



Advisory Editorial Board:

DOMINIK VAN AAKEN
FREDERIK AHLEMANN
THOMAS BÄHLINGER
CHRISTOPH BODE
ROLF BRÜHL
JOACHIM BÜSCHKEN
LEONHARD DOBUSCH
RALF ELSAS
DAVID FLORYSIK
GUNTHER FRIEDL
WOLFGANG GÜTTEL
CHRISTIAN HOFMANN
KATJA HUTTER
LUTZ JOHANNING
STEPHAN KAISER
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
SABINE URNIK
STEPHAN WAGNER
BARBARA E. WEIßENBERGER
ISABELL M. WELPE
HANNES WINNER
CLAUDIA B. WÖHLE
THOMAS WRONA
THOMAS ZWICK

JUNIOR MANAGEMENT SCIENCE

- Leon Szeli**, UX in AI: Trust in Algorithm-based Investment Decisions 1
- Sonja Schattevoy**, Hätte eine Pre-Mortem-Analyse den Tod von eLWIS verhindern können? – Verhaltensorientierte Ansätze für effektives Kostenmanagement in Großprojekten 19
- Sarah Franziska Kovatsch**, Giving in Unilaterally Risky Dictator Games: A Model of Allocation Decisions Under Existential Threat 35
- Quirin Stockinger**, Stochastic Optimization of Bioreactor Control Policies Using a Markov Decision Process Model 50
- Felix Konstantin Steinhardt**, Value Co-Creation: Eine Analyse von Treibern und Gestaltungsansätzen für B2B- vs. B2C-Kunden 81
- Felix Frauendorf**, Die aktuelle Änderung der umsatzsteuerrechtlichen Behandlung wirtschaftlicher Tätigkeiten von juristischen Personen des öffentlichen Rechts 118



UX in AI: Trust in Algorithm-based Investment Decisions

Leon Szeli

University of Cambridge and Technical University Munich

Abstract

This Thesis looks at investors' loss tolerance with portfolios managed by a human advisor compared to an algorithm with different degrees of humanization. The main goal is to explore differences between these groups (Humanized Algorithm, Dehumanized Algorithm, Humanized Human and Dehumanized Humans) and a potential diverging effect of humanizing. The Thesis is based on prior research (Hodge et al., 2018) but incorporates new aspects such as additional variables (demographics, prior experiences) and a comparison between users and non-users of automated-investment products. The core of this research is an experiment simulating an investment portfolio over time with four different portfolio managers. Subjects were asked to decide if they want to hold or sell a declining portfolio at five points in time to measure their loss tolerance. A cox regression model shows that portfolios managed by the Humanized Human had the highest loss tolerance. Humanizing leads to higher loss tolerance for the human advisor but to lower loss tolerance for algorithmic advisors within the non-user group.

Keywords: Künstliche Intelligenz; Artificial Intelligence; Behavioral Finance; Behavioral Economics; Human-Computer-Interaction; User Experience; Investmententscheidungen; Nutzervertrauen.

1. Introduction

This world is endowed with a limited number of AI (Artificial Intelligence) experts. The average citizen has a cartoonish understanding of AI based on overindulgent media headlines, which in itself is a by-product of a rapidly developing area. These headlines can have positive or negative connotations. Headlines such as the following are abundant in some of the most prestigious newspapers: "News broadcast triggers Amazon Alexa devices to purchase \$170 doll houses", "Robot passport checker rejects Asian applications because eyes are closed", "Robots judge a beauty contest and don't select women with dark skin", or "Microsoft's Twitter chatbot turns anti-feminist and pro-Hitler" (Leaden, 2017). Not all of these examples are caused by AI or algorithmic flaws, but in the general perception that does not matter. These headlines fuel mistrust and are not easily remedied by positive reports.

Nonetheless, trust in AI will be the driving factor for adoption in many industries. A good example to illustrate this is the financial sector. While AI already enjoys a high level of trust in some areas (entertainment, navigation), only half of people trust algorithmic investment advice (Shandwick, 2016). If the trust level does not change, Roboadvisors will never be able to reach mass market. This fact in itself makes for a compelling research project. Whereas, for the past ten

years, most researchers have agreed that humans trust humans more than AI. A recent range of experiments has contradicted that common understanding (Hodge et al., 2018).

Previously, you received advice from a human being, but now, due to rapid technological progress, you can also attain advice from AI systems like Roboadvisors. Roboadvisors are automated, algorithmic investment products. In 2017, Roboadvisors already had \$9.1 billion under management globally (Eule, 2017). Experts predict \$2.2 trillion under Roboadvisor management by 2020 (Epperson et al., 2015). Whether or not these predictions become true depends on whether users adopt the new technology. Whether they take the financial advice of a human being over an algorithm critically depends on trust (Hodge et al., 2018).

Yet, while the Roboadvisor industry is growing, there is little research examining how and when consumers trust the financial advice of algorithms. Which factors, from a psychological point of view, influence whether consumers trust AI? Despite the vast number of factors influencing trust and the variety of AI use cases, this paper focuses only on a selection of them. One factor that could drive trust in these products is the humanization of the algorithm. Does it help to make the algorithm more human and social, e.g. by naming it? AI-based assistants such as Amazon's Alexa, Microsoft's Cortana and Apple's Siri have a humanized interface, although

contemporary research voices doubt the positive influence on trust (Hodge et al., 2018). From this paradox, the research question of this thesis is derived: Does trust (measured as loss tolerance) in the financial advice of an algorithm (versus a human) depend on the degree to which the algorithm is humanized (giving the algorithm a human name)? To answer the research question, we conducted an online survey including experiments with users of the Roboadvisor Ginmon and non-users. Moreover, we looked at additional factors such as demographics and prior investment experiences.

2. State of Research

This chapter sheds light on the current state of research regarding trust. First, we look at the trust in humans – its original sense. Subsequently, trust in technology is examined.

2.1. Defining Trust

Trust as a basis for decision-making in diverse contexts has been studied in various fields. Cho et al. (2015) provide a clear overview of the multidisciplinary meanings of trust (Cho et al., 2011; Gambetta, 2000; James, 2002; Kydd, 2005; Lagerspetz, 1998; Lee and See, 2004; Mayer et al., 1995; Rotter, 1980):

Based on the multitude of meanings of trust, Cho et al. (2015) summarize the concept of trust as follows:

“Trust is the willingness of the trustor (evaluator) to take risk based on a subjective belief that a trustee (evaluatee) will exhibit reliable behavior to maximize the trustor’s interest under uncertainty (e.g., ambiguity due to conflicting evidence and/or ignorance caused by complete lack of evidence) of a given situation based on the cognitive assessment of past experience with the trustee.” (p. 28:5)

When individuals decide if they trust an entity, we call this process trust assessment. Cho et al. (2015) summarize this process as “Trustor *i* assesses Trustee *j*’s trust if *j* can perform Task *A*” (p. 28:5). According to game theoretic approaches, trustor *i* is defined as someone who maximizes their interest (or utility) from the relationship with the trustee (Chin, 2009). Trustee *j* is someone who can cause impact on a trustor’s utility with his/her behavior (Castelfranchi and Falcone, 2010).

Task *A* is a crucial factor in the decision if *i* trusts *j* as the importance of the tasks influences the risk assessment and potential outcomes. Afterwards, *i* adjusts trust according to whether the decision was right or wrong (Cho et al., 2015).

2.2. Trust in Information Technology

This chapter explains trust between humans and information technology. The concept of trust is not only applicable to human-to-human relationships but also to human-to-technology relationships (Coeckelbergh, 2012). Nevertheless, there are differences between the two relationships

(Mcknight et al., 2011). Firstly, regardless whether it is about trust in people or trust in technology, it involves risk and uncertainty. In the case of humans, you lack total control. You depend on the trustee to fulfill expected responsibilities which s/he might not fulfil – (un)intentionally. In the case of machines, you also do not have the control as a user since the technology might not demonstrate the expected capability (i.e., without intention). For example, when you trust Dropbox to save your data, you are exposed to the risk of data transmission over the internet and storing confidential data on a third-party server. Secondly, people have moral agency and volition whereas technology is amoral and non-volitional. For example, when you use a word processing program like Grammarly it will correct misspelled words and grammatical errors. But a benevolent human copyeditor might alter your text (reflecting his/her willingness and therefore volition) in order to help you improve even though it is not part of his/her job. A technology will only follow instructions programmed. Thirdly, the trustor’s expectations regarding the object of dependence might be different when comparing humans to machines. When trusting people, you expect them to fulfil a task for you in a competent way. When trusting a technology, you want it to demonstrate possession of functionality. A human helps you if s/he cares for you and is benevolent towards you. A machine’s helpfulness is not rooted in moral agency, but you still expect effective help (e.g. a help menu). When trusting humans, we hope for integrity, reliability and consistence in their actions. Due to humans’ free will, this is a risk. In a machine, we are looking for reliability. It should operate consistently without failing. Potential failures are caused by a bug and not by deliberate actions (Mcknight et al., 2011).

As described in chapter 2.2.3, trust in technology can be influenced by many factors. Siau and Wang (2018) structures these in multiple dimensions: Human characteristics (Personality, Ability), Environment Characteristics (Culture, Task, Institutional Factors) and Technology Characteristics (Performance, Process, Purpose) (Siau and Wang, 2018). For example, someone with a rather trusting personality (Human Characteristic) is likely to trust the technology in the task of filesharing (Environment Characteristic) no matter whether Google Drive or Dropbox is used (Technology Characteristic).

2.3. Trust in Automated and Digital Domains (e-Trust)

So far, we have been describing trust as a general concept and in the context of Information Technology. This chapter focuses on trust in more digital and automated domains.

2.3.1. Defining Trust in AI, HCI and Automation

Cho et al. (2015) defined key trust components for the three domains AI (Artificial Intelligence), HCI (Human-Computer Interaction) and Automation. All of the three applications fall in the category of “e-trust” which is trust occurring in digital contexts (Taddeo, 2010). They are structured based on four kinds of trust. First, Communication Trust which can be measured objectively (i.e. network connectivity). Second, Information Trust which includes quality

Table 1: Multidisciplinary Definitions of Trust according to Cho et al. (2015)

Discipline	Meaning of Trust	Source
Sociology	Subjective probability that another party will perform an action that will not hurt my interest under uncertainty and ignorance	Gambetta (2000)
Philosophy	Risky action deriving from personal, moral relationship between two entities	Lagerspetz (1998)
Economics	Expectation upon a risky action under uncertainty and ignorance based on the calculated incentives for the action	James (2002)
Psychology	Cognitive learning process obtained from social experiences based on the consequences of trusting behaviors	Rotter (1980)
Organizational Management	Willingness to take risk and being vulnerable to the relationship based on ability, integrity, and benevolence	Mayer et al. (1995)
International Relations	Belief that the other party is trustworthy with the willingness to reciprocate cooperation	Kydd (2005)
Automation	Attitude that one agent will achieve another agent's goal in a situation where imperfect knowledge is given with uncertainty and vulnerability	Lee and See (2004)
Computing & Networking	Estimated subjective probability that an entity exhibits reliable behavior for particular operation(s) under a situation with potential risks	Cho et al. (2011)

of information and credibility. Third, Social Trust which describes trust between humans in a social network. And fourth, Cognitive Trust which refers to accumulated knowledge regarding reliability and competence of the trusted party.

As you can see from Table 2, there are many overlaps in the different areas of trust and applications (availability in Communication Trust, belief in Information Trust, importance in Social Trust, expectation in Cognitive Trust). Nonetheless, there are differences and you have to be sure if you are talking about just one of the domains, an overlap of two or a mixture of multiple (Cho et al., 2015).

