Welche Bedeutung spielen Fitness Influencer auf Instagram bei der Generation Y und deren Sportaktivität?“. Anhand einer quantitativen Online Umfrage sollen Personen zwischen 18 und 38 Jahren über dieses Thema ihre Meinung anonym äußern, sodass die Forschungsfrage beantwortet werden kann.

Die vorliegende Arbeit ist in verschiedene Kapitel unterteilt. In Kapitel zwei wird zunächst versucht, die Sozialen Netzwerke und die zunehmende Popularität von Influencern in die Theorie der Kommunikation einzuordnen. Im Kapitel drei wird dann auf die Begrifflichkeiten Sport und Generationen näher eingegangen. Der Begriff Sport kann auf verschiedene Arten verstanden werden, sodass es keine festgelegte Definition gibt (vgl. Preuß et al., 2012, S. 37; Güllich und Krüger, 2013, S. 339). Jedoch sollten für diese Arbeit Sport und Sportaktivität eingegrenzt werden, um das aktuelle Verständnis von Sport darzustellen. Daraufhin werden der Begriff Generation Y und deren Eigenschaften erklärt. Diese Generation stellt eine wesentliche Zielgruppe für die Untersuchung der Arbeit dar, da sie sich täglich mehrere Stunden mit den Sozialen Medien beschäftigt (vgl. Frees und Koch, 2018; Grabs et al., 2018, S. 39). Das vierte Kapitel erläutert das Thema Soziale Medien sowie deren Nutzung. Vor allem die Funktionsweise von Instagram und der auf dieser Plattform entstandene Fitnessstrend werden beschrieben. Ein weiterer Punkt in diesem Kapitel stellen Influencer dar. Dabei werden

der Begriff, dessen Entstehung und Eigenschaften erläutert. Insbesondere Fitness Influencer sind Teil dieses Kapitels. Mit Kapitel fünf wird die methodische Vorgehensweise zur Beantwortung der Forschungsfrage dargelegt. Es wird die Fragebogenkonstruktion, die Stichprobe und die Durchführung beschrieben. Für die Forschungsfrage werden unterstützende Hypothesen aufgestellt. Das Kapitel sechs besteht aus der Darstellung der Forschungsergebnisse und der Überprüfung der Hypothesen. Die Hypothesen und die relevanten Ergebnisse der vorherigen Darstellung werden dafür zur Diskussion herangezogen. Anschließend werden die Limitierungen und weitere Forschungsmöglichkeiten der Arbeit herausgearbeitet. Das letzte Kapitel besteht aus einem Fazit, welches die wichtigsten Ergebnisse und Empfehlungen zusammenfasst.

## 2. Veränderte Kommunikationsmodelle

Die Kommunikation ist für das gemeinschaftliche Leben ein wesentlicher Bestandteil (vgl. Bonfadelli et al., 2012, S. 113). Der Begriff Kommunikation bedeutet „etwas mitteilen“ und es sind mindestens zwei Personen daran beteiligt (vgl. Seeger und Kost, 2019, S. 17). Darauf aufbauend sind verschiedene Kommunikationsmodelle entwickelt worden, die die Vorstellung eines Kommunikationsprozesses vereinfacht darstellen (vgl. Bonfadelli et al., 2012, S. 121). Nach Bentele und Beck (1994) gibt es vier Kommunikationssituationen (siehe Abb. 1). Jede Situation benötigt einen Sender, der eine Nachricht (Sachinformation, Idee, Überzeugung, Wunsch, Gefühl, Erwartung) über einen Kanal einem Empfänger übermittelt (vgl. Polzin und Weigl, 2009, S. 65). Die erste Situation beinhaltet keine Kommunikation (siehe Abb. 1). Die intentionale Abgabe von Informationen wird vom Empfänger nicht beachtet (vgl. Bonfadelli et al., 2012, S. 126). Bei der zweiten Situation kann nicht kommuniziert werden (siehe Abb. 1). Jedes Verhalten ist Teil der Kommunikation (vgl. Polzin und Weigl, 2009, S. 65). Dies bedeutet, dass Kommunikation sowohl verbal, sprachlich als auch nonverbal über Mimik, Gestik, Schrift, Bild usw. stattfinden kann (ebd.). In der dritten Situation kann eine Information vom Empfänger angenommen und verstanden werden (vgl. Bonfadelli et al., 2012, S. 126; siehe Abb. 1). Die vierte Situation ist ein zweiseitiger Informationsaustausch (ebd.). Der Empfänger nimmt die Information auf und gibt ein Feedback als Rückkopplung zum Sender (ebd.).

Im Laufe der Zeit haben sich die Kommunikation sowie die Kommunikationswege (zum Beispiel Brief, TV, Radio, Print usw.) verändert (vgl. Jahnke et al., 2018, S. 2; Gerstenberg und Gerstenberg, 2018, S. 3; Deges, 2018, S. 8). Durch die Entwicklung des Internets und neuer technologischer Geräte, wie des Smartphones, haben sich die Möglichkeiten zu kommunizieren vervielfacht (vgl. Schmidt und Taddicken, 2017, S. 4, 7). Heutzutage sind die Menschen jederzeit erreichbar und können schnell und einfach Informationen, Meinungen und Erfahrungen austauschen (vgl. Kobilke, 2019, S. 12; Deges, 2018, S. 2). Es können Nachrichten über aktuelle Ereignisse auf online Websites abgerufen, weitergeleitet oder kommentiert werden (vgl. Clement und

<sup>1</sup>Zur Verbesserung der Lesbarkeit werden in dieser Arbeit Personenbezeichnungen in der männlichen Form verwendet; gemeint sind dabei in allen Fällen Frauen und Männer.

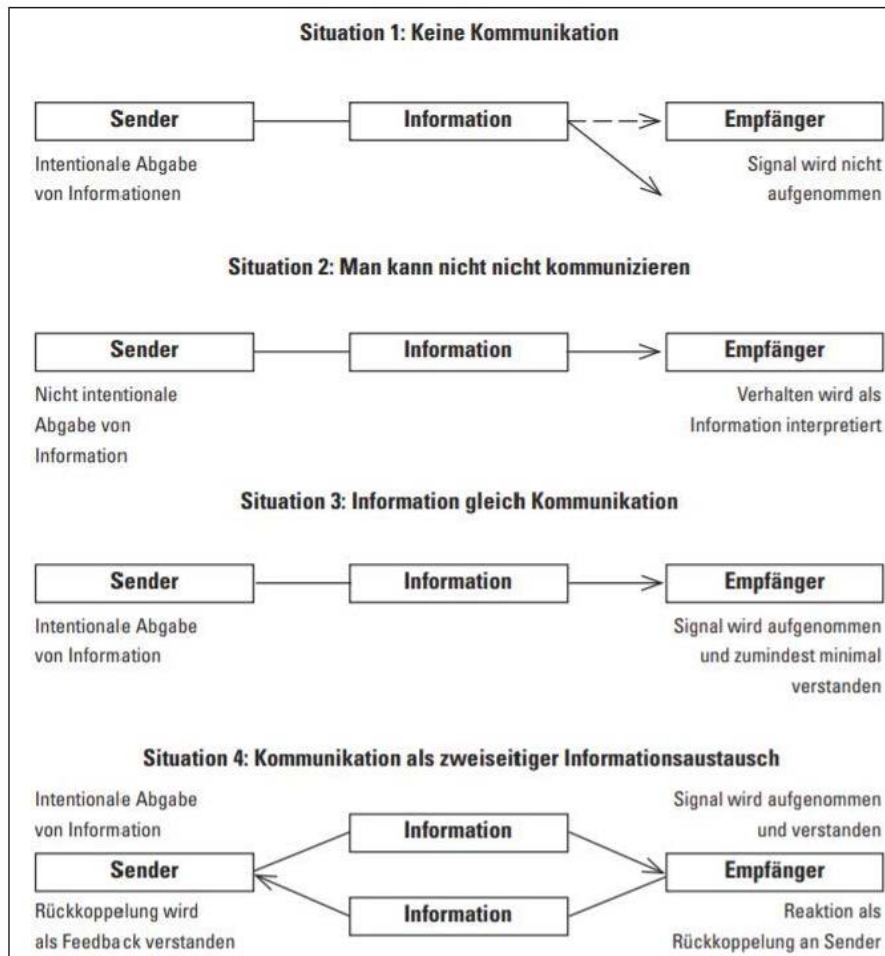


Abbildung 1: Kommunikationsmodell nach Bentele und Beck (1994): 23 (Quelle: Bonfadelli et al., 2012, S. 125)

Schreiber, 2013, S. 424). Mit dem Internet entwickelten sich Soziale Netzwerke/Medien (vgl. Schmidt und Taddicken, 2017, S. 4). Unter Sozialen Medien werden „Apps und Websites, in denen Nutzer Inhalte (z. B. Fotos, Videos, Artikel) kreieren und mit anderen Nutzern teilen können.“ verstanden (Jahnke et al., 2018, S. 275). Dazu zählen die Netzwerke Twitter, Facebook, Snapchat und Instagram (vgl. Jahnke et al., 2018, S. 275). Durch die Sozialen Medien („Digitale Kanäle“) ist die klassische unidirektionale Kommunikation zu einer mehrdirektionalen geworden (vgl. Seeger und Kost, 2019, S.15). Es sind weitere Kommunikationskanäle hinzugekommen, die von mehreren Personen individuell genutzt werden (vgl. Schach und Lommatzsch, 2018, S. 253). Im Zuge dieser Veränderungen haben sich die klassischen Kommunikationsmodelle weiterentwickelt (vgl. Gerstenberg und Gerstenberg, 2018, S. 3). Die Sender und Empfänger treten in Interaktion und die Kommunikation erfolgt zeitnah über die Sozialen Medien (vgl. Deges, 2018, S.9; Gerstenberg und Gerstenberg, 2018, S. 13). Dementsprechend erhält der Sender sofort eine Rückmeldung, wenn andere Netzwerkmitglieder die Botschaften des Senders kommentieren, bewerten oder ergänzen (vgl. Deges, 2018, S. 2; Clement und Schreiber, 2013, S. 424). Dadurch entsteht eine wech-

selseitige Kommunikation und die Botschaft erreicht anstatt einer Person mehrere (vgl. Gerstenberg und Gerstenberg, 2018, S. 4). Der Empfänger kann durch das Weiterleiten der Botschaften an andere Personen selbst zu einem Sender werden (vgl. Deges, 2018, S. 2-9). Dieser Austausch zwischen Sendern und Empfängern kann die Meinungsbildung zu verschiedenen Themen beeinflussen (ebd.). Die Empfehlungen von anonymen Personen aus dem Netz gelten als wahr (ebd.). Sogenannte Meinungsführer („Opinion Leader“) sind Personen, „die häufig um Rat und Meinung gebeten werden und dadurch Einfluss auf andere Personen haben“ (Schach und Lommatzsch, 2018, S. 29). In den Sozialen Medien werden Meinungsführer heutzutage als Influencer (Beeinflusser) bezeichnet (vgl. Schach und Lommatzsch, 2018, S. 76). Influencer erreichen durch ihre tägliche Kommunikation über ein oder mehrere bestimmte Themen viele Interessenten und werden als glaubwürdige und authentische Informationsquellen angesehen (vgl. ebd., S.149; Deges, 2018, S. 1). Schon im Jahre 1940 stellten Lazarsfeld, Berelson und Gaudet bei der US-amerikanischen Präsidentschaftswahl fest, dass Meinungsführer einen starken Einfluss auf das Verhalten und die Einstellungen der Empfänger haben (vgl. Schach und Lommatzsch, 2018, S. 6). Das entstandene Kommunika-

tionsmodell Two-Step-Flow-of-Communication verdeutlicht die Position der Meinungsführer gegenüber der Massenkommunikation (vgl. Bonfadelli und Friemel, 2011, S. 212-215; siehe Abb. 2). Dieses Kommunikationsmodell bildet den Grundstein der Meinungsbildung und kann auf Influencer übertragen werden (vgl. Schach und Lommatzsch, 2018, S.6). Demnach gelangen die Ideen auf einer ersten Stufe von den Medien zu den Influencern und von diesen dann zu den meinungsübernehmenden Konsumenten (ebd.). Robert Cialdini (2001) vertieft das Phänomen der Beeinflussung durch bestimmte Personen. Er geht davon aus, dass die Menschen leicht beeinflussbar und die Handlungen bis zu einem gewissen Grad steuerbar sind (vgl. Nirschl und Steinberg, 2018, S. 5). Demnach verlassen sich die Menschen auf den Rat von Influencern (ebd.).

Es wird deutlich, dass sich die Kommunikationsmittel als auch das Kommunizieren zwischen den Personen verändert haben. Eine Veränderung konnte ebenfalls beim Sporttreiben festgestellt werden (vgl. Grimmer, 2019, S. 4).

### 3. Sport bei der Generation Y

Heutzutage sind sportliche Aktivitäten ein zentrales Element in der Freizeitgestaltung (vgl. Preuß et al., 2012, S. 21). In Deutschland sind es rund 11,81 Millionen Menschen, die mehrmals wöchentlich Sport treiben (vgl. VuMa, Arbeitsgemeinschaft Verbrauchs- und Medienanalyse, 2018a). Mit dem Sport wird eine positive Auswirkung auf die Gesundheit verbunden und die dadurch gesteigerte Lebensqualität (vgl. Hottenrott und Neumann, 2016, S. 12). Im Vordergrund steht der Spaß an Bewegung (vgl. Preuß et al., 2012, S. 95). Vor allem junge deutsche Erwachsene im Alter von 18 bis 38 Jahren, die Generation Y, treiben mehrmals in der Woche Sport (vgl. Mills, 2019). Die Gründe für den Zuspruch am Sporttreiben dieser Generation liegen daran, dass sie sowohl ihre Gesundheit als auch ihren Körper in Form halten und verbessern wollen (ebd.). Für die Generation Y schafft der Sport einen Ausgleich zum Alltag und bietet die Möglichkeit, regelmäßig soziale Beziehungen zu pflegen (ebd.). Heutzutage steht der Sport für Spaß, Gesundheit und Ausgleich zum Alltag. Jedoch hat sich die Bedeutung des Begriffs Sport im Laufe der Zeit verändert (vgl. Schütte, 2015, S. 18).

#### 3.1. Sport im Wandel

In Deutschland basiert der traditionelle Sport auf Leistung, Disziplin, Wettkampf, Geselligkeit und Ehrenamtlichkeit (vgl. Güllich und Krüger, 2013, S. 388). In den vergangenen Jahren rückte der Leistungs- und Wettkampfgedanke in den Hintergrund. Stattdessen gilt der Sport für Erlebnis, Geselligkeit, Spaß, Gesundheit und Wohlbefinden (vgl. Schütte, 2015, S. 18). Hinzu kommt, dass sich das organisierte Sporttreiben im Verein zu einem selbstorganisierten Sporttreiben in der Freizeit löste (vgl. Krüger et al., 2004, S. 430). Sport wird individuell, je nach Bedarf und ohne Verbindlichkeiten betrieben (vgl. Güllich und Krüger, 2013, S. 387f.). Das klassische Sportangebot verändert sich (ebd.). Jedes Jahr werden

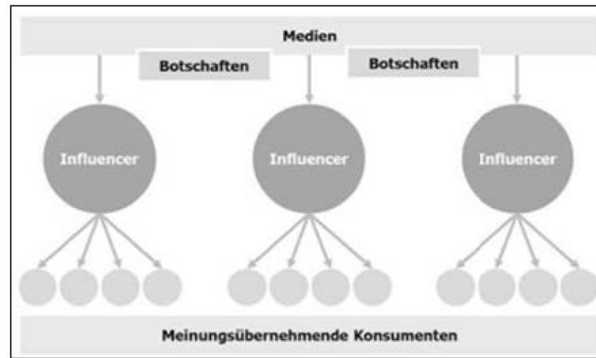
auf Deutschlands größter Sportmesse, der ISPO, neue Trends in der Sportbranche vorgestellt (ISPO Munich, 2018). Die sogenannten Trendsportarten sind innovative Bewegungsmethoden und grenzen sich von den klassischen Sportarten ab (vgl. Krüger et al., 2004, S. 432). Es existieren eine Vielzahl von neuen Sportarten wie Inlineskaten, Le Parkour, BMX-Radfahren usw. (vgl. Güllich und Krüger, 2013, S. 532). Im Vordergrund stehen die Werte „Spaß, Gesundheit, Selbstverwirklichung, Erlebnis, Ästhetik, Fit- und Schönsein“ (Güllich und Krüger, 2013, S. 388). Es wird deutlich, dass Sport ein vielseitiger, äußerst komplexer und sich stetig verändernder Begriff ist (vgl. Preuß et al., 2012, S. 37). Mit der Zeit wurde der facettenreiche Begriff Sport von verschiedenen Personen im Alltag unterschiedlich verwendet (vgl. Güllich und Krüger, 2013, S. 339; Schütte, 2015, S. 18). Daher gibt es heute keine festgelegte Definition (ebd.). Jedoch kann Sport im engeren Sinne verstanden werden

„[...] als Bewegungsverhalten, das nach kulturell vorgegebenen Deutungsmustern von Menschen in sozialen Kontexten durchgeführt und als sportbezogen erlebt wird. [...] Menschen versuchen, unter gleichen Bedingungen einen ungleichen Ausgang zu erzielen – der Bessere möge gewinnen. Chancengleichheit ist die Bedingung dafür, dass das Leistungs- und Wettkampfbzw. Konkurrenzprinzip zur Entfaltung kommen kann.“ (Güllich und Krüger, 2013, S. 339).

Demnach beinhaltet Sport eine körperliche Bewegung, die je nach Kultur unterschiedlich ausgeprägt ist. Die engere Definition kommt dem traditionellen Sportverständnis nahe, dass der Leistungsvergleich im Wettkampf bedeutsam ist (ebd., S. 388). Im weiteren Sinne bedeutet Sport:

„Menschen ihrem bewegungsbezogenen Handeln subjektiv eine sportliche Sinnperspektive zuweisen, die durchaus vom Sport im engeren Sinn abweichen kann, jedoch einen Bezug zu Werten hat, die aus dem ursprünglichen englischen Sport stammen, sei es Fairplay, Leistungsbezogenheit, das Sich vergleichen oder auch nur Freude am Bewegen und Spielen, an der demonstrativen Verausgabung von Kräften oder an der systematischen Verbesserung eigener sportlich physischer Leistungen über die Zeit oder im Vergleich mit anderen.“ (Güllich und Krüger, 2013, S. 339).

Die Bedeutung von Sport wird erweitert und beinhaltet sowohl den Leistungsvergleich im Wettkampf als das Bewegen in der Freizeit. Der Sport kann, ohne das Ziel, an einem Wettkampf teilzunehmen, mit Freude betrieben werden. Weiterhin kann der Sport als Mittel zur sozialen Präsentation genutzt werden (vgl. Güllich und Krüger, 2013, S. 339). Um die Forschungsfrage dieser Arbeit beantworten zu können, müssen zunächst die Begriffe Sportaktivität und Generation definiert werden.



**Abbildung 2:** Two-step-flow-Modell nach Lazarsfeld, Berelson und Gaudet (1944) (Quelle: Nirschl und Steinberg, 2018, S. 29)

### 3.2. Sportaktivität

Unter der sportlichen Aktivität wird die Ausübung von Sport verstanden (vgl. Weltgesundheitsorganisation, 2019). Dies kann der körperlichen Aktivität zugeordnet werden (ebd.). Unter körperlicher Aktivität wird „jede körperliche Bewegung der Skelettmuskulatur, die einen hohen Energieaufwand erfordert“ verstanden (Weltgesundheitsorganisation, 2019). Sie umfasst unter anderem körperliche Aktivitäten am Arbeitsplatz, im Haushalt sowie in der Freizeit (vgl. Weltgesundheitsorganisation, 2019). Die sportlichen Aktivitäten können als Überbegriff für unterschiedliche motivationale Orientierungen wie zum Beispiel Abenteuer-, Leistungs-, Gesundheits-, Extrem-, Kinder- und Integrationssport stehen (vgl. Hottenrott und Neumann, 2016, S.15f.). In den letzten Jahren erhielt der Fitnesssport mehr Aufmerksamkeit (vgl. Güllich und Krüger, 2013, S. 388; Hottenrott und Neumann, 2016, S. 20). „Als Fitness wird die Fähigkeit des Menschen bezeichnet, den Anforderungen der Umwelt und des täglichen Lebens (Berufsfähigkeit) zu entsprechen“ (Hottenrott und Neumann, 2016, S. 20). Demnach trägt der Sport dazu bei, alltägliche Aufgaben wie Treppensteigen, Gartenarbeit usw. zu bewältigen (vgl. Weltgesundheitsorganisation, 2019; Güllich und Krüger, 2013, S. 532). Laut dem DOSB kann „[...] Fitness durch regelmäßige, maßvolle und freudvolle sportliche Betätigung [...], erworben werden“ (DOSB, zitiert nach Hottenrott und Neumann, 2016, S. 20). Somit passt der Fitnesssport in das Verständnis von Sport im weiteren Sinne (siehe Kapitel 3.1). Beim Fitnesssport stehen neben funktionalen und gesundheitlichen Gründen ebenso körperperformende und gewichtsregulierende Intentionen eine entscheidende Rolle (vgl. Güllich und Krüger, 2013, S. 442). Der Fokus des Fitnesssports liegt nicht nur auf der Bewältigung von körperlichen Aktivitäten, sondern durch das Sporttreiben soll ein bestimmtes Körperbild geformt werden (vgl. Lange & Baschta, 2013, zitiert nach Hottenrott und Neumann, 2016, S. 20). Dem Fitnesssport können eine Vielzahl von Sportarten zugeordnet werden wie zum Beispiel Yoga, Pilates, Aerobic usw. (vgl. ebd.; Dilger, 2008, S. 407). Dieser Sport kann von jeder Ziel- und Altersgruppe betrieben werden (vgl. Hottenrott und Neumann, 2016, S. 20). Vor allem bei jungen Erwachsenen ist der Fitnesssport beliebt (ebd.). Zusätzlich gibt es

zahlreiche Online Fitness-Plattformen, oder -Apps (vgl. Mills, 2019). In Form von Videos können Trainingseinheiten übernommen werden. Der Fitnessmarkt boomt und wird von verschiedenen Generationen gelebt (vgl. SAZ Sport, 2018).

### 3.3. Generation

Die Menschen können in verschiedene Generationen nach Geburtsjahr eingeordnet werden (vgl. Mangelsdorf, 2017, S. 12f.). Für die Einteilung finden sich verschiedene Definition in der Literatur wieder. In Deutschland werden Menschen, die in den Jahren 1922 bis 1945 geboren sind, als die Traditionalisten bezeichnet (ebd.). Danach folgen die sogenannten Babyboomer, die in den Jahren von 1946 bis 1964 geboren sind (ebd.). Die Generation X ist in den Jahren von 1965 bis 1979 geboren (ebd.). Zur Generation Y, auch die Millennials genannt, gehören die in den Jahren 1980 bis 1995 geborenen Menschen (vgl. ebd.; Jahnke et al., 2018, S. 17). Heutzutage findet sich eine weitere Generation in der Gesellschaft wieder, die Generation Z. Diese bezeichnet alle Geburtsjahre ab 1996 (ebd.). Mangelsdorf (2017, S. 12) sind Generationen mehr als eine Einordnung von Menschen gleichen Alters. Es handelt sich vielmehr um ein Wertemuster und die daraus entstehende soziale Orientierung und Lebensauffassung der Menschen (vgl. Mangelsdorf, 2017, S. 12f.). Eine relevante Generation für diese Arbeit ist die „Generation Y, Millennials, Digital Natives, Generation me“ (Lorenz, 2019). Für Krampf und Weichert (2017, S.9) gehören zu dieser Generation Y die in den Jahren 1981 bis 2000 (zwischen 18 und 38 Jahre) geborenen Menschen. Diese Definition wird für diese Arbeit verwendet. Das Y wurde von dem englischen Wort „Why“ abgeleitet und bedeutet „warum“ (vgl. Würzburger, 2016, S. 3). Da sie bestimmte Sachverhalte und Werte stets hinterfragen (ebd.). Diese Generation ist an neuen Trends wie beispielsweise aktuelle Trendsportarten, neue Trainingsmethoden oder -geräte interessiert (vgl. Mills, 2019; AS&S Radio GmbH, 2016). Im Gegensatz zu den vorherigen Generationen, ist die Generation Y mit dem Internet groß geworden (vgl. Würzburger, 2016, S. 2). Diese Personen werden als „Digital Natives“ bezeichnet, da sie mit dem digitalen Zeitalter aufgewachsen und mit digitalen Technologien vertraut sind (ebd.). Für diese Generation spielt das

Internet eine wesentliche Rolle (ebd.). Schließlich bilden sie eine große Nutzergruppe in der Onlineaktivität ab (vgl. Deloitte, 2015). Über 80 Prozent der Millennials sind täglich mehrere Stunden im Internet aktiv (vgl. VuMa, Arbeitsgemeinschaft für Verbrauchs- und Medienanalyse, 2018c). Das digitale Netzwerk wird für Nachrichten, Recherchearbeit und Unterhaltung genutzt (vgl. Deloitte, 2015; Krampf und Weichert, 2017, S. 21). Vor allem die Generation Y konsumiert Nachrichten und erhält Informationen mobil auf ihrem Smartphone und beschäftigt sich täglich mehrere Stunden mit dem Thema Social Media (vgl. ebd., S. 102; Krampf und Weichert, 2017, S. 9; Frees und Koch, 2018; Grabs et al., 2018, S. 39).

#### 4. Soziale Medien

Im Laufe des 21. Jahrhunderts bildeten sich Soziale Medien bzw. Netzwerke wie Facebook, Instagram, Snapchat, Xing, Twitter usw. heraus (vgl. Deges, 2018, S. 2,14). Die Sozialen Medien ermöglichen den Nutzern, sich über Informationen, Eindrücke, Meinungen oder Erfahrungen auszutauschen (vgl. BVDW 2009: 5, zitiert nach Goldapp, 2015, S. 29). Durch die aktive Teilnahme der Mitglieder in Form von Kommentieren, Bewerten oder Empfehlen der Inhalte, können soziale Beziehungen entstehen (ebd.). Laut der YouGov-Studie können sich 47 Prozent aller Deutschen nicht vorstellen, einen Tag lang auf Soziale Medien zu verzichten (ebd., S.31). Die Menschen nutzen die Sozialen Medien „[...] zur Kommunikation, zur Inspiration und Interaktion mit Freunden oder Internetbekanntschaften.“ (Grabs et al., 2018, S. 27). Die Nachrichten können direkt nach dem Aufwachen im Bett auf mobilen Endgeräten nachgeschlagen werden und bis abends vor dem Schlafengehen verfolgt werden (ebd.). Bei der Benutzung der Sozialen Medien verbergen sich Gefahren wie Depression oder Schlafmangel (vgl. Royal Society for Public Health, 2017; Zingg, 2018). Die Social-Media-Plattformen lenken vom Alltag ab, sodass keine Zeit mehr zum Sporttreiben bleibt (vgl. Vaterlaus et al., 2015; Lüdemann, 2018). Daher ist der Wunsch nach einer Auszeit oder Einschränkung der Sozialen Netzwerke in der Gesellschaft vorhanden (vgl. Vogl, 2019). Die Social-Media-Kanäle bestimmen mit 66 Prozent die Freizeit der deutschen Bevölkerung (vgl. VuMa, Arbeitsgemeinschaft Verbrauchs- und Medienanalyse, 2018b; Deges, 2018, S. 2). Vor allem die Generation Y, zwischen 18 und 38 Jahren, stellt einen Großteil der Instagram Nutzer in Deutschland (vgl. Grabs et al., 2018, S. 287; We Are Social & Hootsuite & DataReportal, 2019). Instagram bietet dieser Generation eine kurzweilige Unterhaltung und kann mobil über eine App abgerufen werden (vgl. Grabs et al., 2018, S. 279; Deges, 2018, S. 54f.).

##### 4.1. Instagram

Die audiovisuelle Social-Media-Plattform Instagram wurde 2010 in San Francisco (USA) gegründet (vgl. Jahnke et al., 2018, S. 71). In Deutschland nutzen rund 15 Millionen Menschen Instagram (vgl. Firsching, 2019). Es ist davon auszugehen, dass die Nutzungszahl bis 2021 auf 16,3 Millionen

weiter ansteigen wird (vgl. Emarketer, 2017). Instagram ist eine mobile App. Der Nutzer erhält Zugang zu den Inhalten der Plattform, indem er die App auf seinem Smartphone installiert (vgl. Jahnke et al., 2018, S. 71; Deges, 2018, S. 55). Die Mitglieder von Instagram („Community“) können als User, Follower, Abonnenten, Fans oder Freunde bezeichnet werden (vgl. Deges, 2018, S. 28; Schach und Lommatzsch, 2018, S. 49). Das Grundkonzept basiert darauf, Fotos und Videos zu veröffentlichen (vgl. ebd.; Kobilke, 2019, S. 26). Es werden Inhalte zu verschiedenen Themenbereichen wie Sport, Reisen, Interieur, Kosmetik, Food, Fashion etc. publiziert (vgl. Jahnke et al., 2018, S. 72). Die veröffentlichten Inhalte können durch vielseitige Möglichkeiten der Bildbearbeitung in Form von Effekten und Filtern gestaltet und mit der Community geteilt werden (vgl. Deges, 2018, S. 55). Der Fokus liegt auf einer künstlerischen und ästhetischen Darbietung der Bilder (vgl. Faßmann und Moss, 2016, S. 27f.). Die veröffentlichten und meistens inszenierten Inhalte werden zur Inspiration für verschiedene Bereiche von den Usern genutzt (ebd.). Die Bilder oder Videos können mit Hashtags kommentiert, externen Internetseiten verlinkt und Orten („Geotagging“) oder Personen markiert werden (vgl. Kobilke, 2019, S. 27; Grabs et al., 2018, S. 283). Unter einem Hashtag wird „[...] ein Rautenkreuz # gekennzeichnete Markierung und Verschlagwortung von Inhalten, die dadurch gebündelt werden“ verstanden (Deges, 2018, S. 124). Diese Funktion dient dazu, ein Bild mit einem Wort zu verknüpfen (vgl. Deges, 2018, S. 56). Das Bild mit dem genannten Hashtag kann in der Suchfunktion von Instagram gefunden werden (ebd.). Durch die genannten Funktionen ist der Inhalt jederzeit auffindbar (vgl. Kobilke, 2019, S. 26ff.). Die Community kann wiederum diesen Inhalt sowohl mit „Gefällt mir“ („Liken“) bewerten und kommentieren als auch teilen (ebd.). Der Fokus liegt auf der Startseite, dort werden nacheinander die Neuigkeiten von den abonnierten Personen angezeigt. Neben dem Homefeed auf der Startseite können Instagram Stories und Live Videos aufgenommen sowie über die Messenger Funktion Textnachrichten an eine oder mehrere Personen gesendet werden (ebd.). Instagram vereint mehrere Funktionen, die den heutigen Zeitgeist treffen (vgl. Grabs et al., 2018, S. 279). Neben der privaten Nutzung eines Instagram Profils, können Unternehmen oder Personen des öffentlichen Lebens ein Instagram Business Profil erstellen. Die Nutzer können Firmen wie Coca-Cola, Ikea, adidas usw. folgen und werden stets über aktuelle Produkte, Kampagnen etc. informiert (vgl. Kobilke, 2019, S. 16). Das Business Profil unterscheidet sich von einem privaten Profil lediglich durch Statistiktools und weitere Interaktionsmöglichkeiten mit den Followern (ebd. S. 27f.). Im Zuge der Beliebtheit dieser Plattform und des Fitnesssports bildeten sich Fitnesstrends auf Instagram heraus (vgl. Tiggemann und Zaccardo, 2015).

##### 4.2. Fitnesstrend auf Instagram

Der bewusste Umgang mit dem Körper spiegelt den heutigen Zeitgeist wider (vgl. Güllich und Krüger, 2013, S. 388). In den vergangenen Jahren hat sich im englischen Sprachraum bei Instagram ein Sporttrend herauskristallisiert. Un-

ter dem Hashtag, „Fitspiration“ auch genannt „fitspo“, beschreibt die Fitnessinspiration, veröffentlichten Personen von Instagram Bilder und Videos mit Ernährungs- und Sporttipps (vgl. Carrotte et al., 2017; Boepple et al., 2016; Tiggemann und Zaccardo, 2015). Fitspiration soll jeden einzelnen zu Bewegung und Gesundheit anregen (ebd.). Die Fitspo-Inhalte präsentieren ein muskulöses Körperbild, sowohl von Frauen als auch Männern (vgl. Carrotte et al., 2017). Diese Inhalte unterscheiden sich von „Thinspiration“ auch genannt „thinspo“, dort werden schlanke und dünne Körper präsentiert (ebd.). Im Laufe der Zeit veränderte sich das Schönheitsideal von einem ehemals dünnen zu einem trainierten Körperbild (vgl. Deighton-Smith und Bell, 2018). Anhand einer Figuren-Bewertungs-Skala von Stunkard et al. (1983) für Dünn und Dick und einer modifizierten Körper-Bewertungs-Skala für Muskulös von Lynch und Zellner (1999) können Körperbilder von Frauen und Männern eingeteilt werden (vgl. Carrotte et al., 2017; siehe Abb. 10, 11). Diese Skalen zeigen einfache Zeichnungen von männlichen und weiblichen Figuren, um die Dicke und Muskulatur von Männern und Frauen auf Skalen von 1 (sehr dünner oder sehr wenig sichtbarer Muskel) bis 9 (sehr übergewichtig oder sehr muskulös) zu beurteilen (vgl. Carrotte et al., 2017; Grossbard et al., 2011; siehe Abb. 10, 11). Die veröffentlichten Fitspiration-Inhalte von Frauen vermitteln ein gesundes, fittes und starkes Bild (vgl. Kite, 2017). Bei den Männern ist das muskulöse Körperideal beliebt (vgl. Carrotte et al., 2017; Grossbard et al., 2011). Die veröffentlichten „fitspo“ Medien haben das Potenzial, positiv zum Sporttreiben anzuregen (vgl. Carrotte et al., 2017). Des Weiteren animieren die fitnessbezogenen Inhalte zu einem gesünderen Lebensstil (vgl. Santarossa et al., 2019). Neben der positiven Wirkung von Fitnessbeiträgen bestehen auch negative Gefühle bei der Betrachtung von trainierten Körpern (vgl. Raggatt et al., 2018; Easton et al., 2018). Die gezeigten Bilder und Videos von perfekten Körperbildern können zu Essstörungen, verzerrten Schönheitsidealen und Depressionen führen (vgl. Tiggemann und Zaccardo, 2015; Holand und Tiggemann, 2017). Vor allem bei weiblichen Usern kann der Vergleich mit anderen Körperbildern zu Essstörungen führen (vgl. Holand und Tiggemann, 2017). Daraufhin entwickelte sich ein Gegentrend „body positivity“, der den schlanken, makellosen Fitness-Körpern entgegentritt (vgl. Gerling et al., 2018, S. 69). Dieser Trend steht dafür, dass jede Körperform schön ist und akzeptiert werden sollte (ebd.). In Deutschland gibt es unter dem Hashtag „FitFamGermany“, der eine Kombination aus Fitness und Familie darstellt, Beiträge zum Fitnesssport auf Instagram (vgl. Klapp und Klotter, 2019). Dort gehören die Kategorien „schlank“ und „muskulös“ zum normierten Körperideal (ebd.). Unter dem Hashtag „NotHeidisGirl“ werden Instagram Beiträge mit den Themenschwerpunkte „Fitness und Bodypositives“ veröffentlicht (ebd.). Diese Beiträge spielen auf die deutsche TV-Castingshow „Germany’s Next Topmodel“ von Heidi Klum an, wo häufig schlanke weibliche Personen als Idealtyp dargestellt werden. Mit „NotHeidisGirl“ wird zu mehr Diversität auf Instagram aufgerufen (ebd.). Im Zuge der visuellen Darstellung auf Instagram ist das Inszenieren des Körperbildes

immer wichtiger geworden (vgl. Kobilke, 2019, S. 16). Ein wesentlicher Beitrag für die entstehenden Trends auf Instagram tragen Influencer (vgl. Jahnke et al., 2018, S. 215).

#### 4.3. Das Phänomen Influencer

Instagram bietet den Mitgliedern eine kunstvolle und unterhaltsame Inszenierung in Form von Bildern, Videos und Bewegtbildern an (vgl. Grabs et al., 2018, S. 281f.). In den letzten Jahren sind Personen aus dem Alltag durch ihre ästhetisch ansprechenden Bilder, witzige Stories und authentische Blogs bekannt geworden (vgl. Deges, 2018, S. 24; Grabs et al., 2018, S. 282). Vor allem zeichnen sich die Blogs durch die Persönlichkeit des Verfassers aus, da sie einem Tagebucheintrag ähneln (vgl. Nirschl und Steinberg, 2018, S. 19). Die Blogbeiträge können neben persönlichen Texten auch Themen- oder Firmenblogs sein (ebd.). Die Reichweite als auch die Glaubwürdigkeit und das Vertrauen der Community zu der Person nehmen zu (ebd.). Es sind Personen, die zu Influencern (Beeinflussern) werden (vgl. Deges, 2018, S. 13). Der Ausdruck Influencer kann aus dem Englischen „influence“ (Einfluss) abgeleitet werden (vgl. Schach und Lommatzsch, 2018, S.31).

„Ein Influencer ist eine natürliche Person, die einen Blog und/oder einen Account in sozialen Netzwerken mit hochwertigen Informationen zu einem Thema bespielt und dem ihm größtenteils unbekannte Personen als Fans/Follower/Abonnenten folgen, weil sie seinen Inhalten ein hohes Interesse entgegenbringen und ihn als glaubwürdigen und authentischen Experten wahrnehmen.“ (Deges, 2018, S.16).

Influencer teilen ihr privates Leben mittels Bildern, Videos oder Live-Übertragungen mit ihrer Instagram Community (vgl. Deges, 2018, S.14-20, 61). Durch die regelmäßige Kommunikation der Influencer wird der Follower zum alltäglichen Begleiter und Teil einer Community (ebd.). Der Follower vertraut seinem Vorbild (vgl. Sokolova und Kefi, 2019). Dadurch erreicht der Influencer eine hohe Glaubwürdigkeit mit seinen Inhalten und Botschaften (vgl. Kobilke, 2019, S. 223; Nirschl und Steinberg, 2018, S. 28). Influencer bauen zu ihrer Community eine emotionale Verbindung auf, sodass es zu parasozialen Beziehungen kommen kann (vgl. Deges, 2018, S. 15, 20; Sokolova und Kefi, 2019).

„Der Begriff beschreibt „scheinbar zwischenmenschliche Beziehungen“, die durch wiederholte Einzelkontakte und Interaktionen zustande kommen und im Grunde die Phänomene Empathie, sozialer Vergleich oder Identifikation umfassen können.“ (Döring, 2013, S. 303, zitiert nach Schach und Lommatzsch, 2018, S.14).

Laut dem Bundesverband für Digitale Wirtschaft und Influx (2017) haben fast ein Drittel der Social Media Nutzer ein empfohlenes Produkt von einem Influencer gekauft. Mit der Professionalisierung dieser Branche können Influencer

in verschiedene Typen eingeteilt werden (vgl. Grabs et al., 2018, S. 306; Deges, 2018, S. 13).

#### 4.3.1. Influencertypen

Es gibt eine Vielzahl von Personen, die den Beruf des Influencer ausüben (vgl. Jahnke et al., 2018, S. 9). Daher teilt Deges (2018, S. 13) Influencer nach verschiedenen Typen ein (siehe Abb. 3). Eine Kategorisierung ist der Social-Media-Kanal (YouTube, Instagram, Facebook, Snapchat, Music.ly), wo Influencer aktiv kommunizieren (vgl. Deges, 2018, S.21f.; Jahnke et al., 2018, S. 111). Die Meinungsführer können, je nach Autor, anhand ihrer Reichweite in Macro-, Micro-, Key-, Local- und Potential-Influencer eingeteilt werden (ebd.). Unter einem Macro-Influencer wird eine Person mit mehr als 100.000 Followern verstanden (vgl. Jahnke et al., 2018, S.111). Damit besitzt diese Kategorie von Influencern eine hohe Reichweite in den Sozialen Medien (vgl. Kobilke, 2019, S. 226; Grabs et al., 2018, S. 128). Als Micro-Influencer werden Personen unter 100.000 Followern bezeichnet (vgl. Jahnke et al., 2018, S. 111). Diese Personen verfügen über eine geringere Reichweite, aber dafür zeichnen sich Micro-Influencer durch ihr hohes Engagement aus (vgl. Kobilke, 2019, S. 226; Seeger und Kost, 2019, S. 31). Des Weiteren ist der Inhalt dieser Personen spezialisiert auf ein Nischenthema (Key-Influencer) und/oder auf eine Region (Local-Influencer) (vgl. Deges, 2018, S. 23f.). Der gesellschaftliche Status von Influencern nimmt stetig zu, sodass sie zu sogenannten Promi-Influencern werden (ebd.). Heutzutage werben nicht nur Prominente aus den Bereichen Sport, Musik, Mode, Film, Fernsehen und Film für etablierte Marken, sondern die Aufgaben werden zunehmend von Meinungsmachern aus den Sozialen Medien übernommen (vgl. Jahnke et al., 2018, S.1; Kilian, 2016; Deges, 2018, S. 19). Daraus sind drei Sonderformen Peer-, Corporate- und Pet-Influencer entstanden (vgl. Deges, 2018, S. 20-27; siehe Abb. 3). Diese Personen arbeiten mit oder für Unternehmen (ebd.). Durch die soziodemografischen Merkmale Alter/Geschlecht, Nationalität, Sprache, Professional- oder Hobby Influencer, können verschiedene Zielgruppen angesprochen werden (ebd.). Die Sprachauswahl der Influencer kann eine nationale oder internationale Zielgruppe ansprechen oder je nach Geschlecht verschiedene Themen beinhalten (ebd.).