2.3.2. Human vs. Machine: Algorithm Aversion vs. Algorithm Appreciation

As initially stated, algorithms perform better than humans in many cases. Nonetheless, many humans rather trust a human prediction than an algorithmic prediction (Diab et al., 2011; Eastwood et al., 2012). In literature, this phenomenon is called algorithm aversion. Experiments support that – even if the algorithm outperforms the human (Dietvorst et al., 2015). Humans are more tolerant if a human is mistaken than if it is an algorithm (Dietvorst et al., 2018). Additionally, humans put more weight on human statements (Önkal et al., 2009; Promberger and Baron, 2006).

Multiple scholars investigated potential reasons for algorithm aversion: The human desire for perfection (Dawes, 1979; Einhorn and Hogarth, 1988; Highhouse, 2008), the assumption that the human learns based on past experience (Highhouse, 2008), the missing humaneness (Dawes, 1979; Grove and Meehl, 1996) and ethical concerns (Dawes, 1979).

Algorithm aversion has been the status quo, but a recent set of experiments conducted by researchers from Har-

vard Business School contradicts these findings (Logg et al., 2019). They found empirical evidence for Algorithm Appreciation which means that the humans relied more on algorithmic advice than on human advice. They ran six experiments with the following findings: a) people relied on algorithmic advice rather than on humans when estimating a person's body weight based on a picture (Experiment 1A), forecasting the popularity of songs (chart ranking) and romantic matches (attractiveness of the opposite gender) (Experiments 1B and 1C), b) algorithm appreciation persisted no matter if the advice appeared jointly or separately (Experiment 2) and c) Algorithm appreciation was reduced when people could choose between an algorithm's estimate and their own (versus an external advisor's; Experiment 3) and when they had experience in forecasting (Experiment 4).

As a conclusion, we can say that there is no clear answer to the questions whether humans prefer to trust humans or algorithms. There is additional research required in all kinds of domains.

2.3.3. Relevant Trust Factors in AI-based Products

Various factors influencing trust have been identified in past research. The last decades were mainly focused on trust in automation and robotics. In the past years, also studies about trust in digital products and, more specifically, AI have been published. To get an overview, the identified factors are summarized in this chapter. Tables with all factors can be found in the appendix (Table 13, 14 and 15). The sources are mainly meta analyses which represent multiple other papers (Adams et al., 2003; Cho et al., 2015; Hancock et al., 2011; Siau and Wang, 2018). Factors which are not applicable for Roboadvisors were neglected. Highly overlapping factors and factors which are only differing due to denota-

Table 2: Key trust components in different domains according to Cho et al. (2015)

	Communication Trust	Information Trust	Social Trust	Cognitive Trust
Artificial Intelligence	Cooperation, availability	Belief, experience, uncertainty	Importance, honesty	Expectation, confidence, hope, fear, frustration, disappointment, relief, regret
Human-Computer Interaction	Reliability, availability	Belief, experience, uncertainty	Importance, integrity	Expectation, frustration, regret, experience, hope, fear
Automation	Reliability, availability	Belief, experience	Integrity, importance	Expectation, confidence

tion of the authors were summarized as one. The factors can be separated into three categories: Factors regarding the Trustee (the institution/technology (provider) – in our case Ginmon/Ginmon’s algorithm), Trustor (the human user) and Environment (contextual, situational) (Adams et al., 2003). To summarize, we can say that the Trustor is looking for a reliable and adaptable product that is transparent in its actions. The Trustee’s trust in AI depends on his/her demographics, personality and prior touchpoints with the topic. The environmental factors concern the task, legal aspects, risk as well as influence from other people or the current state of mind. Due to the multitude of trust factors, we decided to focus on one which seems particularly interesting: humanization. This decision, as well as the factor itself, will be elaborated further in chapter 2.4.4 and 2.4.5.

Researchers came up with a wide range of different theoretical models regarding trust in technology and/or automation incorporating some of the mentioned factors. The main frameworks which were developed between 1989 and today are summarized in a Table 12 in the appendix.

We can observe the following patterns among trust models: a) most focus on automation and include an operator that receives recommendations but still has the decision power, b) numerous factors are involved due to the complexity of trust as a concept and c) most models are generic and not tailored towards a specific technology (e.g. algorithmic investment decisions).

There is no model (yet) whose application would help us answer our research question on AI-based investment products. Therefore, we follow a rather exploratory approach which is not based on a theoretical model but nonetheless incorporates empirical findings from the past.

2.4. Trust in Automated Investment Decisions

Trust plays an important role in finance, no matter if it is on the global, institutional level or on the people-to-people level (Bottazzi et al., 2016). Giving another entity money involves risk and vulnerability. For digital finance products, trust is even more crucial since there is less human face-to-face interaction (Greiner and Wang, 2010). Especially, new

and innovative products (fintechs) are key as there is less experience value (Van Thiel and Van Raaij, 2017).

2.4.1. Trust in Roboadvisors

This research is focusing on so-called Roboadvisors which came up after the financial crisis in 2008. “Robo-Advisors provide investing advice, wealth management services, sometimes in addition to data aggregation. These Fintech companies provide investment advice and trading services that are automated using algorithms and artificial intelligence” (Gold and Kursh, 2017, p.140). In 2017, Roboadvisors managed \$200 billion in assets (Eule, 2017). Roboadvisors have two main benefits: cost saving due to automation and transparency due to the user interface (Salo, 2017). Therefore, also people who did not have the financial resources to pay a human financial advisor to manage their money have the chance to invest in a comparable manner. There is no human intervention other than the user deciding to liquidate the portfolio managed by the algorithm. How can trust be built if there is no human involved? A similar case is e-banking where ease of use, usefulness, perceived privacy and perceived security are considered trust builders (Yousafzai et al., 2003). In the case of Roboadvisors, price and trustworthiness (initial and ongoing trust) additionally come into play when it comes to building trust (Lee et al., 2018).

Roboadvisors have different customer segments with different expectations. Salo (2017) identified four groups of Roboadvisor users based on their technoliteracy and financial literacy. The “delegators” want to outsource their investing to someone. Therefore, they are looking for an easy, available, neutral solution with low costs. The “optimizers” look for the most efficient (easier/cheaper) solution. They usually have a little more knowledge and net worth than the delegators and would like to be involved in the investment process, if it could increase returns. The third group is called “Do-It-Yourself”. These users want to make investment decisions themselves and are only seeking for advice to consider. The advisor plays a smaller role than in the other two groups (Carré et al., 2016; Salo, 2017). When we talk about building users trust, we have to be clear if we are talking about

all potential users, current users or a specific user group. Since there is only very little research on users of Roboadvisors present, this work is looking at all potential users as well as current users. Looking at a specific user group might be a good idea in a second step.

2.4.2. Trust as a Reaction to Financial Losses

As described in earlier chapters, trust is closely related to risk: “Trust is the willingness of the trustor (evaluator) to take risk [...]” (Cho et al., 2015, p 28:5). Nonetheless, it is not the same (Houser et al., 2010). Traditional finance theory and psychological literature have different understandings of risk. In traditional finance theory, risk is defined as the variance of the expected distribution of returns (Bernstein, 1996; Haugen, 1995). For psychologists, risk is synonymous to loss. The higher your trust in the decision-maker, the bigger your loss tolerance (Payne, 1975; Shapira, 1995; Slovic and Lichtenstein, 1968; Teigen, 1996). Also, you are more likely to forgive a mistake (which is a bad trade in the case of Roboadvisors) (Dietvorst et al., 2018). In this thesis, we take the psychological perspective.

In addition to risk, vulnerability was mentioned multiple times in the different definitions of trust. Trust is only required, if you have “something to lose”. Furthermore, trust is an important factor in investment decisions, such as buying and selling stocks (Guiso et al., 2008). Therefore, we decided to simulate an investment portfolio in our experiment, which will be described later. To find out how the respondents react to losses, we decided to have a constantly declining portfolio. We should keep in mind that we are not measuring trust directly but loss tolerance (Will they liquidate the portfolio?)¹.

Behavioral Economics literature suggests that two biases might occur in this experiment: loss aversion and sunk costs. According to prospect theory (Kahneman and Tversky, 1979), individuals prefer avoiding losses to making equal gains. In the context of stock markets, this can lead to holding stocks which are losing in value for too long instead of selling them despite loss (Odean, 1998). Another effect that leads to following through with your portfolio, even if it is declining, is called sunk costs. Sunk costs are costs that have already been incurred and cannot be reversed. In our case it means that investors already have spent time (e.g. keeping track of portfolio's value). Additionally, costs that will irrevocably incur in the future are included (Arkes and Blumer, 1985). Since the portfolio of our participants in our experiment is declining and involves sunk costs, we have to keep the two biases in mind.

2.4.3. The Case of Ginmon

This work looks at one Roboadvisor – Ginmon – in specific. The Frankfurt-based startup was founded in 2014 and is among the leading Roboadvisors in Germany. They offer

an automated investment portfolio that is managed by an algorithm called Apeiron. Ginmon charges you 0,39% of your invested amount and 10% of your gains.

Your portfolio is allocated based on your risk class (from 1 to 10). To determine your risk class, you have to answer 9 questions – just like with a human advisor. The questions can be found in the attachment. After being assigned to a risk class there is no other human intervention needed. The only intervention possible is liquidating the portfolio. The algorithm takes care of everything else (buying and selling, rebalancing). The customer can check the current value of the portfolio at any time.

The minimum investment amount is 1000€ , which is much less than traditional advisors require. Nonetheless, wealthy individuals are more attractive customers as they invest more money and therefore the share Ginmon receives is bigger. Generally speaking, people with high net worth are older (Krause and Schäfer, 2005). Also, these people have more reservations towards digital products (Millward, 2003). This paradox explains the company's interest in the question how to build trust in their product.

2.4.4. Anthropomorphism

As mentioned earlier, humanization of algorithms or depicting them as social can influence users' trust. The technical term for that phenomenon is called Anthropomorphism. According to Guthrie (1993) Anthropomorphism describes perceiving humanlike characteristics such as physical appearance, emotional or mental states and motivation in nonhuman agents (Guthrie, 1993). Humanizing nonhuman agents in order to increase user acceptance has been investigated in the past (André et al., 2018; Epley et al., 2008). Examples include making a chatbot more social by imitating human-like behavior, giving it a face/avatar or a name.

2.4.5. Research Question & Hypotheses

Humanizing technology demonstrated benefits in some domains (e.g. IBM's “Watson” or Amazon's “Alexa”) but the effects are not clear in financial advising (Hodge et al., 2018). Human-Computer-Interaction research in the past suggests that humanization of technology leads to positive feelings towards the technology and increases the likelihood to use the technology (Burgoon et al., 2000; Chaminade et al., 2007; Eyssel et al.; Gong, 2008; Venkatesh, 2000). However, that does not necessarily mean that the technology is perceived more trustworthy or persuasive (Nan et al., 2006; Riegelsberger et al., 2005). It has been shown that naming a technology can decrease its credibility, which also decreases trust. A technology can be perceived simple and unable to complete complex tasks if it is named (Hafer et al., 1996; Riegelsberger et al., 2005). Naming a human on the other hand, is beneficial. Sharing personal information increases the trust level because the advisor is willing to risk his reputation (Garner, 2005; O'Keefe, 1990; Pornpitakpan, 2004).

Hodge et al. (2018) have shown that humanizing a Roboadvisor decreases the likelihood of subjects to follow

¹Limitations caused by the dependent variable are addressed in Chapter 5. Interpretation, Limitations and Future Research

investment recommendations. For human advisors the opposite is the case. The Humanized Human was the advisor whose recommendations were followed the most (73,6%). Second was the Dehumanized Roboadvisor (69,7%), third the Humanized Roboadvisor (59,2%) and fourth the Dehumanized Human (54,7%) (Hodge et al., 2018).

We are interested in the threshold where people lose trust (which means sell) in their portfolio – their loss tolerance. Is there a difference in users' trust depending on the type of advisor (human/algorithm) and the level of humanization (humanized/dehumanized)?

There are many definitions of losing trust but in our case, we understand it as a point in time where the trustee (user) loses hope in the ability of the trustor (Roboadvisor) to reach the desired outcome (increase in portfolio value)². We assume that each individual has a certain "pain threshold" of loss and as soon that is exceeded, they lose trust and sell (liquidate the portfolio). Based on the mentioned findings, we formulate our hypotheses:

H1: Respondents' loss tolerance is the highest for portfolios managed by Humanized Humans ("Anlageberater Charles").

H2: Respondents' loss tolerance is the second highest for portfolios managed by Dehumanized Algorithms ("Algorithmus").

H3: Respondents' loss tolerance is the third highest for portfolios managed by Humanized Algorithms ("Algorithmus Charles").

H4: Respondents' loss tolerance is the lowest for portfolios managed by Dehumanized Humans ("Anlageberater").

H5: Respondents' loss tolerance is higher for portfolios managed by Humanized Human than for a Dehumanized Human.

H6: Respondents' loss tolerance is higher for portfolios managed by Dehumanized Algorithm than for a Humanized Algorithm.

While our own hypotheses are mainly based on Hodge's findings, there are some distinct differentiations to his work. Firstly, Hodge focuses on investors' recommendations and whether people follow advice. For Roboadvisors, you usually just invest money and the Roboadvisor has full control afterwards. This means that the user is not involved in decision-making anymore due to automation and therefore does not receive advice. As a result, we decided to focus on the question if and how long people trust the algorithm with their money instead of "following recommendations". Secondly, Hodge's sample consists of 108 MBA students which means they have a) a lot of background knowledge in finance and b) are rather young. Reaching older users with less knowledge about digital products and finance is crucial for mass adoption of Roboadvisors. Therefore, we were aiming for a more diverse sample. Thirdly, Hodge gave participants extensive background information about the potential investment decisions. AI-based products and Roboadvisors specifically cannot provide this in real life and, therefore, "blind" trust is required from users. As a consequence, we decided to give

subjects less background information. Fourth, Hodge is looking at one-time investment decisions while we are looking at a portfolio over time. Fifth, Hodge et al. (2018) analyse their data with t-tests and ANOVAs while we took a different approach for the data analysis (cox regression). Sixth, Hodge et al. (2018) did not look at demographic factors or prior investment experiences which we do by controlling for them and comparing users of automated investment products to non-users. We try to thereby contribute to validating the findings while also expanding them.

3. Method

This chapter provides additional information on the sample and survey participants, data collection procedure, and the instruments.

3.1. Sample and Participants

The survey, conducted in December 2019, was sent to German Ginmon customers as well as non-customers. The goal was to gain insights on users and non-users of AI-based products. A total of 258 people took part in the survey with an almost equal split between users and non-users. 82% are male, 18% female. The average age was 36,6, the youngest participant was 18, the oldest 80. When interpreting results, we have to keep in mind that this sample does not represent the German population³.

3.2. Data Collection Procedure

The survey contained 13 closed questions – and took about five minutes to complete. Ginmon users were contacted using Ginmon's newsletters. The non-users were mainly recruited using the authors' social media profiles. These efforts lead to 258 started surveys out of which 223 were completed.

3.3. Instruments

The first question was an experiment where respondents were randomly assigned to one of four groups. The participants then observed a portfolio which was either managed by a) the human advisor named Charles ("Anlageberater Charles"), b) the unnamed human advisor ("Anlageberater"), c) the algorithm named Charles ("Algorithmus Charles") or d) the unnamed algorithm ("Algorithmus").

The information about the portfolio manager was followed by a graphic representation of the portfolio over time. The portfolio was initially worth 10.000€ . The participants were asked at 5 points in time (February until July) whether they wanted to either hold or sell. If they always chose to hold, it looked as follows:

The goal of this question was to find out when participants lose trust in their respective portfolio manager and

²Limitations caused by the dependent variable are addressed in Chapter 5. Interpretation, Limitations and Future Research

³Limitations caused by the sample will be elaborated in the Chapter 5. Interpretation, Limitations and Future Research

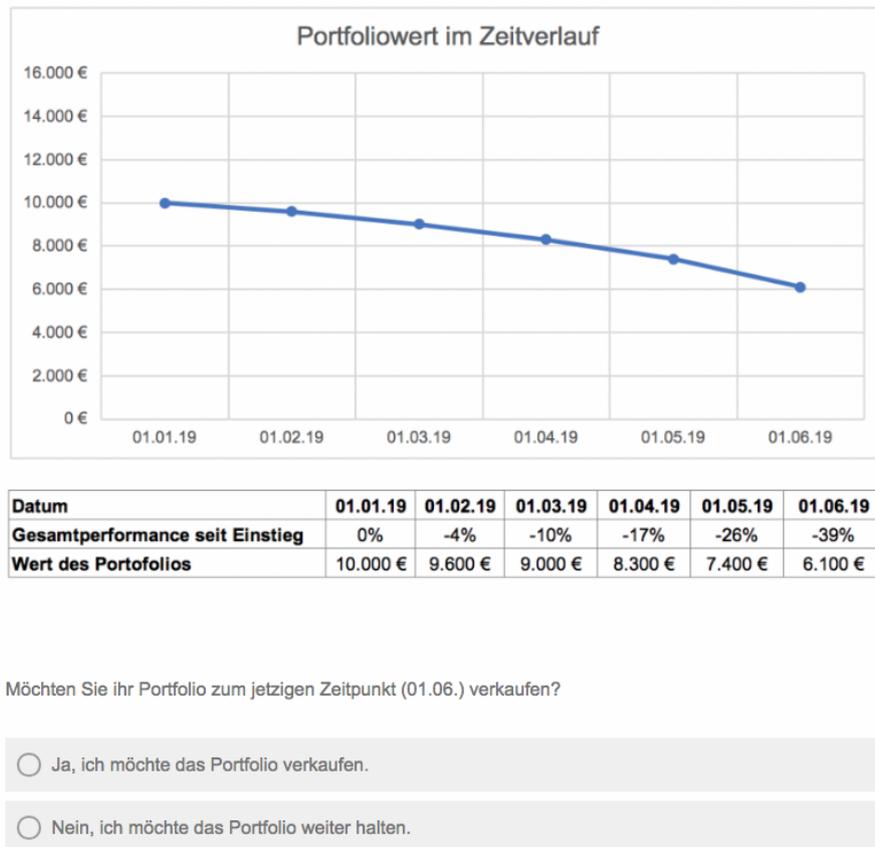


Figure 1: Simulated portfolio in the survey

sell. We wanted to conduct an experiment which is realistic. Therefore, it was also possible to never sell the portfolio. This means a cox regression was the way to analyze the data as it is censored (the event of selling does not occur for every participant). When designing the experiment, past research with similar settings was taken into account (Glaser and Walther, 2013). We are convinced that this approach is a good choice since scholars in the past showed that asking for trust directly in self-report measures does not have strong validity and reliability (Cho et al., 2015; Dzindolet et al., 2003; Mcknight et al., 2011).

In the second question, participants had to distribute 10.000€ between an algorithm as an investor and a human advisor. This joint evaluation is a complement to the separate evaluation in question one. Attributes are easier to evaluate when advisors are presented jointly due to increased information (Hodge et al., 2018). The third section was about demographics (age, gender) and if they ever invested money. If they replied no, the survey was over. If they replied yes, they were asked about their investment experience thus far. How much experience did they have? Did they have experience with automated investments? Did they ever use Ginmon? How happy were they with their returns?

4. Results

The survey had 258 respondents (82% male⁴, $M_{Age} = 36,6$, $SD_{Age} = 15,7$) out of which 233 finished the survey. The sample size for the following analyses varies due to 35 dropouts. Our sample consists of 110 people (70% male, $M_{Age} = 27,2$, $SD_{Age} = 13,5$) who have never used the Roboadvisor Ginmon and 113 people (93% male, $M_{Age} = 45,6$, $SD_{Age} = 18,2$) who use Ginmon. The sample is not representative for the average German⁵. Due to our sampling method, that was not expected in the first place.

4.1. Experiment I: Simulating a Portfolio with Four Different Portfolio Managers

First, we will look at the results of Experiment I which helps us compare the four different types of portfolio managers.

4.1.1. Behavior of Overall Sample

To get a first overview of the data collected throughout the experiment, the following flow chart shows when and

⁴The impact of the gender distribution on representativeness will be discussed in Chapter 5. Interpretation, Limitations and Future Research.

⁵The limitations for generalization associated with this sampling will be discussed in Chapter 5. Interpretation, Limitations and Future Research.

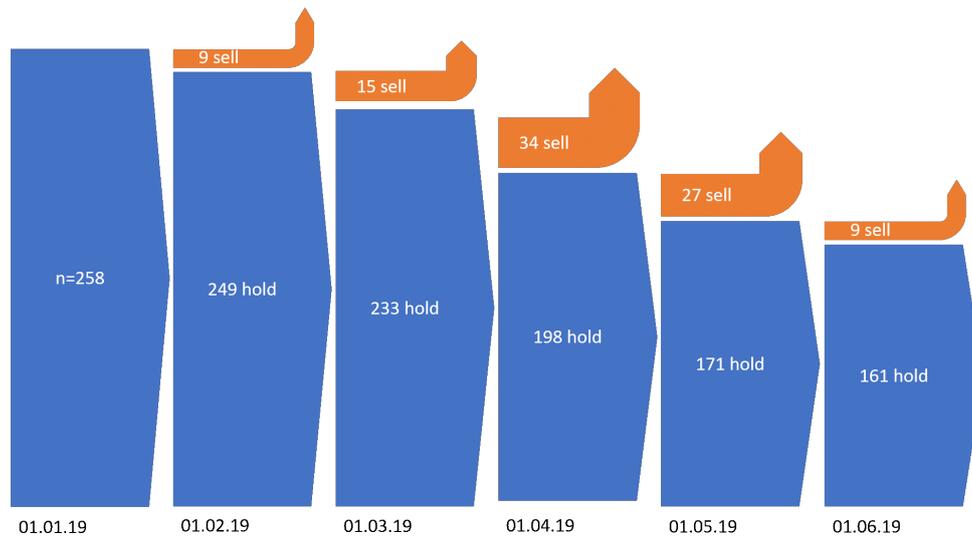


Figure 2: Flow chart depicting sales over time (n = 258)

if participants liquidated their portfolio. It contains the total sample, including all four conditions (Humanized Algorithm, Dehumanized Algorithm, Dehumanized Human and Humanized Human).

As seen in Figure 2, out of 258, 161 (62,4%) never sold their portfolio. 97 respondents (37,6%) sold it. The decline in the portfolio increases over time, nonetheless most people (37) sell in the third month – not in the fifth (9 sales). In general, the respondents held their portfolio for too long. A potential explanation for this phenomenon could be the earlier mentioned loss aversion and sunk costs fallacy. These biases explain why people hold their declining investment portfolios for longer than a rational person should. The respondents are averse to losses which means they tend to not realize their losses by liquidating and “wait it out” instead. In addition, they invested time and energy by tracking the portfolio changes over time. They cannot get the lost time back, so they think it is wasted if they do not manage to sell the portfolio without loss. As a result, they hold the portfolio. Rational behavior would be to only let future costs affect their decisions.