Die Influencer Szene erstreckt sich über nahezu alle Themenbereiche wie Beauty, Fitness, Fashion, Food, Travel, Kunst, Design usw. (vgl. Jahnke et al., 2018, S. 78; Deges, 2018, S. 22; Seeger und Kost, 2019, S. 38). Vor allem Influencer aus der Fitnessbranche sind bei den Followern beliebt (vgl. Statista, 2019a).

#### 4.3.2. Fitness Influencer

Fitness Influencer auf Instagram besitzen eine hohe Reichweite (vgl. Firsching, 2015). Sowohl weibliche als auch männliche Fitness Influencer sind populär in Deutschland (ebd.). Zu den bekanntesten deutschen Meinungsmachern der Fitnessbranche auf Instagram gehören Pamela Reif (4,3 Millionen Follower, Reif, Pamela, 2019) und Daniel Fuchs

(1,9 Millionen Follower, Fuchs, Daniel, 2019) (vgl. Scherkamp, 2017). Die Meinungsmacher aus dem Fitnessbereich zeichnen sich durch ihre strenge Trainingsroutine und reglementierte Nahrungsaufnahme aus, um ein bestimmtes Körperbild zu erreichen oder zu erhalten (vgl. Pilgrim und Bohnet-Joschko, 2019). Dabei präsentieren sie sich im Fitnessstudio mit enganliegender Sportbekleidung und geben Einblicke in ihre vergangenen oder zukünftigen Trainingspläne mittels Bilder und Videos (ebd.). Die Begeisterung der Community an den veröffentlichten Inhalten der Meinungsführer ist groß (vgl. Döring, 2015). Eine quantitative Inhaltsanalyse von 25 deutschsprachigen Lifestyle- und Fitnesskanälen auf der Plattform YouTube zeigt die positive Wirkung der Videos auf den Follower (ebd.). Die kostenlosen Videos der Fitness Influencer dienen zur regelmäßigen Inspiration und Motivation für einen gesunden Lebensstil (ebd.). Die Abonnenten nutzen die Trainingsvideos für den Frühsport (ebd.). In der Fitnessbranche existieren zahlreiche online Fitnessprogramme von Fitness Influencern, in denen Übungen vorgeführt oder kommentiert werden (vgl. Deges, 2018, S. 59). Das Fitnessprogramm von Anne Kissner (229.000 Follower, Kissner, Anne, 2019) wird in erster Linie für Vorher-Nachher-Vergleiche der Teilnehmer eingesetzt, die den Trainingserfolg bzw. die geglückte Transformation durch eine Gewichtsreduktion präsentieren (vgl. Gerling et al., 2018, S. 66). Außerdem werden von Influencern sogenannte „Challenges“ gestartet, die ihre Reichweite und die Sichtbarkeit auf Instagram erhöhen (vgl. Kobilke, 2019, S. 210). Gleichzeitig wird die Fitness-Community dazu bewegt, an den Challenges teilzunehmen (ebd.). Zum Beispiel können unter dem Hashtag „UPPERBODYCHALLENGE“ von der australischen Fitness Influencerin Kayla Itsines (11,8 Millionen Follower, Itsines, Kayla, 2019), die Follower ihre Trainingserfolge veröffentlichen (vgl. Gerling et al., 2018, S. 66; Frank, 2018). Bei dieser Challenge geht es darum, so viele Runden von verschiedenen Übungen in zehn Minuten zu schaffen (vgl. Itsines, Kayla, 2019). Die Inszenierung der perfekten Körper von Fitness Influencern ruft die Community zur Nachahmung auf (vgl. Pilgrim und Bohnet-Joschko, 2019). Für die Follower sind Personen aus dem Fitnessbereich Vorbilder. Die Influencer geben das Gefühl, durch einen gesunden und starken Körper, ein glückliches Leben führen zu können (vgl. ebd.; Döring, 2015). Fitness Influencer sind Menschen aus dem Alltag, die dem Follower durch ihre tägliche Kommunikation private Einblicke geben. Folglich identifiziert sich der Follower mit seinem Vorbild (vgl. Deges, 2018, S. 20). Die Beiträge von Influencern sind für die Follower vertrauenswürdiger als von Prominenten (vgl. Raggatt et al., 2018; Schouten et al., 2019).

„I have to make sure that the pages I follow are realistic. I don't follow any fashion models or people that will make me feel shit about myself. I like to follow people who are honest about how hard it can be to lose weight and to stay healthy. (Female, 25)“ (Raggatt et al., 2018, S. 8).

Influencer beeinflussen nicht nur das Kaufverhalten für



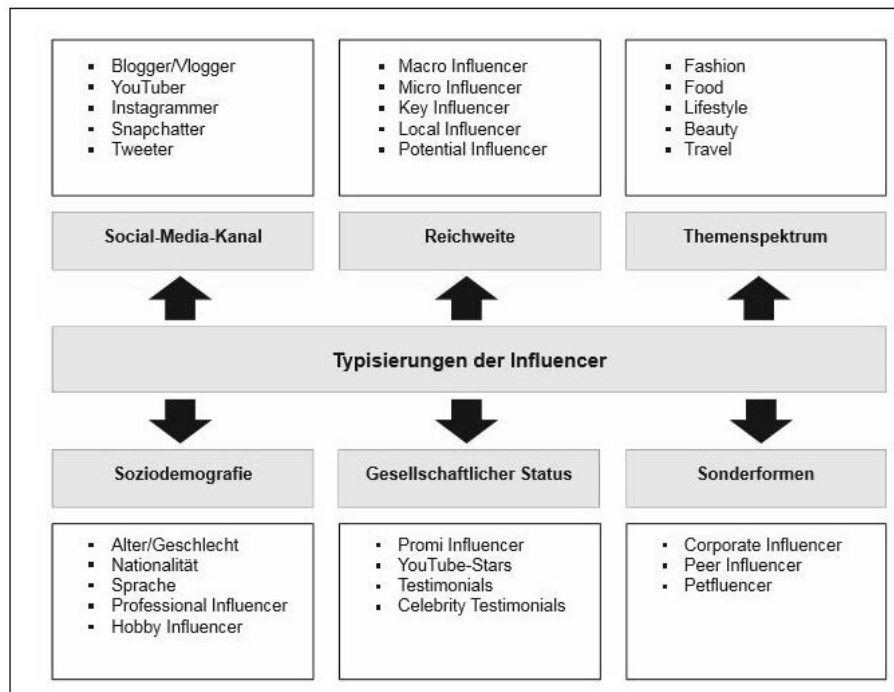


Abbildung 3: Influencer-Typisierungen nach Deges (2018, S. 22)

bestimmte Produkte, sondern auch das soziale Verhalten der Follower (vgl. Bundesverband für Digitale Wirtschaft (BWVD) e. V., 2019, S. 12; Pilgrim und Bohnet-Joschko, 2019). Daher stellt sich die Frage, inwiefern Fitness Influencer die Sportaktivität der Follower beeinflussen.

In der Literatur konnte mittels qualitativer Inhaltsanalysen und Interviews anhand verschiedener Fitnesstrends gezeigt werden, dass Fitpiration-Inhalte und Sportaktivitäten zusammenhängen (siehe Kapitel 4.2). Vor allem kann die Motivation zum Sporttreiben erhöht werden, durch die visuelle Darstellung von Fitspo-Körperbildern (ebd.). Ein Forschungsdefizit besteht darin, dass keine Studien die kausale Wirkung der Fitnessinhalte auf das Gesundheits-, oder Trainingsverhalten untersucht haben (vgl. Carrotte et al., 2015b, und Carrotte et al., 2017; Raggatt et al., 2018; Easton et al., 2018; Döring, 2015). Vor allem wie die Influencer aus dem Fitnessbereich durch ihre Glaubwürdigkeit das Sporttreiben der Follower beeinflussen können. Es konnte lediglich mittels einer qualitativen Inhaltsanalyse gezeigt werden, dass Influencer vertrauenswürdige Personen sind und Autorität in der Gesundheitskommunikation auf Instagram besitzen (vgl. Pilgrim und Bohnet-Joschko, 2019). Vor allem bei der Generation Y spielen der Fitnesssport und die Sozialen Medien eine bedeutsame Rolle in der Freizeit (vgl. Mills, 2019; Frees und Koch, 2018; Grabs et al., 2018, S. 287). Daher ist für diese Arbeit folgende Forschungsfrage entstanden: „Welche Bedeutung spielen Fitness Influencer auf Instagram bei der Generation Y und deren Sportaktivität?“

## 5. Methodik

Die Methodik stellt „[...] den Weg des wissenschaftlichen Vorgehens [...]“ dar, um die Forschungsfrage beantworten und die unterstützenden Hypothesen prüfen zu können (Döring, 2015, S. 365). In den folgenden Kapiteln wird die in dieser Arbeit verwendete Methode vorgestellt und erläutert.

### 5.1. Forschungsdesign

Das Forschungsdesign soll die methodische Vorgehensweise dieser Arbeit darlegen. Mittels einer quantitativen Befragung soll die Forschungsfrage beantwortet werden (vgl. Bortz und Döring, 2015, S. 405). „Bei der quantitativen Forschung steht das standardisierte, strukturierte „Messen“ festgelegter Inhalte im Vordergrund der Erhebung, um Hypothesen zu prüfen“ (Voss, 2015, S. 42). Dabei soll die deduktive Vorgehensweise genutzt werden, um aus einer Theorie Hypothesen abzuleiten und empirisch zu überprüfen (vgl. Hussy et al., 2013, S. 191). Es handelt sich dabei um eine Querschnittsstudie, da einmalig eine Befragung der Personen stattfindet (vgl. Bortz und Döring, 2015, S. 210). Mittels einer Online Umfrage soll die Forschungsfrage beantwortet werden. Anschließend werden die Verteilungen bestimmter Merkmale der Stichprobe dargestellt und die Zusammenhänge geprüft.

### 5.2. Hypothesen

Die vorliegenden Hypothesen sollen der Beantwortung der Forschungsfrage helfen. Es handelt sich dabei um Zusammenhangshypothesen, die die Zusammenhänge zwischen Merkmalen überprüfen (vgl. Bortz und Döring, 2015, S.

677). Dabei wird zwischen der unabhängigen (Ursache) und abhängigen Variable (Wirkung) unterschieden (vgl. Bortz und Schuster, 2010, S. 7). Die jeweiligen Hypothesenpaare werden von der Alternativ-, und Nullhypothese gebildet. Die Alternativhypothese ( $H_1$ ) postuliert den erwarteten Effekt und die Nullhypothese ( $H_0$ ) negiert in Richtung gar kein oder gegengerichteter Effekt (vgl. Bortz und Döring, 2015, S. 614).

Die erste Zusammenhangshypothese orientiert sich an der Studie von Braumüller (2018). Mittels eines Interviews wurden zehn junge Erwachsene von 19 bis 26 Jahren zu ihren sportlichen Aktivitäten und ihrem sportbezogenen Handeln in den Sozialen Medien befragt. Demnach fördern die Sozialen Netzwerke die Sportaktivität. In dieser Arbeit soll nun überprüft werden, ob ein Zusammenhang zwischen dem Folgen von Fitness Influencern (unabhängige Variable) und dem Sportmachen (abhängige Variable) besteht.

$H_1$ : Es gibt einen Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und dem Sportmachen

$H_0$ : Es gibt keinen Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und dem Sportmachen

Die Basis der zweiten Zusammenhangshypothese bilden die Studien von Vaterlaus et al. (2015) und Raggatt et al. (2018). Es ist davon auszugehen, dass sportbezogene Fitness-Inhalte auf Instagram das Bedürfnis nach sportlicher Aktivität verstärken können (ebd.). Dabei wird die Lust nach dem Sporttreiben als ein angenehmes Gefühl empfunden (vgl. Güllich und Krüger, 2013, S. 295). Dieses Gefühl könnte durch die veröffentlichten Bilder und Videos der Meinungsmacher vermittelt werden. Es kann folgende Hypothese mit Bezug auf die Opinion Leaders aufgestellt werden. Dabei bilden die Fitness Influencer die unabhängige Variable und Lust auf Sporttreiben die abhängige Variable.

$H_1$ : Es gibt einen Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und der Lust nach Sporttreiben durch die Bilder und Videos der Fitness Influencer

$H_0$ : Es gibt keinen Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und der Lust nach Sporttreiben durch die Bilder und Videos der Fitness Influencer

Bei der dritten Zusammenhangshypothese wird die Studie von Döring (2015) als Grundlage verwendet. Es nutzen Personen die Sportvideos auf YouTube für ihre Sportaktivitäten (vgl. Döring, 2015). Vor allem Meinungsmacher aus der Fitnessbranche veröffentlichen ihre Trainingsprogramme und -übungen auf Instagram (vgl. Pilgrim und Bohnet-Joschko, 2019). Diese können von der Community zum Sporttreiben genutzt werden. Daher soll überprüft werden, inwiefern Fitness Influencer (unabhängige Variable) und deren Fitnessprogramme (abhängige Variable) zusammenhängen.

$H_1$ : Es gibt einen Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und dem Ausprobieren eines Fitnessprogramms eines Fitness Influencers

$H_0$ : Es gibt keinen Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und dem Ausprobieren eines Fitnessprogramms eines Fitness Influencers

Bei der vierten Zusammenhangshypothese stellen die Studien von Pilgrim und Pilgrim und Bohnet-Joschko (2019) die Grundlage. Die Fitness Influencer sind Vorbilder und vermitteln ein bestimmtes Körperideal bei dem Follower (vgl. Pilgrim und Bohnet-Joschko, 2019). In der aufgestellten Hypothese stellen die Fitness Influencer die unabhängige Variable und das Aussehen wie ein Fitness Influencer die abhängige Variable.

$H_1$ : Es gibt einen Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und genauso aussehen zu wollen wie ein Fitness Influencer.

$H_0$ : Es gibt keinen Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und genauso aussehen zu wollen wie ein Fitness Influencer

### 5.3. Erhebungsinstrument

In Form eines teilstandardisierten Fragebogens mittels einer Online Umfrage können Teilnehmer die Fragen beantworten (vgl. Raab-Steiner und Benesch, 2012, S.47).

„Bei einer Online-Befragung („online survey“, „web survey“) wird ein elektronischer Fragebogen per Internet (vor allem Web) verbreitet und meist an stationären Computern bearbeitet. Die gewählten Antworten der Befragungspersonen werden auf einem Befragungsserver gespeichert und stehen als digitale Datensätze zur Verfügung.“ (Bortz und Döring, 2015, S. 414).

Die halboffenen Antwortformate sind nominal- oder ordinalskaliert, um eine hohe Rücklaufquote zu erzielen (ebd., S. 58). Bei den Fragen ist eine Einfach- oder Mehrfachauswahl möglich. Das offene Antwortformat wäre ebenfalls eine passende Strukturierung für den Fragebogen, jedoch sinkt die Bereitschaft der Teilnehmer, wenn sie etwas selbst ausfüllen müssen (vgl. Baur und Blasius, 2014, S. 715). Die Online Umfrage erspart Kosten und Zeit im Vergleich zu einem Interview (ebd., S. 662). Es wird kein Interviewer benötigt und die Teilnehmer können jederzeit den Fragebogen ausfüllen (ebd.).

### 5.3.1. Erstellung des Online Fragebogens

Der Fragebogen orientiert sich sowohl an den drei Testgütekriterien Objektivität, Reliabilität und Validität und an dem standardisierten Aufbau eines Fragebogens (vgl. Bortz und Döring, 2015, S. 195 und Döring und Bortz, 2016, S. 406). Bei der Konstruktion des Online Fragebogens wurde auf eine zielgerichtete Ansprache der jungen Erwachsenen und eine klare Formulierung und Strukturierung der Items geachtet (vgl. Raab-Steiner und Benesch, 2012, S. 49 ff.). Der Fragebogen setzt sich aus 22 Fragen zusammen und die Beantwortungszeit beträgt ungefähr fünf Minuten. Es gibt vier wesentliche Interessenbereiche: 1) Instagram 2) allgemeines zur Sportaktivität 3) Körperbild und 4) Fitness Influencer. Zunächst erhielten die Teilnehmer eine Vorstellung der Verfasserin, des Themas und über den zeitlichen Umfang der Online Umfrage (siehe Anhang A, 1). Die Verfasserin versicherte den Teilnehmern, mit den Daten vertraulich umzugehen und sie zu anonymisieren. Im Anschluss erschienen drei Einstiegsfragen, die gleichzeitig als Filterfragen genutzt werden konnten (siehe Anhang A, 2-4). Die Teilnehmer erhielten den Hinweis, dass sie sich im Alter zwischen 18 und 38 Jahren befinden sollten. Eine Filterfrage bezog sich darauf, ob die Teilnehmer einen Instagram Account besitzen. Daran anknüpfend folgten drei Fragen zum ersten Interessenbereich, die Nutzungshäufigkeit, die Gründe und die beliebten Themengebiete bei Instagram (siehe Anhang A, 5; Abb. 6,7). Diese Fragen basieren auf der VuMa, Arbeitsgemeinschaft für Verbrauchs- und Medienanalyse (2018c), Frees und Koch (2018) und Statista (2019b). Anschließend wurden zwei Fragen zur möglichen Einschränkungzeit von Instagram und deren Gründen gestellt (vgl. Royal Society for Public Health, 2017; Vaterlaus et al., 2015; Zingg, 2018; Vogl, 2019; siehe Anhang A, 6,7). Der Interessenbereich Zwei enthielt Fragen zum allgemeinen Sporttreiben der Teilnehmer, um eine Einschätzung vornehmen zu können, wie oft und warum die Generation Y Sport betreibt (siehe Anhang, 8,9; Abb. 8,9). Die Fragen setzten sich aus der Studie von Preuß et al. (2012) sowie aus der Mills (2019) Befragung zusammen. Um das aktuelle Schönheitsideal der Teilnehmer zu ermitteln, werden die Körperbewertungs-Skalen von Stunkard et al. (1983) und die modifizierte Skala von Lynch und Zellner (1999) herangezogen (siehe Abb. 10,11). Diese beinhalten dünne und muskulöse Körperbilder, für beide Geschlechter auf verschiedene Arten nach Zahlen geordnet (vgl. Carrotte et al., 2017). Der darauffolgende Interessenbereich Vier stellte den Kern des Fragebogens dar. Zunächst wurde auf Basis vorausgegangener Literatur ein Influencer definiert, um ein einheitliches Verständnis von dem Begriff sicher zu stellen (siehe Anhang A, 12). Der Interessenbereich Vier setzte sich aus acht Fragen zum Thema Fitness Influencer zusammen (siehe Anhang A 13-15; C 1-4; Abb. 12). Die Zusammenstellung der Fragen konnte an folgenden Studien abgeleitet werden: Vaterlaus et al. (2015), Raggatt et al. (2018), Carrotte et al. (2015b) und Carrotte et al. (2017), Easton et al. (2018) und Pilgrim und Bohnet-Joschko (2019).

### 5.3.2. Auswahl der Befragten

Die Stichprobe bezieht sich auf die deutsche Generation Y, im Alter von 18 bis 38 Jahren, da diese eine wichtige Zielgruppe im digitalen Zeitalter darstellt und die meiste Zeit mit Sozialen Medien verbringt (vgl. Jahnke et al., 2018, S.17; Vu-Ma, Arbeitsgemeinschaft für Verbrauchs- und Medienanalyse, 2018c; Frees und Koch, 2018; Raggatt et al., 2018). Die Befragten sollen einen Instagram Account besitzen, da nur Fitness Influencer von Instagram Teil dieser Arbeit sind. Demnach wird nur eine Auswahl von Fällen aus der Zielpopulation (Teilerhebung) untersucht und keine Vollerhebung aller Personen der Generation Y durchgeführt (vgl. Bortz und Döring, 2015, S. 292f.). Eine Mehrfachteilnahme war aufgrund der vorgenommenen Einstellungen beim Onlineportal ausgeschlossen.

### 5.3.3. Durchführung

Der Fragebogen konnte mittels einer Umfragesoftware als Online Umfrage veröffentlicht werden. Die Beantwortung der Fragen war anonym auf mobilen Smartphones, Laptops oder Tablets möglich, die von der Generation Y innerhalb von fünf Minuten durchgeführt werden konnte. Vorab wurde ein Pre-Test durchgeführt, der zur Überprüfung der Funktionalität des Online Portals diente. Die Online Umfrage ist am 4. September 2019 gestartet und wurde am 18. September 2019 beendet. Mithilfe eines Hyperlinks konnte die Umfrage über Instagram, Facebook und WhatsApp geteilt werden. Zusätzlich veröffentlichte eine Fitness Influencerin, deren Name anonym bleiben soll, den Hyperlink in ihrer Instagram Story. Außerdem wurde die Online Umfrage über die Verteilungsemail der Sportuniversität Mainz verschickt.

## 6. Ergebnisse

Die Auswertung der Daten erfolgte durch das Statistikprogramm SPSS Statistic 23 V5. Für die deskriptive Statistik konnten Häufigkeitsverteilungen in Tabellenform von SPSS dargestellt werden (siehe Anhang A). Um den Zusammenhang verschiedener Variablen zu prüfen, konnten Kreuztabellen erstellt und der Chi-Quadrat-Test angewendet werden (siehe Anhang B, C). Diese Methoden sind Teil der Inferenzstatistik und dienen der Überprüfung von Hypothesen (vgl. Janssen und Laatz, 2013, 40-44; Raab-Steiner und Benesch, 2012, S. 17). Je nach Skalenniveau werden verschiedene Messwerte verwendet, die die Stärke des Zusammenhangs angeben (vgl. Janssen und Laatz, 2013, S. 266; Bortz und Döring, 2015, S. 681). Bei zwei nominalskalierten Variablen in einer 2x2 Tabelle wird der Phi-Koeffizient ( $\phi$ ) verwendet (ebd.). Die Stärke des Zusammenhangs kann nach Cohen (1988, S. 227) ab einem Phi-Koeffizienten von 0,10 als kleiner Effekt, ab 0,30 als mittlerer Effekt und ab 0,50 als großer Effekt interpretiert werden (vgl. Bortz und Döring, 2015, S. 821). Das Signifikanzniveau liegt bei gleich oder kleiner als 5% ( $p=0,05$ ). Demnach wird entweder die Nullhypothese ( $H_0$ ) bei  $p$  größer als 0,05 oder die Alternativhypothese ( $H_1$ ) bei  $p$  kleiner oder gleich 0,05 beibehalten (vgl.

Bortz und Döring, 2015, S. 614). Die asymptotische Signifikanz gibt den signifikanten Wert  $p$  an. Um das Zusammenhangsmaß eines signifikanten Ergebnisses der Alternativhypothese aufzeigen zu können, wird der Phi-Koeffizient herangezogen (Anhang B, C). Zur Veranschaulichung der Forschungsergebnisse wurden die Grafiken mit Microsoft Excel erstellt.

### 6.1. Stichprobenkonstruktion

Es haben an der Online Umfrage insgesamt 304 Personen teilgenommen. Davon konnten 32 Personen nicht mit in die Auswertung einfließen, da sie nicht Teil der angegebenen Altersklasse 18 bis 38 Jahre waren oder den Fragebogen nicht vollständig (über 50 Prozent) ausgefüllt haben (vgl. Bortz und Döring, 2015, S. 588). Von 272 Teilnehmern haben 240 (88,24 Prozent) Personen einen Instagram Account und 32 Personen (11,76 Prozent) keinen Instagram Account (siehe Anhang A, 4). Da sich diese Arbeit hauptsächlich mit Fitness Influencern auf Instagram beschäftigt, werden nur Personen mit einem Instagram Account für die Auswertung der Befragung mit einbezogen. Demnach setzt sich die Stichprobe insgesamt aus 240 Teilnehmern ( $N=240$ ) zusammen. Der prozentuale Geschlechteranteil ist in Abbildung 4 in einem Kreisdiagramm dargestellt. Der Anteil der weiblichen Teilnehmer liegt bei 62,08 Prozent ( $n=149$ ) und der der männlichen Teilnehmer bei 37,92 Prozent ( $n=91$ ) (siehe Abb. 4).

Das durchschnittliche Alter lag bei 23,85 Jahren (Standardabweichung = 3,35) (siehe Anhang A, 3). In der Abbildung 5 wird die Generation Y in zwei Altersklassen unterteilt. Es bilden die 18- bis 26-jährigen Teilnehmer 85,83 Prozent ( $n=206$ ) und die 27- bis 38-jährigen Teilnehmer 14,17 Prozent ( $n=34$ ) der Umfrage ab (siehe Abb. 5).

### 6.2. Interessenbereiche

Der erste Interessenbereich beschäftigt sich mit dem Thema Instagram. Das Kreisdiagramm zeigt die Nutzung der 18- bis 38-jährigen Teilnehmer mehrmals am Tag (88,33 Prozent;  $n=212$ ), in der Woche (7,92 Prozent;  $n=19$ ), im Monat (2,92 Prozent;  $n=7$ ) und keine Nutzung (0,83 Prozent;  $n=2$ ) von Instagram (siehe Abb. 6). Darauf aufbauend wurden die Teilnehmer der Umfrage nach den Gründen ( $N=240$ ) und beliebten Themengebieten ( $N=239$ ) befragt. Am häufigsten dient Instagram zur Freizeitbeschäftigung (69,17 Prozent;  $n=166$ ), Inspiration (64,58 Prozent;  $n=155$ ) und „in Kontakt mit den Freunden stehen“ (46,25 Prozent;  $n=111$ ) (siehe Anhang A, 5).

Neben der Kontaktpflege bietet Instagram eine Informationsquelle für neues Wissen (35,83 Prozent;  $n=86$ ) oder wird zur Motivation (33,33 Prozent;  $n=80$ ) genutzt (ebd.). Die Teilnehmer nutzen Instagram, um selbst ihre Erlebnisse zu teilen (28,75 Prozent;  $n=69$ ) und neue kreative Inhalte zu kreieren (14,58 Prozent;  $n=35$ ) oder folgen ihren Lieblingsstars (20,00 Prozent;  $n=48$ ) (ebd.). Weitere Gründe wurden im Fragebogen unter Sonstiges zusammengefasst (4,17 Prozent;  $n=10$ ) und 1,67 Prozent ( $n=4$ ) gaben keine Gründe zur Instagram Nutzung an (ebd.). In der Abbildung

7 sind die beliebten Themengebiete prozentual in einem Balkendiagramm dargestellt. Ein beliebtes Thema der Befragten auf Instagram ist Sport/Fitness mit 85,36 Prozent ( $n=204$ ) (siehe Abb. 7). Danach folgen die Themengebiete Reisen (61,09 Prozent;  $n=146$ ), Ernährung (56,90 Prozent;  $n=136$ ), Mode (43,10 Prozent;  $n=103$ ), Gesundheit (34,73 Prozent;  $n=83$ ), Wohnen (32,22 Prozent;  $n=77$ ), Politik (22,59 Prozent;  $n=54$ ), Beauty (22,18 Prozent;  $n=53$ ), Kunst (12,97 Prozent;  $n=31$ ), Technik (7,53 Prozent;  $n=18$ ) und Sonstiges (11,30 Prozent;  $n=27$ ) (ebd.). 2,09 Prozent ( $n=5$ ) machten keine Angabe zu der Frage (ebd.).

Bei der Frage, ob die Teilnehmer ihre Instagram Aktivität einschränken möchten, haben 55,83 Prozent ( $n=134$ ) mit Ja und 44,17 Prozent ( $n=106$ ) mit Nein geantwortet (siehe Anhang A, 6). Ein Grund für die Einschränkung von Instagram ist, dass die Teilnehmer sich weniger von Instagram ablenken lassen wollen (87,67 Prozent;  $n=128$ ) (siehe Anhang A, 7). Es folgen weitere mehrmals genannte Gründe für die Instagram Einschränkung wie „abends früher ins Bett gehen“ (32,19 Prozent;  $n=47$ ), „keinen unrealistischen und unerreichbaren Lebensstilen nacheifern“ (28,08 Prozent;  $n=41$ ), „mehr Zeit für Freunde und Familie haben“ (26,71 Prozent;  $n=39$ ), „keine inszenierten Urlaubsbilder sehen“ (13,01 Prozent;  $n=19$ ) und „nicht mit anderen Körperbildern vergleichen wollen“ (17,81 Prozent;  $n=26$ ) (ebd.). Weitere Gründe für die Einschränkung wurden unter Sonstiges eingeordnet (6,16 Prozent;  $n=9$ ) (ebd.).

Beim Interessenbereich Zwei wurden die Teilnehmer zu ihrer Sportaktivität befragt. Demnach betreiben 95,00 Prozent ( $n=228$ ) der Befragten Sport und 5,00 Prozent ( $n=12$ ) der Befragten keinen Sport (siehe Anhang A, 8). Die Abbildung 8 zeigt Häufigkeit des Sporttreibens in Prozent. Es treiben 85,42 Prozent ( $n=205$ ) mehrmals in der Woche Sport und 9,17 Prozent ( $n=22$ ) mehrmals im Monat Sport und 5,42 Prozent ( $n=13$ ) machen nie Sport (siehe Abb. 8.) Mit 64,29 Prozent ( $n=153$ ) ordnen die Teilnehmer ihren Sport im Bereich Fitnesssport ein (siehe Anhang A, 9). Danach folgt der Freizeit- und Breitensport (57,98 Prozent;  $n=138$ ), Leistungssport (21,43 Prozent;  $n=51$ ), Erlebnissport (15,13 Prozent;  $n=36$ ), Gesundheitssport (14,29 Prozent;  $n=34$ ), Integrationssport (2,52 Prozent;  $n=6$ ) und Sonstiges (2,52 Prozent;  $n=6$ ) (ebd.). Es gaben 1,26 Prozent ( $n=3$ ) keinen Sportbereich an (ebd.).

Die Abbildung 9 zeigt ein Balkendiagramm mit den Gründen für das Sporttreiben in Prozent. Die Teilnehmer treiben Sport, um ihr Wohlbefinden und ihre Gesundheit zu verbessern (83,75 Prozent;  $n=201$ ) oder als Ausgleich zum Alltag (81,67 Prozent;  $n=196$ ) (siehe Abb. 9). Es folgen Gründe wie „Erhöhung der sportlichen Leistungsfähigkeit“ (68,33 Prozent;  $n=164$ ), „Fit auszusehen“ (61,25 Prozent;  $n=147$ ), „Muskelaufbau“ (52,08 Prozent;  $n=125$ ), „Gewichtsreduktion“ (32,92 Prozent;  $n=79$ ) oder „Menschen kennenlernen“ (14,17 Prozent;  $n=34$ )“ (ebd.). Weitere genannte Gründe konnten in der Kategorie Sonstiges (8,33 Prozent;  $n=20$ ) genannt werden und 1,25 Prozent ( $n=3$ ) machten keine Angabe zu der Frage (ebd.).

Beim Interessenbereich Drei wurde anhand der Figu-

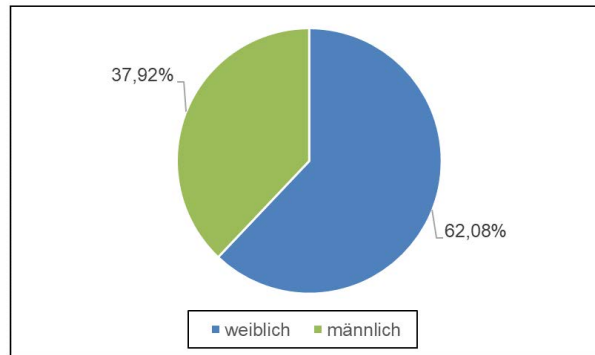


Abbildung 4: Verteilung der Geschlechter in Prozent (N=240)

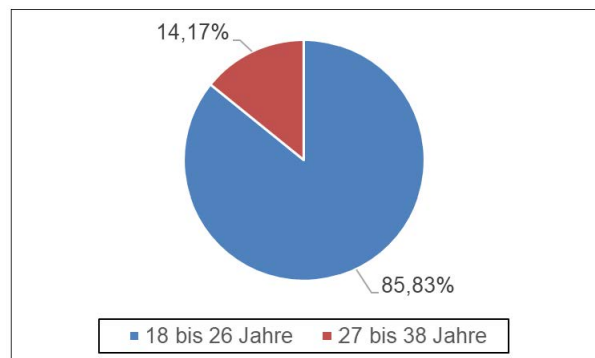


Abbildung 5: Verteilung der Altersklassen in Prozent (N=240)

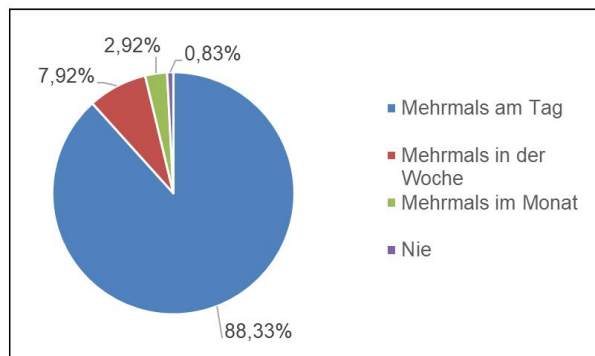


Abbildung 6: Nutzung von Instagram in Prozent (N=240)

ren/ Körper-Bewertungs-Skala von Stunkard et al. (1983) und Lynch und Zellner (1999), das aktuelle Schönheitsideal jeweils von Frauen und Männern befragt. Abbildung 10 zeigt die Figuren/Körper-Bewertungs-Skala einer Frau nach Zahlen geordnet.

Der Mittelwert liegt bei 7,77 mit einer Standardabweichung von 5,25. Die weiblichen Teilnehmerinnen wollen signifikant genauso aussehen wie ein Fitness Influencer (siehe Anhang B, 2.). Da der p-Wert (0,001) kleiner als 0,05 ist (ebd.). Diese Ergebnisse beziehen sich ausschließlich auf die Antworten von weiblichen Teilnehmerinnen (N=149) der Befragung. Die Ergebnisse der Häufigkeitstabellen von der Frauenfigur ergaben, dass mit rund 32 Prozent (n=76) das

Bild drei von Frauen und Männern gewählt wurde (siehe Anhang A, 10). Die Abbildung 11 zeigt die Figuren/Körper-Bewertungs-Skala eines Mannes nach Zahlen geordnet. Der Mittelwert liegt bei 13,68 mit einer Standardabweichung von 2,97. Auch Männer wollen signifikant genauso aussehen wie ein Fitness Influencer (siehe Anhang B, 4). Denn p (0,046) ist kleiner oder gleich 0,05 (ebd.). Diese Ergebnisse beziehen sich ebenfalls nur auf die Beantwortung des männlichen Geschlechts (N=91). Die Häufigkeitstabellen von dem Männerbild zeigen, dass Bild 13 von beiden Geschlechtern (26,25 Prozent; n=63) am häufigsten genannt wurde (siehe Anhang A, 11).

Der Interessensbereich Vier beschäftigt sich mit dem The-

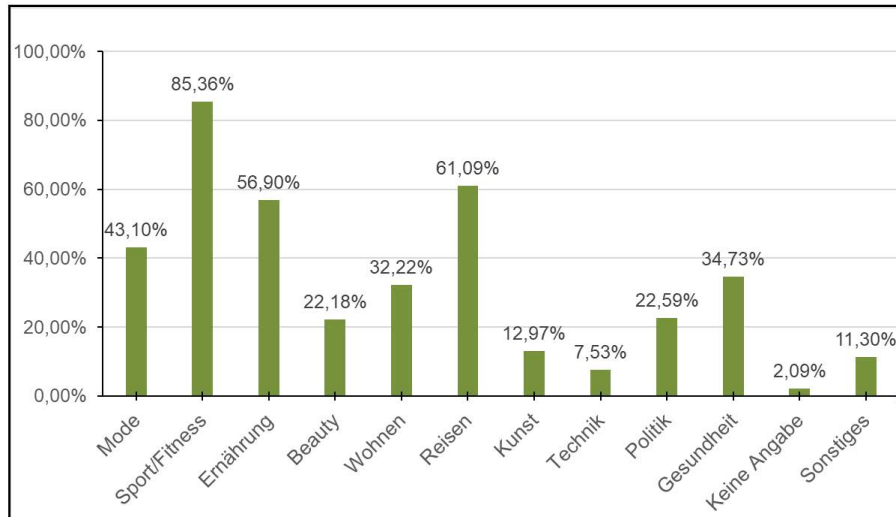


Abbildung 7: Beliebte Themengebiete auf Instagram in Prozent (N=239)

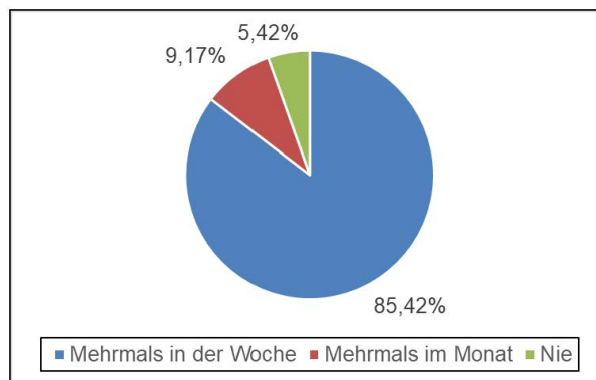


Abbildung 8: Häufigkeit des Sporttreibens in Prozent (N=240)

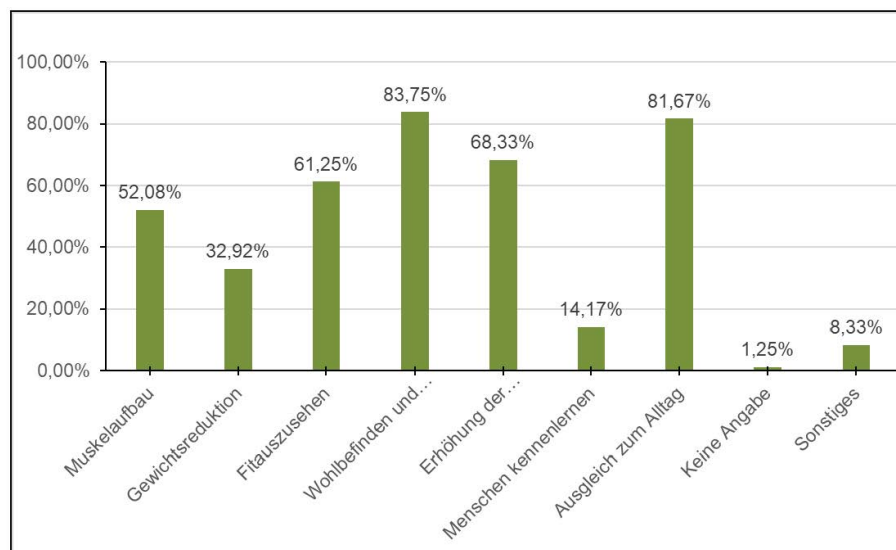
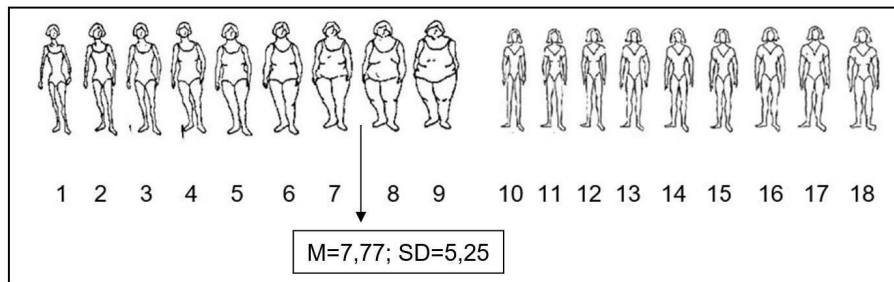
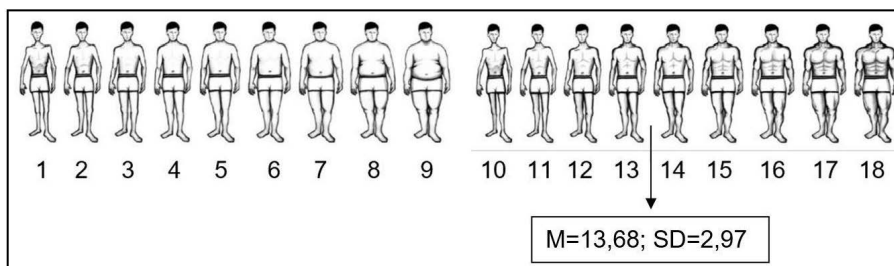


Abbildung 9: Gründe für das Sporttreiben in Prozent (N=240)



**Abbildung 10:** Body rating scale nach Stunkard et al. (1983) und figure rating scale nach Lynch und Zellner (1999) für das weibliche Geschlecht (Quelle: Grossbard et al., 2011) (N=149)



**Abbildung 11:** Body rating scale nach Stunkard et al. (1983) und figure rating scale nach Lynch und Zellner (1999) für das männliche Geschlecht (Quelle: Ralph-Nearman und Filik, 2018) (N=91)

ma Fitness Influencer auf Instagram. Es folgen 68,75 Prozent (n=165) einem Fitness Influencer und 31,25 Prozent (n=75) keinem Fitness Influencer auf Instagram (siehe Anhang A, 13). In einem Balkendiagramm sind die Themengebiete in Prozent aufgelistet, mit denen die Meinungsmacher der Fitnessbranche die Aufmerksamkeit des Followers weckt (siehe Abb. 12). Die Themen sind Ernährung (58,73 Prozent; n=111), Alltagsleben (51,85 Prozent; n=98), Fitnessübungen zum Muskelaufbau (51,32 Prozent; n=97), Regelmäßige Trainingsprogramme (46,56 Prozent; n=88), Tipps zur Gewichtsreduktion (27,51 Prozent; n=52) und Sonstiges (8,47 Prozent; n=16) (ebd.). Es machten 4,23 Prozent (n=8) keine Angabe zu der Frage (ebd.). Bei der Frage „Beschäftigst du dich mehr mit dem Thema Gesundheit und Fitnesport, seit du einem Fitness Influencer folgst“ gaben 28,75 Prozent (n=69) Ja und 71,25 Prozent (n=171) nein an (siehe Anhang A, 14). Bei der Frage „Betreibst du mehr Sport oder hast du angefangen mit Sporttreiben, seit du einem Fitness Influencer folgst“ beantworteten 16,25 Prozent (n=39) mit ja und 83,75 Prozent (n=201) mit nein (siehe Anhang A, 15).