4.1.2. Comparing the Four Portfolio Managers

To analyze the data, we used a cox regression, also called survival analysis. First, we performed a cox regression which compares the experimental groups to the group Humanized Human. We chose the reference group based on Hodge et al. (2018) – we assume this group is the one with the highest loss tolerance and therefore the least sales. Additionally, it makes sense to benchmark the other groups against the Humanized Human since this is still the standard case. When you interact with your bank, you usually talk to a human. Control variables will be added to the model in a later step.

As you can see from Table 3, participants were 2.93 more likely to sell their portfolio when it was managed by a Humanized Algorithm (vs. a Humanized Human) (95% CI,

1.518 - 5.655; $P < 0.01$). When the portfolio was managed by a Dehumanized Algorithm (vs. a Humanized Human), however, participants were (only) around two times (2.093) more likely to sell their portfolio (95% CI, 1.118-3.914; $P < 0.05$). Finally, when the portfolio was managed by a Dehumanized Human (vs. Humanized Human), participants are still around 1.3 times more likely to sell (95% CI, 0.701-2.481; $P > 0.05$ ⁶).

If we create a ranking of the four groups for “loss tolerance” from highest (least sales) to lowest (most sales) and compare them to findings of Hodge et al. (2018)⁷, it looks as follows:

In Table 5, we summarized the results of our Hypotheses 1 to 4: First, in line with H1, we find that participants tolerate losses in their portfolio the most when their portfolio is managed by a Humanized Human. The other three groups have a higher likelihood to sell than the reference group (Humanized Human). Second, our results do not support H2. We assumed respondents’ loss tolerance is the second highest for portfolios managed by Dehumanized Algorithms (“Algorithmus”) but it ranked third. Third, our results do not support H3. We expected the third highest loss tolerance for the group Humanized Algorithm, but the loss tolerance was the lowest. Fourth, H3 is also not supported by our results. We expected the lowest loss tolerance for the group Dehumanized Human, but loss tolerance was the second highest.

4.1.3. Interaction Effect of Humanizing

H5 and H6 are stating that humanizing has a positive effect (in terms of loss tolerance) for human advisors but a negative effect for algorithmic advisors. Our ranking above sup-

⁶The significance level is discussed in Chapter 4.1.5 Exploratory Analysis.

⁷The differences between both studies regarding the Dependent Variable will be discussed in Chapter 5. Interpretation, Limitations and Future Research.

Table 3: Cox regression (n=258)

Group	B	HR	95% CI	P Values
Humanized Human				
Dehumanized Human	.277	1.319	0.701 2.481	.390
Dehumanized Algorithm	.738	2.093	1.118 3.914	.021
Humanized Algorithm	1.075	2.930	1.518 5.655	.001

Table 4: Comparison of group ranks in two studies

Group	Ranked by loss tolerance (Szeli 2019)	Ranked by likelihood to follow recommendation (Hodge et al., 2018)
Humanized Human	1	1
Dehumanized Human	2	4
Dehumanized Algorithm	3	2
Humanized Algorithm	4	3

Table 5: Results of Hypotheses 1 to 4

Hypotheses	Result
H1: Respondents' loss tolerance is the highest for portfolios managed by Humanized Humans ("Anlageberater Charles").	True
H2: Respondents' loss tolerance is the second highest for portfolios managed by Dehumanized Algorithms ("Algorithmus").	Not true
H3: Respondents' loss tolerance is the third highest for portfolios managed by Humanized Algorithms ("Algorithmus Charles").	Not true
H4: Respondents' loss tolerance is the lowest for portfolios managed by Dehumanized Humans ("Anlageberater").	Not true

ports this reasoning; the Humanized Human (Rank 1) outperforms the Dehumanized Human (Rank 2), but the Humanized Algorithm (Rank 4) underperforms the Dehumanized Algorithm (Rank 3). To test these predictions quantitatively, we computed a 2 (advisor: human vs. algorithm) x 2 (humanization: humanized vs. dehumanized) cox regression model (Table 6).

The main effect of the type of advisor is marginally significant ($p=.064$). Given the b-value of $-.461$ and the coding of our variables, we can say that portfolios managed by Human Advisor were sold less than portfolios managed by an Algorithmic Advisor. The interaction effect between the factors humanization (named vs. unnamed) and advisor (human vs. algorithm) was not significant ($p=.133$). That is, the effect of humanization on selling is not statistically different for a human (vs. algorithm). According to the ranking reported in Table 4 as well as our Hypotheses 5 and 6, we assumed a different finding. We expected a significant interaction effect since the Humanized Human had the highest loss tolerance while the Humanized Algorithm had the lowest loss tolerance (Table 4). Thus, we will take a closer look at the changes in significance in later chapters, if we split the data set (4.1.5 Comparing users and non-users) or increase sample size (4.1.6 Exploratory analysis).

4.1.4. Prior Investment Experiences and Demographics

In this chapter, we analyze variables related to prior investment experiences and demographics. 31 respondents (14%) have never invested money. The remaining 190 (86) consider their experience with investing money and automated investments as follows:

Most subjects consider themselves "advanced" investors, only a few edge cases are "newbies" or "experts".

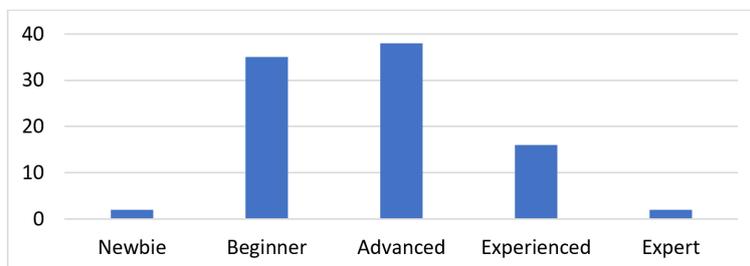
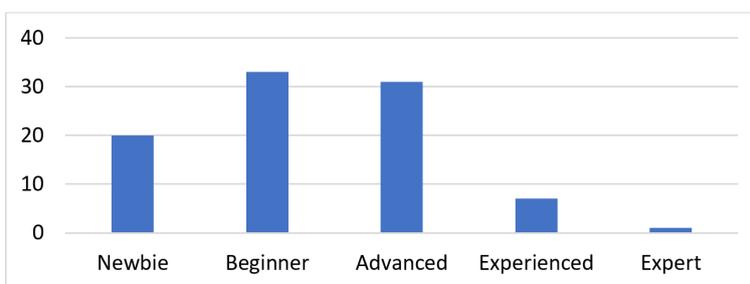
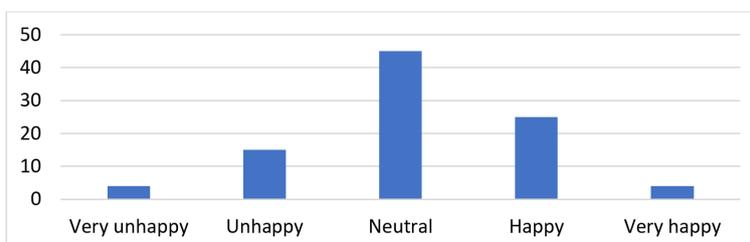
In automated investing, more subjects are newbies and beginners:

The majority (Figure 5) had at least neutral experiences with the returns on their investment in the past. Only a few are very unhappy. One possible explanation for that can be that many respondents are in their twenties which means they have been investing in a bull market (since 2009) all their adult life.

Earlier studies have shown that demographic factors and past investment experiences play an important role in people's future investment behavior (Grinblatt and Keloharju, 2000; Schnell, 2011; Statman, 1999). Therefore, we decided to calculate an extension of the initial cox regression model (Table 3) by controlling for the following factors: Age, Gender, Satisfaction with past returns, Experience with investing

Table 6: Main effects and interaction effects (Cox Regression; n=256)

Effect	df	B	SE	HR	P Values
Humanization (named/unnamed)	1	.277	.322	1.319	.390
Advisor (Algorithm/Human)	1	-.461	.249	.630	.064
Humanization x Advisor	1	-.614	.414	.541	.138

**Figure 3:** Experience with investing money (n=190)**Figure 4:** Experience with automated investments (n=190)**Figure 5:** Satisfaction with past returns

and Experience with automated investing (see Table 7).

Table 7 shows that all of the added factors are not significant. Moreover, we find that the coefficient of the four different portfolio managers do not change substantially, if we include this set of control variables. Thus, these findings suggest that the results reported in 4.1.2 hold true controlling for demographical variables (e.g. gender, age) and for previous financial experiences (e.g. experiences with automated investing).

4.1.5. Comparing Users and Non-users

In this chapter, we compare Ginmon users (N = 110, 93% male, $M_{Age} = 45,6$, $SD_{Age} = 18,2$) versus Ginmon non-users (who have never used any other Roboadviser) (N = 113, 70%

male, $M_{Age} = 27,2$, $SD_{Age} = 13,5$)⁸.

First, we ran the cox regression (Table 3) again – but this time with a split dataset. Table 8 reports differences between users and non-users.

As one can see, results across Ginmon users versus non-users (Table 8) are very similar to the results without splitting the data (Table 3), since the ranking of the groups relative to the reference group Humanized Human is the same.

If we compare users to non-users, particularly striking is that the p-values for all groups are lower for non-users than for users while all the b-values are higher. That is, users are

⁸The differences between in age and gender will be discussed in Chapter 5. Interpretation, Limitations and Future Research.

Table 7: Extended Cox regression (n=258)

Group	B	HR	95% CI	P Values
Humanized Human				
Dehumanized Human	.098	1.104	0.535 2.276	.790
Dehumanized Algorithm	.651	1.918	0.943 3.899	.072
Humanized Algorithm	1.020	2.773	1.304 5.897	.008
Age	-.004	.996	.982 1.010	.575
Gender	-.440	.644	.341 1.217	.175
Satisfaction with past returns	-.191	.826	.621 1.101	.192
Experience with investing	.176	1.192	.867 1.639	.280
Experience with automated investing	.029	1.030	.763 1.389	.849

Table 8: Cox regression for users (n=113) and non-users (n=110)

Cox regression for users (n=113)				
Group	B	HR	95% CI	P Values
Humanized Human				
Dehumanized Human	.384	.681	.242 1.914	.466
Dehumanized Algorithm	.826	.439	.160 1.203	.110
Humanized Algorithm	.916	.400	.143 1.122	.082
Cox regression for non-users (n=110)				
Humanized Human				
Dehumanized Human	.475	.622	.313 1.237	.176
Dehumanized Algorithm	.823	.439	.216 .892	.023
Humanized Algorithm	1.531	.216	.098 .477	.000

less likely to sell their portfolio across all four conditions. Compared to the original sample (n=256), for non-users we found lower p-values despite reduced sample size (n=110)⁹.

Second, we ran the cox regression with a 2x2 design testing for main and interaction effects (like in Table 6). But this time, the analysis was split by users and non-users (Table 9).

For non-users, the interaction effect (p=.015) is significant. Thus, for nonusers, humanizing has a different effect on the time of selling the portfolio depending on whether the advisor is a human or an algorithm. Humanizing the human leads to less sales while humanizing the algorithm leads to more sales (compared to the non-humanized counterpart). For users, however, we cannot make the same statement since the effect of humanizing was not significant (p=.454). Given Table 9 and the results of our ranking, we can say that the following hypotheses hold true for nonusers:

H5: Respondents' loss tolerance is higher for portfolios managed by Humanized Human ("Anlageberater Charles") than for a Dehumanized Human.

H6: Respondents' loss tolerance is higher for portfolios managed by Dehumanized Algorithm ("Algorithmus") than for a Humanized Algorithm ("Algorithmus Charles").

⁹Possible explanations for the differences between users and non-users and ideas for future research are reported in Chapter 5. Interpretation, Limitations and Future Research.

4.1.6. Exploratory Analysis

The interaction effect in the overall sample (Table 6) as well as in the non-user group (Table 9) was not significant. To explore whether and how this would change with increased sample size, we (artificially) duplicated our dataset (Table 10).

Given the larger sample size, the interaction effect (p=.036) and the main effect of the advisor type (p=.009) are significant. Based on the b-value (-.461) and our coding, we can say that there is a main effect which means that portfolios managed by a human were less likely to be sold than portfolios managed by the algorithm. Based on these findings, one could speculate that rerunning the experiment — with an increased sample size — could yield a significant interaction effects of the factor Advisor and Humanization x Advisor (and therefore H5 and H6 would hold true)¹⁰. There needs to be additional empirical validation in the future. We suggest a bigger sample size or a research design which allows to test with an ANOVA, so significant effects can be found more easily due to uncensored data.

Since none of the control variables were significant, we also decided to rerun a cox regression including factors which

¹⁰More details on potential future research can be found in Chapter 5. Interpretation, Limitations and Future Research.

Table 9: Main effects and interaction effects for users (Cox Regression; n=113) and non-users (n=110)

Main effects and interaction effects for users (Cox Regression; n=113)					
Effect	df	B	SE	HR	P Values
Humanization (named/unnamed)	1	.385	.528	1.469	.466
Advisor (Algorithm/Human)	1	-.370	.370	.645	.235
Humanization x Advisor	1	-.477	.636	.621	.454
Main effects and interaction effects for non-users (Cox Regression; n=110)					
Humanization (named/unnamed)	1	.475	.351	1.608	.176
Advisor (Algorithm/Human)	1	-.348	.288	.706	.226
Humanization x Advisor	1	-1.183	.488	.306	.015

Table 10: Main effects and interaction effects with doubled sample size (Cox Regression; n=512)

Effect	df	B	SE	Exp(B)	P Values
Humanization (named/unnamed)	1	.277	.322	1.319	.224
Advisor (Algorithm/Human)	1	-.461	.249	.630	.009
Humanization x Advisor	1	-.614	.414	.541	.036

were closest to marginal significance level. Again, we doubled the sample size (Table 11) to explore, if our sample size is the reason the control variables are not significant.

Interestingly, gender is now significant ($p=.04$). Also, prior investment experience ($p=.108$) and satisfaction with returns ($p=.078$) have p-values close to the marginal significance level. Given the negative b-values of Gender ($B = -.445$; Coding: 1 = female, 2 = male) and Satisfaction with returns ($B = -.180$; Coding: 1 = very unhappy, 5 = very happy), we can conclude that females and respondents who were unhappy with the past performance of their portfolio were more likely to sell than their counterparts. The positive b-value of investment experience ($B = .163$; Coding: 1 = Newbie, 5 = Expert), means that more experienced people were more likely to sell than less experienced people. This supports our findings where we compared users to non-users. As stated earlier, these results should encourage future researchers to run an additional study with larger sample size.

4.2. Experiment II: Allocation of the Budget Between Human and Algorithm

In one question we asked in the survey, respondents had a choice between the human and the algorithm (allocating an investment of 10.000€ between both of them). The following pie charts represent the allocation of the 10.000€ between the two advisors. The first interesting finding is that most (70%) of the money was given to the algorithm.

This is an extreme deviation from the actual distribution of investments managed by human advisors versus algorithmic advisors (7,7 Million EUR in 2019 (Statista, 2018)) in Germany. This is also an interesting finding in the light of the results of Experiment I. Respondents had a higher loss

tolerance with human advisors but still claim that they would allocate most of their money to an algorithm¹¹.

Second, we distinguish between people who do not use Ginmon and people who use Ginmon. There is substantial difference: Ginmon users allocate a lot more money (80%) to the algorithmic advisor than non-users (59%). People with more experience in algorithmic were willing to give more money to the algorithm. A Z-proportions test supports this difference ($Z = 2.868$, $p = 0.002$).

5. Interpretation, Limitations and Future Research

This chapter will cover interpretation of our results, limitations of this research and suggestions for future researchers.

5.1. Experiment I: Simulating a Portfolio with Four Different Portfolio Managers

Some of our findings (H1, H5, H6) were coherent with Hodge et al. (2018), while others were not (H2, H3, H4). Regarding the coherent findings in H1, it remains to be said that Humanized Humans are still the most trusted decision-makers with the highest loss tolerance among investors. This finding cumpers the mass adoption for Roboadvisors, as they will always be benchmarked against the human equivalent. As a result, many companies started to humanize their advisors – probably as an attempt to gain some of the trustworthiness that is attributed to Humanized Humans. As H5 and H6 show, doing that can have the opposite of the hoped effect –

¹¹Possible explanations for the observed differences in budget allocation (human versus algorithm) between in Experiment II, real world behavior and Experiment II are reported in Chapter 5. Interpretation, Limitations and Future Research.

Table 11: Extended Cox regression (n=516)

Group	B	HR	95% CI	P Values
Humanized Human				
Dehumanized Human	.093	1.097	0.659 1.829	.721
Dehumanized Algorithm	.627	1.872	1.140 3.074	.013
Humanized Algorithm	.982	2.669	1.586 3.490	.000
Gender	-.445	.641	.419 .981	.040
Satisfaction with past returns	-.180	.835	.684 1.020	.078
Experience with investing	.163	1.117	.965 1.436	.108

Allocated budget (n=223)**Figure 6:** Allocation of 10.000€ (n=223)**Figure 7:** Allocation of 10.000€ split in non-users (n=110) and users of Ginmon (n=113)

at least for people who have never used a Roboadvisor. Humans seem to trust algorithms for their non-human traits (rational, neutral, high processing power) and are not looking for the same traits that are attributed to humans (empathy, accountability). When interpreting the results of H5 and H6, it makes sense to call oneself the following in mind: A technology can be perceived simple and unable to complete complex tasks, if it is named (Hafer et al., 1996; Riegelsberger et al., 2005). For human advisors, the trust level is increased, if they share personal information, as it demonstrates their willingness to risk their reputation (Garner, 2005; O'Keefe, 1990; Pornpitakpan, 2004). Note that a high trust level in itself is neither positive nor negative, but neutral. For an investor, it can also be beneficial to lose trust early, as we have seen in our case with the declining portfolio.

5.2. Differences to Hodge et al. (2018)

Regarding H2, H3 and H4, we have to say that ranks two to four are not coherent with Hodge et al. (2018) findings. Especially, Dehumanized Humans ranked higher in our study.

When comparing both studies, we have to keep in mind that the dependent variables were not the same. We measured, if and when they sell their portfolio, while the other study asked for their likelihood to follow the investment advice. Another possible explanation for the deviations: In our experiment the advisor was also the one making the investment decision, whereas in Hodge's experiment it was just an investment recommendation. In definitive and binding investment decision-making people seem to still prefer humans over algorithms. However, when it is just about recommendations, they might prefer an algorithm over a human. Future research should look deeper into that thesis. In addition to that, the sample (MBA students versus users and non-users of Roboadvisors) could be a reason for the diverging results. Furthermore, Hodge provided extensive background information while in our case the subjects had to "trust blindly". Most of the existing research focuses on AI as a recommendation engine which requires much lower trust levels than in settings with an algorithm as the final decision-maker. There-

fore, we encourage other scholars to study experimental settings which allocate more power to the algorithm. A future study could test, if a higher amount of background information about the investment decision leads to higher trust in algorithmic decisions. This would help exploring, if the differences in ranking placements two to four are caused by the diverging background information provided.

5.3. Experiment II: Allocation of the Budget Between Human and Algorithm

Respondents said they would allocate 70% of their budget to the algorithm, even though in reality they behave differently. There seems to be a difference between people's intentions and their actions. Experiment I has shown that the portfolios managed by the algorithm are sold earlier, while in Experiment II, respondents claim to trust the algorithm more than the human. One reason could be social desirability. Maybe some respondents think trusting AI is what the researchers expect as a result (Rosenthal effect). Especially, since parts of the sample were recruited using the author's personal social media account.

The result of 70% of budget allocated to the algorithm is also an extreme deviation from the real world: In 2019, Roboadvisors had only 7,7 Million EUR under management in Germany (Statista, 2018). A possible explanation for the difference could be certain patterns in our sample (young, many users of Roboadvisors) which prevent generalizations. Another reason could be that we forced respondents to compare between two options, which is not a realistic scenario. Future research could dig deeper into the divergence between intentions (allocating money to the algorithm) and actions (trusting humans more).

Additionally, it needs to be mentioned that the question from which this data originates was asked after Experiment I, which simulated a declining portfolio. Depending on who managed the declining portfolio, it might affect the answers to the subsequent question. Since our respondents are randomly assigned to groups with equal sizes, this effect can be slightly reduced.

5.4. Limitations Caused by the Sample

The overall sample ($N = 223$, 82% male, $M_{Age} = 36,6$, $SD_{Age} = 15,7$) is not representative for the "average German" for three reasons.

Firstly, our sample is too young since the non-user group ($N = 110$, 70% male, $M_{Age} = 27,2$, $SD_{Age} = 13,5$) mainly consisted of students. The median age of Germans is 44,3 years (UN, 2013). The age difference should not be neglected, since age is a factor in trust in AI and digital products in general (Scopelliti et al., 2005; Evers et al., 2008; Ho et al., 2005; Hancock et al., 2011). Secondly, our sample is male-dominated (82%). This might be representative for users of Roboadvisors or investors in general (Schnell, 2011), but not for the general population. If the goal is to find out how Roboadvisor could reach broader adoption, it is recommended to also look at potential users which are underrepresented in the current user base (e.g. females). Thirdly,

users were overrepresented in our sample compared to the German population. Since this is an exploratory work, the goal was not to have a representative sample for "the average German". Nonetheless, the limitations caused need to be mentioned.

An additional limitation comes with the limited sample size. The explorative analyses showed that a larger sample size could be beneficial. For example, a significant p-value for the Dehumanized Human group in the cox regression (Table 3) might be found with an increased sample size.

5.5. Comparing Users to Non-users

When we compare the two subgroups of users and non-users, we need to keep in mind that the non-user group is a lot younger and contains less females. These factors could be intervening variables when we try to make statements about the effects of being a user of a Roboadvisor.

When opposing users to non-users, the main conclusion is that we found lower p-values and higher b-values for the cox regression comparing all four advisor types (Table 8) as well as for the cox regression with a 2x2 design (Table 9).

An explanation could be that the Ginmon users are more used to portfolio volatility – no matter who manages it. They are already familiar with a somewhat humanized algorithm trading for them (Ginmon's algorithm is named Apeiron). They have more experience and therefore were harder to manipulate with our experimental stimulus. In contrast, non-users, who most likely interact with a humanized algorithm for the first time, might be more prone to manipulation. Consequently, the effect of humanization and advisor type is weakened compared to the non-users. You could say users are more immune to the effect of humanization.

In other words, to investigate effects of humanization we recommend future researchers taking a closer look at non-users. You could choose to observe people who have no experience with Roboadvisors to gain a deeper understanding of the differences between the influence of the four portfolio managers on loss tolerance. Also, you could gain a deeper understanding of diverging effect of humanizing humans versus algorithms.