### 6.3. Überprüfung der Hypothesen

Die wesentlichen Ergebnisse aller aufgestellten Hypothesen sind in der Tabelle 1 dargestellt. Es werden die Antworten von 240 Teilnehmern hinzugezogen (N=240). Die erste Hypothese überprüft, ob ein oder kein Zusammenhang zwischen dem Folgen von Fitness Influencern und dem Sportmachen besteht. Da eine Zelle eine erwartete Häufigkeit kleiner fünf hat und eine 2x2 Tabelle vorliegt, wird der Wert des exakten Tests nach Fisher (zweiseitig) verwendet (vgl. Janssen

und Laatz, 2013, S. 43,781). Es handelt sich um eine ungerichtete Hypothese, somit wird die Exakte Signifikanz nach Fisher zweiseitig (0,200) verwendet. Demnach ist p größer als 0,05 und die Nullhypothese wird beibehalten. Es besteht kein Zusammenhang zwischen „einem Fitness Influencer auf Instagram folgen“ und dem Sportmachen.

Bei der zweiten Hypothese soll geprüft werden, ob ein oder kein Zusammenhang zwischen „Fitness Influencern folgen“ und „Lust auf Sporttreiben“ besteht, wenn der Follower Bilder und Videos von einem Fitness Influencer sieht. Es haben null Zellen eine erwartete Häufigkeit kleiner fünf, sodass der Chi-Quadrat-Wert nach Pearson (18,791) hinzugezogen wurde (vgl. Bühl, 2014, S. 306f.). Es besteht ein signifikantes Ergebnis (p ist kleiner als 0,05) und ein kleiner Effekt (Phi-Koeffizient=0,280). Demnach kann die Alternativhypothese angenommen werden. Es besteht ein Zusammenhang zwischen Fitness Influencern und der Lust auf Sporttreiben.

Anhand der dritten Hypothese soll geprüft werden, ob ein oder kein Zusammenhang zwischen „einem Fitness Influencer folgen“ und dem Ausprobieren eines Fitnessprogramms besteht. Da null Zellen eine erwartete Häufigkeit kleiner fünf haben, kann der Chi-Quadrat-Wert nach Pearson (16,892) interpretiert werden. Dabei ist p (0,000) kleiner als 0,05, somit kann die Alternativhypothese beibehalten werden. Es besteht ein kleiner Effekt (Phi-Koeffizient=0,265) zwischen Fitness Influencern und dem Ausprobieren eines Fitnessprogramms.

Bei der Hypothese Vier soll geprüft werden, ob ein oder kein Zusammenhang zwischen „Fitness Influencern folgen“ und „genauso aussehen zu wollen wie ein Fitness Influencer“ besteht. Der Chi-Quadrat-Wert kann nach Pearson (14,683) interpretiert werden, da null Zellen eine erwartete

**Tabelle 1:** Nennung der wichtigsten Ergebnisse aller Hypothesen (eigene Darstellung)

	Anzahl (N)	Kontinuitätskorrektur	Chi-Quadrat-Wert nach Pearson	Asymptotische Signifikanz (zweiseitig)	Exakte Signifikanz nach Fisher	Phi-Koeffizient ( $\phi$ )
Erste Hypothese	240	1,250 <sup>a</sup>	2,067 <sup>c</sup>	0,151	0,200 <sup>d</sup>	0,093
Zweite Hypothese	240	17,437 <sup>a</sup>	18,791 <sup>b</sup>	0,000 <sup>e</sup>	0,000	0,280 <sup>f</sup>
Dritte Hypothese	240	15,723 <sup>a</sup>	16,892 <sup>b</sup>	0,000 <sup>e</sup>	0,000	0,265 <sup>f</sup>
Vierte Hypothese	240	13,635 <sup>a</sup>	14,683 <sup>b</sup>	0,000 <sup>e</sup>	0,000	0,247 <sup>f</sup>

a Wird nur für eine 2x2-Tabelle berechnet

b 0 Zellen (,0%) haben eine erwartete Häufigkeit kleiner 5.

c Zellen (25,0%) haben eine erwartete Häufigkeit kleiner 5

d nicht signifikant bei  $p > 0,05$

e signifikant bei  $p < 0,05$

f kleiner Effekt bei  $\omega = 0,10$



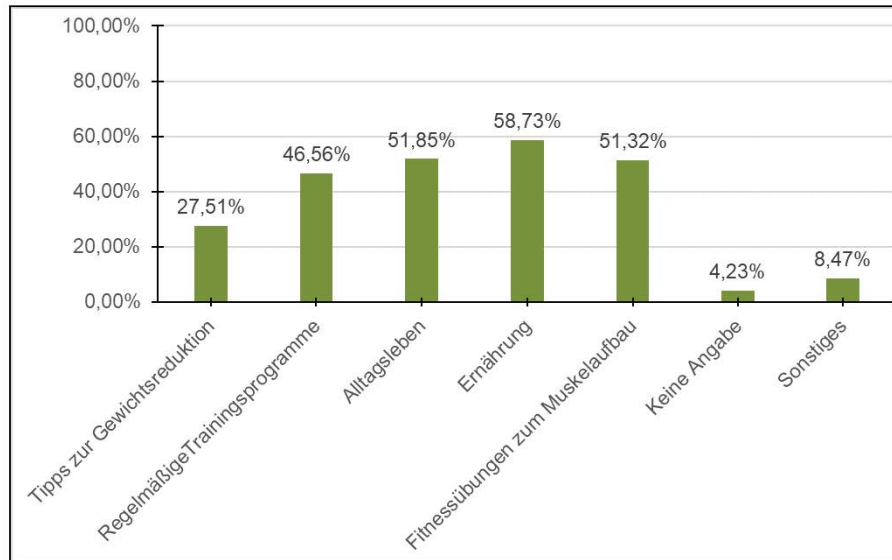


Abbildung 12: Beliebte Themen bei Fitness Influencern (N=189)

te Häufigkeit von kleiner fünf haben. Der p-Wert (0,000) ist kleiner als 0,05, sodass die Alternativhypothese beibehalten werden kann. Es besteht ein kleiner Effekt (Phi-Koeffizient=0,247).

## 7. Diskussion

Im Folgenden werden die Ergebnisse der Online Umfrage analysiert und diskutiert. Es ist bekannt, dass Instagram ein beliebtes Medium bei der Generation Y ist (vgl. Grabs et al., 2018, S. 287). Die Einteilung in zwei Altersklassen zeigt, dass mehr 18- bis 26-jährige (85,83 Prozent; n=206) als 27- bis 38-jährige (14,17 Prozent; n=34) an der Befragung teilgenommen haben (siehe Abb. 5). Die Ergebnisse stützen die Erkenntnis, dass ein Großteil (88,24 Prozent; n=240) der Befragten einen Instagram Account hat (Anhang A,4). Die Plattform wird mit 88,33 Prozent (n=212) mehrmals am Tag von der untersuchten Stichprobe genutzt (siehe Abb. 6). Instagram dient sowohl zur Freizeitbeschäftigung und dem Austausch mit Freunden als auch zur Inspiration und Wissensbeschaffung (siehe Anhang A,5). Die Ergebnisse dieser Arbeit können die Erkenntnis vorausgegangener Literatur bestätigen, dass der Social-Media-Kanal Instagram ein beliebtes Medium zur Freizeitbeschäftigung ist und als Inspirationsquelle genutzt werden kann (siehe Kapitel 4, 4.1). Laut der Stichprobe ist das Thema Sport/Fitness (85,36 Prozent; n=204) gern gesehen (siehe Abb. 7). Die Ergebnisse deuten darauf hin, dass das Thema Sport in den Sozialen Netzwerken populär ist (vgl. Grimmer, 2019, S. 181f.). Gleichzeitig ist zu bedenken, dass über 50 Prozent (n=134) der Teilnehmer die tägliche Nutzung von Instagram einschränken wollen (siehe Anhang A, 6). Dies stimmt mit dem aktuellen Trend, der bewussten Einschränkung dieser App, überein (vgl. Vogl, 2019). Die am häufigsten genannten Gründe dafür bei der Umfrage waren, dass sich die Teilnehmer

weniger ablenken lassen, Abends früher ins Bett gehen und sich nicht mit unrealistischen und unerreichbaren Lebensstilen vergleichen wollen (siehe Anhang A, 7). Die Ergebnisse dieser Arbeit zeigen die negativen Seiten von Instagram. Dies könnte einer der Gründe sein, dass die sportlichen Aktivitäten durch die Nutzung der Sozialen Medien sinken könnten (vgl. Vaterlaus et al., 2015). Dennoch betreiben 85,42 Prozent (n=205) der untersuchten Stichprobe mehrmals in der Woche Sport (siehe Abb. 8). Vor allem ordnen die Teilnehmer ihren Sport im Bereich Fitnesport ein (siehe Anhang A, 9). Der Sport wird genutzt, um das Wohlbefinden und die Gesundheit zu verbessern, einen Ausgleich zum Alltag zu haben oder zur Erhöhung der sportlichen Leistungsfähigkeit (siehe Abb. 9). In der Studie von Mills (2019) konnten diese Gründe ebenfalls dem Sporttreiben der Generation Y zugeordnet werden. Die Ergebnisse der Umfrage spiegeln sowohl die Begeisterung am Fitnesport als auch das Verständnis von Sport im weiteren Sinne wider. Demnach steht der Sport sowohl für Leistung als auch für eine freudvolle Betätigung in der Freizeit, um das physische und psychische Wohl zu verbessern (siehe Kapitel 3, 3.1). Außerdem folgen 68,75 Prozent (n=165) einem Fitness Influencer auf Instagram (siehe Anhang A, 13). Der Follower ist am Alltagsleben und der Ernährung der Influencer interessiert (siehe Abb. 12). Weiterhin erregen die Influencer die Aufmerksamkeit der User mit Fitnessübungen zum Muskelaufbau, regelmäßigen Trainingsprogrammen oder Tipps zur Gewichtsreduktion (ebd.). Allerdings zeigt Hypothese Eins, dass es keinen Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und dem Sportmachen gibt. (siehe Kapitel 6.3). Es beschäftigen sich lediglich 28,75 Prozent (n=69) der Stichprobe mehr mit dem Thema Gesundheit und Fitnesport und 16,25 Prozent (n=39) betreiben mehr Sport oder haben mit Sporttreiben angefangen, seit sie einem Fitness Influencer folgen (siehe Anhang A, 14, 15). Dies könnte daran liegen, dass die teil-

nehmenden Personen der Stichprobe meist aus sportlichen Einrichtungen (Universität Sportwissenschaft, Fitnessstudio) den Fragebogen erhalten haben. Daher könnte ein gewisses Grundverständnis von Sport vorhanden sein und die Teilnehmer betreiben seit der Kindheit Sport.

In der vorgestellten Literatur aus Kapitel 4.3.2 konnte festgehalten werden, dass Fitness Influencer einen wesentlichen Beitrag zum Sporttreiben leisten. Dies konnte durch die aufgestellte Hypothese Zwei bestätigt werden. Demnach gibt es einen Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und der Lust auf Sport (siehe Kapitel 6.3). Zu diesen Ergebnissen kommen ebenfalls die Studien von Vaterlaus et al. (2015) und Raggatt et al. (2018), dass Fitness-Inhalte auf Instagram positiv zum Sporttreiben animieren können. Es liegt die Vermutung nahe, dass ebenfalls die Bilder und Videos der Meinungsmacher die Motivation nach Sport verstärken können. Des Weiteren bestätigt Hypothese Drei, dass Fitness Influencer bei der untersuchten Generation Y die sportliche Aktivität beeinflussen können (siehe Kapitel 6.3). Demnach gibt es einen Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und dem Ausprobieren eines Fitnessprogramms eines Influencers (ebd.). Die veröffentlichten Fitnessprogramme tragen zur sportlichen Aktivität bei. Es konnte lediglich gezeigt werden, dass die Generation Y bereits ein Fitnessprogramm eines Fitness Influencers ausprobiert hat (ebd.). Es können keine Aussagen getroffen werden, inwiefern die Fitnessprogramme von dem Follower nachhaltig zum Sporttreiben genutzt werden. Dennoch wecken die Fitness Influencer bei den Befragten die Aufmerksamkeit mit regelmäßigen Trainingsprogrammen (siehe Abb. 12). Allerdings sind die Informationen über das sportliche Training der Fitness Influencer in den meisten Fällen nicht wissenschaftlich fundiert (vgl. Döring, 2015). Dadurch werden lediglich Behauptungen oder eigene Erfahrungen der Meinungsmacher an den Follower weitergegeben (ebd.). Es ist bekannt, dass die Community ihrem Vorbild vertraut, sodass es zu Fehlinformation kommen kann (ebd.). Damit geht einher, dass die Bilder der Influencer oft inszeniert und mit Bildbearbeitungseffekten veröffentlicht werden (vgl. Faßmann und Moss, 2016, S. 27). Anhand der vorliegenden Befragung können keine Aussagen gemacht werden, inwiefern die Inhalte der Influencer transparent und zuverlässig sind. Dennoch geben die Ergebnisse der Befragung den Hinweis, dass Instagram zur Aneignung neuen Wissens bei der Stichprobe genutzt wird (Anhang A, 5).

Durch die Kommunikation der Fitness Influencer auf Instagram verstärkt sich der Wunsch bei den Followern nach einem trainierten Körperbild (vgl. Carrotte et al., 2017). Dies kann durch Hypothese Vier bestätigt werden (siehe Kapitel 6.3). Demnach gibt es einen Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und dem Bedürfnis, genauso auszusehen wie ein Fitness Influencer. Einerseits können die gezeigten Körper von Fitness Influencern den Follower zum Nacheifern anregen (vgl. Schach und Lommatzsch, 2018, S. 14; Pilgrim und Bohnet-Joschko, 2019). Denn durch die alltägliche Präsentation des Trainings geben

sie ihm das Gefühl, ebenfalls einen athletischen Körper formen zu können (vgl. Raggatt et al., 2018). Dieses Körperbild kann durch eine gesunde Ernährung und tägliche Bewegung erreicht werden und vermittelt der Community ein glückliches Leben (vgl. ebd.; Pilgrim und Bohnet-Joschko, 2019). Die Aufmerksamkeit beim Sporttreiben liegt somit beim eigenen Körper (vgl. Dimitriou und Ring-Dimitriou, 2019, S. 65). Andererseits steigt der soziale Druck in der Gesellschaft sowie der ständige Zwang nach sportlicher Betätigung, um das ideale Körperbild zu formen oder beizubehalten (vgl. Braumüller, 2018). Die Körperform-Skala nach Stunkard et al. (1983) und Lynch und Zellner (1999) gibt Ausschnitte des aktuellen Schönheitsideals für Frauen und Männer (siehe Abb. 10, 11). Bei Frauen liegt der Mittelwert des eigenen Schönheitsideals bei 7,77 mit einer Standardabweichung von 5,25 (Abb. 10). Damit gibt es eine große Streuung um den Mittelwert (ebd.). Durch die Angabe der Häufigkeit lässt sich erkennen, dass das angegebene Bild Drei mit 31,67 Prozent (n=76) den aktuellen Körpertrend einer Frau dieser Stichprobe widerspiegelt (siehe Anhang A, 10). Dieses Körperbild wird mit den Begriffen Dünnheit und Gewichtsverlust verbunden (vgl. Carrotte et al., 2017). Die starke Streuung um den Mittelwert könnte mit der Zusammenlegung der zwei Skalen entstanden sein. Dadurch entsteht bei den Frauen ein vielseitiges Körperideal. Dennoch wollen Frauen signifikant genauso aussehen wie ein Fitness Influencer (siehe Anhang B, 2.). Eine weibliche Fitness Influencerin steht für einen muskulösen Körper mit wenig Körperfettanteil und könnte dem Bild 14 zugeordnet werden (vgl. Pilgrim und Bohnet-Joschko, 2019; siehe Abb. 10). Der Mittelwert bei dem Körperbild der Männer liegt bei 13,68 mit einer geringeren Standardabweichung von 2,97 als bei den Frauen (siehe Abb. 11). Dieses Körperbild (Bild 13) stellt einen muskulösen Mann dar (vgl. Carrotte et al., 2017). Die Ergebnisse deuten darauf hin, dass Männer wie Frauen das Schönheitsideal eines Fitness Influencers anstreben (siehe Anhang B, 4). Durch die tägliche Präsenz der Fitness Influencer auf Instagram können negative Erscheinungsbilder wie Depressionen oder Essstörungen bei dem Follower hervorgerufen werden (siehe Kapitel 4.2). Diese Arbeit bietet lediglich einen Hinweis darauf, dass ein Teil der Generation Y ebenfalls einen trainierten Körper wie ein Fitness Influencer haben will. Dies könnte ein Grund sein, dass die Mehrheit (55,83 Prozent; n=134) der Teilnehmer die Instagram Aktivität einschränken möchten, weil sie sich nicht mit anderen Körperformen vergleichen wollen (siehe Anhang A, 6, 7).

### 7.1. Limitierung der Untersuchung

Die vorliegende Forschungsarbeit ist teilweise limitiert. Die Akquise der Teilnehmer ist durch die passive Rekrutierung beschränkt, um eine optimale Repräsentativität der Umfrage zu gewährleisten. Die Online Befragung wurde mithilfe des Umfragelinks über die Sozialen Netzwerke und einer Verteilungsmail des Sportinstitutes Mainz veröffentlicht. Es resultiert eine Selbstselektion, die die Repräsentativität der Stichprobe einschränkt. Einerseits konnte das Alter (18 bis 38) bei der Stichprobe eingeschränkt werden, sodass eine

merkmalspezifisch-repräsentative Stichprobe garantiert ist. Andererseits können, bedingt durch den sportlichen Kontext der Forscherin und deren Selbstselektion, keine optimalen repräsentativen Ergebnisse gewährleistet werden. Dies lässt sich aus der hohen sportlichen Betätigung der Teilnehmer ableiten (siehe Abb. 8; Anhang A,8). Des Weiteren ist der Anteil der Geschlechter ungleich. Es gibt mehr weibliche (62,08 Prozent; n=149) als männliche (37,92 Prozent; n=91) Teilnehmer, sodass es zu geschlechtsspezifischen Unterschieden bei der Beantwortung kommen kann (siehe Abb. 4). Der Online Fragebogen konnte mithilfe des Befragungsservers umfrage-online erstellt werden. Der zeitliche Aufwand der Teilnehmer und der Forscherin wurde anhand der fehlenden Einstellung von Filterfragen bei dem Server (Alter und Instagram Account) vergrößert. Da es sich um eine anonyme Befragung gehandelt hat, gibt es auch keine Auskunft darüber, inwiefern die Teilnehmer die Fragen ehrlich beantwortet haben. Bei der Gestaltung der Fragen ergeben sich weitere Grenzen, die die Auswertung der Ergebnisse einschränken. Da es sich häufig um dichotome Fragen mit einer Nominalskalierung handelt, konnten lediglich Häufigkeiten aufgezeigt und signifikante Zusammenhänge mithilfe des Chi-Quadrat-Tests geprüft werden. Um weitere Zusammenhänge und Richtungstendenzen aufzeigen zu können, wären Kardinalskalen notwendig gewesen. Um ein eindeutiges Körperideal von Frauen und Männern zu erhalten, wäre die Zerlegung der zwei Figuren-Bewertungs-Skalen (vgl. [Stunkard et al., 1983](#); [Lynch und Zellner, 1999](#)) von Vorteil gewesen.

## 7.2. Implikation weiterer Forschung

Aus der Forschungsarbeit ergeben sich weitere Möglichkeiten, den Einfluss von Fitness Influencern auf die Sportaktivität zu untersuchen. Der Fokus könnte auf weniger sportaktive Menschen gelegt werden, um die vorliegende Forschungsarbeit zu vertiefen. Um den Stellenwert von Fitness Influencern bei der Generation Y herauszufinden, könnten weitere Fragen hinzugezogen werden. Beispielsweise könnte der Fragebogen mit Hilfe einer Ratingskala ergänzt werden, inwiefern der Follower den Ratschlägen der Fitness Influencer vertraut oder die Meinungsmacher als Vorbild ansieht. Des Weiteren könnte der sportliche Hintergrund näher erfragt werden zum Beispiel Sportart oder wo und wann die Personen mit Sporttreiben angefangen haben. Ein experimentelles Studiendesign könnte den nachhaltigen Aspekt von Sport durch das Folgen eines Fitness Influencers herausfinden, indem über einen zeitlichen Verlauf sportliche Aktivitäten und Körpermessindexe der Teilnehmer überprüft werden. Außerdem könnte der kritische Diskurs über das vermittelte Körperideal und der Zwang nach sportlicher Betätigung durch Fitness Influencer erforscht werden.

## 8. Fazit

Bei der vorliegenden Forschungsarbeit konnte, mithilfe einer Online Umfrage, der Einfluss von Fitness Influencern bei Instagram auf die sportliche Aktivität untersucht werden.

Die Stichprobe bezog sich auf die Generation Y, die einen Instagram Account besitzt. Zusammenfassend lässt sich festhalten, dass Instagram täglich von der untersuchten Generation Y zur Freizeitbeschäftigung genutzt wird. Ein beliebtes Thema bei dieser Plattform ist Sport/Fitness. Vor allem folgen die Mehrheit (68,75 Prozent; n=165) der Teilnehmer einem Fitness Influencer auf Instagram. Die Meinungsmacher erregen mit ihrem Alltagsleben, Fitnessübungen-, Programmen etc. die Aufmerksamkeit der Follower. Allerdings bedarf es weiterer Forschung über die Qualität der Empfehlungen von Fitness Influencern im Bezug auf sportliches Training. Es lässt sich der Trend beobachten, dass die tägliche Nutzung von Instagram vom Alltag ablenkt, sodass die untersuchte Generation Y ihre zeitliche Nutzung dieser Plattform einschränken möchte. Dies könnte darauf hinweisen, dass Instagram die sportliche Aktivität negativ beeinflussen könnte. Aus den Ergebnissen der Umfrage geht hervor, dass die meisten Personen mehrmals in der Woche Sport treiben. Obwohl Sport ein vielseitiger Begriff ist und unterschiedlich verstanden wird, lässt sich anhand der Gründe für das Sporttreiben feststellen, dass Sport für Wohlbefinden, Gesundheit, Ausgleich zum Alltag, Erhöhung der sportlichen Leistungsfähigkeit usw. steht. Es konnte jedoch kein Zusammenhang zwischen dem Folgen von Fitness Influencern auf Instagram und dem Sportmachen festgestellt werden. Dafür könnten Teilnehmer mit weniger sportlichen Hintergründen ausgewählt werden. Dennoch zeigen die Ergebnisse der Hypothese Zwei den Zusammenhang zwischen Fitness Influencern und der Lust auf Sporttreiben durch die Bilder und Videos der Meinungsmacher. Außerdem gibt es einen Zusammenhang zwischen dem Folgen eines Fitness Influencers und dem Ausprobieren von Fitnessprogrammen eines Fitness Influencers. Um den Einfluss von Fitness Influencern auf das langfristige Sporttreiben des Followers zu testen, wäre eine experimentelle Studie möglich. Die Fitness Influencer vermitteln ein athletisches Körperideal, das zum Sporttreiben anregen könnte. Dadurch können negative Gefühle bei der Betrachtung dieser Körperideale entstehen. Dennoch geht aus der Studie hervor, dass Fitness Influencer auf Instagram die sportliche Aktivität bei der Generation Y beeinflussen können. Die Sportaktivität wird sowohl durch die visuelle Darstellung der Körper als auch durch die Veröffentlichung von Trainingsprogrammen der Fitness Influencer angeregt. Dadurch könnten Fitness Influencer auf Instagram einen positiven Beitrag zur sportlichen Aktivität bei der Bevölkerung leisten.

## Literatur

- AS&S Radio GmbH. AS&S Radio Deutschland-Kombi Young & Online - den Digital Natives näherkommen, 2016. URL [https://www.ard-werbung.de/fileadmin/user\\_upload/radio/PDF-Dokumente/Generation\\_Y.pdf](https://www.ard-werbung.de/fileadmin/user_upload/radio/PDF-Dokumente/Generation_Y.pdf). Zugriff am 8.9.2018.
- Baur, N. und Blasius, J. *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden, Springer Fachmedien, 2014.
- Boepple, L., Ara, R., Rum, R., und Thompson, K., J. Strong is the new skinny: A content analysis of fitspiration websites. *Body Image*, 17:132–135, 2016. doi: 10.1016/j.bodyim.2016.03.001.
- Bonfadelli, H. und Friemel, T. *Medienwirkungsforschung*. Konstanz, UTB (UTB, 3451: Medien- und Kommunikationswissenschaft, Pädagogik, Psychologie, Soziologie, 2011).
- Bonfadelli, H., Jarren, O., und Siegert, G. *Einführung in die Publizistikwissenschaft*. 3. Aufl. Stuttgart, [s.l.]: UTB GmbH; Haupt, 2012.
- Bortz, J. und Döring, N. *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Limitierte Sonderausgabe. 4., Aufl. 2006. Kartierte Sonderausgabe 2015. Berlin, Springer Berlin (Springer-Lehrbuch), 2015.
- Bortz, J. und Schuster, C. *Statistik für Human- und Sozialwissenschaftler. Mit 163 Tabellen*. 7., vollst. überarb. und erw. Aufl. Berlin, Springer, 2010.
- Braumüller, B. Sportbezogenes Handeln in virtuellen sozialen Netzwerken. *German Journal of Exercise and Sport Research*, 48(1):79–88, 2018. doi: 10.1007/s12662-017-0466-z.
- Bühl, A. *SPSS 22. Einführung in die moderne Datenanalyse*. 14., aktualisierte Aufl. Hallbergmoos, Pearson (ST, Scientific Tools), 2014.
- Bundesverband für Digitale Wirtschaft und Influry. Bedeutung von Influencer Marketing in Deutschland 2017, 2017. URL [https://www.bvdw.org/fileadmin/bvdw/upload/studien/171128\\_IM-Studie\\_final-draft-bvdw\\_low.pdf](https://www.bvdw.org/fileadmin/bvdw/upload/studien/171128_IM-Studie_final-draft-bvdw_low.pdf). Zugriff am 22.10.2019.
- Bundesverband für Digitale Wirtschaft (BWVD) e. V. Digital Trends – Umfrage zum Umgang mit Influencern, 2019. URL [https://www.bvdw.org/fileadmin/user\\_upload/190404\\_IM-Studie\\_BVDW\\_2019.pdf](https://www.bvdw.org/fileadmin/user_upload/190404_IM-Studie_BVDW_2019.pdf). Zugriff am 22.10.2019.
- Carrotte, E. R., Vella, A. M., und Lim, M. S. C. Predictors of 'Liking' Three Types of Health and Fitness-Related Content on Social Media: A Cross-Sectional Study. *Journal of Medical Internet Research*, 17(8, e205), 2015a. doi: 10.2196/jmir.4803.
- Carrotte, E. R., Prichard, E., und Lim, M. S. C. Predictors of 'Liking' Three Types of Health and Fitness-Related Content on Social Media: A Cross-Sectional Study. *Journal of Medical Internet Research*, 17(8, e205), 2015b. doi: 10.2196/jmir.4803.
- Carrotte, E. R., Prichard, E., und Lim, M. S. C. 'Fitspiration' on Social Media: A Content Analysis of Gendered Images. *Journal of Medical Internet Research*, 19(3, e95), 2017. doi: 10.2196/jmir.6368.
- Clement, R. und Schreiber, D. *Internet-Ökonomie. Grundlagen und Fallbeispiele der vernetzten Wirtschaft*. 2. vollst. überarb. u. erweiter. Aufl. 2013. Berlin, Heidelberg, Springer Gabler, 2013.
- Cohen, J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, N.J., L. Erlbaum Associates, 1988.
- Dege, F. *Quick Guide Influencer Marketing. Wie Sie durch Multiplikatoren mehr Reichweite und Umsatz erzielen*. Wiesbaden, Springer Gabler, 2018. URL <http://dx.doi.org/10.1007/978-3-658-22163-8>.
- Deighton-Smith, N. und Bell, B. T. Objectifying fitness: A content and thematic analysis of #fitspiration images on social media. *Psychology of Popular Media Culture*, 7(4):467–483, 2018. doi: 10.1037/ppm0000143.
- Deloitte. Datenland Deutschland – Die Transparenzlücke, 2015. URL <https://www2.deloitte.com/content/dam/Deloitte/de/Documents/deloitte-analytics/DAI-Datenland-Deutschland-Generationsluecke-s.pdf>. Zugriff am 23.19.2019.
- Deloitte. European Health & Fitness Market Report 2019, 2019a. URL [http://www2.deloitte.com/content/dam/Deloitte/de/Documents/consumer-business/EuropeActive\\_Deloitte\\_EHFM2019\\_DE.pdf](http://www2.deloitte.com/content/dam/Deloitte/de/Documents/consumer-business/EuropeActive_Deloitte_EHFM2019_DE.pdf). Zugriff am 9.6.2019.
- Dilger, E. *Die Fitnessbewegung in Deutschland. Wurzeln, Einflüsse und Entwicklungen*. Schorndorf, Hofmann (Beiträge zur Lehre und Forschung im Sport, Bd. 164), 2008.
- Dimitriou, M. und Ring-Dimitriou, S. H. *Der Körper in der Postmoderne. Zwischen Entkörperlichung und Körperwahn*. Wiesbaden, Springer, 2019.
- Döring, N. *Gesundheitskommunikation auf YouTube: Fallstudien zu 25 Fitness- und Lifestyle-Kanälen*. In: Markus Schäfer, Oliver Quiring, Constanze Rossmann, Matthias R. Hastall und Eva Baumann (Hg.): *Gesundheitskommunikation im gesellschaftlichen Wandel*. Baden-Baden, Nomos, 2015.
- Döring, N. und Bortz, J. *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Unter Mitarbeit von Sandra Pöschl-Günther. 5. vollständig überarbeitete, aktualisierte und erweiterte Auflage. Berlin, Heidelberg: Springer, 2016. doi: <http://dx.doi.org/10.1007/978-3-642-41089-5>.
- Easton, S., Morton, K., Tappy, Z., Francis, D., und Dennison, L. Young People's Experiences of Viewing the Fitspiration Social Media Trend: Qualitative Study. *Journal of Medical Internet Research*, 20(6, e219), 2018. doi: 10.2196/jmir.9156.
- Emarketer. Instagram Users and Penetration Worldwide, 2016-2021 (millions, change and URL <https://www.emarketer.com/Chart/Instagram-Users-Penetration-Worldwide-2016-2021-millions-change-of-social-network-users/214458>. Zugriff am 16.6.2019.
- Faßmann, M. und Moss, C. *Instagram als Marketing-Kanal: Die Positionierung ausgewählter Social-Media-Plattformen*. Wiesbaden, Springer Science and Business Media, 2016.
- Firsching, J. Influencer Marketing: Wer sind diese Instagram Influencer überhaupt?, 2015. URL <https://www.futurebiz.de/artikel/wer-sind-instagram-influencer/#.XbQ77egzY2x>. Zugriff am 22.10.2019.
- Firsching, J. Instagram Statistiken für 2019: Nutzerzahlen, Instagram Stories, Instagram Videos & tägliche Verweildauer, 2019. URL <http://www.futurebiz.de/artikel/instagram-statistiken-nutzerzahlen/>. Zugriff am 15.6.2019.
- Frank, G. Kalya Itsines easy bikini body workout: exercises from Bikini Body Guide, 2018. URL <https://www.today.com/series/start-today/kayla-itsines-easy-bikini-body-workout-exercises-bikini-body-guide-t130950>. Zugriff am 22.10.2019.
- Freese, B. und Koch, W. ARD/ZDF-Onlinestudie 2018: Zuwachs bei medialer Internet-nutzung und Kommunikation. In: *Media Perspektiven* (9), 2018. URL [http://www.ard-zdf-onlinestudie.de/files/2018/0918\\_Freese\\_Koch.pdf](http://www.ard-zdf-onlinestudie.de/files/2018/0918_Freese_Koch.pdf). Zugriff am 28.6.2019.
- Fuchs, Daniel. Instagram-profil, 2019. URL [https://www.instagram.com/magic\\_fox/](https://www.instagram.com/magic_fox/). Zugriff am 23.10.2019.
- Gerling, W., Holschbach, S., und Löffler, P. *Fotografische Praktiken in der digitalen Kultur*. Bielefeld, Transcript Verlag, 2018.
- Gerstenberg, F. und Gerstenberg, C. *Quick guide social relations. PR-Arbeit mit Bloggern und anderen Influencern im Social Web*. Wiesbaden, Springer Gabler, 2018.
- Güllich, A. und Krüger, M. *Sport. Das Lehrbuch für das Sportstudium*. Berlin, Springer Spektrum, 2013.
- Goldapp, N. *Medienunternehmen im Social Web. Erkenntnisse zur reichweitenstarken Content-Generierung*. Wiesbaden, Springer Fachmedien (EBL-Schweitzer), 2015.
- Grabs, A., Bannour, K.-P., und Vogl, E. *Follow me! Erfolgreiches Social Media Marketing mit Facebook, Instagram, Pinterest und Co*. Bonn, Rheinwerk Verlag (Rheinwerk Computing), 2018.
- Grimmer, C. G. *Sportkommunikation in digitalen Medien. Vielfalt, Inszenierung, Professionalisierung*. Wiesbaden, Springer Fachmedien, 2019.
- Grossbard, J. R., Beighors, C., und Larimer, M. E. Perceived norms for thinness and muscularity among college students: what do men and women really want? *Eating Behaviors*, 12(3):192–199, 2011. URL [10.1016/j.eatbeh.2011.04.005](https://doi.org/10.1016/j.eatbeh.2011.04.005).
- Holand, G. und Tiggemann, M. 'Strong beats skinny every time': Disordered eating and compulsive exercise in women who post fitspiration on Instagram. *The International Journal of Eating Disorders*, 50(1):76–79, 2017. doi: 10.1002/eat.22559.
- Hottenrott, K. und Neumann, G. *Trainingswissenschaft. Ein Lehrbuch in 14 Lektionen*. 3., überarbeitete Auflage. Hg. v. Wolf-Dietrich Brettschneider. Aachen, Meyer & Meyer, 2016.
- Hussy, W., Schreier, M., und Echterhoff, G. *Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor*. Berlin, Springer, 2013.
- ISPO Munich. Alles, was zählt – Zahlen und Fakten zur ISPO, 2018. URL <https://www.ispo.com/sites/default/files/2017-10/ispo-munich-zahlen-fakten.pdf>. Zugriff am 23.10.2019.
- Itsines, Kayla. Instagram-profil, 2019. URL [https://www.instagram.com/kayla\\_itsines/](https://www.instagram.com/kayla_itsines/). Zugriff am 23.10.2019.
- Jahnke, M., Brix, R., Bruce, A., und Fuchs, T. *Influencer Marketing für Unternehmen und Influencer. Strategien, Plattformen, Instrumente, rechtlicher*

- Rahmen: mit vielen Beispielen. Wiesbaden, Springer Gabler, 2018.
- Janssen, J. und Laatz, W. *Statistische Datenanalyse mit SPSS für Windows: Eine anwendungsorientierte Einführung in das Basissystem version 8 und das Modul Exakte tests*. Berlin, Springer, 2013.
- Kilian, K. Influencer sind die neuen Promis. In: absatzwirtschaft (7/8). Medien & Kommunikation, 2016. URL [http://www.markenlexikon.com/texte/asw\\_kilian\\_influencer-sind-die-neuen-promis\\_7-8\\_2016.pdf](http://www.markenlexikon.com/texte/asw_kilian_influencer-sind-die-neuen-promis_7-8_2016.pdf).
- Kissner, Anne. Instagram-profil, 2019. URL [https://www.instagram.com/anne\\_bodykiss/](https://www.instagram.com/anne_bodykiss/). Zugriff am 23.10.2019.
- Klapp, A.-L. und Klotter, J. C. #fitness vs. #BodyPositive – die unterschiedliche soziale Konstruktion von Gesundheit auf Instagram. *Prävention und Gesundheitsförderung*, 14(4):362–367, 2019.
- Kobilke, K. *Marketing mit Instagram. Das umfassende Praxis-Handbuch*. Frechen, MITP-Verlags GmbH & Co. KG, 2019.
- Krampf, L. und Weichert, S. *Der Millennial Code. Junge Mediennutzer verstehen - und handeln*. Leipzig, Vistas, 2017.
- Krüger, A., Dreyer, A., Czech, M., Freyer, W., Görlich, S., Horbel, C., Ostermann, F., und Pfaff, S. *Sportmanagement. Eine themenbezogene Einführung*. Hg. v. Arnd Krüger und Axel Dreyer. München, De Gruyter, 2004. URL [http://www.degruyter.com/search?f\\_0=isbnissn&q\\_0=9783486781861&searchTitles=true](http://www.degruyter.com/search?f_0=isbnissn&q_0=9783486781861&searchTitles=true).
- Lorenz, S. Generation Y, Millennials, Digital Natives, Generation me, 2019. URL <https://www.manager-magazin.de/unternehmen/karriere/generation-y-millennials-digital-natives-generation-me-podcast-a-1264969.html>. Zugriff am 8.9.2019.
- Lüdemann, D. Drei Stunden am Tag sind normal. In: Zeit Online, 2018. URL <https://www.zeit.de/digital/internet/2018-03/social-media-dak-studie-instagram-whatsapp-sucht-jugendliche>. Zugriff am 22.10.2019.
- Lynch, S. M. und Zellner, A., Debra. *Sex Roles*, 40(9/10):833–843, 1999. URL [10.1023/A:1018868904239](https://doi.org/10.1023/A:1018868904239).
- Mangelsdorf, M. *Von Babyboomer bis Generation Z. Der richtige Umgang mit unterschiedlichen Generationen im Unternehmen*. Offenbach, Gabal Verlag (Whitebooks), 2017.
- Mills, L. Les Mills Global Consumer Fitness Survey 2019 - Die wichtigsten Erkenntnisse, 2019. URL <https://www.lesmills.com/2019-consumer-survey/>. Zugriff am 8.9.2019.
- Nirschl, M. und Steinberg, L. *Einstieg in das Influencer Marketing. Grundlagen, Strategien und Erfolgsfaktoren*. Wiesbaden, Springer Gabler, 2018. URL <http://dx.doi.org/10.1007/978-3-658-19745-2>.
- Pilgrim, K. und Bohnet-Joschko, S. Selling health and happiness how influencers communicate on instagram about dieting and exercise: mixed methods research. *BMC Public Health*, 19(1):1054, 2019. doi: 10.1186/s12889-019-7387-8.
- Polzin, B. und Weigl, H. *Führung, Kommunikation und Teamentwicklung im Bauwesen. Grundlagen, Anwendung, Praxistipps; mit zahlreichen Beispielen*. Wiesbaden, Vieweg + Teubner, 2009.
- Preuß, H., Alfs, C., und Ahlert, G. *Sport als Wirtschaftsbranche. Der Sportkonsum privater Haushalte in Deutschland*. Wiesbaden, Springer Gabler, 2012.
- Raab-Steiner, E. und Benesch, M. *Der Fragebogen. Von der Forschungsidee zur SPSS-Auswertung*. Wien, Facultas.wuv (UTB, 8406), 2012.
- Raggatt, M., Wright, C. J. C., Carrotte, E., Jenkinson, R., Mulgrew, K., Pritchard, I., und Lim, M. S. C. 'i aspire to look and feel healthy like the posts convey': engagement with fitness inspiration on social media and perceptions of its influence on health and wellbeing. *BMC Public Health*, 18(1):1002, 2018. doi: 10.1186/s12889-018-5930-7.
- Ralph-Nearman, C. und Filik, R. New Body Scales Reveal Body Dissatisfaction, Thin-Ideal, and Muscularity-Ideal in Males. *American Journal of Men's Health*, 12(4):740–750, 2018. doi: 10.1177/1557988318763516.
- Reif, Pamela. Instagram-profil, 2019. URL [https://www.instagram.com/pamela\\_rf/](https://www.instagram.com/pamela_rf/). Zugriff am 23.10.2019.
- Reifegerste, D. und Baumann, E. *Medien und Gesundheit*. Wiesbaden, Springer, 2018. URL <https://doi.org/10.1007/978-3-658-20013-8>.
- Royal Society for Public Health. Instagram ranked worst for young people's mental health, 2017. URL <https://www.rsph.org.uk/uploads/assets/uploaded/d125b27c-0b62-41c5-a2c0155a8887cd01.pdf>. Zugriff am 23.10.2019.
- Santarossa, S., Coyne, P., Lisinski, C., und Woodruff, S. J. #fitspo on Instagram: A mixed-methods approach using Netlytic and photo analysis, uncovering the online discussion and author/image characteristics. *Journal of Health Psychology*, 24(3):376–385, 2019. doi: 10.1177/1359105316676334.
- SAZ Sport. So ticken Millennials in Bezug auf Sport, 2018. URL <https://www.sazsport.de/markt-sortimente/zielgruppe/so-ticken-millennials-in-bezug-sport-1545135.html>. Zugriff am 8.9.2013.
- Schach, A. und Lommatzsch, T. *Influencer Relations. Marketing und PR mit digitalen Meinungsführern*. Wiesbaden, Springer Gabler, 2018. URL <https://dx.doi.org/10.1007/978-3-658-21188-2>.
- Scherkamp, H. Vom Maschinenbauer zum erfolgreichen Instagram-Model, 2017. URL <https://www.gruenderszene.de/karriere/daniel-fuchs-instagram-magic-fox>. Zugriff am 22.10.2019.
- Schmidt, J.-H. und Taddicken, M. *Handbuch Soziale Medien*. Wiesbaden, Springer Fachmedien, 2017.
- Schouten, A. P., Janssen, L., und Verspaget, M. Celebrity vs. Influencer endorsements in advertising: the role of identification, credibility, and Product-Endorser fit. *Journal of Advertising*, 23(7):1–24, 2019. doi: 10.1080/02650487.2019.1634898.
- Schütte, N. *Sportmanagement. Historie, Themen, Berufsbild*. Stuttgart, UTB, 2015.
- Seeger, C. und Kost, J. F. *Influencer Marketing. Grundlagen, Strategie und Management*. München, UVK Verlag, 2019.
- Sokolova, K. und Kefi, H. Instagram and YouTube bloggers promote it, why should I buy? How credibility and parasocial interaction influence purchase intentions. *Journal of Retailing and Consumer Services*, 2019. doi: DOI:10.1016/j.jretconser.2019.01.011.
- Statista. Welche der folgenden Instagram-Influencer, die mit Fitness-Inhalten bekannt geworden sind, kennen Sie zumindest dem Namen nach? . Statista, Zugriff am 23.10.2019, 2019a. URL <https://de.statista.com/prognosen/958398/umfrage-zur-bekanntheit-von-social-media-influencern-Fitness>.
- Statista. Warum nutzen sie instagram? Statista, 2019b. URL <https://de.statista.com/prognosen/957928/umfrage-in-deutschland-zu-gruenden-aus-denen-man-instagram-nutzt>. Zugriff am 23.10.2019.
- Stunkard, A. J., Sørensen, T., und Schulsinger, F. Use of the Danish Adoption Register for the study of obesity and thinness. *Research publications - Association for Research in Nervous and Mental Disease*, 60, 1983.
- Tiggemann, M. und Zaccardo, M. 'Exercise to be fit, not skinny': The effect of fitpiration imagery on women's body image. *Body Image*, pages 61–67, 2015. doi: 10.1016/j.bodyim.2015.06.003.
- Vaterlaus, J. M., Patten, E. V., Cesia, R., und Young, J. A. #Gettinghealthy: The perceived influence of social media on young adult health behaviors. *Computers in Human Behavior*, 45:151–157, 2015. doi: 10.1016/j.chb.2014.12.013.
- Vogl, M. Social Snippet Digital Detox: Der Wunsch nach der digitalen Auszeit, 2019. URL <https://www.brandwatch.com/de/blog/digital-detox/>. Zugriff am 23.10.2019.
- Voss, R. *Wissenschaftliches Arbeiten*. Konstanz, UTB, 2015.
- VuMa, Arbeitsgemeinschaft für Verbrauchs- und Medienanalyse. Millennials in Deutschland nach Häufigkeit der Nutzung von Social Media (z.B. Facebook, Twitter, Instagram, WhatsApp) im Jahr 2018. Statista, 2018c. URL <https://de.statista.com/statistik/daten/studie/712857/umfrage/umfrage-unter-millennials-zur-haeufigkeit-der-nutzung-von-social-media/>. Zugriff am 5. Juli 2019.
- VuMa, Arbeitsgemeinschaft Verbrauchs- und Medienanalyse. Bevölkerung in Deutschland nach Häufigkeit des Sporttreibens in der Freizeit von 2015 bis 2018 (Personen in Millionen). 2018a. URL <https://de.statista.com/statistik/daten/studie/171911/umfrage/haeufigkeit-sport-treiben-in-der-freizeit/>. Zugriff am 23.10.2019.
- VuMa, Arbeitsgemeinschaft Verbrauchs- und Medienanalyse. Beliebteste Freizeitbeschäftigungen, Aktivitäten und Sportarten (mindestens mehrmals im Monat) in Deutschland in den Jahren 2015 bis 2018. Statista, 2018b. URL <https://de.statista.com/statistik/daten/studie/171601/umfrage/mehrmals-pro-monat-ausgeuebte-freizeitaktivitaeten/>. Zugriff am 5.7.2019.
- We Are Social & Hootsuite & DataReportal. Anzahl der Instagram-Nutzer nach Altersgruppen und Geschlecht weltweit im Januar 2019 (in Millionen). Statista, 2019. URL <https://de.statista.com/statistik/da>

- ten/studie/809703/umfrage/instagram-nutzer-nach-alter-und-geschlecht-weltweit/. Zugriff am 23.10.2019.
- Weltgesundheitsorganisation. Physical Activity, 2019. URL <https://www.who.int/dietphysicalactivity/pa/en/>. Zugriff am 8.9.2019.
- Würzburger, T. *Key Skills für die Generation Y: Die wichtigsten Tipps für eine erfüllte Karriere*. Wiesbaden, Springer, 2016.
- Zingg, R. Immer auf Instagram! Wie geht man sinnvoll mit der App um?, 2018. URL <https://www.welt.de/icon/partnerschaft/article174943467/Social-Media-Instagram-nervt-So-macht-die-App-endlich-wieder-Spass.html>. Zugriff am 22.10.2019.