5.6. Further Limitations and Future Research

Given the results of our exploratory analysis (increased sample size, control variables) and user versus non-user comparison, we can conclude that future researcher should try to recruit a bigger sample size, incorporate control variables and focus on non-users without prior automated investment experiences.

Trust in Artificial Intelligence is a rather broad topic. For the sake of feasibility, we had to focus on one specific manifestation of trust (loss tolerance) and one application of AI (Roboadvisors).

Many researchers tried to operationalize trust but did not arrive at an overarching conclusion. Therefore, for the dependent variable we decided to focus on one key aspect which made sense in the context of investing: How do investors react to losses? Do they sell or do they hold? This

comes with the limitation that we did not directly operationalize and measure trust. Future research could try to look at the bigger picture of trust, e.g. by surveying multiple items which cover more than loss reaction and loss tolerance. The study can be based on other scholars' attempts to quantify trust. For example, Cho et al. (2015) developed a scale ranging from complete distrust to complete trust including undistrust and untrust (Cho et al., 2015).

Also, our choice of the independent variable leads to certain shortcomings. As mentioned in the literature review, there are dozens of factors which potentially influence user trust in AI. In order to not exceed the limits of this research, we manipulated just one variable: humanization. Humanization is of great interest, as many companies in the space name their technologies or represent them humanlike (avatars, chatbots) – especially in finance. Since the effect on users is controversial in research, we decided to contribute to that discussion. Within the factor humanization, we – again – had to narrow it down. We named the decision-maker, but you could also humanize in other ways, e.g. give him/her a face or a voice. Future research could look at other factors which could influence user trust and at other interesting ways of humanization.

Another limitation is caused by the scope of our study. We cannot make general statements about AI as a whole. The domain is too broad due to the manifold of applications in different industries. We had to focus on AI-based finance products, more specifically Roboadvisors. Humanizing decision-makers could have the opposite effect (to what we found) in other domains such as entertainment, navigation or healthcare. Conducting a study in other domains is recommended. In general, due to the context-dependency of trust, studies in non-finance environments are also beneficial.

In our survey, we simulated the development of a portfolio over time. Obviously, the ideal and more realistic study would be a longitudinal study with real money to lose. Due to time- and budget-constraints this was not feasible, but maybe it is for other scholars. To give a concrete recommendation: We would conduct a similar survey, just in the mentioned realistic setting. Instead of just asking, if the subjects would like to hold or sell, we would add more questions on trust. For example, you could ask a question such as: Do you consider the decision-maker trustworthy? You could even offer them to switch to another decision-maker (human/algorithm) to allow for joint evaluation. If respondents answer the questions monthly, and not just directly one after another, it could lead to interesting insights and we could learn more about the change in trust over time.

Additionally, the simulation of the declining portfolio as the first part of the survey biases the questions afterwards. We still decided to put it first since the question was the core of our hypotheses and we wanted to maximize the sample size for the cox regression. Putting other questions first would have distorted subsequent answers as well. The bias should be bearable, as the portfolio development was identical across all four experimental groups. Nonetheless, future studies could run separate surveys to prevent that shortcom-

ing.

6. Implications

Humanizing AI seems not to be a sufficient solution to increase users trust in AI – at least in the finance domain. Also, the trust in AI (in terms of allocated investments) differs across demographics. You can ask yourself, if increasing the trust level should always be the end goal. In our experiment, as in many investment decisions in real life, you were better off (financially), if you lost trust sooner. The findings of this work have several implications on three different levels: Regulation, User Experience and Technology.

Regulation is an important topic, given the observed willingness of some consumers to allocate a substantial amount of money to algorithms. But regulators were not able to keep up with the rapid technological progress. Regulating AI in general is most likely not suitable since there are plenty of different applications and domains. Nonetheless, there is consensus that a global ethical standard for AI companies would be beneficial. The challenge is to protect fundamental rights and freedom while still encourage innovation of AI technologies. Especially, in our context of humanization and investments, there are issues arising when the borders between human and machine become blurred. Formal advancements by the EU are concerned with anti-discrimination, biases and fairness of algorithms. Also, in the finance domain, this can be an issue (e.g. credit scoring). Thelisson (2017) proposes four solutions: First, a Code of Conduct that one or multiple companies bind themselves to. Second, Quality Labels and Audits. This means the company either discloses the code publicly (issues with interpretability and IP remain) or hires another authority that looks into the code and certifies it with some kind of quality seal. Third, transparency in the data chain. And fourth, de-biasing datasets and algorithms which can be broken down to discrimination detection and discrimination prevention (Thelisson et al., 2017). The GDPR constitutes that individuals have the right to object to decisions even if they are made purely on the basis of automation. In addition, they also have the right to obtain information about the existence of an automated decision making system as well as the “logic involved” and its consequences (Thelisson, 2017). How this right of explanation plays out practically (e.g. depth of explanation, technical feasibility) will be decided in court in the future. The technological aspects of the explainability will be analyzed in the last paragraph of this chapter.

User experience also plays an important role to build trust in AI-based finance products. While humanization was not beneficial in our case, it still needs to be explored for other products or to other degrees (e.g. avatars). Other scholars found out that users prefer conversational agents which are young, match their ethnicity and show non-verbal cues while gender does not matter (Cowell and Stanney, 2003). These agents do not only humanize the product but also increase personalization and familiarity. There is empirical evidence

that these factors increase trust and product adoption (Komiak and Benbasat, 2006). From a UX perspective, it could also be beneficial to give users the option to interact with and try the algorithm. Dietvorst et al. (2018) showed that giving users the opportunity to modify the algorithm reduces algorithm aversion. Giving a share of control back to the human may sound like a contrast to the idea of a Roboadvisor (full automation), but even slight modifications increase the users' preference for that algorithm (Dietvorst et al., 2018). Giving the users the possibility to make optional slight alterations (e.g. changing the risk class after the initial investment) could be a solution. Another factor is accountability. Who is responsible if the algorithm loses your money? This can be considered as UX since the product needs to convey a sense of security (e.g. to make up for lack of accountability). One solution for startups, which do not have a well-known brand yet, is to offer the Roboadvisor as a whitelabel solution of an established bank. But the accountability issue also involves technological aspects (error-proneness) and legal aspects (Who can be hold accountable by law?). Additionally, UI-decisions, such as design, should also have the goal to increase the Roboadvisors' trustworthiness. An option to overcome the lack of trust in Roboadvisors is to slowly move customers from a conventional bank with a human advisor to a Roboadvisor. This allows them to downsize advisory operations. High net-worth individuals can still have their human account while others can also have the option to take part in economic growth by investing without expert knowledge. This is especially crucial for risk-averse, inexperienced, low-budget individuals (Jung et al., 2018).

Humanization is not a technological challenge in itself and also not beneficial according to the conducted study. Therefore, the technological implications evolve mainly around transparency of algorithms. Many users would be interested in understanding how the algorithm arrives at its conclusions by explaining its reasoning. For example, if your Roboadvisor loses money, you would like to know how it happened and seek a justification for the action that lead to losses (e.g. a specific trade) (Gregor and Benbasat, 1999). This kind of explainable AI would make decision-making more transparent and predictable and therefore easier to trust.

7. Conclusion

To conclude, we can say that we arrived at three main findings. Firstly, for non-users humanizing a human decision-maker increases loss tolerance, while humanizing an algorithm decreases loss tolerance. Given the slight manipulation (difference of one word), this finding is astonishing. Secondly, the humanized Human is still the most trusted financial advisor. Thirdly, despite the second finding: Respondents were willing to allocate much more money to the algorithmic advisor compared to the human advisor. This leaves us with many opportunities for future research since the scope and therefore potential generalization was limited in our study.

On a theoretical level, the conclusion is that there are many theories concerning trust in AI but most of them focus on automated recommendations for action and the human as the final decision maker. A structured, comprehensive model that includes the high number of trust factors and the algorithm as the decision maker is missing. With the current raise in automation and the complexity of trust, researchers from different disciplines shall collaborate to build a holistic model.

The question determining the next decades is not about technological feasibility. The amount of data available grows every day. The crucial question is, if users will trust in the decisions made by the algorithms. Given the current levels of users trust in Roboadvisors, a hybrid model of virtual advice in wealth management that combines human and algorithm seems to be realistic (Cocca, 2016).