# The Effect of Perceived CEO Overconfidence on Consumers' Product and Brand Evaluations

Lara Katharina Hilgert

WHU - Otto Beisheim School of Management

## Abstract

This research examines the impact of CEO overconfidence as a mediator of consumer attitudes towards products and brands as well as its influence on perceived brand personality. The results of the study show that revealed verbally communicated overconfidence leads to lower ratings of the product, a lower general attitude towards the brand, lower trust in the brand and lower ratings of descriptiveness regarding the brand personality trait competence. The negative impact of CEO overconfidence is most influential on the participants' trust in the brand. The study additionally observed participants' evaluations of perceived CEO overconfidence and their brand attitudes regarding well-known brands. The results show that participants' general attitude towards a brand and trust in a brand were lower when they perceived the CEO as more overconfident. The insights gained through this study have important implications for the literature on the impact of overconfidence on an individual's status and the literature on CEO endorsement. Additionally, this research can act as a guidance for CEOs and executives on how his or her trait of overconfidence affects consumers' and how the company's performance may be affected by this relationship.

**Keywords:** Overconfidence; leadership; consumer; perception; brand.

## 1. Introduction

In the literature, the various negative and positive consequences that CEO overconfidence may have for companies have been extensively studied. For example, research has found that overconfident CEOs destroy value as they overestimate the returns their investments can generate (Heaton, 2002). On the other hand, overconfident CEOs create greater employee and supplier commitment (Phua et al., 2018). However, an aspect that has not been thoroughly examined is the relationship between CEO overconfidence and consumer attitudes. This gap in the literature was a significant motivating factor for the current research.

### 1.1. Background

Elon Musk is an example of a CEO which some may classify as overconfident. He is the co-founder and CEO of Tesla and SpaceX and is known for setting, and in part reaching, radical goals. One example which reflects his extreme level of confidence is his compensation plan at Tesla (Edmans, 2018). It equals other executives' compensation plans in that the payout is linked to the achievement of specific performance targets. However, the targets of Musk's compensation plan

are set at very high levels. For Musk to receive his payout the market value of Tesla must reach \$100 billion. Further targets are set at \$150 billion, continuing up to \$650 billion. This may seem very ambitious if not even overconfident as the current market valuation is at around \$45 billion (Bloomberg, 2019). Thus, the high targets in Musk's compensation plan implicitly underline his high level of confidence in his personal and his company's ability to succeed.

Elon Musk's overconfidence can be observed even more explicitly in his public statements and the publications of the company's performance goals. For example, in 2018 Musk admitted that one cause for issues in the production process of Tesla's cars was that they had become too confident in their ability to produce batteries (Thompson, 2018). Also, SpaceX had announced the goal to make the first trial flight with an in-cabin crew using one of their spacecrafts this summer (Dunn, 2019). However, a recent accident which occurred during a crew-less test flight has delayed this milestone to an unknown time in the future.

Will it affect consumer attitudes towards Tesla's cars if Elon Musk is overly confident about the timeframe in which he will make private spaceflight possible? My research is con-

cerned with how CEO overconfidence impacts consumer attitudes towards products and the brand of a company. I have thoroughly explored this topic using the existing literature on overconfidence and consumer attitudes as a starting point and extending the previous findings in literature with an empirical study concerning the impact of CEO overconfidence on consumer attitudes.

## 1.2. Outline

First, I will give a general overview of the different types and categories of overconfidence and I will elaborate on the different possible causes of overconfidence that are discussed in the literature. Then I will frame this information specifically in a management context to show why the overconfidence bias especially affects managers and how managerial overconfidence has been shown to impact businesses financial results as well as relationships with stakeholders. In section four I will review literature that explores the psychological processes by which CEO personality traits such as overconfidence influence consumer attitudes. Lastly, I will discuss the concept of consumer attitudes, specifically brand trust, and how these affect consumer's decisions and behavior towards brands.

Following the review of relevant literature, I will introduce the study which I conducted to explore the impact of CEO overconfidence on consumer attitudes empirically. The research was designed to explore the central hypothesis that verbally communicated CEO overconfidence would negatively affect consumer's attitudes towards products and the brand of a company. As the study consisted of two main parts, I will present the method, results, and discussion of the results separately. For both parts, I will discuss the essential findings and frame them in the context of the relevant literature. I will explore what meanings the two parts of the study have for each other and how their findings are related. To conclude my research, I will examine the implications the findings have for the previous and future literature related to the relationship of CEO overconfidence and consumer attitudes. Finally, I will discuss the implications the study holds for CEOs and executives today.

## 2. Overconfidence

In 1995, De Bondt and Thaler said that "Perhaps the most robust finding in the psychology of judgment is that people are overconfident" (De Bondt and Thaler, 1995, p. 389). The overconfidence bias has been thoroughly explored in literature, and the scope of application and the consequences of overconfidence are far-reaching. As I aim to observe the impact of CEO overconfidence on consumer attitudes, I will narrow my view to a specific type of overconfidence.

### 2.1. Concept of Overconfidence

In literature, it is common to differentiate between three types of overconfidence (Moore and Healy, 2008) and more recently the distinction between verbally and non-verbally

communicated overconfidence has been introduced (Tenney et al., 2019). A specific context in which overconfidence finds a critical application is the field of managerial overconfidence. In the following, I will give a general overview of the different definitions of overconfidence and explore the specific field of managerial overconfidence. These distinctions play an essential role in the design of the current study which explores the impact of exposed CEO overconfidence on consumers.

#### 2.1.1. Overestimation, Overplacement, Overprecision

Overconfidence is described as an incorrect, overly positive estimation regarding personal ability and knowledge (Moore and Healy, 2008). This broad description applies to a variety of cognitive processes. According to Moore and Healy (2008) it is useful to distinguish between three types of overconfidence: overestimation of one's actual performance, overplacement of one's performance relative to others, and excessive certainty regarding the precision of one's beliefs. Moore and Healy (2008) define overestimation as an "overestimation of one's actual ability, performance, level of control, or chance of success" (p. 502). Most studies are concerned with this type of overconfidence. One Example of such a study showed that college students overestimated the starting salary that they were going to earn after graduation and their performance on exams (Shepperd et al., 1996). Overplacement occurs when people believe that they are more skilled than others. It is related to the "better-than-average" effect which describes people's belief that they perform above average on different social or intellectual tasks (Burson et al., 2006). One example of overplacement is a study by Klein and Kunda (1993) in which they gave one group of participants the information of how often people on average engage in specific health-threatening behavior and the second group did not receive this information. Then, they asked participants how often they engaged in these types of behaviors. The participants from the group which had received the information about the peer's average frequency of specific behavior indicated that they engaged less frequently in these types of behaviors. The other group that had no averages to compare their evaluation to, indicated a higher engagement in those behaviors. Lastly, Moore and Healy (2008) define overprecision as excessive certainty regarding the accuracy of one's beliefs. When examining overprecision participant are usually asked to complete questions with quantifiable answers and to give confidence intervals around their statement.

Differentiating between these different types of overconfidence is important for several reasons. Firstly, there may be inconsistencies between the different types of overconfidence which can lead to inconsistencies in interpretation (Moore and Healy, 2008). For example, participants may show a high level of underestimation, but at the same time a high level of overplacement when asked to estimate performance on easy tasks. This phenomenon occurs as when tasks are easy, people will underestimate their performance, but underestimate others performance even more, and therefore believe



that they are better than others (Moore and Small, 2007). Secondly, different types of overestimation apply in different situations (Moore and Healy, 2008). As a consequence, in certain situations, it may be more appropriate to observe one type of overconfidence over another. In my study, I choose to focus on the overconfidence type of overestimation.

### 2.1.2. Verbally and Non-Verbally Expressed Overconfidence

A study by Tenney et al. (2019) differentiates between verbally and non-verbally expressed overconfidence. According to this research people's perception of the overconfident person depends on the channel by which overconfidence is expressed. The study finds that non-verbally expressed overconfidence yields reputational benefits over verbally expressed confidence. The results showed that if one member of a group either verbally stated their confidence about their ability to succeed at a task or just acted very confident about their ability to succeed they were awarded a higher status in the group. However, when it was revealed that the person's actual performance was worse than their previous level of confidence suggested the ones who had made the verbal statement received a lower social status while those who had non-verbally communicated their overconfidence did not receive a social punishment in the form of lower status. Thus, the differentiation between verbal and non-verbal expression of overconfidence has important implications for the research on the consequences of managerial overconfidence. It suggests that the channel by which a CEO or executive communicates overconfident statements will impact how stakeholders evaluate the manager's level of competence if their overconfidence is revealed.

Most research on CEO overconfidence uses an overconfidence measure which was first introduced by Malmendier and Tate (2005) that is based on the observation of the CEOs behavior regarding the holding and exercising of his or her company stock options. This portfolio-based measure of overconfidence considers the amount of in the money stock options the CEO holds. As the compensation of CEOs often includes a large number of stocks and option grants, CEOs are naturally in a situation where their personal portfolio is under-diversified, and they are exposed to a large amount of company-specific risk. Therefore, it would be rational to increase diversification by exercising these stock options. However, an overconfident CEO will overestimate the future performance and valuation of the company and thus will irrationally choose a higher risk exposure by holding options for longer. Following this logic, an overconfident CEO is identified if he or she exercises the options within one year of expiration and if the options were in the money to a certain degree one year before the expiration date. This measure does not take the CEO's channel of expression into account by which they communicate overconfidence to others. Therefore, to gain full insights into the reputational and social effects of CEO overconfidence the measure by Malmendier and Tate is not appropriate. Here a method that takes the difference between verbally and non-verbally communicated overconfidence into account should be used.

### 2.2. Reasons for the Persistence of Overconfidence

Individuals who are overconfident struggle to make correct self-evaluations and cannot correctly assess the outcomes of certain actions as they misperceive their abilities or how their abilities compare to those of others (Anderson et al., 2012b). Thus, they may make decisions that are sub-optimal or even harmful. Still, the bias has persisted in the behavior of humans. Research explains that possible reasons for the persistence of overconfidence are psychological and social advantages which result from overconfidence.

#### 2.2.1. Psychological Benefits

In literature, many examples of psychological benefits from confidence have been explored. As a consequence of overconfidence, people have a biased-self view which occurs in the form of a self-serving bias (Anderson et al., 2012b). This bias causes people to have altering perceptions of causality so that they attribute the cause of their success to their abilities while they blame failures on external factors (Miller and Ross, 1975). The bias is universal and affects most age groups (Mezulis et al., 2004). The self-serving bias has proven to yield psychological benefits in the form of improved mental health, for example, lower levels of depression (Kuiper, 1978). Another example of consequences of the bias is found in an experiment by Alicke (1985) in which he asked students to rate how various traits characterized themselves and to which degree they applied to the average college student. He found that the students rated themselves above average regarding traits which they considered highly desirable. The opposite was true for traits of low desirability. Alicke explained that the high degree of confidence enables the maintenance of a positive self-concept and leads to high self-esteem.

Another manifestation of overconfidence is the illusion of control bias which describes the phenomenon that people tend to overestimate their personal ability to influence future events while they underestimate the impact of chance on a situation's outcome (Langer, 1975). When exploring this phenomenon, Langer found that even in situations where the outcome depended on chance, such as a lottery, people behaved as if they had personal control in determining the outcome. Taylor and Brown (1988) summarize findings from literature which show that people with strong positive illusions of control have a stronger motivation to succeed and longer persistence when working on tasks.

#### 2.2.2. Social Benefits

Besides personal psychological benefits, overconfidence can also yield social benefits. Anderson et al. (2012b) conducted a study concerning the status-enhancing effect of overconfidence and explored the hypothesis that overconfidence could increase a person's social status. The authors based their theory on the observation that overconfident individuals are perceived as more competent by their peers even if they lack the ability to justify their level of confidence. This bias is the consequence of competence not being a visible trait

or behavior which people can easily assess. Therefore, individuals rely on observing specific patterns of behavior, also known as competence cues, to judge a person's competence (Anderson and Kidluff, 2009). One example of such a cue is speaking in a fluent and assertive manner (Reynolds and Gifford, 2001). Research suggests that individuals who think highly of themselves, such as an overconfident person, show these types of behavior more often and thus are evaluated by peers as more competent (Baumeister, 1982). During the study by Anderson et al. (2012b), participants had to give a self-evaluation regarding how they expected to perform on a particular task. Afterwards, they had to complete the task on their own, so that it became apparent whether they had accurately estimated their ability or whether they had been overconfident. Participants were then placed in groups in which they had to complete the same task jointly. Following the task, the participants were asked to rate their peer's competence and social status in the group. The results showed that those participants who had overestimated their ability when completing the task alone were perceived as more competent and were assigned higher social status by their peers. The results confirmed the author's hypothesis that peers perceived overconfident individuals to be more competent and that as a result they were granted a higher social status in groups.

A weakness of the experimental design of the study was that peers were unlikely to discover that his or her level of ability did not justify the individual's level of confidence. Kennedy et al. (2013) argued that in reality, it is likely that peers will be able to assess the actual ability of colleagues and peers. Therefore, they conducted a study in which they explored whether peers would punish overconfident individuals using status demotions if the overconfidence was revealed. Kennedy et al. used the experiment of Anderson et al. (2012b), which was designed to observe the status-enhancing effect of overconfidence, as a starting point. This experiment was modified so that the status of individuals in a group was assessed before and after it became clear who in the group had been overconfident. The results showed that even after it was revealed which of the individuals had been overconfident, they were still awarded the higher status in the group. These findings supported the status-enhancement theory in that they provided evidence that even after overconfidence was exposed being overconfident still led to an overall enhanced social status for an individual.

In contrast, a study by Tenney et al. (2019) came to a different conclusion when examining the social effects of exposed overconfidence. In this study, they differentiated between verbally and non-verbally expressed overconfidence. They examined whether peers would punish overconfidence, once it was revealed not to be justified by ability and the degree to which such punishment depended on the channel by which overconfidence had been expressed. They used actors to model non-verbal overconfidence. The actors were either instructed to show expansive body language and a confident tone of voice or the exact opposite. The verbal overconfidence condition was imitated by using written state-

ments which included confident statements about the belief regarding personal performance. The results showed that non-verbally expressed overconfidence yielded reputational benefits over verbally expressed confidence. The authors explained this outcome with the concept of plausible deniability. According to this concept, people who verbally express overconfidence are held accountable to their claims when it is exposed that they had overestimated their ability or knowledge. However, non-verbally expressed overconfidence is harder to unveil and thus the denial of the overconfidence is more plausible.

### 3. Managerial Overconfidence

#### 3.1. The Concept of Managerial Overconfidence

Overestimation is greatest when dealing with challenging tasks, when making forecasts in areas where the outcome is characterized by low predictability, and during actions where feedback is not immediate or clear (Barber and Odean, 2001). These conditions usually apply in the environments in which managers and CEOs have to make decisions. Thus, they are particularly exposed to situations where overestimation is likely. Bertrand and Schoar (2003) found that individual characteristics of CEOs and top executives are significantly related to corporate decisions. While they only examined the specific characteristic of having an MBA and the birth year their research implies that other individual characteristics such as the level of overconfidence possibly also impact corporate practices and performance. Thus, managerial overconfidence is critical as managers are often in situations in which they are prone to overconfidence. Additionally, when they are overconfident, this trait may have consequences for the company as a whole.

Malmendier and Tate (2015, p. 46) define overconfidence in a managerial context as "the overestimation of the value a manager believes he or she can create". This thought pattern affects managers and CEOs in two major ways. Firstly, the overconfident manager will believe that the market undervalues the company. Secondly, the overconfident manager will overestimate the value he or she can create through future investment decisions. Investments may be internal such as capital expenditures and research and development investments or external such as mergers and acquisitions.

Overconfidence will also affect the communication of managers. Overconfident managers will overestimate their abilities and their company's chances of success (Banerjee et al., 2018). As a result, they may be overly optimistic when publicly communicating their views about the company or fail to disclose negative information as they are convinced to be able to reverse the poor performance.

#### 3.2. Consequences of Managerial Overconfidence for Businesses

The behavioral bias of overconfidence can influence corporate performance and stakeholder relations due to its effect on the CEO's decision-making process and the impact

of overconfidence on the perception of others. There are a variety of consequences attributed to managerial overconfidence. However, there is no universal consensus whether the overall impact of CEO overconfidence is positive or negative for a company.

### 3.2.1. Financial Impact

Studies have shown that CEO overconfidence can impact various business areas. Many studies argue that overconfident CEOs destroy value. For example, [Heaton \(2002\)](#) proposed that managerial overconfidence is a cause for corporate misinvestment decisions as overconfident managers overestimate the returns their investments will generate. [Malmendier and Tate \(2008\)](#) examined the role CEO overconfidence plays in mergers and acquisition. They explained that as overconfident CEOs believe their companies are undervalued, they are reluctant to use external sources of financing. However, they found that overconfident CEOs were still more likely to conduct mergers and even more so when internal financing was available. They also found that investors reacted more negatively to merger announcements from companies led by overconfident CEOs which indicates that these mergers may be more likely to destroy value.

On the other hand, there are also defendants of the positive financial impact that CEO overconfidence can have. [Alberto and Timothy \(2011\)](#) explored the effect of managerial overconfidence on corporate innovation. They base their hypothesis that CEO overconfidence is associated with a higher propensity to innovate on a career concern model. According to the model, overconfident CEOs are more likely to innovate as they are rewarded for successful innovation and they underestimate the risk of failure. Galasso and Simcoe measured the level of innovation using U.S. patent data and measured CEO overconfidence using Malmendier and Tate's portfolio-based measure. Their results suggested that not only is CEO overconfidence positively related to the number of patents but also to the probability that these technological developments will lead the company in a new direction.

### 3.2.2. Stakeholder Relations

Literature has shown how CEO overconfidence affects market participants and stakeholders of a company. [Phua et al. \(2018\)](#) focused on how the leadership of an overconfident CEO affects the relationship with suppliers and employees. Specifically, they were interested in observing whether companies led by overconfident CEOs benefited from higher stakeholder commitment. They based their assumption in part on the findings of the previously discussed research which found that peers grant overconfident individuals status-enhancing benefits, as they are evaluated as being more competent ([Anderson et al., 2012b](#)).

The results suggest that CEO overconfidence supports the development of relationships with suppliers as they showed that CEO overconfidence was related to higher rates of initiation and expansion of the company's network of dependent suppliers ([Phua et al., 2018](#)). To assess whether CEO overconfidence has an impact on employee commitment, Phua

et al. observed the variables of employee turnover and the amount of the employer's stock held in the retirement benefit plans of employees. They found that employee turnover was lower for companies with overconfident CEOs. They also found a positive relationship between CEO overconfidence and the amount of employee stock holdings in the company. Therefore, both measures indicated that overconfident CEOs increased employee commitment. [Kramer and Liao \(2016\)](#) examined how the CEOs individual characteristic of overconfidence affected the reporting of analysts. The relationship between CEO overconfidence and the likelihood of optimistic analyst reports has significant implications as analyst reports are often used by other market participants, such as investors, to make critical financial decisions. Kramer and Liao found that analyst forecasts of firms led by overconfident CEOs were more likely to be optimistic, meaning that these reports overestimated earnings more often than they underestimated them than those for companies that were not led by overconfident CEOs.

This research proves that the CEO's trait of overconfidence affects certain stakeholders in a way that alters their behavior regarding the company. The research of [Phua et al. \(2018\)](#) shows how the status-enhancement theory can set the foundation for the assumption that CEO overconfidence influences stakeholders. It thus also has relevancy for the exploration of the impact of CEO overconfidence on the stakeholder group of consumers which has not yet been explored similarly.

## 4. From CEOs to Consumers

While no research has examined the impact of the CEO's trait of overconfidence on consumer attitudes before, there is literature which provides models that have been designed to explain the general connection between the traits of brand representatives and consumer attitudes towards products and brands ([McCracken, 1989](#)). In the following, I will focus on the endorsement theory by McCracken as the theory is widely accepted in the celebrity endorsement literature (e.g., [Scheidt et al., 2018](#); [Tripp et al., 1994](#)). It can help explain why CEO personality traits such as overconfidence impact consumer attitudes.

### 4.1. Endorsement Theory

[McCracken \(1989\)](#) approach to endorsement is based on the idea that the endorsement process itself depends on the "symbolic properties of the (...) endorser". Consequently, he designed a meaning transfer model which explains the process by which brand endorsement influences consumer attitudes. The model suggests that the meanings people associate with a person, who represents the brand, influence consumer attitudes towards products and brands. While McCracken created his endorsement theory in the context of brand endorsement by celebrities, a CEO may also be understood as the endorser of a company. Thus, this theory can set the foundation for the research on CEO endorsement. [Scheidt et al. \(2018\)](#) connected McCracken's meaning

transfer model and CEO endorsement to show that specific attributes of CEOs will transfer meaning to a brand. They found a connection between CEO attributes and consumer attitudes which can explain the impact of CEO overconfidence on the product and brand attitudes of consumers.

#### 4.1.1. Meaning Transfer Model

Brand endorsement has been shown to alter people's brand attitudes and their behavior towards a brand (Amos et al., 2008). This phenomenon can be explained by the meaning transfer model which describes the process by which meaning is transferred from a person to a product and then from the product to the consumer (McCracken, 1989). In the first step of the process, advertisers find a person that carries a cultural meaning which they want their product to represent. This meaning is then transferred from the person to the product by means of advertising. The consumer then transfers the meaning from the product to his or her personal life by finding a personal interpretation of the meanings which reside in the product. A mechanism which consumers may use to do so is by assigning personality traits to brands (Aaker, 1997). The brand personality that they identify then serves as a symbol or in a self-expressive way (Keller, 1993). This model shows that there is a direct link between consumer associations with a brand endorser and their attitudes towards a product or brand.

#### 4.1.2. Celebrity Endorsement

McCracken specifically focusses on celebrity endorsement, and examines how the representation of brands through celebrities affects consumers (McCracken 1989). Celebrities form a strong set of meanings around their personality due to the role they have in society and the way they are portrayed through different media outlets. McCracken argues that endorsement through celebrities instead of anonymous models can improve the meaning transfer process, as they deliver precise demographic information and other meanings that, for example, may be related to their lifestyle due to their publicity. People will associate information such as age, gender, and cultural background and more vague meanings, for example, a humble personality, with celebrities. Characteristics such as trustworthiness, expertise, and attractiveness have shown the strongest positive impact on the effectiveness of celebrity endorsement measured in purchase intention and brand attitudes of customers (Amos et al., 2008).

The meaning transfer model by McCracken (1989) is characterized by a high degree of complexity as the various meanings which are associated with celebrities are taken into account. However, a limitation of the model is its restriction of defining conventional advertising as the only channel of meaning transfer. O'Reilly criticizes that this is not the only mean by which meanings are projected onto products (O'Reilly, 2005). Today, celebrities can adopt strategies of personal media management as they have access to a variety of new channels of communication to promote themselves.

#### 4.1.3. CEO Endorsement

Celebrity status is no longer only held by famous actors and musicians (Moulard et al., 2015). Today also other groups of people have reached celebrity status including influential business personalities such as CEOs. Thus, the meaning transfer model can also find application in the context of CEO endorsement. According to the meaning transfer theory by McCracken (1989), the endorser's personal traits and the meanings consumers attach to the endorser's personal brand may influence consumer brand and product perception.

The study by Scheidt et al. (2018) observed which CEO attributes provided meaning transfer to brands. They found that one of the CEO traits that predicted CEO endorsement was the trait of credibility. This finding supports the prevailing consensus in the literature that source credibility plays an essential role in the effectiveness of celebrity endorsement (Ohanian, 1990). Through the meaning transfer process, the perceived CEO credibility will influence the perceived brand credibility. In the context of brand credibility, Erdem and Swait (2004, p. 192) defined credibility as the "believability of the product information contained in a brand, which requires that consumers perceive that the brand ha[s] the ability (...) and willingness (...) to continuously deliver what has been promised". A study by them showed that brand credibility affects consumers' brand consideration and brand choice. Additionally, they explored the impact of the two subdimensions, trustworthiness and expertise, of credibility. They found that trustworthiness, in general, has a stronger effect on consideration and choice than expertise. This finding stresses the importance of brand trustworthiness which will be discussed in more detail later on.

#### 4.2. Brand Personality

McCracken (1989) explains that there are different ways by which consumers can transfer the meaning from a brand or product to their personal life. One concept which can explain this phenomenon is brand personality. The literature shows that consumers humanize and personalize brands which facilitates the formation of a bond between them (Fournier, 1998). Consumers do so by assigning personality qualities to them (Aaker, 1997). Aaker defines brand personality as "the human characteristics that consumers attribute to brands" (p. 347). It helps consumers to express their perception of the brand and is measured along five dimensions: sincerity, excitement, competence, sophistication, and ruggedness. In the context of CEO overconfidence and the status-enhancement theory, the brand personality trait of competence has a key position. The trait measures how well people believe characteristics such as reliability, intelligence, and success describe the brand.

Consumer's association of personality traits with a brand is directly influenced by the people who are associated with the company (McCracken, 1989). Fournier (1998, p. 345) discussed the idea that "Brands are possessed by the spirit of a past or present other". These people that consumers associate with the brand may be models which are part of the user

imagery portrayed in the brand's advertising (Keller, 1993), celebrities that act as brand endorsers (McCracken, 1989), the employees and the CEO of a company (Aaker, 1997). Therefore, it is possible that people project personality traits that they associate with the CEO of a company onto the brand itself.

As previously discussed, the opinions in literature diverge on the question of whether exposed overconfidence leads to an overall higher or lower evaluation of an individual's level of competence. Therefore, it is unclear, whether exposed CEO overconfidence will lead to a higher or lower perception of CEO competence by consumers. The endorsement theory by McCracken (1989) suggests that the lower or higher evaluation of the CEO's competence will influence consumers propensity to mentally attach personality traits related to competence with the brand.

## 5. Consumer Attitudes

The attitude theory by Fischbein and Ajzen (1975) suggests that a person's beliefs are the only mediators of attitudes. Therefore, the core proposition of the model is that beliefs cause attitudes. Fishbein and Ajzen define a belief as an association between two separate concepts. It may be formed based on direct observations or information received by outside sources. For example, according to the theory of Fishbein, a belief may be formed based on a report in a digital news outlet which suggests an association between a CEO and the trait overconfidence. The meanings a consumer associates with a particular CEO will depend on his or her beliefs about the CEO.

The endorsement theory explains the process by which consumer attitudes towards a product or brand are affected by the associations that consumers have with individuals who represent the company. As a consequence of this connection, consumer attitudes will alter depending on the associations a consumer has with an overconfident CEO compared to a CEO which he does not evaluate as overconfident. Consumer attitudes will impact a variety of consumer decisions and behaviors (Cacioppo et al., 1997). For this reason, consumer attitudes have played a central role in the consumer and social psychology research. In the context of brands, consumer attitudes are understood as the degree of positivity or negativity towards a brand (Whan Park et al., 2010).

The combination of the attitude theory (Fischbein and Ajzen, 1975) and the endorsement theory (McCracken, 1989) can help explain the connection between consumers beliefs about a CEO and their attitudes concerning the company's products and brand. As exposed CEO overconfidence will alter the consumer's beliefs about the CEO, this will alter the meanings he or she attaches to the CEO. Consequently, the consumer's attitude towards the products and brand will be affected. In the following, I will explore why especially a consumer's trust in a brand may be affected by exposed overconfidence. A positive or negative impact on consumer trust has important implications as literature has found that

higher brand trust and brand affect will increase brand loyalty which in turn improves the performance of a company.

### 5.1. Brand Trust

Moorman et al. (1992, p. 315) defined trust as "a willingness to rely on an exchange partner in whom one has confidence". They believe that for trust to be present two conditions must be fulfilled. Firstly, a person must have the belief of trust, meaning they believe the other person to be trustworthy due to their characteristics such as reliability. Secondly, they have a behavioral intent to trust the other person in situations of uncertainty in which they rely on the trustee. Literature shows that trust can influence the quality of a relationship as it, for example, improves the communication between parties (Mohr and Nevin, 1990).

Based on the definition by Moorman et al. (1992), Chaudhuri and Holbrook (2001) defined brand trust as the willingness of a person to trust in the brand's ability to deliver what it promises. The critical role that customer trust, which is similar to brand trust, plays in building strong consumer-firm relationships is widely recognized in literature (Sirdeshmuck et al., 2002). Sirdeshmuck, Singh, and Sabol explored the causes of customer trust and created a model of trustworthy behaviors and practices in the context of businesses which operate in experience services, such as retail companies. Specifically, they tested their model using firms from the clothing retail industry and airline travel industry. They defined three dimensions of trustworthy behavior which increase customer trust regarding management policies and practices: operational competence, operational benevolence, and problem-solving orientation.

#### *Operational competence*

The authors summarize the findings of different streams of research which suggest that a consumer develops trust when his or her expectation of a partner's "consistently competent performance" is fulfilled (Sirdeshmuck et al., 2002, p. 17). Based on this research, as well as empirical support for the influence of different dimensions of competence on trust, the authors include the dimension of operational competence in their model of trustworthiness. They define operational competence as the observable behavior that indicates the competence of a person. They focus on observable behavior and not on true competence that resides within a person, as a customer will only judge competence based on indicators that are visible to him or her.

#### *Operational benevolence*

The second dimension is characterized by behavior that indicates the manager's motivation to place the customer-interests ahead of their interests. Here again, the authors stress that the intention has to be observable to customers. Morgan and Hunt (1994) suggest that opportunistic behavior will decrease the other party's trust. Consequently, benevolent behavior, which entails that a person refrains from opportunistically self-serving actions, promotes trust in a relationship (Sirdeshmuck et al., 2002).

### *Problem-solving orientation*

The final dimension, problem-solving orientation, is defined as the consumer's perception of the management's motivation to identify and solve problems that occur during a service exchange (Sirdeshmuck et al., 2002). The authors were able to show that the problem-solving behavior is distinct from benevolence and competence as it requires different cognitive judgements. This dimension is specific to service providers.

Even though Sirdeshmuck et al. (2002) designed their model in the context of consumer-service provider exchanges I believe that especially the first dimension can also be transferred to other fields of consumer trust. The dimension of competence orientation can explain the increase or deterioration of consumer trust as a consequence of exposed CEO overconfidence. It also creates an essential link between credibility and competence. As previously described trustworthiness is defined as a dimension of credibility (Erdem and Swait, 2004) while operational competence is one of the dimensions of trustworthy behavior (Sirdeshmuck et al., 2002). It is thus likely that if perceived competence is affected by exposed overconfidence, perceived credibility will be affected in the same way. Figure 1 visualizes the relationship of the two concepts.

### 5.2. Brand Affect

Chaudhuri and Holbrook (2001, p. 82) define brand affect as the ability of a brand to trigger "a positive emotional response in the average consumer as a result of its use". They believe that while the process of brand trust development is more calculated and carefully considered, the development of brand affect is more spontaneous and less controlled. Dick and Basu (1994) discuss four types of affect that they believe can precede and cause loyalty. They differentiate between emotions, moods, primary affect, and satisfaction. These factors influence decision making and may be better predictors of an individual's behavior than cognitive evaluations.

### 5.3. Brand Loyalty

Chaudhuri and Holbrook (2001) propose that a consumer's brand trust and positive brand affect are the critical determinants of his brand loyalty. They introduce the concepts of purchase loyalty and attitudinal loyalty which are two separate types of brand loyalty. Chaudhuri and Holbrook define purchase loyalty as the willingness of consumers to repurchase products from a brand and attitudinal loyalty as the commitment of consumers to a brand. They found that both brand trust and brand affect independently influence both types of brand loyalty. Brand trust is related to brand loyalty as it fosters an exchange relationship which is sustained by the loyalty of customers (Morgan and Hunt, 1994). Dick and Basu (1994) suggest that brand affect in the form of positive emotions, mood and affect is linked to higher levels of brand loyalty. Both purchase and attitudinal loyalty influence a company's performance (Chaudhuri and Holbrook, 2001).

Purchase loyalty will lead to a higher market share as a result of the increased number of repurchases by loyal consumers (Chaudhuri and Holbrook, 2001). However, this may not be the case for every product category and thus may also depend on the segment in which the company operates. Chaudhuri and Holbrook suggest that attitudinal loyalty increases the price consumers are willing to pay for a product compared to the price for products of competitors. This assumption draws from the theory of brand equity which Aaker (1991, p. 15) defines as "a set of brand assets and liabilities linked to a brand, its name, and symbol, that add or detract from the value provided by a product or service to a firm". He defined brand loyalty as one of five classes of brand assets that determine brand equity. Keller suggests that consumers will have a higher willingness to pay a premium price for products if they have a positive attitude towards the brand (Keller, 1993).

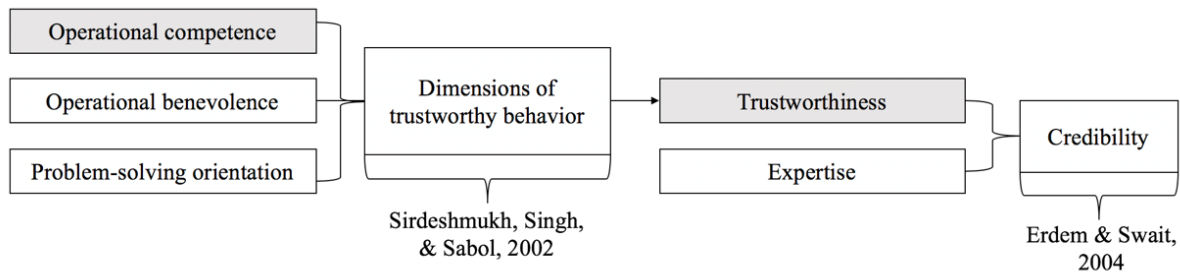
A higher market share and the ability to charge a higher relative price will have a positive effect on the profitability of the company. Therefore, brand loyalty is a connector that links positive brand trust and affect to improved brand performance outcomes. Literature has also linked brand loyalty to marketing advantages such as positive word of mouth effects, decreasing behavior of information searching for alternative brands and stronger resistance to counter-persuasion strategies by competitors (Dick and Basu, 1994).

## 6. Study

In the following, I will describe the study I conducted to observe the impact of CEO overconfidence on consumer attitudes. The study consists of two parts which I will analyze separately. The first part consists of an experiment with an overconfidence manipulation and the second part considers participants' perceptions of CEO overconfidence and their attitudes towards the brands of well-known companies. Both parts are designed to test my main hypothesis that exposed CEO overconfidence will negatively affect consumer attitudes towards products and brands.

### 6.1. Part One – CEO Overconfidence Manipulation

The first part of the study addresses whether participants' evaluation of a product and brand are affected by the information that the CEO has been overconfident. Considering the different types of overconfidence found in the literature (Moore and Healy, 2008; Tenney et al., 2019), the experiment specifically examines the impact of verbally communicated overestimation. The participants were divided into two groups of which one received a report which revealed that the CEO had made a verbal statement in which he overestimated the performance of his company. As I want to observe how this type of exposed CEO overconfidence affects consumers, I focus on the impact it has on a few specific consumer attitudes. I observed the impact on the participant's general attitude towards the product, general attitude towards the brand and his or her trust in a brand (see Figure 2). I chose



**Figure 1:** Connecting the Concepts of Competence and Credibility. Concept of the dimensions of trustworthy behavior from “Consumer Trust, Value, and Loyalty in Relational Exchanges,” by D. Sirdeshmukh, J. Singh, and B. Sabol, *Journal of Marketing*, 66. Concept of credibility from “Brand Credibility, Brand Consideration, and Choice,” by T. Erdem and J. Swait, *Journal of Consumer Research*, 31. Source: Own figure

these attitudes based on their influence on business performance or their connection to CEO overconfidence.

My first hypothesis is based on the endorsement theory (McCracken, 1989) and the study by Tenney et al. (2019) which provided evidence of the status deteriorating effect of verbally communicated overconfidence. I believe that the group which receives the revealed CEO overconfidence condition (G1) will evaluate the CEO as less competent. Thus, they will project this negative image onto the product and brand. I expect that participants from G1 will on average give a lower evaluation of the product, lower brand attitude scores and have lower trust in the brand.

H1: Participants from G1 will on average give a lower product evaluation, lower brand attitude scores and have lower trust in the brand. As I expect that participants of G1 will evaluate the CEO as less competent, my second hypothesis is that exposed CEO overconfidence will impact the perception of brand personality traits (see Figure 2). Specifically, I observe the impact on the brand personality trait competence from the five-factor model by Aaker (1997). I expect that the participants of G1 will rate the personality trait competence to be less descriptive of the brand.

H2: Participants from G1 will evaluate the personality trait of competence to be less descriptive of the brand. My third hypothesis is that certain personality traits and demographic factors of the individual who is judging the product and brand will influence the degree to which overconfidence affects his or her attitudes regarding the product and brand (see Figure 2). Specifically, I am interested in the mediating effect of ambition on the impact of exposed overconfidence on consumer attitudes. Additionally, I observe whether factors such as gender and educational or professional background are related to the perception of overconfidence.

H3: The personality traits of extraversion and

ambition as well as demographic factors as gender and educational or professional background in business will influence the impact exposed overconfidence has on an individual's attitudes towards a product and brand.

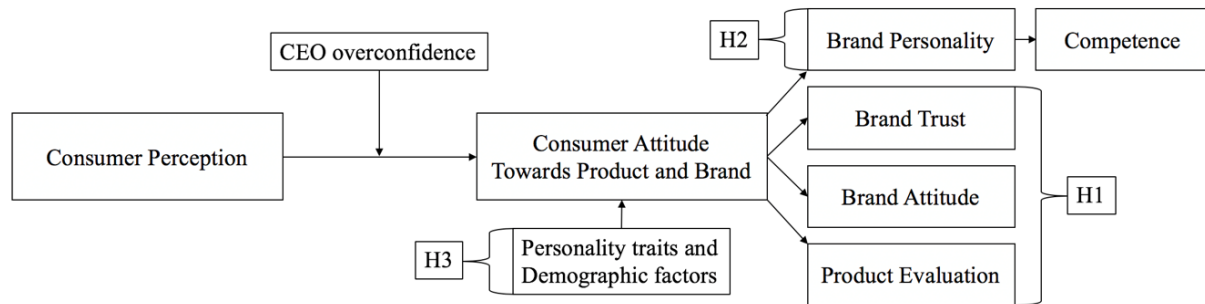
#### 6.1.1. Method

##### *Participants*

My sample consisted of 120 participants aged 18 to 65 with an average age of 29 years old. 42% of participants were female and 58% male. Sixty percent of participants had completed or were pursuing a bachelor's degree while 5% had only completed their high school diploma and the rest had either completed or were pursuing a postgraduate degree. Eighty-three percent of participants had an educational or professional background in business. The participants of the survey were randomly assigned to one out of two groups by the survey software. All participants took part in the first and second part of the study. I divided the participants into two experimental groups. Group one which was exposed to the overconfidence condition and consisted of 55 participants (23 women, 32 men, Mage=27.9, age range: 18-59) and the second group which was the control group consisted of 65 participants (27 women, 38 men, Mage=28.9, age range: 18-65).

##### *Procedure*

To test the hypothesis that revealed CEO overconfidence would lead to a more negative product and brand evaluations, I gave the two groups the same information about a fictional phone and the producing company but differed the information regarding the CEO of the company (see Appendix A). Both groups received all of the information in text form, to imitate a verbally communicated overconfident statement. The information about the product outlined a premium segment phone with modern features similar to the current models from market-leading companies such as Apple and Samsung. The more general information about



**Figure 2:** Summary and Visual Presentation of the Main Hypotheses. Source: Own figure

the company revealed that it supports ethical and sustainable practices and that it is a desired employer. Both groups received confident statements from the CEO regarding the company's future profitability margins, a product launch unrelated to the phone, and a sales target. The two groups received different information concerning how accurate these statements by the CEO had been (see Appendix B). One group received the information that the announcements had been overconfident, and that the company did not perform in the way the CEO had proclaimed (G1). The statements were formulated in ways that indicated the CEO had overestimated his company's ability to succeed, and that the cause of the discrepancies was not merely regular fluctuations that occur in business forecasts. The other group received the information that the company's performance was in line with the CEO's announcements (G2). Following this information, the participants answered several questions regarding their attitude towards the phone and the brand of the company (see Appendix C). The questions were selected to address the three hypotheses about how exposed verbally communicated CEO overconfidence influences consumers' product evaluation, brand attitude, brand trust and the perception of the brand personality trait competence.

#### Product evaluation

To assess the impact of overconfidence on a participant's product evaluation he or she had to indicate their overall evaluation of the product. The participants had to rate their product attitude on a five-item, seven-point semantic differential measure of aspects related to the quality of the product. The items used were based on a scale developed for the product evaluation of high-tech goods (Roehm and Sternthal, 2001). As the items correlated with each other ( $\alpha .85$ ), I combined them in a measure of the general evaluation of the products ( $M = 5.5$ ,  $SD = .95$ ). The purchase intention was measured using a three-item, seven-point Likert-type scale to assess the degree to which participants intended to buy the product in the future (Baker and Churchill Jr, 1977). The items were again correlated ( $\alpha .87$ ) and were thus aggregated to a general measure of purchase intention ( $M = 4.4$ ,  $SD = 1.4$ ). The participants were also asked to indicate their willingness to pay for the phone. This is a single item measure that may be influenced by factors (age, in-

come) independently from the consumers attitudes towards the product and brand. Thus, the explanatory power of this measure may be limited. Still, it can be an indicator for the participant's overall evaluation of the product.

#### Brand attitude

Participants were then asked about their general attitude towards the brand using a five-item, seven-point semantic differential measure (Ahluwalia et al., 2000). These items showed a strong correlation with each other ( $\alpha .91$ ) and were combined as a measure of the general attitude towards the brand ( $M = 5.4$ ,  $SD = .96$ ). Participants gave their assessment of the trustworthiness of the brand by answering a four-item, seven-point Likert-type scale based on a measurement model introduced by Erdem and Swait (2004). These were again aggregated to a measure of brand trust ( $M = 4.5$ ,  $SD = 1.2$ ) as they showed high correlation ( $\alpha .91$ ). Lastly, participants indicated to which degree they believed the brand could be described using different personality-like characteristics related to competence. The items used in this step were based on the scale for the brand personality trait of competence developed by Aaker (1997). Participants rated to which extent they believed the specific characteristics were descriptive of the brand on a scale of one (Not at all descriptive) to seven (Extremely descriptive). As these were strongly correlated ( $\alpha .88$ ), they were combined to form a measure of the brand personality trait competence ( $M = 4.9$ ,  $SD = 1.1$ ).

#### Participant personality traits

In addition to demographic information such as age, gender and whether they came from an educational or professional business background, participants were also asked to evaluate themselves concerning the character trait of ambition. They had to indicate the degree to which they believed certain statements suited their personality on seven-point Likert-type scales. The statements which measured the level of ambition were based on the ambition scale by Duckworth and her colleagues (Duckworth et al., 2007). Here again, I aggregated the statements in a single measurement of ambition ( $M = 5.4$ ,  $SD = .84$ ) as they showed correlation ( $\alpha .74$ ).



### 6.1.2. Results

The results from the analysis of the first part of the survey support the hypothesis that verbally communicated overconfidence influences consumer's product and brand perception. All five areas of consumer attitudes: overall product evaluation, purchase intention, brand attitude, brand trust and willingness to pay, showed a significant negative relationship with CEO overconfidence (see Appendix D). Participants were also less likely to rate characteristics related to the personality trait competence as descriptive for the brand.

#### *Product evaluation*

As expected, exposed CEO overconfidence predicted lower product evaluation ( $\beta = -.52$ ,  $t(118) = -3.1$ ,  $p < .01$ ), lower intention to purchase ( $\beta = -.67$ ,  $t(118) = -2.65$ ,  $p < .01$ ), and lower willingness to pay ( $\beta = -.91$ ,  $t(118) = -2.59$ ,  $p < .05$ ). As shown by Table 1, participants of G1 gave on average a lower evaluation of the product compared to participants of G2. Table 1 also shows that they indicated a lower purchasing intention and an overall lower willingness to pay for the phone.

**Brand attitude.** The results showed that exposed CEO overconfidence led to a more negative attitude towards the brand ( $\beta = -.84$ ,  $t(118) = -5.27$ ,  $p < .001$ ). As a consequence, the average attitude towards the brand of participants of G1 was lower than that of G2 participants (see Table 1). The negative effect of exposed CEO overconfidence on brand trust was even stronger ( $\beta = -1.11$ ,  $t(118) = -5.69$ ,  $p < .001$ ). This result indicated that CEO overconfidence affects the perception of specific attributes of the brand. This finding is further supported by the result that exposed CEO overconfidence predicts lower scores on the perceived brand personality trait competence ( $\beta = -.73$ ,  $t(118) = -3.81$ ,  $p < .001$ ). Table 1 shows that participants of G1 on average indicated that they believed characteristics related to competence to be less descriptive of the brand than G2.

#### *Personality traits and demographics*

The general personality trait measures for ambition showed no significant impact on the different variables which measured participants' product evaluation and attitude towards the brand. However, the rating on the specific item "I am ambitious" from the ambition scale showed significant interaction effects on all four measures of product and brand evaluation (product evaluation:  $\beta = .29$ ,  $t(116) = 1.72$ ,  $p < .1$ ; brand trust:  $\beta = .46$ ,  $t(116) = 2.34$ ,  $p < .05$ ; brand attitude:  $\beta = .27$ ,  $t(116) = 1.7$ ,  $p < .1$ ; brand personality:  $\beta = .4$ ,  $t(116) = 2.08$ ,  $p < .05$ ). In all cases the interaction effect was positive, meaning that if the CEO was verbally overconfident about the company's performance a higher evaluation of personal ambition decreased the negative effect that the overconfidence condition generally had on the attitudes towards the product and brand. This finding suggests that people who consider themselves to be ambitious are less sensitive to the negative impact that exposed CEO overconfidence has on product and brand evaluation.

When considering the interaction effect of the independent variable of especially high ambition (one standard deviation higher than the mean) the main effect of the overconfidence condition on the evaluation of the product and the perception of the brand personality trait competence is no longer significant for brand personality and product evaluation (see Table 2).

The results showed that a difference in gender did not impact the evaluation of any of the aspects measuring product evaluation and attitude towards the brand. Whether participants had an academic or professional background in business was also irrelevant for the evaluation along the previously mentioned brand attitude dimensions. The only area in which it showed a significant impact was the general product evaluation ( $\beta = -1.15$ ,  $t(116) = -2.33$ ,  $p < .05$ ). As the interaction term is negative, this indicates that if the CEO is verbally overconfident about performance a professional or educational background in business will increase the negative effect that the overconfidence condition has on product evaluations.

### 6.1.3. Discussion

In the following, I will discuss the implications of the results regarding the hypotheses about the influence of revealed CEO overconfidence on consumer attitudes towards products and brands. I will examine how exposed CEO overconfidence influences consumer attitudes and what the resulting implications for managers are. I will also examine how the results are linked to the existing research on overconfidence and consumer research. I base my analysis on the results of the experiment as well as the literature previously discussed.

The connection between CEO overconfidence and consumers. The results show that exposed CEO overconfidence will influence the attitudes a consumer forms about a product and brand. This finding supports the meaning transfer theory by McCracken (1989). The treatment of the two experimental groups only differed with regard to the information the participants received about the accuracy of the CEO's predictions regarding the company's performance. G1 received the information that the CEO had been overconfident about the company's performance, while G2 was informed that his predictions had been accurate. The areas of performance which the CEO commented on were unrelated to the quality of the product. Thus, the participant's evaluation of the product should not have been affected by the CEO's overconfident statements if his or her attitudes were formed only based on his or her beliefs about the product. However, the results showed that participants from the two groups gave significantly different product and brand evaluations which proves that CEO overconfidence influenced their evaluation. This finding confirms the meaning transfer model as it suggests that the meanings a consumer associates with the CEO influence his or her attitudes towards the products and brand. The difference in product and brand evaluations indicates that the meanings associated with the CEO altered between the two groups.

**Table 1:** Summary of the Ratings Regarding the Variables of Consumer Attitudes and a Comparison Between the Experimental Groups

		Group 1	Group 2	Difference
		Overconfidence	No Overconfidence	MG2 – MG1
Product Evaluation	M	5.19	5.71	0.52
	SD	1.04	.79	
Purchase Intention	M	4.04	4.71	0.67
	SD	1.4	1.37	
Willingness to Pay	M	468.55	559.95	91.4
	SD	195.76	190.22	
Brand Attitude	M	4.96	5.79	0.83
	SD	.94	.8	
Brand Trust	M	3.94	5.05	1.11
	SD	1.02	1.12	
Brand Personality Trait Competence	M	4.47	5.2	0.73
	SD	1.03	1.05	

Note. M = mean, SD = standard deviation. All scales were seven-point Likert-type scale except for willingness to pay which was asked for in an integer. Source: Own table

**Table 2:** Interaction Effects of Ambition Regarding the Relationship of CEO Overconfidence and Consumer Attitudes

		Ambition		At the mean		One SD above mean	
		$\beta$	t	$\beta$	t		
Product evaluation		-.52	-3.11**	-.23	-.95(n.s.)		
CEO Overconfidence on	Brand Attitude	-.84	-5.28***	-.56	2.49*		
	Brand Trust	-1.11	-5.76***	-.65	2.37*		
	Brand Personality Trait Competence	-.72	-3.84***	-.32	-1.2(n.s.)		

Note. SD = standard deviation. n=116; n.s.=not significant; (.) p = 0.1, \* p = 0.05, \*\* p = 0.01, \*\*\* p = 0.001

Status-enhancement theory of overconfidence. In addition to participants of the two groups showing significantly different ratings of the product and brand, the results also revealed that ratings from participants of G1 were on average lower on all variables of product and brand evaluation than those of G2 (see Table 1). These results confirm my first hypothesis that participants from G1 will on average give a lower product evaluation, brand attitude scores and have lower trust in the brand. The finding that participants of G1 gave worse product and brand evaluations offers support to the study by Tenney et al. (2019). The results also extend this study by providing proof of this theory in the specific field of CEO overconfidence.

The results implicitly support the findings of Tenney et al. (2019) who showed that exposed verbally communicated overconfidence diminishes social status. This finding is not

directly replicated as participants were not asked to rate the CEO's competence or status. Still, based on the endorsement theory, it is possible to draw a conclusion about the impact of exposed verbally communicated overconfidence on the participant's attitude toward the CEO from the more negative attitudes participants had towards the product and brand. The meaning-transfer model shows how consumers transfer meanings that they associate with a person, who represents a company, to the products and brand of that company (McCracken, 1989). Based on this model, participants in the experiment will transfer meanings they associate with a CEO to the product and brand. Tenney et al. (2019) show that when peers discover the overconfidence they will evaluate the individual as less competent and thus, award them a lower social status. According to this theory participants of G1, who are given the information that the CEO had been overconfident,

will evaluate the CEO as less competent and less deserving of social status. The meanings they associate with the CEO will be more negative than those of participants from G2. This process is illustrated in Figure 3. According to the meaning transfer model, they will transfer the more negative meanings to the product and brand. The current study replicates this process as the results show lower ratings and thus more negative attitudes towards the product and brand from the participants of G1. Thus, the results of the current study support the theory of the social status diminishing effect of exposed verbally communicated overconfidence. The lower evaluation of the product and brand can be explained by a lower evaluation of the CEO's competence.

The study also offers a new point of view on the social consequences of overconfidence as it examines the impact of exposed verbally communicated overconfidence in the context of CEO and customer relationships. Previous studies which dealt with social consequences of overconfidence have mainly examined this effect in the context of teamwork dynamics among peers (Erdem and Swait, 2004; Tenney et al., 2019). As the findings implicitly support the theory that exposed verbally communicated overconfidence has a status diminishing effect, the study proves that this theory can be generalized to other areas of overconfidence. It also extends previous research on CEO overconfidence. It offers an extension to the research on the effects of CEO overconfidence on stakeholder relations.

#### *Brand trust*

The results of the experiment show that exposed CEO overconfidence has the strongest negative impact on the variable of brand trust out of the different product and brand attitude measures (see Table 1, see Appendix D). The customer trust model by Sirdeshmuck et al. (2002) can help to explain the cause of the lower levels of brand trust of participants of G1. When considering the model's three dimensions of trustworthy behavior it becomes apparent that overconfidence may influence the first dimension, operational competence. This dimension encompasses behavior that indicates the company representative's competence. The authors stress that this behavior has to be observable to the public. The accuracy of verbal statements a CEO makes regarding his expectations of the company's performance can be understood as a behavior by which consumers can judge competence. As previously discussed in the context of the status-enhancement theory, this verbally communicated exposed overconfidence leads others to have a lower perception of the individual's level of competence. This finding is supported by the results of the current study. As the CEO who is revealed to be overconfident receives a lower evaluation concerning their operational competence, they also receive a lower level of customer trust than the CEO who accurately predicts the company's performance.

This finding has important managerial implications as consumers trust in a brand will indirectly influence the company's profitability. As described by Erdem and Swait (2004), trustworthiness is one of the dimensions of credibility which is linked to the consumer's brand consideration and brand

choice. Additionally, Chaudhuri and Holbrook (2001) find a direct link between customer trust and brand loyalty which in turn influences the number of repurchases by customers and a higher willingness to pay a premium price for the products of the brand. Thus, brand loyalty does not only determine whether a consumer initiates the relationship with a brand but also how prosperous the relationship is afterwards.

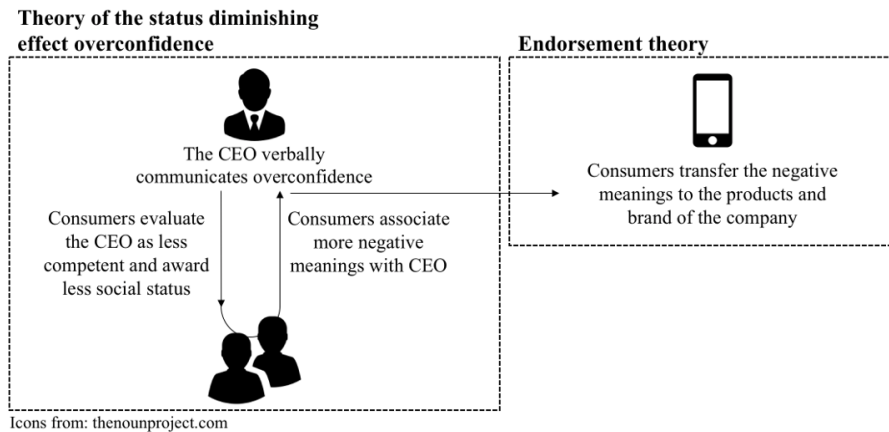
#### *Brand personality*

Participants who were part of G1 indicated that they evaluated the personality trait competence to be less descriptive of the brand than participants from the other group. This finding additionally supports the hypothesis that consumers evaluate the overconfident CEO to be less competent. The literature discusses the phenomenon that brand personality attribution is linked to the people who represent the brand (Keller, 1993). This process suggests that if participants rate the trait competence as less descriptive of the brand, this reflects their perception of the competence of the individual they associate with the brand. As the brand in the experiment was fictional and the only representative of the company that participants were introduced to the CEO, this implies that the perception of the CEO's personality is the basis for participants' evaluation of brand personality. Aaker (1997) already suggested that consumer's associations with a company's CEO directly influences their perception of the brand's personality. The current findings extend this suggestion as the study examined the effect of the specific CEO personality trait of overconfidence on brand personality. It offers deeper insights concerning the connection between specific human personality traits and brand personality.

Impact of participants' level of ambition. Another interesting result was the impact the individual's level of ambition had on his or her attitudes towards the product and brand when exposed to the overconfidence condition. The finding that a higher level of ambition reduces the negative effect that CEO overconfidence has on consumer attitudes suggests that individual characteristics play a role in the perception of overconfidence and the formation of attitudes. The results indicate that someone who is highly ambitious is less likely to punish a CEO's overconfidence by evaluating him or her as less competent. A possible theory to explain this observation may be based on the assumption that highly ambitious people set high goals for themselves similar to overconfident CEOs. As a result, they may relate to the CEO and sympathize when he or she fails to reach overly high goals. If ambitious people are able to form a stronger personal connection with an overconfident CEO, this may positively influence their attitude towards the person. Thus, they also transfer more positive meanings to the product and brand. Further research on the influence of ambition on the perception of overconfidence is necessary to confirm whether this theory can explain the results.

#### *Measure of overconfidence*

The results of the current study raise criticism towards the portfolio-based measure of overconfidence which currently



**Figure 3:** Chain of Effects from CEO Overconfidence to Consumer Attitudes. Connecting the concepts of the status diminishing effect of overconfidence and the endorsement theory. The model refers to the theory of the status diminishing effect of overconfidence from “Is overconfidence a social liability? The effect of verbal versus nonverbal expressions of confidence,” by E. R. Tenney et al., 2019, *Journal of Personality and Social Psychology*, 116. As well as the endorsement theory from “Who Is the Celebrity Endorser? Cultural Foundations of the Endorsement Process,” by G. McCracken, 1989, *Journal of Consumer Research*, 16. Source: Own figure

dominates in the research on CEO overconfidence. Most studies which are concerned with the consequences of CEO overconfidence use the popular portfolio-based measure of managerial overconfidence by Malmendier and Tate (2005). This measure of overconfidence does not distinguish between the channels by which overconfidence is communicated. Therefore, it is not possible to differentiate between CEOs who express overconfidence verbally and those who express it non-verbally. The previously mentioned studies which have examined the influence of CEO overconfidence on the company’s stakeholder relationships also used the portfolio-based measure, and all found that CEO overconfidence increased stakeholder commitment (Phua et al., 2018). These results may indicate that CEO overconfidence positively influences stakeholder’s attitudes. However, another possibility is that these results are biased by the measure of overconfidence used. Especially when assessing the social consequences of CEO overconfidence, the measure by Malmendier and Tate may not be suitable as stakeholders will judge manager’s confidence and whether it is justified by their ability based on a CEO’s public appearances and announcements, and not the CEO’s personal portfolio in company stocks and options. The study did not observe the impact of non-verbally communicated overconfidence, still, the results of Tenney et al. (2019) suggest that the distinction between the channels of communication may be necessary when evaluating the impact of CEO overconfidence on other stakeholders. Thus, a different measure to the one of Malmendier and Tate (2005) is needed.

#### Limitations

While the experimental design allowed me to accurately observe the impact of exposed CEO overconfidence on consumer attitudes regarding the product and brand it also held

some limitations. One major limitation was the laboratory like design of the experiment with an unfamiliar CEO, brand, and product. Due to this setup and the limited time and space to provide information about the company the way participants formed attitudes about the brand and the CEO did not perfectly replicate how this process takes place in reality. There may be significant structural discrepancies between the results of the current study and those that would be found in reality (Laurent, 2013). Laurent argues that this is the case as the experiment observes participant’s “short-run answers to short-run stimuli” (p. 323), in contrast, consumers typically develop decision-making strategies over a longer timeframe and based on multiple, repeated interactions with a stimulus.

A second limitation of the experiment is its focus on a single company in a specific industry. The survey only examined the effect of CEO overconfidence on consumer attitudes in the context of smartphone products. Therefore, the survey has limited generalizability to other companies that operate in different product segments such as commodities or textiles.

#### Further research

This study has brought together the theory of the social consequences of overconfidence and the endorsement theory in the context of CEO overconfidence. It has provided a new perspective on the process of consumer attitude formation. The results may be a starting point for further research in the area of CEO confidence and its impact on consumers.

An important area of further research is the impact of revealed non-verbally communicated CEO overconfidence on consumer attitudes. This study would be able to draw further conclusions about the generalizability of the findings of Tenney et al. (2019), and add to the general discussion in

the literature regarding the consequences of overconfidence for social status. The study would also have important managerial implications and CEOs and executives may use the results as a guide whether they should overall avoid all expression of overconfidence or whether they may benefit from non-verbally expressed overconfidence.

Another interesting area of further research may be the effects of CEO underconfidence on consumer product and brand attitudes. Many instances of underconfidence have been noted in literature where again a separation in underestimation, underplacement and underprecision is possible (Moore and Healy, 2008). The concept of managerial underconfidence is mostly unexplored, therefore, it should first be observed whether underconfident CEOs are found in reality. The measure by Malmendier and Tate (2005) could be used in making this assessment. While it is not well suited as a measure of confidence in assessing the relationship between CEO confidence and consumer attitudes it can be used as a general measure to indicate the level of confidence a CEO has in his company. In contrast to when it is used to determine overconfident CEOs, when looking for underconfident CEOs, it should be used to determine those executives who hold an under proportional amount of company options in their personal portfolio. If it is found that underconfidence affects some CEOs, one possible area of additional studies is the impact of underconfidence on social status, similar to the research concerning the status-enhancing effect of overconfidence (e.g., Anderson et al., 2012b; Kennedy et al., 2013; Tenney et al., 2019). The current study could be used as a starting point and adjusted so that it explores how CEO underconfidence impacts consumers' product and brand attitudes.

As the study has been able to show that CEO traits influence consumer attitudes, it would be interesting to observe whether and in what way other CEO personality traits influence consumer attitudes. It is possible that in contrast to the impact of CEO overconfidence other traits may improve consumer attitudes towards products and the brand. If this was the case, the finding would have important implications for CEOs and executives as they could use the results as a guide for their public communication strategy. It may indicate what personal traits a CEO should highlight in his or her communication and which, such as overconfidence, he or she should avoid.

To determine the generalizability of the results, the experiment should be replicated with different product categories. Using other products than premium segment smartphones may support or contradict the finding that CEO overconfidence leads to lower product and brand evaluations of customers. For example, further research could observe the impact of CEO overconfidence in the context of commodity products or impulse buys.

## 6.2. Part Two - Perceived CEO Overconfidence and Brand Attitudes

One major limitation of the first part of the survey is its partly unrealistic setup. This limitation is caused by the ex-

perimental design being based on a fictional company and conducted in the form of an online survey. To support the findings of the first part of the survey I want to observe the relationship between consumer perceived CEO overconfidence and consumer attitudes towards the brand in a more realistic way. To do so, I used real companies and CEOs in the second part of the survey. I asked participants about their perception of five well-known CEO's level of overconfidence. Specifically, I chose to measure perceived overestimation of the personal and the company's ability to succeed. Then participants were asked about their general attitude and trust towards the companies those CEOs currently lead. In line with my hypotheses concerning the first part of the survey, I expect that participants who give higher ratings of perceived CEO overconfidence will give lower ratings of brand attitude and brand trust.

H4: Participants who perceive a CEO to be more overconfident will have a lower attitude towards the brand and lower trust in the brand.

### 6.2.1. Method

#### *Procedure*

I selected the five CEOs, Mark Zuckerberg, Elon Musk, Jeff Bezos, Satya Nadella, and Tim Cook, based on the level of public awareness, their strong profile, and the expected divergence in the perceived overconfidence by consumers. First participants were asked with which of the CEOs they were familiar so that in the following they only received questions about the CEOs they had claimed to know. The participants had to assess whether and to what degree they believed them to be overconfident and then had to indicate their general attitude towards the brand and their trust in the brand (see Appendix E). In the analysis of the results, I only included the answers from participants of G2, the group which had not been exposed to the CEO overconfidence manipulation in the first part of the study. As the participants of the two groups had previously shown substantial differences in decisions, separating them in the analysis of the results collected in the second part increases the validity of the data. I chose G2 as these participants had not been exposed to overconfidence before. Consequently, I assume them to be less biased than the other group. While I will only refer to the analysis of the data collected from participants of G2 in the following, the results from G1 and the whole sample size show similar patterns that support the findings (see Appendix F).

The participants indicated to which degree they believed the CEO to be overconfident on a four-item, seven-point Likert type scale. The items were selected to specifically observe to what extent participants believed the CEOs overestimated their own and their company's performance. They were formulated based on the definition of overestimation given by Moore and Healy (2008). I chose to observe perceived overestimation as this was the type of overconfidence observed in the first part of the study. For example, one of the items was: "The CEO overestimates his company's success". Participants

had to indicate to which degree they agreed with this statement. As they could give an answer in the range from one (Highly disagree) to seven (Highly agree) any score above the median of four would indicate that the participant perceives the CEO to be overconfident. As the items showed significant correlation ( $\alpha_B .85$ ,  $\alpha_Z .71$ ,  $\alpha_M .88$ ,  $\alpha_C .91$ ,  $\alpha_N .91$ ) I aggregated them into a measure of perceived CEO overconfidence ( $M_B = 2.9$ ,  $SD_B = 1.2$ ,  $M_Z = 4$ ,  $SD_Z = 1$ ,  $M_M = 5.4$ ,  $SD_M = 1.3$ ,  $M_C = 3.3$ ,  $SD_C = 1.3$ ,  $M_N = 3$ ,  $SD_N = 1.2$ ). Afterwards, participants had to indicate their attitude towards the brands on the same five-item, seven-point semantic differential scales previously used in the first part of the survey (Ahluwalia et al., 2000). This measure was aggregated in a single measure of the general attitude towards the brand. Additionally, they gave their level of trust in the brand. The four items used for this assessment were again based on the model by Erdem and Swait (2004). The items were aggregated into a single measure of brand trust.

### 6.2.2. Results

#### *CEO overconfidence*

The range of the average perceived CEO overconfidence lay between 2.9 and 5.4 (see Table 3). The results show that only one of the CEOs was on average evaluated as overconfident by the participants (Elon Musk perceived overconfidence:  $M=5.4 > 4$ ). The means of brand attitudes lay between 3.59 (Facebook) and 5.69 (Apple). The means of brand trust lay between 3.38 (Facebook) and 5.77 (Amazon). For the three companies of which the CEOs were perceived as most overconfident, the brand attitude was higher than the brand trust. The opposite was the case for the two brands of which the CEOs were perceived to be the least overconfident (see Table 3).

#### *Brand attitude*

here is no absolute relationship between perceived CEO overconfidence and brand attitude or brand trust (see Table 3). However, the results show that for individual companies the participant's attitude toward a brand is lower when they perceive the CEO to be more overconfident (see Appendix G). However, this relationship is only significant in the case of two out of the five observed brands (Facebook:  $\beta = -.65$ ,  $t(62) = -4.47$ ,  $p < .001$ ; Tesla:  $\beta = -.27$ ,  $t(60) = -2.21$ ,  $p < .05$ ). These two brands were also the two in which participants perceived the CEOs as the most overconfident out of the group of five (see Table 3).

**Brand trust.** The results show that higher perceived CEO overconfidence is associated with lower ratings of brand trust (see Appendix H). This effect was significant for three out of the five brands (Facebook:  $\beta = -.6$ ,  $t(62) = -4.6$ ,  $p < .001$ ; Tesla:  $\beta = -.57$ ,  $t(60) = -5.29$ ,  $p < .001$ ; Apple:  $\beta = -.31$ ,  $t(59) = -3.32$ ,  $p < .01$ ). This result indicates that the impact of higher perceived CEO overconfidence was stronger on brand trust than on the general attitude towards the brand. Again, the three companies for which this effect could be observed, where the three companies of which the CEOs were

perceived as the most overconfident out of the five.

#### *CEO familiarity*

The results indicated a slight trend of higher familiarity with the CEOs who are perceived to be more overconfident. However, the results do not show a linear relationship (see Appendix I). For example, more than twice as many people were familiar with Jeff Bezos, the CEO of Amazon, than with Satya Nadella, the CEO of Microsoft, who participants rated as more overconfident than Bezos.

### 6.2.3. Discussion

#### *Measurement of perceived overconfidence*

I believe that the measurement of perceived overconfidence offers a new perspective to the literature on overconfidence. This measure does not aim to observe the CEO's actual level of overconfidence as previous measures have tried to do (e.g., Malmendier and Tate, 2005). Instead, it focusses on the consumer's view of overconfidence. This measure has two significant advantages. First, it is a straightforward measure that does not require any in-depth analysis or difficult to access information such as the portfolio-based measure by Malmendier and Tate (2005). Secondly, individual perception plays a central role in the formation of attitudes and thus this measure is more suitable for the cause of investigating the relationship between CEO overconfidence and consumer attitudes. Fischbein and Ajzen (1975) suggests that people form attitudes based on the beliefs they hold about a person. They explain that beliefs are frequently formed based on information provided by an outside source, for example, newspapers, television, and radio. Today, other critical outside sources of information are online news outlets and social media. Thus, a consumer will form his or her beliefs about the CEO based on reporting of his or her statements, performance, and behavior in the media. Based on the study by Tenney et al. (2019) I assume that if people come to believe that an individual is overconfident, no matter whether this is true or not, they will evaluate him or her as less competent. This evaluation will negatively influence the meanings they associate with the CEO which according to the endorsement theory will then negatively affect their attitudes towards the products or brand. As the perception of overconfidence and actual overconfidence may diverge, the measure of perceived overconfidence is, therefore, more relevant in the study of consumer attitudes.

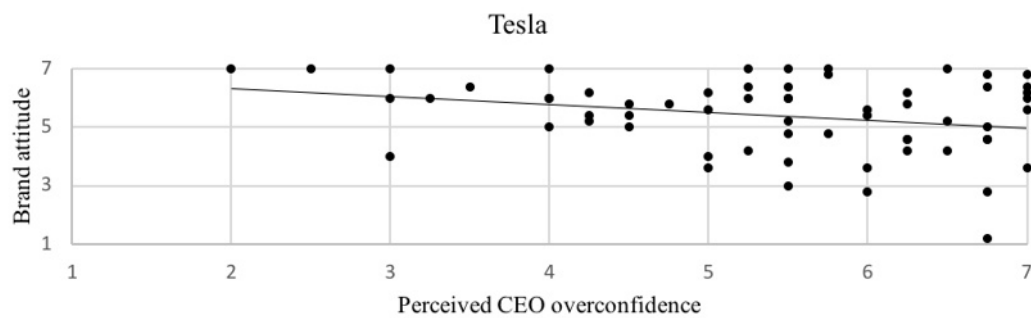
#### *Brand attitude and brand trust*

The findings of the first part of the study suggest that exposed overconfidence predicts lower ratings of the general attitude towards a brand. I assume that ratings of perceived CEO overconfidence yield similar results to those indicated by participants of the overconfidence condition group in the first part of the study. The results from the second part of the study confirm my fourth hypothesis. The results show that in two out of five cases higher perceived overconfidence is found to be connected with lower ratings of attitude towards

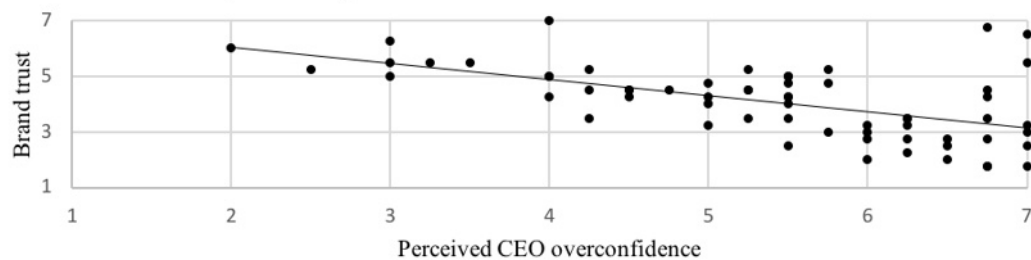
**Table 3:** Summary of Means and Standard Deviations for the Variables Perceived CEO Overconfidence, Brand Attitudes, and Brand Trust

	n	Perceived CEO Overconfidence		Participant Brand Attitude		Participant Brand Trust		
		Mean	SD	Mean	SD	Mean	SD	
Elon Musk	62	5.4	1.3	Tesla	5.41	1.27	4.08	1.29
Mark Zuckerberg	64	4	1	Facebook	3.59	1.31	3.38	1.18
Tim Cook	61	3.3	1.3	Apple	5.69	1.27	5.64	0.99
Satya Nadella	21	3	1.2	Microsoft	5.32	0.67	5.65	0.91
Jeff Bezos	57	2.9	1.2	Amazon	5.32	1.24	5.77	1.24

Note. SD = standard deviation. Results from G2 participants. Source: Own table



Note.  $\beta = -.27$ ,  $t(60) = -2.21$ ,  $p < .05$

**Figure 4:** Linear Regression Analyses for the Brand Tesla. Source: Own figure

the brand. Similar to the results of the first part of the study, the results of the second part show that higher levels of perceived overconfidence are related to lower ratings of brand trust. This relationship is stronger than that between perceived CEO overconfidence and brand attitude. These findings are especially strong for the company Tesla (see Figure 4).

The companies which showed a significantly negative relationship between perceived overconfidence and general attitude towards a brand as well as brand trust were the companies in which the average ratings of perceived CEO overconfidence were the highest out of the group of CEOs. In reality, it can be challenging to assess whether a person has been overconfident or merely confident. Thus, whether or not a CEO is overconfident may be a controversial question. The results indicate that in a situation where this is the case, those consumers who feel like the CEO is overconfident will have a lower attitude towards the brand than those who per-

ceive the CEO as merely confident.

#### *Familiarity with a CEO*

When observing the impact of CEO overconfidence on consumer attitudes in reality, familiarity with the CEO plays an important role. For CEO overconfidence to impact consumer attitudes in any way, consumers must associate the CEO with a company for meaning transfer to occur. Thus, if consumers are not familiar with a CEO or do not associate him or her with the company, there will be no impact of CEO overconfidence on consumer attitudes. There may be a relationship between CEO familiarity and CEO overconfidence. The results of the study show a slight trend which indicates a higher rate of familiarity with CEOs who are perceived as more overconfident and lower rates of familiarity with CEOs who are perceived as less overconfident (see Appendix H). However, this relationship is not linear. Banerjee et al. (2018) suggest that CEO overconfidence will affect the CEOs communica-

tion with the public as they will make more overly optimistic predictions. Additionally, I assume that it is likely that overconfident CEOs make a larger quantity of public statements which increases their publicity and thus familiarity among the public. This assumption could explain the higher rate of familiarity with CEOs who are perceived as more overconfident, however, further research is necessary to confirm this relationship.

#### *Limitations*

A severe limitation to the explanatory power of the data collected during the second part of the study is the lack of control for other mediating factors and the missing causality between perceived CEO overconfidence and brand attitude. The participants' attitudes towards the company and their trust in the brand may be affected by a variety of other factors next to their perception of the CEO's level of overconfidence. For example, the study by [Mitchell and Olson \(1981\)](#) suggests that a consumer's beliefs about a product are an important mediator of their attitude towards the brand. However, the current study did not control for individual beliefs participants held about the company's products.

A higher rating of perceived CEO overconfidence may also be the consequence of a more negative view of the brand. This effect is the opposite of what the endorsement theory describes. However, the phenomenon could be explained by the Halo effect which is a bias that causes an individual's overall impression of an object to influence his or her evaluation of specific attributes concerning that object ([Thorndike, 1920](#)). Research has found that rationally formed consumer attitudes may be distorted by affective overtones which occur as a consequence of the Halo effect ([Holbrook, 1983](#)). Whether the individual has a systematically positive or negative view about the attributes of an object thus depends on the general attitude towards it. Following the Halo effect theory, it is possible that the general impression a participant has of the brand influences his or her specific evaluation of the company's CEO. If the participant has an overall negative impression of the brand, it is likely that he or she will view the CEO in a more negative light and ascribe negative traits such as overconfidence to them.

#### *Further research*

While there are severe limitations to the explanatory power of the survey, it holds the advantage of observing realistic impressions of perceived CEO overconfidence and consumer attitudes towards brands. Further studies should thus focus on observing the relationship between perceived CEO overconfidence and consumer brand attitudes in realistic settings. However, they need to ensure a causal relationship between perceived CEO overconfidence and the attitudes observed. Further research could also measure other consumer attitudes next to brand attitude and trust.

A further area of research may be the exploration of the question of whether there is a link between the CEO's personality characteristic of overconfidence and the rate of CEO familiarity among the public. As previously suggested a foun-

dition for the exploration of this relationship may be the assumption that overconfident CEOs make more overly positive public statements which increases their publicity and thus familiarity among the public.

## 7. Conclusion

The results of the second part of the study support the results of the first part of the study which show that exposed verbally communicated CEO overconfidence leads to lower evaluations of the product, the attitude towards the brand, trust in the brand and the brand personality trait of competence. Overall the study indicates that exposed or perceived CEO overconfidence negatively affects consumer attitudes. The study builds on the connection of research concerning the social impact of overconfidence and the endorsement theory. In this section, I will summarize the implications the study carries for previous and future research. Finally, I will conclude what CEOs and executives can learn from the study and how they can benefit from the findings.

### 7.1. Implications for Academic Research

The current study has offered a new angle to the previous literature regarding the impact of overconfidence on status and the endorsement theory. As a consequence, different areas of further research are recommended to extend the findings of the study. In the following I will summarize the implications the current research has regarding previous and future literature.