References

- Adams, B., Bruyn, L., Houde, S., and Angelopoulos, P. Trust in automated systems literature review. Defense research and development Canada Toronto No. Technical report, CR-2003-096, 2003.
- André, E., Gimpel, H., and Olenberger, C. An investigation of the effects of anthropomorphism in collective human-machine decision-making. Technical report, 2018.
- Arkes, H. R. and Blumer, C. The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, 35(1):124–140, February 1985.
- Bernstein, P. L. *Against the gods: the remarkable story of risk*. New York, John Wiley & Sons, 1996.
- Bottazzi, L., Da Rin, M., and Hellmann, T. The importance of trust for investment: evidence from venture capital. *The Review of Financial Studies*, 29(9):2283–2318, 2016.
- Burgoon, J. K., Bonito, J. A., Bengtsson, B., Cederberg, C., Lundeberg, M., and Allspach, L. Interactivity in human-computer interaction: a study of credibility, understanding, and influence. *Computers in Human Behavior*, 16(6):553–574, November 2000.
- Carré, O., James, E., and Witz, P. Keynote presentation automation in financial advice: Trends, opportunities and implications. *The Luxembourg Bankers' Association event Robo-Advisors and Digital Investments*, 2016.
- Castelfranchi, C. and Falcone, R. *Trust theory: a socio-cognitive and computational model*, volume 18. Chichester, John Wiley & Sons, 2010.
- Chaminade, T., Hodgins, J., and Kawato, M. Anthropomorphism influences perception of computer-animated characters' actions. *Social Cognitive and Affective Neuroscience*, 2(3):206–216, September 2007.
- Chin, S. H. On application of game theory for understanding trust in networks. In *2009 International Symposium on Collaborative Technologies and Systems*, pages 106–110. IEEE, 2009.
- Cho, J., Swami, A., and Chen, I. A survey on trust management in mobile ad hoc networks. *IEEE Communications Surveys Tutorials*, 13(4):562–583, Fourth 2011. ISSN 2373-745X. doi: 10.1109/SURV.2011.092110.00088.
- Cho, J.-H., Chan, K., and Adali, S. A survey on trust modeling. *ACM Computing Surveys (CSUR)*, 48(2):1–40, 2015.
- Cocca, T. D. Potential and limitations of virtual advice in wealth management. *The Capco Institute Journal of Financial Transformation*, 44:45–57, 2016.
- Coeckelbergh, M. Can we trust robots? *Ethics and information technology*, 14(1):53–60, 2012.
- Cowell, A. J. and Stanney, K. M. Embodiment and interaction guidelines for designing credible, trustworthy embodied conversational agents. In *International workshop on intelligent virtual agents*, pages 301–309. Springer, 2003.
- Dawes, R. M. The robust beauty of improper linear models in decision making. *American psychologist*, 34(7):571, 1979.
- Diab, D. L., Pui, S.-Y., Yankelevich, M., and Highhouse, S. Lay perceptions of selection decision aids in us and non-us samples. *International Journal of Selection and Assessment*, 19(2):209–216, 2011.
- Dietvorst, B. J., Simmons, J. P., and Massey, C. Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1):114, 2015.
- Dietvorst, B. J., Simmons, J. P., and Massey, C. Overcoming algorithm aversion: people will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, 64(3):1155–1170, 2018.
- Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., and Beck, H. P. The role of trust in automation reliance. *International journal of human-computer studies*, 58(6):697–718, 2003.
- Eastwood, J., Snook, B., and Luther, K. What people want from their professionals: attitudes toward decision-making strategies. *Journal of Behavioral Decision Making*, 25(5):458–468, 2012.
- Einhorn, H. J. and Hogarth, R. M. Decision making under ambiguity: a note. In *Risk, decision and rationality*, pages 327–336. Springer, 1988.
- Epley, N., Waytz, A., Akalis, S., and Cacioppo, J. T. When we need a human: Motivational determinants of anthropomorphism. *Social Cognition*, 26(2):143–155, April 2008.
- Epperson, T., Hedges, B., Singh, U., and Gabel, M. Hype vs. reality: The coming waves of “robo” adoption. Technical report, A.T. Kearney, 2015.
- Eule, A. Rating the robo-advisors. barron's-the dow jones business and financial weekly, 2017. URL <https://www.barrons.com/articles/rating-the-robo-advisors-1501303316>.
- Evers, V., Maldonado, H., Brodecki, T., and Hinds, P. Relational vs. group self-construal: untangling the role of national culture in hri. In *2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 255–262. IEEE, 2008.
- Eyssel, F., Ruitter, L. D., Kuchenbrandt, D., Bobinger, S., and Hegel, F. 'if you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. URL <https://aiweb.techfak.uni-bielefeld.de/files/p125.pdf>.
- Gambetta, D. 'Can We Trust Trust?', in Gambetta, Diego (ed.) *Trust: Making and Breaking Cooperative Relations*, electronic edition, Department of Sociology, University of Oxford, chapter 13, pp. 213-237, 2000. URL <http://www.sociology.ox.ac.uk/papers/gambetta213-237.pdf>.
- Garner, R. What's in a Name? Persuasion Perhaps. *Journal of Consumer Psychology*, 15(2):108–116, 2005.
- Glaser, M. and Walther, T. Run, Walk, or Buy? Financial Literacy, Dual-Process Theory, and Investment Behavior. *SSRN Electronic Journal*, 2013.
- Gold, N. A. and Kursh, S. R. Counterrevolutionaries in the financial services industry: Teaching disruption-a case study of roboadvisors and incumbent responses. *Business Education Innovation Journal*, 9(1):139–146, 2017.
- Gong, L. How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior*, 24(4):1494–1509, July 2008.
- Gregor, S. and Benbasat, I. Explanations from intelligent systems: Theoretical foundations and implications for practice. *MIS Quarterly*, 23(4):497, December 1999.
- Greiner, M. E. and Wang, H. Building consumer-to-consumer trust in e-finance marketplaces: An empirical analysis. *International Journal of Electronic Commerce*, 15(2):105–136, 2010.
- Grinblatt, M. and Keloharju, M. The investment behavior and performance of various investor types: a study of finland's unique data set. *Journal of Financial Economics*, 55(1):43–67, January 2000.
- Grove, W. M. and Meehl, P. E. Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, public policy, and law*, 2(2):293, 1996.
- Guiso, L., Sapienza, P., and Zingales, L. Trusting the stock market. *The Journal of Finance*, 63(6):2557–2600, December 2008.
- Guthrie, S. E. *Faces in the Clouds: A New Theory of Religion*. Oxford, Oxford University Press, 1993.
- Hafer, C. L., Reynolds, K. L., and Obertynski, M. A. Message comprehensibility and persuasion: Effects of complex language in counterattitudinal appeals to laypeople. *Social Cognition*, 14(4):317–337, December 1996.
- Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., and Parasuraman, R. A meta-analysis of factors affecting trust in human-robot interaction. *Human factors*, 53(5):517–527, 2011.
- Haugen, R. A. *The new finance: the case against efficient markets*. Henglewood Cliffs, Prentice Hall, 1995.
- Highhouse, S. Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology*, 1(3):333–342, 2008.
- Ho, G., Wheatley, D., and Scialfa, C. T. Age differences in trust and reliance of a medication management system. *Interacting with Computers*, 17(6):690–710, 2005.
- Hodge, F. D., Mendoza, K. I., and Sinha, R. K. The effect of humanizing robo-advisors on investor judgments. Available at SSRN 3158004, 2018.
- Houser, D., Schunk, D., and Winter, J. Distinguishing trust from risk: an anatomy of the investment game. *Journal of Economic Behavior & Organization*, 74(1-2):72–81, May 2010.
- James, H. S. The trust paradox: a survey of economic inquiries into the nature of trust and trustworthiness. *Journal of Economic Behavior & Organization*, 47(3):291 – 307, 2002. ISSN 0167-2681. doi: [https://doi.org/10.1016/S0167-2681\(01\)00214-1](https://doi.org/10.1016/S0167-2681(01)00214-1). URL <http://www.sciencedirect.com/science/article/pii/S0167268101002141>.
- Jung, D., Dorner, V., Weinhardt, C., and Puzmaz, H. Designing a robo-advisor for risk-averse, low-budget consumers. *Electronic Markets*, 28(3):367–380, August 2018.
- Kahneman, D. and Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, 1979.
- Komiak, S. Y. X. and Benbasat, I. The effects of personalization and familiarity on trust and adoption of recommendation. Technical report, 2006.
- Krause, P. and Schäfer, A. Verteilung von Vermögen und Einkommen in Deutschland: Große Unterschiede nach Geschlecht und Alter. *DIW*

- Wochenbericht, 72(11):199–207, 2005.
- Kydd, A. H. *A blind spot of philosophy. Trust and Mistrust in International Relations*. Princeton University Press, 2005.
- Lagerspetz, O. *Trust: the tacit demand*, volume 1. Luxembur, Springer Science & Business Media, 1998.
- Leaden, R. 9 of the Funniest and Most Shocking AI Fails, 2017. URL <https://www.entrepreneur.com/slideshow/289621>.
- Lee, J. D. and See, K. A. Trust in automation: designing for appropriate reliance. *Human factors*, 46(1):50–80, 2004.
- Lee, S., Choi, J., Ngo-Ye, T. L., and Cummings, M. Modeling trust in the adoption decision process of robo-advisors: an agent-based simulation approach. In *Proceedings of the Tenth Midwest Association for Information Systems Conference, Saint Louis, MO, USA*, pages 17–18, 2018.
- Logg, J. M., Minson, J. A., and Moore, D. A. Algorithm appreciation: people prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151:90–103, 2019.
- Mayer, R. C., Davis, J. H., and Schoorman, F. D. An integrative model of organizational trust. *Academy of management review*, 20(3):709–734, 1995.
- Mcknight, D. H., Carter, M., Thatcher, J. B., and Clay, P. F. Trust in a specific technology: an investigation of its components and measures. *ACM Transactions on management information systems (TMIS)*, 2(2):1–25, 2011.
- Millward, P. The 'grey digital divide': perception, exclusion and barriers of access to the internet for older people. *First Monday*, 8(7), July 2003.
- Nan, X., Anghelcev, G., Myers, J. R., Sar, S., and Faber, R. What if a Web Site can Talk? Exploring the Persuasive Effects of Web-Based Anthropomorphic Agents. *Journalism & Mass Communication Quarterly*, 83(3): 615–631, September 2006.
- Odean, T. Are investors reluctant to realize their losses? *The Journal of Finance*, 53(5):1775–1798, October 1998.
- O'Keefe, D. J. *Persuasion: theory and research*. Thousand Oaks, SAGE Publications, 1990.
- Önköl, D., Goodwin, P., Thomson, M., Gönül, S., and Pollock, A. The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making*, 22(4):390–409, 2009.
- Payne, J. W. Relation of perceived risk to preferences among gambles. *Journal of Experimental Psychology: Human Perception and Performance*, 1(1): 86–94, 1975.
- Pornpitakpan, C. The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology*, 34(2): 243–281, February 2004.
- Promberger, M. and Baron, J. Do patients trust computers? *Journal of Behavioral Decision Making*, 19(5):455–468, 2006.
- Riegelsberger, J., Sasse, M. A., and McCarthy, J. D. The mechanics of trust: A framework for research and design. *International Journal of Human-Computer Studies*, 62(3):381–422, March 2005.
- Rotter, J. B. Interpersonal trust, trustworthiness, and gullibility. *American psychologist*, 35(1):1–7, 1980.
- Salo, A. Robo advisor, your reliable partner? Building a trustworthy digital investment management service. Master's thesis, 2017.
- Schnell, C. Wenn's ums geld geht – männersache. Retrieved from <https://www.handelsblatt.com/finanzen/maerkte/boerse-inside/privatvermoegen-wenns-ums-geld-geht-maennersache-seite-2/5784664-2.html>, 2011.
- Scopelliti, M., Giuliani, M. V., and Fornara, F. Robots in a domestic setting: a psychological approach. *Universal access in the information society*, 4(2): 146–155, 2005.
- Shandwick, W. Ai-ready or not: Artificial intelligence here we come! - weber shandwick, 2016. URL <https://www.webershandwick.com/news/ai-ready-or-not-artificial-intelligence-here-we-come/>.
- Shapira, Z. *Risk taking: a managerial perspective*. New York, Russel Sage Foundation, 1995.
- Siau, K. and Wang, W. Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, 31(2):47–53, 2018.
- Slovic, P. and Lichtenstein, S. Relative importance of probabilities and payoffs in risk taking. *Journal of Experimental Psychology*, 78:1–18, 1968.
- Statista. Robo-Advisors Deutschland. Technical report, 2018.
- Statman, M. Behavioral finance: Past battles and future engagements. *Financial Analysts Journal*, 55(6):18–27, November 1999.
- Taddeo, M. Modelling trust in artificial agents, a first step toward the analysis of e-trust. *Minds and machines*, 20(2):243–257, 2010.
- Teigen, K. H. Risk-taking behavior. *Journal of Behavioral Decision Making*, 9:73–74, March 1996.
- Thelisson, E. Towards Trust, Transparency and Liability in AI/AS systems. In *IJCAI*, pages 5215–5216, 2017.
- Thelisson, E., Padh, K., and Celis, L. E. Regulatory mechanisms and algorithms towards trust in AI/ML. In *Proceedings of the IJCAI 2017 Workshop on Explainable Artificial Intelligence (XAI), Melbourne, Australia*, 2017.
- UN. World population prospects: The 2012 revision. Technical report, 2013.
- Van Thiel, D. and Van Raaij, F. Explaining customer experience of digital financial advice. *Economics*, 5(1):69–84, 2017.
- Venkatesh, V. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2):186–204, December 2000.
- Yousafzai, S. Y., Pallister, J. G., and Foxall, G. R. A proposed model of e-trust for electronic banking. *Technovation*, 23(11):847–860, 2003.