Firstly, the study has significant implications for the literature regarding the impact of overconfidence on status. The discussion of the results has shown how the findings of the survey implicitly support the findings of [Tenney et al. \(2019\)](#). Thus, this study offers further support to the theory of the status diminishing effect of overconfidence and finds no evidence for the status-enhancement theory of overconfidence ([Anderson et al., 2012b](#)). Further research is needed to explore whether non-verbally communicated overconfidence yields status-enhancing benefits for CEOs. Some studies have already explored the impact that CEO overconfidence has on the relationship with stakeholders. Similar to the current study, this research also used the status-enhancement theory of overconfidence as a foundation to explain the impact that CEO overconfidence has on stakeholders such as employees and suppliers ([Phua et al., 2018](#)). The current study adds to this research in two ways. Firstly, it offers a new angle of view as it focusses on consumers, which has not been done before. Secondly, contrary to the previous literature, the current study finds a negative relationship between CEO overconfidence and consumer attitudes.

Secondly, the research has important implications for the literature in the field of the endorsement theory. The current study has confirmed the validity of the meaning transfer process in the context of CEO endorsement. This finding is an extension to previous literature where the attention has mostly been on brand endorsement through the use of celebrities in



advertising. As CEOs usually do not endorse specific products but represent the company as a whole, it is an interesting finding that through the meaning transfer process not only the general attitude towards the brand and brand trust are affected but also the attitude towards a specific product.

Lastly, the study suggests the necessity of future research concerning the debiasing of CEO overconfidence. The results of the current study indicate that revealed verbally communicated overconfidence by the CEO negatively influences consumers' attitudes towards products and brands and thus indirectly negatively influences the company's performance. As a consequence, methods for the debiasing of CEO overconfidence should be developed and tested so that CEOs can decrease the degree to which they are perceived as overconfident by consumers. For CEOs to be able to draw a complete conclusion regarding the optimal display and communication of confidence, further research concerning the impact of non-verbally communicated CEO overconfidence and CEO underconfidence is necessary.

## 7.2. Managerial Implications

The study has important implications for the communication strategies and general behavior of CEOs. The findings suggest that CEOs should avoid making statements that consumers may interpret as overconfident. CEOs should be aware of the process of meaning transfer by which consumers transfer meanings they associate with the CEO to the products and brand of the company. While the association of the trait overconfidence leads to more negative attitudes towards the products and brand, CEOs can also try and find ways to use this process to their advantage. In contrast to overconfidence, other CEO personality traits may transfer positive meanings to the products and brand of a company. This effect needs to be confirmed by further research.

When applying the findings of the study to the case of Elon Musk, the CEO of Tesla, it is suggested that he should reduce the amount of publicly communicated overconfident statements he makes, as this negatively affects consumers' attitudes towards the Tesla products and brand. However, for him to actively make these changes in his communication, it is required that he develops the ability to identify statements and behaviors that may be interpreted as overconfident. Jim Cantrell, co-founder of Space X, has said the following about Musk in 2014: "The one major important distinction that sets him apart is his inability to consider failure. It simply is not even in his thought process". If Musk is convinced that he will succeed in reaching the goals he has set himself, no matter how high they are, he will not be able to consider his statements and behavior as overconfident. As this mindset may be common among CEOs the greatest difficulty for them will be to develop the ability of a targeted self-reflection by which they can identify which of their statements could be considered as overconfident. In conclusion, the finding that CEO overconfidence negatively affects consumer attitudes sets the foundation for CEOs to understand how the communication of overconfident goals and forecasts takes an active part in the formation of consumer attitudes.

However, CEOs will only be able to make use of this finding if they can develop tools which help them identify how and when they are overconfident. Only then will they also be able to act accordingly and reduce their own verbally communicated overconfidence.

## References

- Aaker, D. A. *Managing Brand Equity: Capitalizing on the Value of a Brand Name*. N.Y., New York, Free Press, 1991.
- Aaker, J. L. Dimensions of Brand Personality. *Journal of Marketing Research (JMR)*, 34(3):347–356, 1997. doi: 10.2307/3151897.
- Ahluwalia, R., Burnkrant, R. E., and Unnava, H. R. Consumer Response to Negative Publicity: The Moderating Role of Commitment. *Journal of Marketing Research (JMR)*, 37(2):203–214, 2000. doi: 10.1509/jmkr.37.2.203.18734.
- Alberto, G. and Timothy, S. S. CEO Overconfidence and Innovation. *Management Science*, 57(8):1469–1484, 2011. doi: 10.1287/mnsc.1110.1374.
- Alicke, M. D. Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology*, 49(6):1621–1630, 1985. doi: 10.1037/0022-3514.49.6.1621.
- Amos, C., Holmes, G. E., and Strutton, D. Exploring the relationship between celebrity endorser effects and advertising effectiveness. *International Journal of Advertising*, 27:209 – 234, 2008. doi: 10.1080/02650487.2008.11073052.
- Anderson, C. and Kidluff, G. J. Why do dominant personalities attain influence in face-to-face groups? The competence-signaling effects of trait dominance. *Journal of Personality and Social Psychology*, 96(2):491–503, 2009. doi: 10.1037/a0014201.
- Anderson, C., Brion, S., Moore, D. A., and Kennedy, J. A. A status-enhancement account of overconfidence. *Journal of Personality and Social Psychology*, 103(4):718–735, 2012a. doi:10.1037/a0029395.
- Anderson, C., Brion, S., Moore, D. A., and Kennedy, J. A. A status-enhancement account of overconfidence. *Journal of Personality and Social Psychology*, 103(4):718–735, 2012b. doi: 10.1037/a0029395.
- Baker, M. J. and Churchill Jr, G. A. The Impact of Physically Attractive Models on Advertising Evaluations. *Journal of Marketing Research (JMR)*, 14(4): 538–555, 1977. doi: 10.2307/3151194.
- Banerjee, S., Humphery-Jenner, M., Nanda, V., and Tham, M. Executive Overconfidence and Securities Class Actions. *Journal of Financial and Quantitative Analysis*, 53(6):2685–2719, 2018. doi: 10.1017/S0022109018001291.
- Barber, B. M. and Odean, T. Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment. *The Quarterly Journal of Economics*, 116(1): 261–292, 2001. doi: 10.1162/003355301556400.
- Baumeister, R. F. Self-esteem, self-presentation, and future interaction: A dilemma of reputation. *Journal of Personality*, 50(1):29–45, 1982. doi: 10.1111/j.1467-6494.1982.tb00743.x.
- Bertrand, M. and Schoar, A. Managing with Style: The Effect of Managers on Firm Policies. *Quarterly Journal of Economics*, 118(4), 2003. doi: 10.1162/003355303322552775.
- Bloomberg, L. P. Market capitalization of Tesla Inc on 4/5/2019, 2019. URL <https://www.bloomberg.com/quote/TSLA:US>. Retrieved May 4, 2019.
- Burson, K. A., Larrick, R. P., and Klayman, J. Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology*, 90(1):60–77, 2006. doi: 10.1037/0022-3514.90.1.60.
- Cacioppo, J. T., Gardner, W. L., and Berentson, G. G. Beyond Bipolar Conceptualizations and Measures: The Case of Attitudes and Evaluative Space. *Personality & Social Psychology Review (Lawrence Erlbaum Associates)*, 1(1):3–25, 1997. doi: 10.1207/s15327957pspr0101\_2.
- Cantrell, J. How Did Elon Musk Learn Enough About Rockets To Run SpaceX? Jim Cantrell Answers, 2014. URL <https://www.forbes.com/sites/quora/2014/07/16/how-did-elon-musk-learn-enough-about-rockets-to-run-spacex-cofounder-jim-cantrell-answers/>. Retrieved May 5, 2019.
- Chaudhuri, A. and Holbrook, M. B. The Chain of Effects from Brand Trust and Brand Affect to Brand Performance: The Role of Brand Loyalty. *Journal of Marketing*, 65(2):81–93, 2001. doi: 10.1509/jmkg.65.2.81.18255.
- De Bondt, W. F. M. and Thaler, R. H. Financial decision-making in markets and firms: A behavioral perspective. *Chapter 13. In Handbooks in Operations Research and Management Science*, 9, 1995. doi: 10.1016/S0927-0507(05)80057-X.
- Dick, A. S. and Basu, K. Customer Loyalty: Toward an Integrated Conceptual Framework. *Journal of the Academy of Marketing Science*, 22(2):99–113, 1994. doi: 10.1177/0092070394222001.
- Duckworth, A. L., Peterson, C., Matthews, M. D., and Kelly, D. R. Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6):1087–1101, 2007. doi: 10.1037/0022-3514.92.6.1087.
- Dunn, M. SpaceX suffers serious setback with crew capsule accident. Washington Post, 2019. URL <https://www.washingtonpost.com/national/health-science/spacex-suffers-serious-setback-with-crew-capsule-accident/2019/04/22/e91bba54-650f-11e9-a698-2a8f808c9cfb>. Retrieved May 5, 2019.
- Edmans, A. Why Elon Musk's Compensation Plan Wouldn't Work for Most Executives. Harvard Business Review Digital Articles, 2-4, 2018. URL <https://hbr.org/2018/01/why-elon-musks-compensation-plan-wouldnt-work-for-most-executives>.
- Erdem, T. and Swait, J. Brand Credibility, Brand Consideration, and Choice. *Journal of Consumer Research*, 31(1):191–198, 2004. doi: 10.1086/383434.
- Fischbein, M. and Ajzen, I. Chapter 5: Belief Formation. In Reading, Mass.:Addison-Wesley Pub. Co., *Belief, attitude, intention, and behavior: an introduction to theory and research*, 1975. pp. 131–215.
- Fournier, S. Consumers and Their Brands: Developing Relationship Theory in Consumer Research. *Journal of Consumer Research*, 24(4):343–370, 1998. doi:10.1086/209515.
- Heaton, J. B. Managerial Optimism and Corporate Finance. *Financial Management*, 31(2):33–45, 2002. doi: 10.2307/3666221.
- Holbrook, M. B. Using a Structural Model of Halo Effect to Assess Perceptual Distortion due to Affective Overtones. *Journal of Consumer Research*, 10(2):247–252, 1983. doi: 10.1086/208963.
- Keller, K. L. Conceptualizing, Measuring, Managing Customer-Based Brand Equity. *Journal of Marketing*, 57(1):1–22, 1993. doi: 10.1177/002224299305700101.
- Kennedy, J. A., Anderson, C., and Moore, D. A. When overconfidence is revealed to others: Testing the status-enhancement theory of overconfidence. *Organizational Behavior and Human Decision Processes*, 122(2): 266–279, 2013. doi:10.1016/j.obhdp.2013.08.005.
- Klein, W. and Kunda, Z. Maintaining Self-Serving Social Comparisons: Biased Reconstruction of One's Past Behaviors. *Personality and Social Psychology Bulletin*, 19(6):732–739, 1993. doi: 10.1177/0146167293196008.
- Kramer, L. A. and Liao, C. M. The spillover effects of management overconfidence on analyst forecasts. *Journal of Behavioral and Experimental Finance*, 12:79–92, 2016. doi: 10.1016/j.jbef.2016.10.004.
- Kuiper, N. A. Depression and causal attributions for success and failure. *Journal of Personality and Social Psychology*, 36(3):236–246, 1978. doi: 10.1037/0022-3514.36.3.236.
- Langer, E. J. The Illusion of Control. *Journal of Personality & Social Psychology*, 32(2):311–328, 1975. doi: 10.1037/0022-3514.32.2.311.
- Laurent, G. EMAC Distinguished Marketing Scholar 2012: Respect the data! *International Journal of Research in Marketing*, 30(4):323–334, 2013. doi: 10.1016/j.ijresmar.2013.07.003.
- Malmendier, R. and Tate, G. Behavioral CEOs: The Role of Managerial Overconfidence. *Journal of Economic Perspectives*, 29(4):37–60, 2015. doi: 10.1257/jep.29.4.37.
- Malmendier, U. and Tate, G. CEO Overconfidence and Corporate Investment. *The Journal of Finance*, 60(6):2661–2700, 2005. doi: 10.1111/j.1540-6261.2005.00813.x.
- Malmendier, U. and Tate, G. Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics*, 89(1):20–43, 2008. doi: 10.1016/j.jfineco.2007.07.002.
- McCracken, G. Who Is the Celebrity Endorser? Cultural Foundations of the Endorsement Process. *Journal of Consumer Research*, 16(3):310–321, 1989. doi: 10.1086/209217.
- Mezulis, A. H., Abramson, L. Y., Hyde, J. S., and Hankin, B. L. Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, 130(5):711–747, 2004. doi: 10.1037/0033-2909.130.5.711.
- Miller, D. T. and Ross, M. Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, 82(2):213–225, 1975. doi: 10.1037/h0076486.
- Mitchell, A. A. and Olson, J. C. Are Product Attribute Beliefs the Only Mediator of Advertising Effects on Brand Attitude? *Journal of Marketing Research (JMR)*, 18(3):318–332, 1981. doi: 10.2307/3150973.

- Mohr, J. and Nevin, J. R. Communication strategies in marketing channels: A theoretical perspective. *Journal of Marketing*, 54(4):36–51, 1990. doi: 10.1177/002224299005400404.
- Moore, D. A. and Healy, P. J. The trouble with overconfidence. *Psychological Review*, 115(2):502–517, 2008. doi: 10.1037/0033-295X.115.2.502.
- Moore, D. A. and Small, D. A. Error and bias in comparative judgment: On being both better and worse than we think we are. *Journal of Personality and Social Psychology*, 92(6):972–989, 2007. doi: 10.1037/0022-3514.92.6.972.
- Moorman, C., Zaltman, G., and Deshpande, R. Relationships Between Providers and Users of Market Research: The Dynamics of Trust Within and Between Organizations. *Journal of Marketing Research (JMR)*, 29(3): 314–328, 1992. doi: 10.2307/3172742.
- Morgan, R. M. and Hunt, S. D. The Commitment-Trust Theory of Relationship Marketing. *Journal of Marketing*, 58(3):20–38, 1994. doi: 10.1177/002224299405800302.
- Moulard, J. G., Garrity, C. P., and Rice, D. H. What Makes a Human Brand Authentic? Identifying the Antecedents of Celebrity Authenticity. *Psychology & Marketing*, 32(2):173–186, 2015. doi: 10.1002/mar.20771.
- Ohanian, R. Construction and Validation of a Scale to Measure Celebrity Endorsers' Perceived Expertise, Trustworthiness, and Attractiveness. *Journal of Advertising*, 19(3):39–52, 1990. doi: 10.1080/00913367.1990.10673191.
- O'Reilly, D. Cultural Brands/ Branding Cultures. *Journal of Marketing Management*, 21(5-6):573–588, 2005. doi: 10.1362/0267257054307336.
- Phua, K., Tham, T. M., and Wei, C. Are overconfident CEOs better leaders? Evidence from stakeholder commitments. *Journal of Financial Economics*, 127(3):519–545, 2018. doi: 10.1016/j.jfineco.2017.12.008.
- Reynolds, J. D. J. and Gifford, R. The sounds and sights of intelligence: A lens model channel analysis. *Personality and Social Psychology Bulletin*, 27(2):187–200, 2001. doi: 10.1177/0146167201272005.
- Roehm, M. L. and Sternthal, B. The Moderating Effect of Knowledge and Resources on the Persuasive Impact of Analogies. *Journal of Consumer Research*, 28(2):257–272, 2001. doi: 10.1086/322901.
- Scheidt, S., Gelhard, C., Strotzer, J., and Henseler, J. In for a penny, in for a pound? Exploring mutual endorsement effects between celebrity CEOs and corporate brands. *Journal of Product and Brand Management*, 27(2): 203–220, 2018. doi: 10.1108/JPBM-07-2016-1265.
- Shepperd, J. A., Ouellette, J. A., and Fernandez, J. K. Abandoning unrealistic optimism: Performance estimates and the temporal proximity of self-relevant feedback. *Journal of Personality and Social Psychology*, 70(4):844–855, 1996. doi: 10.1037/0022-3514.70.4.844.
- Sirdeshmuck, D., Singh, J., and Sabol, B. Consumer Trust, Value, and Loyalty in Relational Exchanges. *Journal of Marketing*, 66(1):15–37, 2002. doi: 10.1509/jmkg.66.1.15.18449.
- Taylor, S. E. and Brown, J. D. Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103(2):193–210, 1988. doi: 10.1037/0033-2909.103.2.193.
- Tenney, E. R., Meikle, N. L., Hunsaker, D., Hunsaker, D., Moore, D. A., and Anderson, C. Is overconfidence a social liability? The effect of verbal versus nonverbal expressions of confidence. *Journal of Personality and Social Psychology*, 116(3):396–415, 2019. doi: 10.1037/pspi0000150.
- Thompson, C. Elon Musk: Model 3 production problems stem from Tesla getting 'overconfident' and 'too comfortable'. Business Insider, 2018. URL <https://www.businessinsider.de/elon-musk-tesla-model-3-production-problems-causes-overconfident-batteries-2018-2?r=US&IR=T>. Retrieved May 5, 2019.
- Thorndike, E. L. A constant error in psychological ratings. *Journal of Applied Psychology*, 4(1):25–29, 1920. doi: 10.1037/h0071663.
- Tripp, C., Jensen, T. D., and Carlson, L. The Effects of Multiple Product Endorsements by Celebrities on Consumers' Attitudes and Intentions. *Journal of Consumer Research*, 20(4):535–547, 1994. doi: 10.1086/209368.
- Whan Park, C., MacInnis, D. J., Priester, J., Eisingerich, A. B., and Iacobucci, D. Brand Attachment and Brand Attitude Strength: Conceptual and Empirical Differentiation of Two Critical Brand Equity Drivers. *Journal of Marketing*, 74(6):1–17, 2010. doi: 10.1509/jmkg.74.6.1.



## Over-Confidence Bias in Strategic Decision-Making: Antecedents and Consequences

### Over-Confidence Bias in strategischen Entscheidungsprozessen: Entstehung, Konsequenzen und Lösungsansätze

Jule Neckermann

Heinrich-Heine-Universität Düsseldorf

#### Abstract

Since managers often overestimate themselves and their decision-making ability, wrong decisions and serious crises for companies occur time and again. In this thesis, this distortion caused by overestimation of self, the over-confidence bias, will therefore first be incorporated into the context of strategic decision-making processes. In particular, it is shown that the cognitive bias due to overconfidence does not only refer to one phase of the decision-making process; instead, an overconfidence results from a multitude of other cognitive biases, situational factors, and various motivational and physiological states and thus represents a complex construct in the research. In the context of this thesis, we will therefore also investigate which influencing factors lead to the development of the bias and which opportunities, but also risks, can arise from this. In order to counteract these risks, this thesis also presents possible solutions and debiasing techniques that can tie in with previous research results from psychology and economic research and reduce the over-confidence bias in strategic decision-making processes.

#### Zusammenfassung

Da Manager sich und ihre Entscheidungsfähigkeit häufig überschätzen, kommt es immer wieder zu Fehlentscheidungen und folgenschweren Krisen für Unternehmen. In dieser Arbeit wird diese Verzerrung durch Selbstüberschätzung, der Over-Confidence Bias, daher zunächst in den Kontext strategischer Entscheidungsprozesse eingearbeitet. Hier zeigt sich insbesondere, dass sich die kognitive Verzerrung durch Over-Confidence nicht nur auf eine Phase des Entscheidungsprozesses bezieht; stattdessen resultiert ein übersteigertes Selbstbewusstsein aus einer Vielzahl anderer kognitiver Verzerrungen, situativer Faktoren sowie diversen motivationalen und physiologischen Zuständen und stellt somit ein komplexes Konstrukt in der Forschung dar. Im Rahmen dieser Arbeit wird daher ebenfalls untersucht, welche Einflussfaktoren zur Entstehung des Bias führen und welche Chancen, aber auch Risiken dadurch entstehen können. Um diesen Risiken entgegenwirken zu können, werden in dieser Arbeit ebenfalls mögliche Lösungsansätze und Debiasing-Techniken vorgestellt, die an bisherige Forschungsergebnisse aus der Psychologie und der Wirtschaftsforschung anknüpfen und den Over-Confidence Bias in strategischen Entscheidungsprozessen reduzieren können.

**Keywords:** Overconfidence; strategische Entscheidung; kognitive Verzerrungen; Debiasing.

#### 1. Einleitung

No problem in judgment and decision making is more prevalent and more potentially catastrophic than overconfidence. (Plous, 1993, S. 217)

Der Over-Confidence Bias stellt die Überschätzung der

eigenen Fähigkeiten und des eigenen Wissens durch ein Individuum oder eine Organisation dar und gilt als eine der bekanntesten kognitiven Verzerrungen des strategischen Entscheidungsverhaltens (Russo und Schoemaker, 2016). In der Unternehmenspraxis kommt es in diesem Kontext immer wieder zu strategischen Fehlentscheidungen, die zu

schwerwiegenden Folgen für den Entscheidungsträger, das Unternehmen und ganze Staaten führen, denn der Over-Confidence Bias wird in diesem Kontext als Auslöser für Finanzkrisen, die hohe Insolvenzrate bei neugegründeten Unternehmen, Kriege oder das Scheitern von Unternehmensfusionen verantwortlich gemacht (Bazerman und Moore, 2013).

In der vorliegenden Arbeit steht deshalb die Frage im Vordergrund, inwieweit sich Over-Confidence auf strategische Entscheidungsprozesse auswirkt. Es wird untersucht, welche Besonderheiten hinsichtlich strategischer Entscheidungen bestehen und weshalb diese Entscheidungen nicht komplett rational getroffen werden. Der Over-Confidence Bias wird hinsichtlich seiner Definition und verschiedenen Ausprägungen näher untersucht. Insbesondere im Hinblick auf die Entstehung von Over-Confidence im Entscheidungsprozess wird dessen Bedeutung sowie die damit verbundenen Risiken und Chancen verdeutlicht. Fokussiert wird sich auf die bestehenden Lösungsansätze und die Frage, inwieweit die Verhinderung des Bias durch ein sog. Debiasing notwendig ist. Die vorliegende Arbeit geht dabei über die Deskription des Over-Confidence Bias hinaus und wendet sich der Analyse von Ursachen und Konsequenzen zu. In diesem Rahmen wird hinterfragt, ob ein Debiasing in strategischen Entscheidungsprozessen notwendig ist.

Hierzu wird zunächst ein Überblick über die grundlegenden Begriffe und Konstrukte, den strategischen Entscheidungsprozess sowie den Over-Confidence Bias gegeben. Im Hinblick auf die kognitiven Besonderheiten von Entscheidungen und weiteren Einflussfaktoren wird die Entstehung von Over-Confidence thematisiert. Darauf folgt eine Analyse der Auswirkungen des Bias auf den Entscheidungsprozess, den Entscheidungsträger und das Ergebnis, sowohl unter Betracht der Chancen als auch Risiken. Es werden im Anschluss Lösungsansätze vorgestellt und die zentralen Ansätze beleuchtet. Daraus ergeben sich für das Management relevante Implikationen und ein Veränderungsprozess. Zuletzt werden die Ergebnisse diskutiert, kritisch hinterfragt und ein abschließendes Fazit festgestellt, bevor ein Ausblick auf weitere Forschungen erfolgt. Dabei wird deutlich, dass die Unternehmen angehalten sind, ein neues Bewusstsein für kognitive Verzerrungen zu schaffen. Der Fokus liegt dabei auf den Konsequenzen sowie den vorgestellten Debiasing-Techniken. Als Unterstützung dient ein Veränderungsmodell, anhand dessen eine Integration der Debiasing-Maßnahmen in die Praxis ermöglicht wird. Als Ergebnis der anschließenden Diskussion lässt sich ein maßgeblicher Einfluss des Over-Confidence Bias auf strategische Entscheidungen und die Notwendigkeit eines Debiasing feststellen.

## 2. Relevanz des Over-Confidence Bias in strategischen Entscheidungsprozessen

### 2.1. Ausgestaltung strategischer Entscheidungsprozesse

#### 2.1.1. Begriff der strategischen Entscheidung

In der Unternehmenspraxis sowie dem alltäglichen Leben werden täglich eine Vielzahl von Entscheidungen getroffen.

Seit Jahrzehnten beschäftigt diese Thematik die Betriebswirtschaftslehre, doch auch andere Wissenschaften, wie die Psychologie und Soziologie, beschäftigen sich mit den Tatbeständen und der Definition von Entscheidungen (Göbel, 2014, Pfister, 2017). Eine Entscheidung lässt sich dabei zunächst als Resultat eines Wahlprozesses interpretieren (Göbel, 2014). Dabei wählt der Entscheidungsträger aus einer oft nicht genau quantifizierbaren Anzahl von Alternativen diejenige Option aus, die für die eigene Zielerreichung am besten geeignet ist (Göbel, 2014).

Auf der Managementebene lassen sich operative und strategische Entscheidungen unterscheiden. Strategische Entscheidungen differenzieren sich von den operativen hinsichtlich des Planungszeitraums und der Relevanz für das Unternehmen. Operative Entscheidungen werden demnach regelmäßig und routiniert in Unternehmen getroffen. Die Auswirkungen dieser Entscheidungen sind von kurzer Dauer, was eine relative Sicherheit aus Sicht der Entscheider bewirkt. Hier lassen sich z.B. Entscheidungen bezüglich der Maschinenauslastung während einer speziellen Schicht einordnen. Zu den strategischen Entscheidungen zählen hingegen solche, welche einen längeren Zeitrahmen umfassen und beispielsweise eine Veränderung der Produktlinie oder auch mögliche Kooperationen mit anderen Unternehmen betreffen. Es besteht i.d.R. eine hohe Unsicherheit hinsichtlich des Entscheidungsproblems und das Entscheidungsumfeld ist durch eine erhöhte Komplexität geprägt. Strategische Entscheidungen haben dadurch eine maßgeblich existenzsichernde Bedeutung für den Betrieb und bestimmen die langfristige Ausrichtung. (Göbel, 2014)

Sie sind im Vergleich zu operativen Entscheidungen selten zu treffen, legen allerdings grundlegende Prämissen fest, welche die darauffolgenden operativen Entscheidungen lenken und beeinflussen (Nutt und Wilson, 2010). Strategische Entscheidungen führen zur Verwendung enormer organisationaler Ressourcen sowie zu der Berücksichtigung und Integration externer Akteure in der Entscheidungsfindung (Clark und Maggitti, 2012). Zu den Merkmalen einer strategischen Entscheidung nach Nutt und Wilson (2010) zählen darüber hinaus eine durch die Komplexität der expliziten Sachverhalte erschwerte Problemdefinition und die Existenz mehrerer Lösungsmöglichkeiten; wobei die verschiedenen Interessen der internen und externen Akteure abgewogen werden. Außerdem sind strategische Entscheidungen nur schwer reversibel, sobald die dafür erforderlichen Ressourcen eingesetzt werden (Nutt und Wilson, 2010). Im Zuge der strategischen Entscheidung besteht Ungewissheit bezüglich der Bewertung von Chancen und Risiken der Optionen und es handelt sich somit um mehrdeutige Entscheidungen, da kein genaues Wissen über die exakte Anzahl der Alternativen und Einflussfaktoren existiert (Mintzberg et al., 1976).

#### 2.1.2. Rationales Prozessmodell strategischer Entscheidungen

Die oben genannte Definition der strategischen Entscheidung vernachlässigt allerdings die Dynamik der Entscheidungen und der damit verbunden Systematik. Für die Be-

trachtung des strategischen Entscheidungsprozesses lässt sich diese Definition entsprechend erweitern. Eine Entscheidung lässt sich dabei als mentaler Prozess bestehend aus Beurteilung, Bewertung und Auswahl der Alternativen zusammensetzen (Pfister, 2017). Der strategische Entscheidungsprozess stellt die Durchsetzung einer grundlegenden Entscheidung durch das oberste Management dar (Das und Teng, 1999). Aufgrund der oben erläuterten Komplexität und Neuartigkeit einer strategischen Entscheidung besteht keine verbindliche und einheitliche Vorgehensweise für den Umgang mit Entscheidungsproblemen. Es existieren allerdings zahlreiche Ansätze mit unterschiedlichen Schwerpunkten zur Modellierung der entsprechenden Prozesse (Göbel, 2014). Rationale Entscheidungsmodelle basieren meist auf der Identifikation mehrerer Prozessschritte, durch welche dem Entscheider die Auswahl der optimalen Lösung für das Problem aus mehreren Alternativen ermöglicht wird (Korte, 2003). Sie stellen somit i.d.R. ein sequentielles Vorgehen dar und sind Teil der präskriptiven Entscheidungstheorie, die auf der Rationalität des Entscheiders basiert und sich mit der Empfehlung von Entscheidungen vor dem Hintergrund der bestmöglichen Alternative beschäftigt (Göbel, 2014). Rationalität in diesem Zusammenhang impliziert einen Entscheidungsprozess, welcher gegeben den Werten und der Risikoeinstellung des Entscheiders zu einer optimalen Lösung führt (Bazerman und Moore, 2013).

Die einzelnen Prozessphasen des rationalen Entscheidungsprozessmodells nach Mintzberg et al. (1976) sind in Abbildung 1 grafisch dargestellt. Die erste Stufe in diesem Modell stellt die Problemidentifikation dar. Diese Phase umfasst die zwei Routinen der Entscheidungserkennung und der Diagnose. Die erste Routine umfasst die Erkennung der Notwendigkeit einer Entscheidung anhand des Abwägens von Problemen, Chancen und krisenhaften Entwicklungen. Dafür ist es notwendig, dass ein Unterschied zwischen dem Ist-Zustand und dem erwarteten Zustand erkannt wird. Die Diagnose-Routine führt dazu, dass die verschiedenen Umweltstimuli interpretiert und in Zusammenhang mit der Situation gebracht werden. Dafür werden die notwendigen Ressourcen im Unternehmen mobilisiert, um den Entscheidungsprozess zu initiieren. (Mintzberg et al., 1976)

Darauf aufbauend umfasst die zweite Stufe im Entscheidungsprozessmodell die Generierung von Optionen. Sie bildet den Kern des Entscheidungsprozesses. In dieser Phase werden die meisten Ressourcen benötigt, um eine oder mehrere Lösungen für das Entscheidungsproblem aufzuzeigen. In dieser Entwicklungsphase lassen sich ebenfalls zwei Routinen, die Such- und Designroutine, einordnen, welche die Unterscheidung in konvergentes und divergentes Denken widerspiegeln. In der Suchroutine können vier verschiedene, oftmals hierarchisch und schrittweise aufeinander folgende Formen der Informationssuche unterschieden werden. Neben dem Scannen der bereits existierenden Informationen, der passiven Suche nach unaufgeforderten Alternativen und der Suchanfrage, lässt sich als letzte Form dieser Routine die aktive Suche nach Alternativen durch Fokussierung auf einen Teilbereich oder intensives Scannen beschreiben. Die ande-

re Routine in der Entwicklungsphase stellt die Designroutine dar. Hier lässt sich feststellen, dass Unternehmen auf eigene maßgeschneiderte oder modifizierte Lösungsmöglichkeiten zurückgreifen. Maßgeschneiderte Lösungen sind das Resultat des Entwicklungsprozesses und mit einem höheren Aufwand verbunden. (Mintzberg et al., 1976)

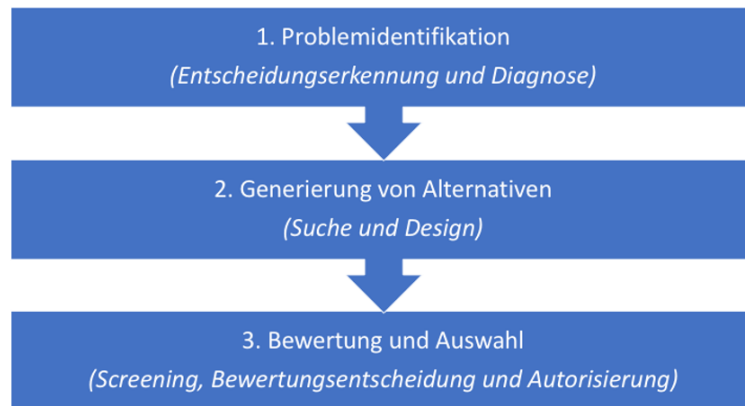
Zuletzt erfolgt die Selektionsphase, welche sich in die Routinen des Screenings, der Bewertungsentscheidung und der Autorisierung unterteilt. Während in der ersten Routine eine Reduzierung der Alternativen stattfindet, werden die selektierten Alternativen dann bewertet und eine Handlung festgelegt. Daraufhin wird in der Autorisierungsroutine die Entscheidung auf einer höheren Hierarchiestufe festgelegt. (Mintzberg et al., 1976)

Es wird sich im Folgenden auf das Modell nach Mintzberg et al. (1976) fokussiert, da es die drei bereits erwähnten notwendigen Grundschritte eines Entscheidungsprozesses, die Beurteilung, Bewertung und Auswahl, umfasst (Simon, 1955, Mintzberg et al., 1976, Pfister, 2017). Weiterführende Modelle beinhalten dieselben Schritte in ähnlicher Form, wobei z.T. noch Schritte ergänzt oder einzelne Phasen weiter unterteilt werden (Schwenk, 1984, Bazerman und Moore, 2013). Es ist übersichtlich und der später untersuchte kognitive Fehler kann in den Prozess eingeordnet werden, was angesichts der im Weiteren analysierten Thematik in Kapitel 3.1.2 von hoher Bedeutung ist (Schwenk, 1984).

### 2.1.3. Kognitive Vorgänge bei Entscheidungen

Sometimes decisions defy purely step-by-step logic. To be effective, companies also should embrace intuitive or action-oriented forms of decision making. (Mintzberg und Westley, 2001, S.89)

Mintzberg und Westley (2001) vertreten hier die These, dass für ein effektives Entscheiden neben der schrittweisen Vorgehensweise auch intuitive und handlungsorientierte Formen genutzt werden sollten. Das Zitat verdeutlicht die Komplexität und Dynamik strategischer Entscheidungsprozesse, welche durch ein rationales, schrittweises Prozessmodell nur begrenzt dargestellt werden können, sodass die Integration verhaltenswissenschaftlicher Erkenntnisse erforderlich ist (Mintzberg und Westley, 2001). Nach den Ansätzen der klassischen Ökonomie wählt ein Individuum als sog. Homo Oeconomicus die Alternative, welche als Ergebnis aus dem rationalen Entscheidungsprozess hervorgeht (Beck, 2014). Ein Wirtschaftssubjekt handelt demnach nutzenmaximierend, rational und emotionsfrei und verfügt über einen fehlerfreien Informationsverarbeitungsprozess (Beck, 2014). Zwar ermöglicht diese Theorie eine mathematische Formalisierung und ökonomische Modellbildung (wie in einem strukturierten rationalen Entscheidungsprozess) jedoch lassen sich diese Ansätze in der Form nicht auf die Realität beziehen, da ein Individuum nur begrenzt über Rationalität, Willenskraft und Eigennutzen verfügt (Beck, 2014). Hier lassen sich die Ansätze der Verhaltensökonomie einordnen, denn dort werden verhaltenswissenschaftliche Aspekte in die Fundierung wirtschaftswissenschaftlicher Arbeiten miteinbezogen (Beck,



**Abbildung 1:** Entscheidungsprozess nach Mintzberg et al. (1976) (eigene Darstellung)

2014). Schon seit den 1950ern scheinen Wirtschaftswissenschaftler von der Behauptung, dass Nutzenmaximierung das primäre Ziel eines Individuums bei Entscheidungsproblemen sei, Abstand zu nehmen. Denn nach Simon (1957) strebt der Mensch bei der Lösung dieses Problems statt der Optimierung lediglich eine Zufriedenstellung an (Simon, 1957). Anstelle der bestmöglichen Lösung begnügt sich der Mensch aufgrund der Umweltzustände und eigenen kognitiven Limitationen mit der ersten, den Bedürfnissen entsprechenden Option – die Nutzenmaximierung als primäres Ziel wird hierbei durch die Zufriedenstellung ersetzt (Simon, 1957).

Das menschliche Denken ist in seiner Rationalität folglich aufgrund intrapersoneller aber auch umweltbedingter Umstände beschränkt (Bazerman und Moore, 2013). Der Begriff der begrenzten Rationalität bezieht sich auf die Verarbeitung von Informationen im Entscheidungsprozess und beschreibt die limitierte Fähigkeit eines Menschen, vollkommen rational und fehlerfrei zu entscheiden (Bazerman und Moore, 2013). Die begrenzte Rationalität des menschlichen Handelns impliziert laut Angaben des Internationalen Controller Vereins e.V. (ICV) die Anwendung kognitiver Vereinfachungsstrategien, sog. Heuristiken. Sie vereinfachen und strukturieren ein Entscheidungsproblem, basierend auf den eigenen Erfahrungen (ICV, Internationaler Controller Verein e.V.). Der Begriff stammt aus dem Griechischen und bedeutet eine verbesserte Lösung von Problemen (Beck, 2014). Durch die Anwendung dieser Faustregeln wird eine schnellere und kognitiv weniger anspruchsvolle Entscheidungsfindung ermöglicht (Göbel, 2014). Es lassen sich je nach Prinzip verschiedene Formen von Heuristiken kategorisieren, wie beispielsweise die Bestätigungsheuristik, Verfügbarkeitsheuristik und Repräsentativitätsheuristik. Bei der Anwendung der Verfügbarkeitsheuristik wird ein Ereignis wahrscheinlicher eingeschätzt, wenn die Umstände dieses Ereignisses bereits im Gedächtnis verfügbar und abrufbar sind (Bazerman und Moore, 2013). Die Repräsentativitätsheuristik bezieht sich auf die bewusste oder unbewusste Implikation von Stereotypen und die Bestätigungsheuristik umfasst die Vereinfachungsstrategie, Informationen, die der eigenen Meinung widersprechen, zu ignorieren (Bazerman und Moore, 2013). Auf Basis dieser

Erkenntnisse wird anstelle des Homo Oeconomicus vermehrt vom Homo Heuristicus gesprochen, welcher nicht rein rational entscheidet, sondern aufgrund von individuellen oder situationsbedingten Gründen zu „mentale[n] Abkürzungen“ (Beck, 2014, S.25) in Form von Heuristiken greift. Der Homo Heuristicus ignoriert Teile der verfügbaren Informationen und weist zahlreiche kognitive Verzerrungen auf (ICV, Internationaler Controller Verein e.V.). Dennoch ist er in der Lage, effizienter als sein Gegenspieler Homo Oeconomicus zu entscheiden, welcher ohne kognitive Verzerrungen, ressourcenintensiver und generalstrategieorientiert handelt (Gigerenzer und Brighton, 2009).

Die Grundlage für diese existierenden Inkonsistenzen zwischen der rational begründeten und tatsächlich realisierten Entscheidung bildet der Zwei-System-Ansatz. Dieser basiert auf der Annahme, dass das menschliche Denken auf zwei verschiedenen, sich gegenseitig beeinflussenden Systemen basiert (Kahneman, 2011). In intuitiven, schnellen und unbewussten Situationen wirkt das kognitive System 1, welches sich vor allem auf automatische Handlungen und intuitive Entscheidungen bezieht. Ergänzend dazu ist das System 2 für aufwendige, bewusste und rationale kognitive Prozesse und Handlungen verantwortlich. Es erfordert ein höheres Maß an Aufwand und ist daher für wichtige und komplexe Entscheidungsprobleme geeignet (Kahneman und Frederick, 2002, Kahneman, 2011). Da er stets rational und nutzenmaximierend agiert, beruht der Homo Oeconomicus seine Entscheidungen auf dem analytischen System 2 (Beck, 2014). Doch die begrenzte Verfügbarkeit der erforderlichen Ressourcen wie z.B. Zeit, finanzielle Mittel und Informationen beeinflusst diesen Entscheidungsprozess. Dadurch wird das Denken von System 1 bei der Entscheidung in der Realität immer wichtiger und beeinflusst den Entscheidungsprozess, sodass der Homo Heuristicus sich bei der Entscheidungstreffung zwischen den zwei Systemen auszurichten hat (Kahneman, 2011). In Tabelle 1 sind einige Charakteristika der zwei Systeme, die sich in der Entscheidungsforschung durchgesetzt haben, im Vergleich dargestellt (Evans und Stanovich, 2013). Basierend auf ihren Untersuchungsergebnissen präferieren Evans und Stanovich (2013)

den Begriff des Zwei-Typen-Ansatzes. Die Verwendung dieses Begriffs stellt sicher, dass mehr als ein einzelnes kognitives oder neurales System an einem Entscheidungsprozess beteiligt sein kann (Evans und Stanovich, 2013).

## 2.2. Grundlagen des Over-Confidence Bias

### 2.2.1. Abgrenzung des Over-Confidence Bias

Durch die Anwendung von Heuristiken werden komplexe und unstrukturierte Entscheidungsprobleme auf kognitiver Ebene vereinfacht. Im Rahmen dieser strategischen Entscheidungen können allerdings auch Abweichungen von dem rationalen Ideal in Form kognitiver Verzerrungen, sog. Bias, entstehen. Sie stellen systematische Fehler dar, welche bei der unsachgemäßen Anwendung von Heuristiken auftreten können (ICV, Internationaler Controller Verein e.V.). Der Begriff umfasst eine Vielzahl von unterschiedlichen kognitiven Verzerrungen, es lassen sich unterschiedliche Gruppen von Bias unterscheiden und den Kategorien der Heuristiken zuordnen. So zählt zu den Bias, welche sich aus der Bestätigungsheuristik ergeben, u.a. der Over-Confidence Bias (Bazerman und Moore, 2013).

Beim Over-Confidence Bias handelt es sich um eine kognitive Verzerrung, die den Glauben eines Individuums oder einer Organisation beschreibt, über bessere Fähigkeiten und ein höheres Maß an Wissen zu verfügen, als in der Realität existieren (Russo und Schoemaker, 2016). Dabei stellt Over-Confidence keine allgemeine Persönlichkeitseigenschaft dar, sondern ist vom jeweiligen Kontext abhängig und kann je nach Situation schwanken (Glaser, 2019). Ebenso lässt sich Over-Confidence von der Selbstdarstellung abgrenzen. Die Selbstdarstellung bezieht sich auf die bewusste und freiwillige Modifizierung des eigenen sozialen Verhaltens nach außen, wobei diese Darstellung nicht zwingend der eigenen Wahrnehmung des Individuums entspricht (Anderson et al., 2012). Over-Confidence kann hingegen auch unbewusst entstehen und umfasst nach von Hippel und Trivers (2011) zwar die tatsächliche, aber fehlerhafte Wahrnehmung der eigenen Fähigkeiten (von Hippel und Trivers, 2011, Anderson et al., 2012). Over-Confidence bildet eine Basis für weitere kognitive Verzerrungen und beeinflusst dadurch das menschliche Entscheidungsverhalten, sodass der Bias auch als „mother of all biases“ (Bazerman und Moore, 2013, S.14) bezeichnet wird. Diese Form der kognitiven Verzerrung gilt als die tiefgreifendste und stärkste Verzerrung, für die das menschliche Urteilsvermögen anfällig ist (Kahneman, 2011, Mannes und Moore, 2013). Das ist einer der Gründe dafür, weshalb Over-Confidence zu den bekanntesten und am meisten untersuchten kognitiven Verzerrungen zählt (Fischhoff, 1982). Eine Messung des Konstrukts bedarf eines Vergleichs zwischen den subjektiven Ansichten des Individuums oder der Organisation und der Realität (Moore und Schatz, 2017).

Bezüglich des Ausmaßes des Over-Confidence Bias können individuelle Unterschiede nachgewiesen werden. So weisen z.B. Unternehmensgründer einen höheren Grad an Over-Confidence als Manager großer Organisationen auf (Busenitz und Barney, 1997). Dies lässt sich mit der Vorteilhaftigkeit

bei der Durchsetzung risikobehafteter Entscheidungen und der damit verbundenen Beeinflussung der Stakeholder begründen (Busenitz und Barney, 1997). Dies gibt auch Hinweis darauf, weshalb Unternehmer z.T. schlechtere Manager ausmachen, denn Manager großer Unternehmen beruhen ihre Entscheidungen eher auf Entscheidungsmethoden und der Betrachtung von Trends der vergangenen Performance als auf ihr Bauchgefühl (Busenitz und Barney, 1997). Darüber hinaus konnten Barber und Odean (2001) geschlechtsspezifische und altersbedingte Unterschiede in Bezug auf die Stärke von Over-Confidence nachweisen. Männer weisen demnach im Vergleich zu Frauen verstärkt Over-Confidence auf, Ältere sind häufiger davon betroffen als junge Menschen (Barber und Odean, 2001). Kamas und Preston (2012) bestätigen den geschlechtsspezifischen Unterschied hinsichtlich der Ausprägungen Überplatzierung und Überpräzision (Kamas und Preston, 2012). Andere Forschungen konnten allerdings keine signifikanten Unterschiede zwischen den Geschlechtern feststellen (Mannes und Moore, 2013, Cueva et al., 2019).

### 2.2.2. Ausprägungen des Over-Confidence Bias

Moore und Schatz (2017) gehen davon aus, dass der Over-Confidence Bias verschiedene Erscheinungsformen aufweist und die obige allgemeine Definition des Over-Confidence Bias einer Spezifizierung bedarf. Wie in Abbildung 2 grafisch dargestellt, lassen sich die drei Ausprägungen Überpräzision, Überschätzung und Überplatzierung differenzieren (Moore und Schatz, 2017).

Die erste Ausprägung umfasst die Überschätzung der eigenen Fähigkeiten und Eigenschaften eines Individuums und dessen Tendenz zur Annahme, besser zu sein als der Realität entspricht, und bezieht sich vor allem auf die Prognose zukünftiger Erfolge (Moore und Schatz, 2017).

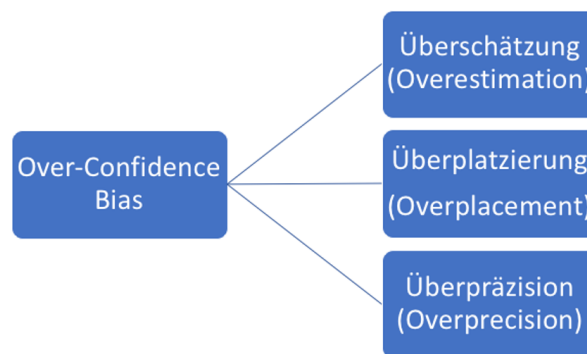
Eine Überschätzung der eigenen Schnelligkeit, Attraktivität, Leistung oder Intelligenz kann dazu führen, dass die benötigten Ressourcen und die Kontrolle über die Situation verzerrt eingeschätzt werden. So lassen sich einige Unterformen der Überschätzung differenzieren, z.B. Selbsterhöhung, Kontrollillusion, Planungsfehler und übermäßiger Optimismus. Die Selbsterhöhung umfasst die Neigung, sich selbst übersteigert positiv wahrzunehmen und davon auszugehen, dass das eigene soziale Umfeld anderen gegenüber höherwertig ist. Die Kontrollillusion beschreibt die falsche Wahrnehmung, eine konkrete Situation beeinflussen und beherrschen zu können, wobei der sog. Planungsfehler die Überschätzung des Arbeitstempos bei Projekten oder Aufgaben beschreibt. Oft wird dieser Begriff im Kontext von großen Bauprojekten verwendet. Des Weiteren existiert in diesem Zusammenhang übermäßiger Optimismus. Dieser bezieht sich auf die Einstellung eines Individuums in Bezug auf sein zukünftiges Leben, insbesondere auf das Eintreten unwahrscheinlicher, positiver Ereignisse. (Bazerman und Moore, 2013)

Die Unterschätzung der eigenen Performance ist hingegen durch Pessimismus geprägt. Individuen überschätzen sich i.d.R. bei schwierigen Aufgaben und unterschätzen sich tendenziell bei simplen Aufgaben (Moore und Healy, 2008).



**Tabelle 1:** Merkmale der kognitiven Prozesstypen 1 und 2 (in Anlehnung an Evans und Stanovich, 2013, S.225)

Typ 1	Typ 2
Schnell	Langsam
Hohe Kapazität benötigt	Geringe Kapazität benötigt
Parallel	Seriell
Unbewusst	Bewusst
Verzerrte Reaktion	Normative Reaktion
Kontextualisiert	Abstrakt
Automatisch	Kontrolliert
Ganzheitlich	Analytisch
Assoziativ	Regelbasiert
Erfahrungsbasierte Entscheidungsfindung	Regelkonforme Entscheidungsfindung
Geringe Anforderungen an die kognitive Leistungsfähigkeit	Hohe Anforderungen an die kognitive Leistungsfähigkeit



**Abbildung 2:** Ausprägungen des Over-Confidence Bias (eigene Darstellung)

Die zweite Ausprägung, Überplatzierung, bezieht sich auf den Vergleich der eigenen Fähigkeiten in Relation zu anderen Individuen und beschreibt die Annahme, besser als andere zu sein (Moore und Schatz, 2017). Insbesondere in Wettbewerbssituationen tendieren Individuen dazu, sich besser als andere einzuschätzen und einen höheren Rang zuzuweisen (Moore und Schatz, 2017). Diese Form des Over-Confidence Bias erklärt fehlerhafte Einschätzungen in Kriegen oder Verhandlungen, in denen die Wehrmacht oder Verhandlungsstärke relativ zu anderen überschätzt wird (Bazerman und Moore, 2013). In diesem Zusammenhang wird häufig auch der Better-than-Average-Effekt (BTA) erwähnt, der sich darauf bezieht, dass die Mehrheit der Menschen einer Gruppe behauptet, besser als der Median zu sein, was mathematisch nicht möglich ist (Logg et al., 2018). In diesem Fall wird ein Bezug der eigenen Wahrnehmung zu der tatsächlichen Leistung einer Gruppe hergestellt. Bei der Überplatzierung wird die subjektive Überschätzung der eigenen Fähigkeiten gegenüber anderen mit der Realität abgeglichen und dadurch genauer gemessen (Logg et al., 2018). Wenn sich Individuen schlechter als andere einschätzen, wird von einer Unterplatzierung der eigenen Fähigkeiten gesprochen und dies tritt insbesondere bei schwierigen Aufgaben auf (Moore und Healy, 2008, Moore und Schatz, 2017).

Die dritte Ausprägung des Over-Confidence Bias um-

fasst die Überpräzision. Hierbei besteht eine übersteigerte Überzeugung, die Wahrheit zu kennen und dass die eigenen Urteile richtig sind (Moore und Schatz, 2017). Dies geschieht in Befragungen insbesondere durch die Angabe von enger Konfidenzintervalle, was impliziert, dass sich derjenige zu sicher ist, die Wahrheit zu kennen (Bazerman und Moore, 2013). Diese Form ist leicht zu beeinflussen, es besteht zudem das geringste wissenschaftliche Verständnis für diese Erscheinungsform, da es auf der Selbsteinschätzung der Probanden basiert, z.B. anhand der Abgleichung der Antwort mit einem vorher angegebenen Konfidenzintervall (Moore und Schatz, 2017). Analog dazu lässt sich in der Over-Confidence-Forschung nur ein kleiner Teil der Überpräzision zuordnen. Im Jahr 2017 bestanden nach Moore und Schatz (2017) 19% der Over-Confidence-Forschungsstudien bei Google Scholar aus Überpräzision, 21% beschäftigten sich mit der Überplatzierung und etwa 60% mit der Überschätzung (Moore und Schatz, 2017). Unterpräzision tritt ebenfalls eher selten auf und entsteht dadurch, dass die Art der Fragestellung und Methodenwahl eine Variabilität in dem Grad der Überpräzision erzeugt (Moore und Healy, 2008). Der Einfluss des Schwierigkeitsgrads von Aufgaben auf Over- bzw. Under-Confidence ist bekannt als sog. Hard-Easy-Effekt und basiert auf den Erkenntnissen von Lichtenstein und Fischhoff (1977). Menschen neigen demnach dazu,

ihre Leistung bei schwierigen Aufgaben zu überschätzen und bei einfachen Aufgaben zu unterschätzen (Lichtenstein und Fischhoff, 1977, Moore und Schatz, 2017).

### 3. Einfluss des Over-Confidence Bias in strategischen Entscheidungsprozessen

#### 3.1. Entstehung von Over-Confidence im Entscheidungsprozess

##### 3.1.1. Ursachen für Over-Confidence in Entscheidungen

Der folgende Abschnitt behandelt die Entstehung des Over-Confidence Bias in strategischen Entscheidungsprozessen. Wie bereits in 2.2.1 erwähnt, ist die Entstehung eines Bias allgemein durch die fehlerhafte Anwendung von Heuristiken zu begründen (ICV, Internationaler Controller Verein e.V.). Nach Bazerman und Moore (2013) entsteht Over-Confidence primär aus der fehlerhaften Anwendung der Bestätigungsheuristik (Bazerman und Moore, 2013). Der Over-Confidence Bias kann jedoch aufgrund seiner weiten Verbreitung vielfältige Ursachen haben (Russo und Schoemaker, 2016). Die Ursachen von Over-Confidence lassen sich in kognitive, motivationale, physiologische und umweltbedingte Faktoren kategorisieren (Russo und Schoemaker, 1992, Russo und Schoemaker, 2016).

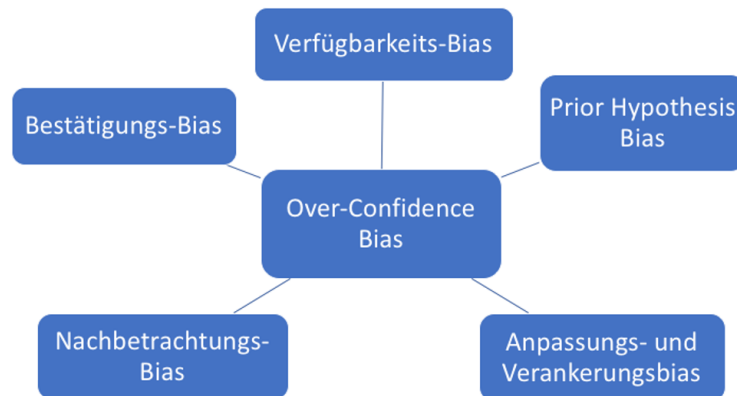
Wie in Abbildung 3 dargestellt, kann der Over-Confidence Bias auf der kognitiven Ebene als Nebenprodukt anderer kognitiver Prozesse interpretiert werden (Chen und Lee, 2003).

Die kognitiven Ursachen für Over-Confidence sind mit dem Verfügbarkeits-Bias, der Anpassung und Verankerung, sowie dem sog. Prior Hypothesis Bias eng verbunden (Chen und Lee, 2003). Eine der Hauptursachen für die Entstehung des Over-Confidence Bias stellt analog zu den Erkenntnissen von Bazerman und Moore (2013) der Verfügbarkeits-Bias dar (Russo und Schoemaker, 1992, Bazerman und Moore, 2013). Aufgrund seiner begrenzten Informationsverarbeitungskapazität durch das System 2 sind dem Entscheidungsträger nicht alle möglichen Optionen bewusst, sodass intuitiv lediglich die Optionen ausgewählt werden, die dem Entscheidungsträger nahe liegen und greifbar für ihn sind (Russo und Schoemaker, 1992). So neigen Entscheidungsträger dazu, sich bei ihrer Entscheidung auf aktuelle oder leicht aus dem Gedächtnis abrufbare Informationen zu verlassen (Chen und Lee, 2003). Wenn Informationen in großen Mengen vorhanden sind, tendieren Entscheidungsträger nach Chen und Lee (2003) ebenfalls dazu, sich an früheren Hypothesen oder Überzeugungen zu orientieren, um nach einer Bestätigung der wahrgenommenen Informationen zu suchen (Chen und Lee, 2003). Diese Verzerrungen, Prior Hypothesis Bias und Bestätigungs-Bias, veranlassen die Entscheidungsträger dazu, wichtige Informationen zu ignorieren und eine Überpräzision bezüglich ihrer Urteile zu entwickeln (Chen und Lee, 2003). Ebenso haben Entscheidungsträger Schwierigkeiten damit, sich alle möglichen Szenarien vorzustellen. Da es ihnen nicht gelingt, sich sämtliche Entwicklungen im Kontext wichtiger zukünftiger Ereignisse vorzustellen, werden Entscheidungsträger unangemessen zuversichtlich hinsichtlich

derjenigen Optionen, die sie in Erwägung ziehen und für wahrscheinlich halten (Russo und Schoemaker, 1992, Chen und Lee, 2003). Der Verankerungs-Bias beschreibt die Tendenz, an einem bestimmten Wert oder einer Idee festzuhalten und sich nicht ausreichend von dieser zu entfernen (Tversky und Kahneman, 1974, Chen und Lee, 2003). Nach Russo und Schoemaker (1992) kann der Nachbetrachtungs-Bias ebenfalls dazu führen, dass Individuen annehmen, ihr Umfeld sei leichter vorherzusagen als es der Realität entspricht. Rückblickend neigt der Entscheidungsträger zu der Überzeugung, den Ausgang eines vergangenen Ereignisses bereits zuvor gut eingeschätzt zu haben (Russo und Schoemaker, 1992).

Darüber hinaus wird davon ausgegangen, dass die Entstehung von Over-Confidence sich auf motivationale Aspekte beziehen kann (Russo und Schoemaker, 2016). Es lassen sich hierfür Ursprünge in der Selbstdarstellung nach außen finden. Erhöhtes Selbstbewusstsein impliziert demnach Vertrauenswürdigkeit, Glaubwürdigkeit und eine hohen Rang, so signalisieren z.B. Politiker wie Donald Trump durch ihr imposantes Verhalten Kompetenz (Moore und Schatz, 2017). Eine intrinsische Motivation lässt sich nur unter bestimmten Bedingungen als Auslöser für Over-Confidence identifizieren, da die Wirkung von Motivation auf die Überschätzung mehr von individuellen Konstrukten der Bewertung als von selbstverstärkender Täuschung bestimmt wird (Logg et al., 2018). Eine mögliche Motivation könnte in diesem Zusammenhang die Erlangung eines höheren sozialer Rangs darstellen (Anderson et al., 2012). Nach Bazerman und Moore (2013) lässt sich die Entstehung von Überpräzision, einer der drei Ausprägung des Over-Confidence Bias, darauf zurückführen, dass Individuen interne Dissonanzen verringern wollen. Um diesen inneren Konfliktzustand zu lösen, neigen Individuen dazu, ihr eigenes Meinungsbild den Umweltbedingungen anzupassen, selbst wenn dies nicht ihrer tatsächlichen Meinung entspricht (Bazerman und Moore, 2013).

Bezüglich der physiologischen Ursachen lässt sich ergänzen, dass Euphorie als Form von Wohlbefinden, ausgelöst durch Erfolgserlebnisse, zu Over-Confidence führen kann (Russo und Schoemaker, 1992). So besteht nach Schilke und Reimann (2007) ein Zusammenhang zwischen der Aktivierung des Belohnungssystems, dem sog. Nucleus Accumbens, im Gehirn und Over-Confidence. Sie nehmen Bezug auf die Forschung von Peterson (2005), wonach die Entwicklung von Over-Confidence bei Investoren mit einer Aktivierung dieser Gehirnregion zusammenhängt. Die Aktivierung dieses Belohnungssystems kann dazu führen, dass ein kognitiver Kreislauf entsteht, da durch eine erfolgreiche Investitionstätigkeit die Aktivierung der entsprechenden Gehirnsektion erfolgt (Peterson, 2005, Schilke und Reimann, 2007). Dadurch entwickelt der Investor Over-Confidence bezüglich seiner Fähigkeit, Risiken und Chancen richtig einzuschätzen und gute Investitionsentscheidungen zu treffen, was sich auch auf zukünftige Entscheidungen auswirkt (Peterson, 2005, Schilke und Reimann, 2007). Darüber hinaus beeinflussen Alkohol, Drogen und einige Stimmungszustände das Entscheidungsverhalten und können zu einem Anstieg von Over-Confidence führen (Russo und Schoemaker, 2016). Sie



**Abbildung 3:** Kognitive Ursachen des Over-Confidence Bias (eigene Darstellung)

verzerren die Wahrnehmung und führen dazu, dass Individuen ihre Kontrolle überschätzen (Russo und Schoemaker, 2016).

Weitere umweltbedingte Faktoren können Over-Confidence erhöhen (Russo und Schoemaker, 2016). So kann z.B. ein hohes Maß an Schmeichelei und Meinungskonformität das Maß an Over-Confidence Bias bei CEOs bezüglich ihrer strategischen Urteils- und Führungsfähigkeit erhöhen, was dann zu einer verzerrten strategischen Entscheidungsfindung führt (Park et al., 2011). So kann ein Individuum Over-Confidence in Folge eines verfälschten Feedbacks seines Umfelds entwickeln (Park et al., 2011). Ebenso besteht ein Zusammenhang zwischen dem Erleben von Macht und dem Treffen von zu selbstbewussten Entscheidungen (Fast et al., 2012). Durch eine hohe Position in Unternehmen neigen Manager dazu, Over-Confidence im Zuge ihrer Laufbahn zu entwickeln, zum Teil weil Selbstbewusstsein eng mit der Signalisierung einer hohen Kompetenz verknüpft ist (Moore und Schatz, 2017). Dieses Erkenntnis hängt dabei eng zusammen mit den später in Kapitel 3.2.1 erläuterten gesellschaftlichen Chancen von Over-Confidence.

### 3.1.2. Einordnung in den strategischen Entscheidungsprozess

Auf Basis dieser Erkenntnisse wird Over-Confidence nun in das Entscheidungsprozessmodell nach Mintzberg et al. (1976) eingeordnet. In der ersten Phase, der Problemdefinition, wird nach der rationalen Entscheidungstheorie angenommen, dass ein Entscheidungsträger die vorliegenden Informationen ohne Fehler sammelt und analysiert (Mintzberg et al., 1976). Während in der Erkennungsroutine Probleme in der Umgebung identifiziert werden, beinhaltet die Routine der Diagnose die Entscheidung darüber, welche Art von Informationen gesammelt werden müssen, um den Prozess der Problemlösung einzuleiten (Mintzberg et al., 1976). In der Realität neigen Individuen dazu, bei der Formulierung und der Analyse der Umwelt Risiken nicht richtig zu identifizieren und zu konkretisieren; die subjektive Erwartung eines Entscheidungsträgers kann die Interpretation und Wahrnehmung von Problemen dabei erheblich beeinflussen (Schwenk, 1984). Schwenk (1984) ordnet u.a. die kogniti-

ven Verzerrungen Prior Hypothesis Bias sowie die Anpassung und Verankerung in diese Phase ein, welche bereits als Ursachen für Over-Confidence definiert wurden (Schwenk, 1984, Chen und Lee, 2003). Informationen werden ignoriert und Risiken nicht ausreichend wahrgenommen, wodurch der weitere Entscheidungsprozess beeinträchtigt wird, sodass ein Individuum in dieser Phase, ausgelöst durch diese vorangehenden Bias, Over-Confidence entwickeln kann (Tversky und Kahneman, 1974, Russo und Schoemaker, 1992).

In der zweiten Phase werden Alternativen generiert, entweder durch die Entwicklung oder das Suchen von Alternativen (Mintzberg et al., 1976). Kognitive Verzerrungen können hier bewirken, dass eine geeignete Alternative eliminiert, ignoriert oder nicht entwickelt wird. Dies kann auch dazu führen, dass lediglich eine einzige Alternative näher betrachtet wird (Schwenk, 1984). Dies kann zum Beispiel die Meinung des Entscheidungsträgers selbst sein, da aufgrund von Over-Confidence zusätzliche Informationen unter Betracht der Bestätigungs-Bias ignoriert werden (Russo und Schoemaker, 1992). Der Bestätigungs-Bias beeinflusst den Suchprozess eines Individuums insofern, als dass Informationen, welche die eigene Meinung unterstützen überbewertet und Gegenargumente ignoriert werden (Russo und Schoemaker, 1992). Es entsteht eine einseitige Sicht, insbesondere bei komplexen und unsicheren Entscheidungen, sodass auch die Gewichtung der Argumente unterschiedlich ausfällt (Russo und Schoemaker, 1992).

In der letzten Phase, der Bewertung und Auswahl, lässt sich Over-Confidence verstärkt einordnen (Schwenk, 1984). Aus den vorher generierten Alternativen wird mittels eines Bewertungsverfahrens die bestmögliche Alternative ausgewählt. Dieser Schritt ist somit abhängig von der Rationalität vorgelagerter Phasen, denn im Falle des Auftretens Bias bei der Generierung von Alternativen ist die eigentlich optimale Option womöglich gar nicht mehr verfügbar (Schwenk, 1984). Ebenso kann die Kontrollillusion, die eine Unterform der Überschätzung darstellt, in diese Prozessphase eingeordnet werden (Schwenk, 1984). Die Einordnung von Over-Confidence in die Phase der Bewertung und Auswahl umfasst somit das Ergebnis aus dem Zusammenspiel verschiedener vorangegangener kognitiver Verzerrungen, wie in Tabelle 2

dargestellt ist. Nach Larrick (2004) können Biases mehrere verschiedene Dimensionen aufweisen, wodurch eine klare Zuordnung zu den entsprechenden Ursachen erschwert wird (Larrick, 2004).

### 3.2. Konsequenzen von Over-Confidence im Entscheidungsprozess

#### 3.2.1. Chancen von Over-Confidence

In einem strategischen Entscheidungsprozess kann Over-Confidence somit durch die Anwendung von Heuristiken, auch als Nebenprodukt anderer kognitiver Verzerrungen auftreten (Chen und Lee, 2003). Im Folgenden werden unter Kenntnis der Entstehungsfaktoren von Over-Confidence die Vorteile auf den Entscheidungsprozess, den Entscheidungsträger und das Unternehmen näher betrachtet. Insbesondere auf der Ebene des Entscheidungsträgers lassen sich, sowohl intra- als auch interpersonell Chancen von Over-Confidence identifizieren (Moore und Schatz, 2017). Auf der intrapersonellen Ebene ist belegt, dass Over-Confidence zu einer Steigerung des allgemeinen Wohlbefindens eines Individuums führen kann (Moore und Schatz, 2017). Moore und Schatz (2017) gehen dabei auf die Forschung von Armor et al. (2008) ein, wonach Individuen davon ausgehen, dass Optimismus an sich das tatsächliche Ergebnis ihres Handelns positiv beeinflussen kann (Armor et al., 2008, Moore und Schatz, 2017). Auch auf zwischenmenschlicher Ebene lassen sich Chancen für den Entscheidungsträger nachweisen, insbesondere im Kontext von Selbstständigkeit und Unternehmensneugründungen. Typischerweise weisen Gründer im Vergleich zu Managern großer Organisationen ein erhöhtes Maß an Over-Confidence auf, wie bereits in Kapitel 2.2.1 erwähnt (Busenitz und Barney, 1997). Für ein erfolgreiches Agieren in den neuen Märkten müssen Gründer in der Lage sein, ihren eigenen Fähigkeiten zu vertrauen um schwer abschätzbare Risiken einzugehen und sich von dem Ideal der rationalen Entscheidung zu distanzieren (Beck, 2014). Darüber hinaus gelingt es Individuen mit übersteigertem Selbstbewusstsein, Kreditgeber und Aktionäre von ihren oft innovativen Ideen zu begeistern, wodurch ein besserer Zugang zum Kapitalmarkt ermöglicht wird (Beck, 2014). Außerdem sind Unternehmensgründer weniger dazu geneigt, andere in ihrer sozialen Gruppe zu imitieren (Bernardo und Welch, 2001). Statt Trends zu Folgen werden zusätzliche Informationen und Ideen generiert, welche die soziale Gruppe bzw. das Unternehmen bereichern können (Bernardo und Welch, 2001).

Darüber hinaus stellen auch Manager, die zu Over-Confidence tendieren, innovativere Führungskräfte dar (Hirshleifer et al., 2012). Das Übervertrauen hängt mit einer erhöhten Risikobereitschaft der Manager zusammen, welche die Implementierung neuartiger Prozess oder Produkte erfordert (Hirshleifer et al., 2012). Dies gilt dann, wenn das Unternehmen an sich als innovativ einzustufen ist (Hirshleifer et al., 2012). Analog dazu geben selbstbewusste Manager deutlich mehr für Forschungsinvestitionen aus, als Manager, die eine gute Kalibrierung aufweisen (Zavertiaeva

et al., 2018). Dies kann sich indirekt positiv auf finanzielle Determinanten des Unternehmens, wie z.B. Liquidität und Profitabilität, auswirken (Zavertiaeva et al., 2018). Galasso und Simcoe (2011) konnten ebenfalls bestätigen, dass CEOs, die Over-Confidence aufweisen das Risiko für ein mögliches Scheitern unterschätzen und dadurch eher dazu geneigt sind, ihre Unternehmen technologisch innovativ auszurichten (Galasso und Simcoe, 2011).

Darüber hinaus erlangen Individuen durch ein übersteigertes Selbstbewusstsein ein höheres Vertrauen und signalisieren ihren Mitmenschen eine höhere Kompetenz, wodurch sie Vorteile bei der Erlangung hoher Positionen, auch in öffentlichen Ämtern, besitzen (Magee und Frasier, 2014, Moore und Schatz, 2017). Ebenso kann Over-Confidence zu einer Verbesserung des sozialen Status führen. Anderson et al. (2012) konnten feststellen, dass Individuen mit einem übersteigerten Selbstbewusstsein von anderen Gruppenmitgliedern ein höherer sozialer Status zugeordnet wird (Anderson et al., 2012). Diese Assoziation von Over-Confidence mit einer sozialen Stellung bezieht sich vor allem auf Tätigkeiten, deren Leistung von anderen der Gruppe nicht genau zu erkennen ist und bei denen somit schwer einzuschätzen ist, ob tatsächlich eine hohe Kompetenz vorliegt (Anderson et al., 2012). Selbst wenn Gruppenmitglieder erkennen, dass ein Gruppenmitglied oder ein Manager zu Selbstüberschätzung neigt, erfolgt nach Anderson et al. (2012) in dem Zusammenhang kein Verlust des sozialen Status (Anderson et al., 2012). Analog dazu besagt die sog. Status-Verbesserungstheorie, dass der soziale Nutzen von Over-Confidence die möglichen Statuskosten überwiegt, wodurch der Nutzen von Over-Confidence in höherrangigen Positionen deutlich wird (Kennedy et al., 2013).

#### 3.2.2. Risiken von Over-Confidence

Der Over-Confidence Bias scheint sich folglich in manchen Situationen als vorteilhaft für einen Entscheidungsträger und das Ergebnis des Prozesses zu erweisen. Dennoch birgt er ein erhebliches Gefahrenpotential für den Entscheidungsträger, ganze Konzerne oder Staaten. Denn Over-Confidence wird u.a. als Ursache für vergangene Kriege, Börsenblasen, unnötige Gerichtsverfahren, hohe Insolvenzraten sowie für das Scheitern von Unternehmensübernahmen gesehen (Bazerman und Moore, 2013). Der Bias gilt aufgrund der Folgen der daraus resultierenden Fehlentscheidungen als „more potentially catastrophic“ (Plous, 1993, S.217) als alle anderen kognitiven Verzerrungen.

Durch Überpräzision neigt der Entscheidungsträger dazu, Ratschläge von anderen nur zögerlich anzunehmen, was dazu führt, dass die Einschätzungen anderer nicht angemessen in ihren Entscheidungen berücksichtigt werden – insbesondere solche, die stark von ihren eigenen abweichen (Minson et al., 2011). Ebenso lässt sich bei der Überpräzision die Gefahr eines möglichen Reputationsverlusts identifizieren. Denn wie bereits zuvor beschrieben, signalisiert ein erhöhtes Selbstbewusstsein Mitarbeitern, Aktionären oder anderen Stakeholdern eine hohe Kompetenz und Vertrauenswürdigkeit und bei einer Nicht-Erfüllung dieser Erwartung

**Tabelle 2:** Einordnung von Over-Confidence und verknüpfter Bias in den Entscheidungsprozess (in Anlehnung an Schwenk, 1984, S.115)

Problemidentifikation	Generierung von Alternativen	Evaluation und Auswahl
Prior Hypothesis Bias Anpassung und Verankerung	Bestätigungs-Bias	Kontrollillusion Over-Confidence (i.e.S.)

kann es dazu führen, dass der Entscheidungsträger an Ansehen verliert (Bazerman und Moore, 2013). Over-Confidence kann darüber hinaus dazu führen, dass in großen Projekten zu wenig Zeit und Ressourcen für die Fertigstellung antizipiert werden, wodurch die Kosten und der damit verbundene Aufwand für diese Projekte stark ansteigen können (Son und Rojas, 2010).

Nach Camerer und Lovallo (1999) neigen Entscheidungsträger dazu, ihre Chancen auf einen relativen Erfolg zu überschätzen und häufiger in Märkte einzutreten. Ihr experimenteller Ansatz erklärt, dass durch die Überplatzierung und einer damit verbundenen Vernachlässigung anderer Wettbewerber ein Teil von Unternehmensgründungen kurz nach dem Markteintritt scheitert (Camerer und Lovallo, 1999). Cain et al. (2015) stellen in diesem Kontext fest, dass Unternehmer den Wettbewerb nicht vernachlässigen, sondern vielmehr unterschätzen (Cain et al., 2015). Deshalb tendieren Unternehmensgründer durch Over-Confidence dazu, trotz Warnungen einen Markt ohne ausreichende Analyse zu bearbeiten und riskieren dadurch ein Scheitern ihres Vorhabens (Cain et al., 2015).

Bei Unternehmenskäufen und Fusionen lässt sich Over-Confidence ebenfalls als Risikofaktor identifizieren. Die sog. Hybris-Hypothese nach Roll (1986) begründet die Unternehmensübernahmen, bei denen der Preis über dem Marktpreis liegt, mit Over-Confidence des Managers des Käuferunternehmens (Roll, 1986). Wenn Over-Confidence besteht, ist das Ausmaß der Fusionen von der vom jeweiligen CEO wahrgenommenen Unterbewertung ihres Unternehmens als eigenständiges Unternehmen und der Überschätzung der zukünftigen Erträge aus dem Unternehmenskauf abhängig (Malmendier und Tate, 2008). So neigen übersteigert selbstbewusste CEOs dazu, qualitativ minderwertige Akquisitionen durchzuführen, insbesondere wenn ihr Unternehmen über umfangreiche interne Ressourcen verfügt (Malmendier und Tate, 2008). Aufgrund des Over-Confidence werden CEOs nach positiven Erfahrungen aus vergangenen Akquisitionen eher wieder akquirieren, selbst wenn diese Geschäfte wertvernichtend sind (Billet und Qian, 2008).

Darüber hinaus führt Over-Confidence nach Daniel und Hirshleifer (2015) in Kapitalmärkten zu risikobehafteten Entscheidungen. Investoren nehmen demnach an, dass die Genauigkeit ihres Wissens über die Vermögenswerte höher als die eines durchschnittlichen Anlegers ist (Daniel und Hirshleifer, 2015). In Folge dessen nehmen sie an, dass sie Renditen über dem Marktdurchschnitt erzielen können und höhere Erfolgchancen besitzen als andere Anleger (Daniel und Hirshleifer, 2015).

Im Bankensektor ist zu beobachten, dass Over-Confidence zu einer erhöhten Vergabe von Krediten durch die jeweilige Bank führt und dabei vor allem die Vorteile der Aktionäre in Betracht gezogen werden (Ho et al., 2016). Dies kann in Krisensituationen dazu führen, dass die jeweilige Bank einen größeren Kapitalverlust erleidet (Ho et al., 2016).

Auch bei der Personalauswahl stellt Over-Confidence nach Roulin et al. (2014) einen entscheidenden Faktor hinsichtlich der Bewertung der Kandidaten dar. Durch ein übermäßiges Vertrauen in ihre Urteile neigen Personaler bzw. Interviewer dazu, die Annahme zu entwickeln, dass sie den Bewerber regelrecht durchschauen können (Roulin et al., 2014). Sie entwickeln Überpräzision hinsichtlich ihres intuitiven Urteils über das Gegenüber und nehmen aufgrund der eigenen Kompetenzen an, keine Verzerrung in der Bewertung aufzuweisen (Roulin et al., 2014).

Auch im Bereich der medizinischen Diagnose lässt sich Over-Confidence als Ursache für Fehlentscheidungen beobachten. Nach Berner und Graber (2008) erfolgt die medizinische Diagnose durch die Nutzung automatischer kognitiver Prozesse zwar zum großen Teil korrekt, Abweichungen lassen sich z.T. allerdings durch Over-Confidence erklären. Ärzte unterschätzen demnach die Wahrscheinlichkeit, dass ihre Diagnosen falsch sind und überschätzen ihre Fähigkeiten, wodurch medizinische Fehldiagnosen erzeugt werden können (Berner und Graber, 2008). Ein Risiko birgt die kognitive Verzerrung ebenso im universitären Kontext. Nach Wüst und Beck (2018) entwickeln Studierende Over-Confidence bei der Einschätzung ihrer eigenen Prüfungsleistung. Dadurch wenden sie zu wenig Zeit für ihre Prüfungsvorbereitung auf und durch die Enttäuschung über ein unerwartet schlechtes Ergebnis kann die Motivation der Studierenden langfristig sinken, was im Zweifelsfall zu einem Abbruch des Studium führen kann (Wüst und Beck, 2018).

#### 4. Potentielle Maßnahmen zur Verhinderung von Over-Confidence

##### 4.1. Vorstellung möglicher Lösungsansätze

Es wird deutlich, dass der Over-Confidence Bias zu erheblichen Risiken in strategischen Entscheidungsprozessen führen kann. Der folgende Unterabschnitt befasst sich deshalb mit möglichen Lösungsansätzen und Methoden für den Umgang mit dem Over-Confidence Bias auf Basis der bestehenden Forschung. Die Entscheidungsforschung befindet sich derzeit in einer Expansionsphase, ausgehend von der

Grundlagenforschung, die sich mit der Identifikation und Aufklärung von Bias beschäftigt, zur Entwicklung und Prüfung der praktischen Lösungsansätzen (Morewedge et al., 2015). Das sog. Debiasing erfordert in diesem Rahmen eine bewusste Entkopplung der intuitiven Verarbeitung durch den Typ 1 und den Übergang zur analytischen Verarbeitung von Informationen durch Typ 2 (Croskerry et al., 2013a). Die Debiasing-Techniken umfassen eine einzelne oder ein Bündel von Strategien, welche dazu dienen, die kognitive Verzerrung oder die daraus resultierenden Effekte zu verhindern oder abzuschwächen (Larrick, 2004, Correia, 2018). Debiasing-Interventionen stellen effektive kostengünstige Methoden für die Verbesserung der Entscheidungsfindung und der Reduktion der kostenintensiven Fehler dar, welche durch die kognitiven Entscheidungsverzerrungen entstehen (Morewedge et al., 2015). Der Over-Confidence Bias hat sich dabei in der Vergangenheit insgesamt als robust und resistent gegenüber verschiedenen Debiasing-Techniken erwiesen (Fischhoff, 1982, Ferretti et al.). Dennoch gilt eine gute Kalibrierung als trainierbare Fähigkeit (Arkes et al., 1987, Russo und Schoemaker, 1992). Durch eine Studie über Debiasing von Over-Confidence in der Finanzbranche von Kaustia und Perttula (2012) konnte hinsichtlich der oben erwähnten Differenzierung des Konstrukts bestätigt werden, dass die unterschiedlichen Formen von Over-Confidence sich durch unterschiedliche Methoden entkoppeln lassen (Kaustia und Perttula, 2012). Nach Soll et al. (2015) kann das Debiasing durch die Modifikation des Entscheidungsträgers oder durch die Modifikation des Umfelds erfolgen. Durch die erste Vorgehensweise soll ein Individuum mit einer Kombination aus Wissen und nützlichen Instrumenten versorgt werden, um es bei der Überwindung von Entscheidungsproblemen zu unterstützen (Soll et al., 2015). Wenn hingegen die Umwelt modifiziert wird, soll dadurch der Kontext, in welchem Entscheidungen getroffen werden, angepasst werden (Soll et al., 2015).

Im Folgenden werden zentrale Debiasing-Techniken genannt, die sich im Umgang mit Over-Confidence als effektiv erwiesen haben. Zu den vorbeugenden Maßnahmen im Rahmen des Debiasing von Over-Confidence lässt sich zunächst die Berücksichtigung von Gegenargumenten einordnen. Bei dieser Strategie geht es darum, Argumente aufzuzeigen, die gegen die eigene Meinung sprechen, wodurch mehrere Sichtweisen kombiniert werden und die Entscheidungsqualität gesteigert wird (Russo und Schoemaker, 1992). Eine Gegenargumentation ermöglicht eine Erweiterung der beschränkten Sichtweise eines Individuums darauf, dass die eigenen Gründe die einzig verfügbaren sind, sodass auch alternative Perspektiven in die Entscheidungsfindung miteinbezogen und die durch Over-Confidence entstandenen Fehler behoben werden (Kahneman, 2011, ICV, Internationaler Controller Verein e.V.). Es besteht hier allerdings die Gefahr eines gegenläufigen Effekts, wenn zu viele Argumente betrachtet werden (Larrick, 2004).

Darüber hinaus kann in Gruppenentscheidungen ein sog. Advocatus Diaboli eingesetzt werden, um die Qualität der Entscheidung durch die Reduktion von Over-Confidence zu

verbessern. Bei der Methode haben ein oder mehrere Mitglieder der Gruppe die Aufgabe, nur Gegenargumente und Zweifel an der favorisierten Entscheidung aufzuführen (Russo und Schoemaker, 2016). Die Tatbestände der Entscheidung sollen durch diesen Advocatus Diaboli kritisch hinterfragt werden (ICV, Internationaler Controller Verein e.V.).

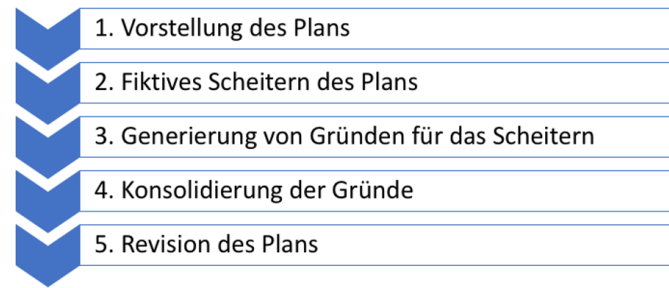
Eine weitere Methode, der vorrausschauende Rückblick, beschreibt die Betrachtung einer zu treffenden Entscheidung in die Zukunft, von der angenommen wird, dass sie gescheitert ist (Russo und Schoemaker, 2016). Dann wird zurückgeblickt um diesen Misserfolg oder Erfolg zu erklären und Gründe dafür zu generieren, wobei festgestellt wurde, dass Individuen i.d.R. in diesem imaginären Rahmen mehr Argumente erzeugen (Mitchell et al., 1989, Russo und Schoemaker, 2016).

Die Pre-Mortem Analyse basiert ebenso auf diesem Perspektivenwechsel und befasst sich mit dem fiktiven Scheitern einer zukünftigen Entscheidung (Klein, 2007). Die Pre-Mortem-Methode unterscheidet sich von anderen Methoden, weil sie davon ausgeht, dass der Plan gescheitert ist, aber trotzdem den vorrausschauenden Rückblick nach Mitchell et al. (1989) berücksichtigt (Mitchell et al., 1989, Veinott et al.).

Im ersten Schritt machen sich die Teammitglieder mit dem Plan bzw. der favorisierten Entscheidungsalternative vertraut. Daraufhin wird den Mitgliedern mitgeteilt, dass die Idee ein Fiasko sei und der Plan nicht funktioniert hat. Die Aufgabe des Teams ist es, in Schritt drei selbständig Gründe für das Scheitern zu generieren und schriftlich zu fixieren. Nach einer anschließenden Konsolidierung der Gründe im Team erfolgt im letzten Schritt eine Überarbeitung des Plans im Hinblick auf die Schwächen. Während der Durchführung des Plans wird anhand der Pre-Mortem-Liste überprüft, ob eines der Szenarios eintritt und weitere Maßnahmen durchzuführen sind. (Veinott et al.)

In Abbildung 4 sind die notwendigen Schritte der Pre-Mortem-Analyse dargestellt, welche i.d.R. vor der Umsetzung einer favorisierten Entscheidungsalternative erfolgt (Veinott et al.).

Im Zusammenhang mit Over-Confidence konnte eine hohe Effektivität festgestellt werden und die Pre-Mortem-Analyse stellt somit ein effektives und simples Planungsinstrument dar, das Entscheidungsträgern sowohl vor als auch während Krisensituationen helfen kann, ihre Pläne zu bewerten (Klein, 2007, Veinott et al.). Nach der Durchsetzung einer Entscheidung kann ein Lernen aus Over-Confidence stattfinden, um wiederum in Zukunft bessere Entscheidungen treffen zu können. Diese Maßnahmen treten nach einer getroffenen Fehlentscheidung auf, in welcher Over-Confidence rückblickend erkannt wurde (Russo und Schoemaker, 2016). Auch der ICV empfiehlt die Implementierung von Feedback und eine darauf aufbauende Analyse der erfolgstreibenden Faktoren, um Over-Confidence zu reduzieren (ICV, Internationaler Controller Verein e.V.). Feedback impliziert dabei eine Rückmeldung und Bewertung der durchgesetzten Entscheidungen. Ein zeitliches Feedback kann dazu dienen, dass ein Individuum begreift, dass die eigenen Prognosen verzerrt



**Abbildung 4:** Schritte der Pre-Mortem-Analyse (eigene Darstellung)

sind (Russo und Schoemaker, 1992). Russo und Schoemaker (1992) gehen davon aus, dass ein zeitnahes und präzises Feedback Over-Confidence verbessern kann. Hingegen konnten andere Forschungen wie Chen et al. (2015) feststellen, dass Unternehmen, die von übersteigert selbstbewussten CEOs geführt werden, weniger auf ein korrigierendes Feedback reagieren, um die Genauigkeit der Prognosen zu verbessern (Chen et al., 2015). Neben den erwähnten Methoden kann auch das alleinige Bewusstsein über die Existenz des Over-Confidence Bias dazu führen, dass Manager ihre Entscheidungen hinterfragen und ihr Maß an Selbstüberschätzung gesenkt werden kann (Russo und Schoemaker, 1992). Bereits das Aussprechen von Warnungen kann nach Kaustia und Perttula (2012) zu einer Reduzierung von Überplatzierung in Form des Better-than-Average-Effekts führen (Kaustia und Perttula, 2012). Analog dazu empfiehlt der ICV Entscheidungen mit Skepsis zu begegnen und bewusst davon auszugehen, dass eine Überschätzung von Daten existiert (ICV, Internationaler Controllerverein e.V.).

#### 4.2. Implikationen für den Umgang mit Over-Confidence

Aufbauend auf der Kenntnis der verschiedenen Lösungsansätze fokussiert sich der folgende Unterabschnitt auf die resultierenden Empfehlungen für die Unternehmenspraxis. In Anbetracht der Häufigkeit von Over-Confidence sollte ein wachsendes Bewusstsein für den richtigen Umgang mit kognitiven Verzerrungen in Unternehmen entstehen, um der Entstehung von Fehlentscheidungen entgegenzuwirken. Es ist erforderlich, dass Unternehmen eine Kultur aufbauen, in der offen mit kognitiven Verzerrungen umgegangen und nötige Veränderungen implementiert werden. Nach Russo und Schoemaker (2016) lässt sich der Schritt, in dem ein Bias identifiziert und erkannt wird, von der Suche nach geeigneten Lösungsansätzen trennen. Denn um der kognitiven Verzerrung entgegenwirken zu können, muss sie zunächst festgestellt werden – „You can't fix it if you can't find it“ (Russo und Schoemaker, 2016, S.9). Ein kognitiver Aufwand ist dabei nach Kahneman (2011) zwingend notwendig, um Entscheidungen zu verbessern. Nach der Identifikation eines Bias sollte der Prozess verlangsamt werden und durch die Nutzung von Typ 2 eine Vermeidung von Fehlern stattfinden (Kahneman, 2011).

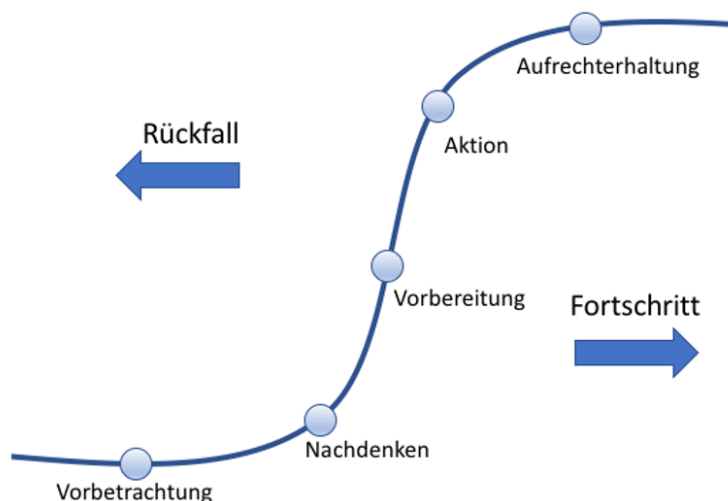
Auf Basis der Erkenntnisse von Prochaska et al. (1992) stellen Croskerry et al. (2013b) ein Modell auf, welches den

für ein effektives Debiasing erforderlichen Veränderungsprozess strukturiert (Prochaska et al., 1992, Croskerry et al., 2013b). Demnach beinhaltet das kognitive Debiasing Veränderungen, die weniger durch ein einzelnes Ereignis, als durch eine Abfolge von Schritten zustande kommen (Croskerry et al., 2013b). Der erste Schritt des Modells, welches in Abbildung 5 grafisch veranschaulicht ist, umfasst die Vor-Betrachtung, in dem einem Entscheidungsträger die kognitive Verzerrung, in diesem Fall Over-Confidence, zunächst bewusst wird (Croskerry et al., 2013b). Die Erkennung von Entscheidungsfehlern vergleicht Kahneman (2011) mit der medizinischen Diagnose, die ebenfalls einer präzisen Sprache erfordert (Kahneman, 2011).

In der zweiten Phase, dem Nachdenken, wird durch den Entscheidungsträger entschieden, ob eine Maßnahme gegen den Bias angewendet wird. Hierbei ist eine Motivation der Organisation bzw. des Entscheidungsträgers für den Veränderungsprozess erforderlich. Die dritte Phase umfasst die Vorbereitung, in der abgeschätzt wird, welchen Einfluss die Verzerrung auf das Entscheidungsverhalten besitzt. Dabei kann sich auf bisherige Debiasing-Interventionen im eigenen oder vergleichbaren Unternehmen bezogen werden. Umgesetzt wird die gewählte Strategie in Schritt vier, in welchem sich für eine bestimmte Methode entschieden wird. In dieser Phase der Aktion wird eine gewählte Maßnahme angewendet, wie eine der hier vorgetragenen Debiasing-Methoden. Es folgt im Idealfall eine Aufrechterhaltung des Debiasing. Bei einer erfolgreichen Implementierung dieses Prozesses wird von einem Fortschritt, bei Abbruch von einem Rückfall der Veränderung gesprochen. (Croskerry et al., 2013b)

## 5. Kritische Diskussion

Diese Arbeit soll einen Beitrag dazu leisten, die Auswirkungen von Over-Confidence in strategischen Entscheidungsprozessen zu bestimmen und mögliche Lösungsansätze ableiten zu können. Daher wird im Folgenden hinsichtlich dieser Zielsetzung und der zentralen Fragestellung, inwieweit Over-Confidence Einfluss auf strategische Entscheidungsprozesse besitzt, diskutiert. Die zur Beantwortung dieser Fragestellung durchgeführte Analyse führt zu dem Ergebnis, dass Over-Confidence einen Risikofaktor für Unternehmen darstellt und ein Debiasing angestrebt werden sollte.



**Abbildung 5:** Veränderungsmodell für ein langfristiges Debiasing (in Anlehnung an [Croskerry et al., 2013b](#), S.ii66)

Zu Beginn dieser Thesis wird deutlich, dass strategische Entscheidungen für das Unternehmen von großer Bedeutung und für den langfristigen Erfolg verantwortlich sind ([Nutt und Wilson, 2010](#)). Hinsichtlich der theoretischen Fundierung der Arbeit lässt sich das gewählte Modell von [Mintzberg et al. \(1976\)](#) näher betrachten. Es eignet sich insofern für die Analyse, als dass es die Kernschritte einer Entscheidung umfasst und sich gut für die Einordnung des im weiteren relevanten Over-Confidence Bias eignet, wie bereits bei [Schwenk \(1984\)](#). Dennoch ist es fraglich, inwiefern dieses Modell in der Praxis durchsetzbar und diese Struktur tatsächlich angesichts der Dynamik der Unternehmensumwelt erkennbar ist. Es ist zweifelhaft, ob die Prozessmodelle des präskriptiven Ansatzes ausreichen, um der realen Entscheidungsfindung gerecht zu werden. Zwar eignen sich präskriptive Modelle als Orientierungshilfe und zur Strukturierung der Sachverhalte, dennoch lässt sich nicht außer Acht lassen, dass es einer Ergänzung mit Erkenntnissen der deskriptiven Entscheidungstheorie bedarf ([Göbel, 2014](#)). Hinsichtlich der Aussagekraft der Theorie ist anzumerken, dass das Verfassungsdatum der hier berücksichtigten Literatur mehrere Jahrzehnte zurückliegt. Dies lässt sich dadurch begründen, dass die Entscheidungs-Forschung auf diesen Erkenntnissen beruht und die damaligen Erkenntnisse die Forschung bis zum jetzigen Zeitpunkt prägen. Neuere Erkenntnisse scheinen eine Annäherung verschiedener Disziplinen wie der Psychologie und Wirtschaftswissenschaften zu beinhalten.

Um das Auseinanderklaffen des menschlichen Denkens zu erklären, wird dabei durch die Forschung von [Kahneman \(2011\)](#) sowie [Evans und Stanovich \(2013\)](#) auf psychologischer Ebene hervorgehoben, dass das menschliche Denken sich in die Typen 1 und 2 aufteilt. Im Kontrast zu dem Typ-2-nutzenden Homo Oeconomicus stellt der Homo Heuristics angesichts dieses Mechanismus eine sinnvolle Weiterentwicklung der ökonomischen Theorie dar, um Abweichungen vom rationalen Ideal einer strategischen Entscheidung nachzuvollziehen.

Einen Ansatz zur Erklärung von Fehlentscheidungen bietet in dieser Arbeit der Over-Confidence Bias, welcher verschiedene Erscheinungsformen besitzt. Die Mehrdimensionalität des Konstrukts bewirkt, dass eine einheitliche Definition des Over-Confidence Bias nicht existiert. Stattdessen werden die Formen von Over-Confidence zum Teil vermischt. Ebenso bedarf es einer Abgrenzung zu ähnlichen Konstrukten, da es in der Literatur immer wieder zu Verwechslungen von z.B. Over-Confidence und Selbstdarstellung kommt.

Eine Besonderheit des Over-Confidence Bias stellt seine Verknüpfung zu anderen Bias des menschlichen Entscheidungsverhaltens dar. Der Bestätigungs-Bias, der Verfügbarkeits-Bias, die Verankerung, der Nachblick, der Prior Hypothesis Bias sowie die Kontrollillusion stellen kognitive Ursachen für Over-Confidence dar und formen ein eng verbundenes Netz, sodass sie sich nur bedingt einzeln betrachten lassen. Im Anbetracht der zahlreichen Ursachen für den Over-Confidence Bias konnte die Erwartung, dass Over-Confidence sich in eine einzelne Phase des Entscheidungsprozessmodells beschränkt, widerlegt werden, da Over-Confidence sich in seinen Unterformen in sämtlichen Entscheidungsphasen einordnen lässt. Die Ergebnisse dieser Arbeit unterstreichen dadurch die Mehrdimensionalität des Konstrukts. Sie unterscheidet sich von anderen Arbeiten, indem sie eine Einordnung der Ursachen in das Modell von [Mintzberg et al. \(1976\)](#) vornimmt und verdeutlicht, dass Over-Confidence aus einer Vielzahl kognitiver Verzerrungen, situativen Faktoren sowie motivationalen und physiologischen Zuständen resultiert. Die Ergebnisse sind mit den Erkenntnissen von [Larrick \(2004\)](#) vereinbar, wonach ein „one-to-one mapping“ ([Larrick, 2004](#), S. 319) von Ursachen zu den entsprechenden Bias nur begrenzt realisierbar ist.

Ebenso ist zu diskutieren, wie sich Over-Confidence grundsätzlich bewerten lässt und ob ein Debiasing überhaupt notwendig erscheint. Im Kontext von Unternehmensgründungen und Innovationen erscheint der Bias essentiell. Over-Confidence scheint eine notwendige Bedingung für



die Erreichung eines hohen sozialen Rangs und Ansehens im Unternehmen zu sein (Anderson et al., 2012). Die sog. Status-Verbesserungs-Theorie besagt zwar, dass der Nutzen durch Over-Confidence mögliche Kosten überwiegt (Kennedy et al., 2013). Dennoch ist zweifelhaft, ob Over-Confidence nicht in diesem Fall auch einen negativen Einfluss auf die Reputation eines Individuums haben kann (Bazerman und Moore, 2013).

Ebenso scheint Over-Confidence dem Individuum einen Nutzen zu stiften, indem es sich gut anfühlt, sich selbst zu überschätzen. Diesem Argument lässt sich begrenzt zustimmen, da anzunehmen ist, dass im Zuge von Optimismus auch schnell eine Enttäuschung erfolgen kann. Ein gutes Beispiel stellt in diesem Zusammenhang die Forschung von Wüst und Beck (2018) dar, die belegt, dass übersteigert selbstbewusste Studierende im Zuge der Enttäuschung über unerwartet schlechte Prüfungsleistungen dazu neigen, ihr Studium abzubrechen. Ebenso wird dadurch deutlich, dass Over-Confidence ein Konstrukt ist, das sämtliche Bereiche des menschlichen Lebens beeinflusst und sich nicht erst in der Berufspraxis nachweisen lässt. Zwar konnten Chancen für Entscheidungsträger erzielt werden, im Zuge strategischer Entscheidungen konnte jedoch kein Nachweis über eine tatsächliche Verbesserung durch Over-Confidence festgestellt werden. Plous (1993) hat meines Erachtens Recht, wenn er den Bias als „more potential catastrophic“ (Plous, 1993, S.217) beschreibt. Es wurden in Kapitel 3.2.2 zahlreiche negative Konsequenzen durch Over-Confidence u.a. bei Unternehmensfusionen, der Personalauswahl, bei Investitionsentscheidungen und großen Projekten nachgewiesen. Die Ergebnisse dieser Arbeit implizieren somit eine partielle Zustimmung der bestehenden Forschungsströmung, die Bias als durchaus schädlich einstufen, auch wenn zu beachten ist, dass unter gewissen Umständen ein Nutzen durch Over-Confidence entstehen kann. In Anbetracht der Erkenntnisse dieser Arbeit sind folglich beide Extrema abzulehnen und eine abwägende Haltung hinsichtlich der Bewertung von Over-Confidence im jeweiligen Kontext einzunehmen.

Demnach ist ein Ergebnis dieser Arbeit, dass ein Debiasing angestrebt werden sollte, um den Risiken entgegenzuwirken. Die zuvor in Kapitel 4.1 vorgestellten Methoden des Debiasing greifen dabei unterschiedliche Aspekte auf und eignen sich in Folge dessen für unterschiedliche Sachverhalte und lassen sich unterschiedlich bewerten.

Die Gegenargumentation konfrontiert den Entscheidungsträger mit Argumenten gegen die eigene Meinung. Dadurch wird auf der einen Seite ein Anstoß gegeben, die eigenen Ansichten zu hinterfragen. Es ist hier jedoch auf der anderen Seite einzuwenden, dass sich diese Methode nur bedingt für die Unternehmenspraxis eignet, da bei der Gegenargumentation durch Teammitglieder insbesondere in hierarchischen Unternehmen womöglich keine ausreichende Bereitschaft besteht, den Plan eines Managers zu kritisieren. Die Vielfalt der Argumente hängt von der Kreativität des Gegenargumentierenden ab. Dennoch bietet diese Methode eine informelle und flexible Möglichkeit des Debiasing, welche im Vergleich zu den anderen hier diskutierten Methoden

mit relativ wenig Aufwand verbunden ist und auch alleine durch den Entscheidungsträger angewendet werden kann.

Der *Advocatus Diaboli* kann dabei als die Weiterentwicklung der Gegenargumentation interpretiert werden. Indem ein organisationaler Rahmen für die Kritik am Vorhaben des Unternehmens geschaffen wird, wird das Debiasing zum Bestandteil einer Entscheidungsfindung. Auf Dauer kann dies jedoch zu einem enormen zeitlichen Verzug führen und zu einem erheblichen Aufwand für Unternehmen. So lässt sich kritisch fragen, inwiefern sich diese Methodik für kleinere Unternehmen eignet. Ein *Advocatus Diaboli* scheint sich unter Anbetracht der Beschaffenheit eher für große Projekte zu eignen, um Planungsfehlern in Gruppenentscheidungen vorzubeugen.

Die Pre-Mortem-Methode findet ebenfalls im Rahmen von Teamentscheidungen statt und ist angesichts der Konsequenzen von Over-Confidence in Entscheidungen dafür geeignet, verschiedene Sichtweisen zu generieren. Dadurch dass jedes Teammitglied angehalten ist, zu diskutieren wird Kreativität und Teamarbeit gefördert. Allerdings ist diese Methode zeit- und personalaufwendig. Angesichts der Stärke, dass möglichst viele verschiedene Sichtweisen generiert werden, scheint sie sich für eine Bandbreite von Problemsituationen zu eignen. Um auf die hier erläuterten Kontexte einzugehen, lässt sich nach eigenem Ermessen die Pre-Mortem-Methode insbesondere hinsichtlich Over-Confidence im Kontext von Unternehmensfusionen, Projekten, möglichen Markteintritten und Investitionsentscheidungen anwenden.

Die Feedback-Methode ist für ein Debiasing eher problematisch einzuschätzen, da sie in der Unternehmenspraxis nur bedingt integriert werden kann. Es ist zu befürchten, dass es in unstrukturierter Form auch schnell zu Missverständnissen und Verzerrungen kommen kann. Feedback ist somit abhängig von den zeitlichen, inhaltlichen und situativen Umständen und für strategische Entscheidungen nur begrenzt geeignet, da es sich i.d.R. um einmalige Situationen handelt und ein Lernprozess demnach nur begrenzt Wirkung zeigt. Hinsichtlich der verzerrten Wahrnehmung durch Over-Confidence lässt sich ein sachliches Feedback jedoch als ein Ansatz bewerten, der dazu dienen kann, Transparenz bezüglich der tatsächlichen eigenen Leistung zu schaffen. Im Zuge der Beschaffenheit und hohen Relevanz von einmaligen strategischen Entscheidungen sind demnach allerdings präventive Methoden, insbesondere die Pre-Mortem-Analyse, vorzuziehen.

Trotz der Bandbreite verfügbarer Gegenmaßnahmen gilt Over-Confidence immer noch als eine der robustesten kognitiven Verzerrungen. Gründe für die mögliche Ineffektivität der Debiasing-Interventionen können z.B. Konflikte zwischen der benötigten Zeit, den entstehenden Kosten und dem Nutzen des Debiasing sein. Ein möglicher Verlust des Expertenstatus und der eigenen Autorität können auf Seiten der übersteigert selbstbewussten Manager eine ablehnende Haltung gegenüber Debiasing-Techniken bewirken. Sie neigen dazu, an ihren Überzeugungen festzuhalten und möchten die alleinige Kontrolle über den Entscheidungsprozess nicht ver-

lieren. Ebenso könnte die Anwendung einer unterstützenden Technik den Entscheidungsträgern komplex erscheinen, da mögliche Vorteile erst verzögert auftreten. Aufgrund des Umfangs dieser Arbeit ist zu berücksichtigen, dass nur ein Bruchteil der existierenden Debiasing-Techniken analysiert wurde und auf eine umfassende Analyse verzichtet werden musste. Aus demselben Grund wurden hier diejenigen Methoden ausgewählt, die einen bedeutsamen Einfluss auf die Entscheidungsfindung besitzen und im Kontext von Over-Confidence bereits auf eine besondere Wirkung untersucht wurden. Denn die hier vorgestellten Methoden gleichen sich hinsichtlich des Merkmals, dass sie eine bewusste Entkopplung der kognitiv intuitiven Verarbeitung vom Typ 1 und den Übergang zur analytischen Verarbeitung des Typ 2 beinhalten, sodass die intuitiven Urteile überprüft werden können (Croskerry et al., 2013a).

Diese Arbeit geht dabei verstärkt auf die langfristige Integration des Debiasing in ein Unternehmen ein, indem das Veränderungsprozessmodell nach Croskerry et al. (2013b) vorgestellt wird. Dieses Modell dient der Orientierung und bietet eine Struktur für Veränderungen in Organisationen hinsichtlich der Integration von Debiasing-Instrumenten. Es bietet den Vorteil eines logischen, sequentiellen Vorgehens, die Effektivität ist auch hier wieder abhängig von der jeweiligen kognitiven Leistung des Entscheidungsträgers und eine Überprüfung der Anwendung in der Praxis bietet Platz für weitere Forschungen.

Zu den praktischen Implikationen dieser Arbeit zählt somit vor allem die Erkenntnis, dass Over-Confidence einen maßgeblichen, überwiegend negativen Einfluss auf das menschliche Entscheidungsverhalten besitzt. Es ist davon auszugehen, dass die Entscheidungsträger eines Unternehmens zu einem gewissen Ausmaß in ihrem Urteilsvermögen von Over-Confidence beeinflusst werden. Um Fehlentscheidungen vorzubeugen, sollten Manager in diesem Rahmen zwischen „deciding and doing“ (Russo und Schoemaker, 1992, S.16) unterscheiden können. Den ersten und wichtigsten Schritt stellt damit ein Bewusstsein für den Bias dar, wie anhand des vorgestellten Veränderungsmodells verdeutlicht wird. Es wird angesichts der damit verbundenen Problematik deutlich, dass es nicht das Ziel eines Unternehmens sein sollte, Over-Confidence komplett zu verhindern, sondern vielmehr Sensibilität für diesen Bias zu schaffen und ein Bewusstsein für mögliche Lösungsansätze herzustellen.

## 6. Fazit und Ausblick für weitere Forschung

In dieser Arbeit wurde der Einfluss von Over-Confidence auf den strategischen Entscheidungsprozess thematisiert und hinsichtlich der Entstehung, der Chancen und Risiken sowie möglicher Lösungsansätze untersucht. Das Ziel war es dabei, das Aufkommen dieser kognitiven Verzerrung nachzuvollziehen, um darauf aufbauend Lösungsansätze und Implikationen für das Management zu formulieren.

Im Rahmen der vorliegenden Arbeit konnte ein erheblicher Einfluss des Over-Confidence Bias auf den strategischen Entscheidungsprozess festgestellt werden. Im Unter-

nehmenskontext prägen Neuartigkeit und Unsicherheit die strategischen Entscheidungen. Eine rationale Entscheidung im Rahmen der normativen Theorie sieht sich angesichts des Wandels des Zwei-Typen-Ansatzes dem neuen Konstrukt des Homo Heuristicus gegenüber. In einem komplexen Geflecht verschiedener kognitiver Prozesse lässt sich die Entstehung von Over-Confidence nachweisen und somit an zahlreichen Stellen des Entscheidungsprozesses einordnen. Auf Basis der erläuterten Konsequenzen wird deutlich, dass Over-Confidence im Kontext von Innovationen, Unternehmensgründungen und sozialem Ansehen einen Vorteil darstellen kann. Dennoch scheinen die Risiken im Kontext von Projektplanungen, Fusionen oder Investitionsentscheidungen zu überwiegen. Bisher liegen keine Ergebnisse bezüglich einer tatsächlich verbesserten Entscheidung durch Over-Confidence vor. Als konkrete Gegenmaßnahmen konnte die Gegenargumentation, die Einführung eines *Advocatus Diaboli*, die *Pre-Mortem-Analyse* und das zeitnahe Feedback vorgestellt werden. Ein Veränderungsprozess impliziert die Integration dieser Ansätze in die Unternehmenspraxis auf Basis eines höheren Bewusstseins für Bias. Das Ergebnis dieser Arbeit verdeutlicht die hohe Relevanz eines integrativen Debiasing von Over-Confidence in strategischen Entscheidungsprozessen.

Aufbauend auf den Erkenntnissen dieser Arbeit lassen sich Ansatzpunkte bezüglich weiterer Forschungen hervorheben. Die bisherige Forschung befasst sich ausführlich mit den theoretischen Grundlagen der Heuristiken und Bias sowie der Strukturierung von Entscheidungsprozessen und hat sich in den letzten Jahrzehnten verstärkt auf die Begrifflichkeiten fokussiert. Es ist nun Zeit, sich intensiver mit den Methoden des Debiasing zu beschäftigen und zu hinterfragen, wie sich ein konkreter Bias in dem jeweiligen Entscheidungskontext bewerten lässt. Insbesondere im Hinblick auf die Anwendung der Debiasing-Instrumente ist das Ausmaß dieser Arbeit nur beschränkt dazu fähig, die Auswirkung und Effektivität der vorgestellten Methoden zu präsentieren und diskutieren. Es wäre von hohem wissenschaftlichen Interesse, die Wirkungsweise der Methoden näher zu betrachten und auf einzelne Individuen und Unternehmen zu beziehen. Es sollten Unterschiede zwischen verschiedenen Gruppen näher betrachtet und eventuell bestehende Besonderheiten unterschiedlicher Hierarchieebenen und Branchen näher erforscht werden. Ebenso wird es in Zukunft von hoher Bedeutung sein, wie Unternehmen in einem immer dynamischer werdenden Umfeld mit der kognitiven Verzerrung des Over-Confidence umgehen. Möglicherweise werden technologische Innovationen wie IT-basierte Entscheidungsunterstützungssysteme in Zukunft an Bedeutung für Unternehmen gewinnen. Neue technologiegestützte Verfahren und moderne Berichtssysteme führen zu einem Anstieg der verfügbaren Informationen in kürzester Zeit und bieten ein erhebliches Potential für die Entscheidungsfindung in Unternehmen. Es stellt sich die wichtige Frage, inwieweit sich der Einfluss von Over-Confidence in diesem Kontext verändert.

Aufgrund des Umfangs dieser Arbeit konnte nur ein Teilbereich des Over-Confidence Bias untersucht werden. Des-

halb sollte im Zuge zukünftiger Forschungsarbeiten die in dieser Arbeit präsentierten Ergebnisse erweitert und vertieft werden. Es sollte sich intensiv mit der Besonderheit der Entstehung von Over-Confidence auseinandergesetzt werden, um die damit verknüpfte Qualität strategischer Entscheidungen in einem immer schneller werdenden dynamischen Umfeld zu gewährleisten. Der Fokus sollte auf einer kontextbasierten Analyse kognitiver Verzerrungen liegen, sodass Entscheidungsträger zukünftig in der Lage sind, Debiasing-Techniken effizient einzusetzen und Fehlentscheidungen vorzubeugen.

## Literatur

- Anderson, C., Brion, S., Moore, D. A., und Kennedy, J. A. A status-enhancement account of overconfidence. *Journal of Personality Social Psychology*, 103(4):718–735, 2012. ISSN 1939-1315.
- Arkes, H. R., Christensen, C., Lai, C., und Blumer, C. Two methods of reducing overconfidence. *Organizational Behavior and Human Decision Processes*, 39(1):133–144, 1987. ISSN 0749-5978.
- Armor, D. A., Massey, C., und Sackett, A. M. Prescribed optimism: Is it right to be wrong about the future? *Psychological Science*, 19(4):329–331, 2008. ISSN 0956-7976.
- Barber, B. M. und Odean, T. Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, 116(1):261–292, 2001. ISSN 1531-4650.
- Bazerman, M. und Moore, D. A. *Judgment in managerial decision making*. New York, N.Y., Wiley, 2013.
- Beck, H. *Behavioral Economics*. Wiesbaden, Springer Gabler, 2014.
- Bernardo, A. E. und Welch, I. On the evolution of overconfidence and entrepreneurs. *Journal of Economics & Management Strategy*, 10(3):301–330, 2001. ISSN 1058-6407.
- Berner, E. S. und Graber, M. L. Overconfidence as a cause of diagnostic error in medicine. *The American Journal of Medicine*, 121(5):S2–S23, 2008. ISSN 0002-9343.
- Billet, M. T. und Qian, Y. Are overconfident managers born or made? Evidence of self-attribution bias from frequent acquirers. *Management Science*, 54(6):1037–1051, 2008.
- Busenitz, L. W. und Barney, J. B. Differences between entrepreneurs and managers in large organizations: Biases and heuristics in strategic decision-making. *Journal of Business Venturing*, 12(1):9–30, 1997. ISSN 0883-9026.
- Cain, D. M., Moore, D. A., und Haran, U. Making sense of overconfidence in market entry. *Strategic Management Journal*, 36(1):1–18, 2015. ISSN 0143-2095.
- Camerer, C. und Lovallo, D. Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89(1):306–318, 1999. ISSN 0002-8282.
- Chen, G., Crossland, C., und Luo, S. Making the same mistake all over again: CEO overconfidence and corporate resistance to corrective feedback. *Strategic Management Journal*, 36(10):1513–1535, 2015. ISSN 0143-2095.
- Chen, J. Q. und Lee, S. M. An exploratory cognitive DSS for strategic decision making. *Decision Support Systems*, 36(2):147–160, 2003. ISSN 0167-9236.
- Clark, K. D. und Maggitti, P. G. TMT potency and strategic decision-making in high technology firms. *Journal of Management Studies*, 49(7):1168–1193, 2012. ISSN 0022-2380.
- Correia, V. Contextual debiasing and critical thinking: Reasons for optimism. *Topoi*, 37(1):103–111, 2018. ISSN 0167-7411.
- Croskerry, P., Singhal, G., und Mamede, S. Cognitive debiasing 1: origins of bias and theory of debiasing. *BMJ Quality & Safety*, 22(Suppl 2):ii58–ii64, 2013a. ISSN 2044-5415.
- Croskerry, P., Singhal, G., und Mamede, S. Cognitive debiasing 2: impediments to and strategies for change. *BMJ Quality & Safety*, 22(Suppl 2):ii65–ii72, 2013b. ISSN 2044-5415.
- Cueva, C., Iturbe-Ormaetxe, I., Ponti, G., und Tomás, J. Boys will still be boys: Gender differences in trading activity are not due to differences in (over) confidence. *Journal of Economic Behavior & Organization*, 160:100–120, 2019. ISSN 0167-2681.
- Daniel, K. und Hirshleifer, D. Overconfident investors, predictable returns, and excessive trading. *Journal of Economic Perspectives*, 29(4):61–88, 2015. ISSN 0895-3309.
- Das, T. und Teng, B. Cognitive biases and strategic decision processes: An integrative perspective. *Journal of Management Studies*, 36(6):757–778, 1999. ISSN 0022-2380.
- Evans, J. S. B. und Stanovich, K. E. Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, 8(3):223–241, 2013. ISSN 1745-6916.
- Fast, N. J., Sivanathan, N., Mayer, N. D., und Galinsky, A. D. Power and overconfident decision-making. *Organizational Behavior and Human Decision Processes*, 117(2):249–260, 2012. ISSN 0749-5978.
- Ferretti, V., Guney, S., Montibeller, G., und von Winterfeldt, D. Testing best practices to reduce the overconfidence bias in multi-criteria decision analysis. In *49th Hawaii International Conference on System Sciences (HICSS)*, pages 1547–1555. IEEE. ISBN 0769556701.
- Fischhoff, B. *Debiasing*, pages 422–444. Cambridge, Cambridge University Press, 1982.
- Galasso, A. und Simcoe, T. S. CEO overconfidence and innovation. *Management Science*, 57(8):1469–1484, 2011. ISSN 0025-1909.
- Gigerenzer, G. und Brighton, H. Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1(1):107–143, 2009. ISSN 1756-8757.
- Glaser, C. *Selbstüberschätzung*. Berlin, Springer, 2019.
- Göbel, E. *Entscheidungen in Unternehmen*. Stuttgart, UTB, 2014.
- Hirshleifer, D., Low, A., und Teoh, S. H. Are overconfident CEOs better innovators? *The Journal of Finance*, 67(4):1457–1498, 2012. ISSN 0022-1082.
- Ho, P.-H., Huang, C.-W., Lin, C.-Y., und Yen, J.-F. CEO overconfidence and financial crisis: Evidence from bank lending and leverage. *Journal of Financial Economics*, 120(1):194–209, 2016. ISSN 0304-405X.
- ICV, Internationaler Controller Verein e.V. Was macht Controller erfolgreich(er)? Auf das Verhalten kommt es an.
- Kahneman, D. *Thinking, fast and slow*. London, Macmillan Publishers, 2011. ISBN 0374275637.
- Kahneman, D. und Frederick, S. *Representativeness revisited: Attribute substitution in intuitive judgment*, pages 49–81. Cambridge, Cambridge University Press, 2002.
- Kamas, L. und Preston, A. The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior & Organization*, 83(1):82–97, 2012. ISSN 0167-2681.
- Kaustia, M. und Perttula, M. Overconfidence and debiasing in the financial industry. *Review of Behavioural Finance*, 4(1):46–62, 2012. ISSN 1940-5979.
- Kennedy, J. A., Anderson, C., und Moore, D. A. When overconfidence is revealed to others: Testing the status-enhancement theory of overconfidence. *Organizational Behavior Human Decision Processes*, 122(2):266–279, 2013. ISSN 0749-5978.
- Klein, G. Performing a project premortem. *Harvard business review*, 85(9):18–19, 2007.
- Korte, R. F. Biases in decision making and implications for human resource development. *Advances in Developing Human Resources*, 5(4):440–457, 2003. ISSN 1523-4223.
- Larrick, R. P. *Debiasing*, pages 316–338. 2004.
- Lichtenstein, S. und Fischhoff, B. Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance*, 20(2):159–183, 1977. ISSN 0030-5073.
- Logg, J. M., Haran, U., und Moore, D. A. Is overconfidence a motivated bias? Experimental evidence. *Journal of Experimental Psychology: General*, 147(10):1445, 2018. ISSN 1939-2222.
- Magee, J. C. und Frasier, C. W. Status and power: The principal inputs to influence for public managers. *Public Administration Review*, 74(3):307–317, 2014. ISSN 0033-3352.
- Malmendier, U. und Tate, G. Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics*, 89(1):20–43, 2008. ISSN 0304-405X.
- Mannes, A. E. und Moore, D. A. A behavioral demonstration of overconfidence in judgment. *Psychological Science*, 24(7):1190–1197, 2013. ISSN 0956-7976.
- Minson, J. A., Liberman, V., und Ross, L. Two to tango: Effects of collaboration and disagreement on dyadic judgment. *Personality and Social Psychology Bulletin*, 37(10):1325–1338, 2011. ISSN 0146-1672.
- Mintzberg, H. und Westley, F. Decision Making: It's not what you think. *MIT Sloan Management Review*, 42(3):89–93, 2001.
- Mintzberg, H., Raisinghani, D., und Theoret, A. The structure of 'unstructured' decision processes. *Administrative Science Quarterly*, 21(2):246–275, 1976. ISSN 0001-8392.
- Mitchell, D., Russo, E., und Pennington, N. Back to the future: Temporal perspective in the explanation of events. *Journal of Behavioral Decision Making*, 2(1):25–38, 1989. ISSN 0894-3257.
- Moore, D. A. und Healy, P. J. The trouble with overconfidence. *Psychological Review*, 115(2):502, 2008. ISSN 1939-1471.
- Moore, D. A. und Schatz, D. The three faces of overconfidence. *Social Personality Psychology Compass*, 11(8):e12331, 2017. ISSN 1751-9004.
- Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H.,

- und Kassam, K. S. Debiasing decisions: Improved decision making with a single training intervention. *Policy Insights from the Behavioral Brain Sciences*, 2(1):129–140, 2015. ISSN 2372-7322.
- Nutt, P. C. und Wilson, D. C. *Crucial Trends and Issues in Strategic Decision Making*, volume 6, pages 3–29. Hoboken, N.J., John Wiley & Sons, 2010. ISBN 1444323121.
- Park, S. H., Westphal, J. D., und Stern, I. Set up for a fall: The insidious effects of flattery and opinion conformity toward corporate leaders. *Administrative Science Quarterly*, 56(2):257–302, 2011. ISSN 0001-8392.
- Peterson, R. L. The neuroscience of investing: fMRI of the reward system. *Brain Research Bulletin*, 67(5):391–397, 2005. ISSN 0361-9230.
- Pfister, H.-R. *Die Psychologie der Entscheidung: Eine Einführung*. Berlin, Springer, 4. aufl. 2017 edition, 2017. ISBN 9783662530375 978-3-662-53038-2 3662530384.
- Plous, S. *The psychology of judgment and decision making*. New York, N.Y., McGraw-Hill Book Company, 1993. ISBN 0070504776.
- Prochaska, J. O., DiClemente, C. C., und Norcross, J. C. In search of how people change: applications to addictive behaviors. *American Psychologist*, 47(9):1102, 1992. ISSN 1935-990X.
- Roll, R. The hubris hypothesis of corporate takeovers. *Journal of Business Venturing*, pages 197–216, 1986. ISSN 0021-9398.
- Roulin, N., Bangerter, A., und Levashina, J. Interviewers' perceptions of impression management in employment interviews. *Journal of Managerial Psychology*, 29(2):141–163, 2014. ISSN 0268-3946.
- Russo, J. E. und Schoemaker, P. J. H. *Overconfidence*, pages 1–11. Hampshire, Palgrave Macmillan UK, 2016. ISBN 978-1-349-94848-2. doi: 10.1057/978-1-349-94848-2\_323-1. URL [https://doi.org/10.1057/978-1-349-94848-2\\_323-1](https://doi.org/10.1057/978-1-349-94848-2_323-1).
- Russo, J. E. und Schoemaker, P. J. Managing overconfidence. *Sloan Management Review*, 33(2):7–17, 1992.
- Schilke, O. und Reimann, M. Neuroökonomie: Grundverständnis, Methoden und betriebswirtschaftliche Anwendungsfelder. *Journal für Betriebswirtschaft*, 57(3-4):247–262, 2007. ISSN 0344-9327.
- Schwenk, C. R. Cognitive simplification processes in strategic decision-making. *Strategic Management Journal*, 5(2):111–128, 1984. ISSN 1097-0266.
- Simon, H. A. A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118, 1955. ISSN 1531-4650.
- Simon, H. A. *Models of man; social and rational*. Hoboken, N.J., Wiley, 1957.
- Soll, J. B., Milkman, K. L., und Payne, J. W. *A user's guide to debiasing*, volume 2, pages 924–951. Hoboken, N.J., Blackwell, 2015.
- Son, J. und Rojas, E. M. Impact of optimism bias regarding organizational dynamics on project planning and control. *Journal of Construction Engineering Management Science*, 137(2):147–157, 2010. ISSN 0733-9364.
- Tversky, A. und Kahneman, D. Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131, 1974. ISSN 0036-8075.
- Veinott, B., Klein, G. A., und Wiggins, S. Evaluating the effectiveness of the PreMortem technique on plan confidence. In *Proceedings of the 7th International ISCRAM Conference*. [http://idl.iscram.org/files/veinott/2010/1049\\_Veinott\\_etal2010.pdf](http://idl.iscram.org/files/veinott/2010/1049_Veinott_etal2010.pdf).
- von Hippel, W. und Trivers, R. The evolution and psychology of self-deception. *Behavioral Brain Sciences*, 34(1):1–16, 2011. ISSN 1469-1825.
- Wüst, K. und Beck, H. "I Thought I Did Much Better"—Overconfidence in University Exams. *Decision Sciences Journal of Innovative Education*, 16(4):310–333, 2018. ISSN 1540-4595.
- Zavertiaeva, M. A., López-Iturriaga, F. J., und Kuminova, E. V. Better innovators or more innovators? Managerial overconfidence and corporate R&D. *Managerial Decision Economics*, 39(4):447–461, 2018. ISSN 0143-6570.