Hätte eine Pre-Mortem-Analyse den Tod von eLWIS verhindern können? – Verhaltensorientierte Ansätze für effektives Kostenmanagement in Großprojekten

Sonja Schattevoy

Heinrich-Heine-Universität Düsseldorf

Abstract

Im Juni 2018 gab der Discounter Lidl bekannt, dass er das in Kooperation mit SAP durchgeführte IT-Investitionsprojekt „eLWIS“ mit sofortiger Wirkung einstellen würde. Die bis zu diesem Zeitpunkt angefallenen Kosten betrug bereits rund 500 Mio. Euro. Damit reiht sich Lidl in eine durchaus beachtliche Liste jener Unternehmen ein, die ein geplantes IT-Großprojekt nicht im vorgesehenen Kosten- und Zeitrahmen abzuwickeln vermochten. Die enorme Verbreitung dieser Problematik über eine Vielzahl von Branchen hinweg gibt Anlass zu einer intensiven Analyse der Erfolgsfaktoren für effektives Kostenmanagement in Großprojekten. Einen vielversprechenden Ansatz stellt in diesem Kontext das verhaltensorientierte Controlling dar, das vor allem Maßnahmen gegen kognitive Fehlleistungen und Könnensdefizite der handelnden Akteure in den Mittelpunkt der Betrachtung rückt. Im Rahmen dieser Arbeit werden ausgewählte kognitive Verzerrungen dargestellt, die im Kontext von Investitionsprojekten zu dysfunktionalem Entscheidungsverhalten führen können. Darauf aufbauend sollen Lösungsmöglichkeiten dargelegt werden, mit deren Hilfe diese Verzerrungen aufgefangen und negative Konsequenzen für das Unternehmen vermieden werden können.

Keywords: Investitionsprojekte; Projektcontrolling; verhaltensorientiertes Controlling; kognitive Verzerrungen.

1. „Over Budget, Over Time, Over and Over Again“¹ - Herausforderungen eines effektiven Kostenmanagements in Großprojekten

„ELWIS [...] 'rockt und rollt' nicht mehr“ (Nickel, 2018). Diese Hiobsbotschaft musste die Lidl Stiftung & Co. KG, mit rund 10.000 Filialen, 140 Logistikzentren (vgl. Schüler, 2018) und einem Jahresumsatz von über 70 Milliarden € einer der größten Lebensmitteldiscounter der Welt (vgl. Lehmann, 2018, S. 1), im Juli 2018 der Öffentlichkeit verkünden (vgl. Nickel, 2018). Gemeint war damit nicht etwa der King of Rock'n'Roll, sondern eines der größten IT-Projekte in der Historie des Unternehmens (vgl. Kannenberg, 2018; Kolf und Kerkmann, 2018). Im Zuge dieses 2011 in Kooperation mit dem Software-Unternehmen SAP gestarteten Projekts mit dem Beinamen ‚eLWIS‘ (kurz für ‚elektronisches Lidl-Warenwirtschaftsinformationssystem‘, vgl. Mumme, 2018) hatte sich Lidl zum Ziel gesetzt, durch die Einführung eines neuen Warenwirtschaftssystems eine weltweite Harmonisierung der Warenwirtschaftsprozesse und eine Vereinheitlichung der Stammdaten in sämtlichen Filialen

des Unternehmens zu erreichen (vgl. Mumme, 2018; Rode, 2015). Obwohl das auf der neuartigen Datenbanktechnologie SAP HANA basierende System (vgl. Schüler, 2018) im Laufe des Jahres 2015 bereits in einigen Landesgesellschaften in den Testbetrieb genommen wurde (vgl. Kannenberg, 2018; Mumme, 2018; Rode, 2015), kam es schließlich im Juli 2018 zum Abbruch des Projekts (vgl. Lehmann, 2018, S. 4). Zu diesem Zeitpunkt hatte das Unternehmen bereits sieben Jahre Entwicklungszeit und schätzungsweise rund 500 Millionen € in eLWIS investiert (vgl. Kolf und Kerkmann, 2018). Den Abbruch begründet Lidl damit, dass „die ursprünglich definierten strategischen Ziele nicht mit vertretbare[m] Aufwand“ zu realisieren gewesen seien (vgl. Stockburger und Fritze, 2018).

Da diese Begründung insbesondere angesichts der bereits investierten Mittel zunächst doch recht vage anmutet, scheint es für das Verständnis einer derartigen Fehlentwicklung erforderlich zu sein, die Charakteristika und die Zieldimensionen eines solchen Projekts näher zu betrachten.

Im betriebswirtschaftlichen Kontext ist unter einem Projekt im Allgemeinen ein einmaliges „Vorhaben mit definierbarem Anfang und Abschluss“ (Madauss, 2017, S. 4) zu verste-

¹Flyvbjerg et al. (2009), S. 171.

hen, das unter Verwendung begrenzter Ressourcen eine klar definierte Zielvorgabe verfolgt (vgl. Fiedler, 2016, S. 2). Diese Zielvorgabe lässt sich dabei in drei untergeordnete Zielkategorien gliedern. Verfolgt werden zunächst stets Kosten- und Terminziele, deren Einhaltung sicherstellen soll, dass das Projekt innerhalb der durch die Ressourcenrestriktionen vorgegebenen Grenzen verbleibt. Außerdem existieren in jedem Projekt Sachziele, die das angestrebte Projektergebnis und die Qualität dieses Ergebnisses umfassen (vgl. Fiedler, 2016, S. 1).

Verantwortlich für das Erreichen dieser Ziele ist das Projektmanagement bzw. der Projektmanager (vgl. Mörsdorf, 1998, S. 2), dem gemäß der Projektmanagementdefinition der DIN 69901 die „Gesamtheit von Führungsaufgaben, -organisation, -techniken und -mitteln für die Initiierung, Definition, Planung, Steuerung und den Abschluss von Projekten“ (Deutsches Institut für Normung, 2009, S. 14) obliegt. Unterstützung erhält er dabei durch das Projektcontrolling, dem die Aufgabe zuteil wird, das Projektmanagement während des gesamten Projekts subsidiär zu begleiten und der Controlling-Definition nach Weber folgend insbesondere auch die Rationalität des Projektmanagements zu gewährleisten (vgl. Fiedler, 2016, S. 12).

Die Begründung der Lidl Stiftung & Co. KG für den Abbruch von eLWIS lässt vermuten, dass es dem Projektmanagement nicht gelang, das vorgegebene Sachziel – also die Entwicklung eines konzernweit einheitlichen Warenwirtschafts-systems – unter Einhaltung der gegebenen Kosten- und Terminziele zu erreichen. Daraus ergibt sich die Frage, ob es sich hierbei um einen Einzelfall von Managementversagen handelt, der durch eine angemessene Projektsteuerung hätte verhindert werden können, oder ob nicht vielmehr ein tiefer greifendes Problem vorliegt, das einer umfassenderen Analyse bedarf.

Ein Blick auf einige vergleichbare Projekte der vergangenen Jahre genügt, um zu erkennen, dass Lidl mit seinen Schwierigkeiten bei der Durchführung eines IT-Großprojekts keineswegs alleine ist. So sahen sich neben Lidl beispielsweise auch der Otto-Konzern, die Deutsche Bank und der Lidl-Konkurrent Edeka mit ähnlichen Problemen bei der Einführung eines neuen SAP-Systems konfrontiert (vgl. Kroker, 2018). Auch bei einer Erweiterung der Betrachtung über die Landesgrenzen hinaus zeigen sich vergleichbare Schwierigkeiten. Der CHAOS-Report aus dem Jahr 2015 belegt, dass rund 43 % der größten IT-Projekte der vergangenen Jahre ebenso wie eLWIS vollständig scheiterten, weitere 51 % konnten lediglich mit deutlichen Kosten- und/oder Terminüberschreitungen fertiggestellt werden. Bemerkenswert ist bei dieser Studie vor allem der Unterschied zwischen Groß- und Kleinprojekten: Bei den Kleinprojekten scheiterten nicht einmal 10 %, Kosten- und Terminüberschreitungen lagen nur bei rund 30 % dieser Projekte vor (vgl. International, 2015, S. 3).

Bei einer Fokussierung auf Großprojekte zeigt hingegen auch der Blick über den IT-Bereich hinaus, dass eine Vielzahl von Projekten mit erheblichen Problemen bezüglich der Erreichung ihrer Sach-, Kosten- und Terminziele zu kämpfen

hat (vgl. Flyvbjerg et al., 2009, S. 171 f). Es lässt sich also eine gewisse Systematik erkennen, die Flyvbjerg pointiert in seinem „iron law of megaprojects“ zum Ausdruck bringt: „Over budget, over time, over and over again“ (Flyvbjerg, 2014, S. 11).

Da dieses Phänomen über eine Vielzahl von Branchen hinweg zu beobachten ist und Flyvbjerg zufolge bereits seit mehreren Jahrzehnten das Management großer Investitionsprojekte prägt (vgl. Flyvbjerg, 2005, S. 1 und S. 3), stellt sich die Frage, welche Faktoren ein derart systematisches Versagen des Projektmanagements begründen und welche Maßnahmen im Rahmen des Projekt- und Kostenmanagements ergriffen werden können, um diesem Trend entgegenzuwirken.

Um der Beantwortung dieser Frage einen Schritt näher zu kommen, soll im Rahmen dieser Arbeit zunächst analysiert werden, welche spezifischen Charakteristika Großprojekte aufweisen, die sie im Vergleich zu ihren kleineren Gegenstücken deutlich anfälliger für eine Verfehlung der Projektziele machen. Im Zuge dessen soll außerdem kurz dargelegt werden, welche Ursachen das Scheitern eines Projekts im Allgemeinen haben kann. Im weiteren Verlauf soll der Fokus speziell auf eine dieser Ursachen, nämlich auf Fehlentscheidungen aufgrund psychologischer Faktoren bzw. kognitiver Begrenzungen der Projektmanager, gelegt werden. Hierfür werden in Kapitel 3 ausgewählte kognitive Verzerrungen betrachtet, die in verschiedenen Phasen der Projektdurchführung zu Entscheidungsfehlern der Manager führen können und die möglicherweise auch im Falle von eLWIS eine Fehlentwicklung des Projekts ausgelöst haben. Darauf aufbauend werden in Kapitel 4 einige konkrete Lösungsmöglichkeiten präsentiert, die insbesondere der Projektcontroller im Rahmen des Kostenmanagements eines solchen Projekts ergreifen kann, um den drei zuvor erläuterten kognitiven Verzerrungen entgegenzuwirken und folglich das Verfehlen der Kosten-, Termin- und Sachziele zu verhindern. Abschließend soll in Kapitel 5 dargelegt werden, inwieweit die in Kapitel 3 aufgeführten Verzerrungen auch im Rahmen von eLWIS aufgetreten sein könnten und inwiefern es plausibel erscheint, dass eine Anwendung der erläuterten Gegenmaßnahmen – also beispielsweise die Durchführung einer Pre-Mortem-Analyse – das Scheitern dieses Projekts hätte verhindern können.

2. Spezifische Charakteristika großer Investitionsprojekte

Auch wenn die zuvor angedeutete Diskrepanz zwischen den Erfolgsaussichten großer und kleiner IT-Investitionsprojekte zunächst verwunderlich erscheinen mag, lässt sich dieser Sachverhalt durch eine kurze Betrachtung der spezifischen Eigenschaften großer Investitionsprojekte plausibel erklären.

Bei dem wohl wichtigsten Merkmal eines Großprojekts, den durch das Projekt verursachten Kosten, lässt sich der Literatur kein eindeutiger Grenzwert entnehmen, der ein Projekt als Großprojekt klassifiziert (vgl. Oliomogbe und Smith, 2012, S. 618). Überwiegend werden Großprojekte jedoch als

diejenigen Projekte charakterisiert, deren Plankosten einen Betrag von 100 Millionen \$ übersteigen (vgl. Flyvbjerg, 2005, S. 1; Flyvbjerg et al., 2009, S. 345; Oliomogbe und Smith, 2012, S. 618).

Diese Projekte weisen eine Vielzahl von Eigenschaften auf, die sie im Vergleich zu ihren kleineren Korrelaten anfälliger machen für deutliche Zeit- und Kostenüberschreitungen. So erstrecken sich Großprojekte meist über einen überdurchschnittlich langen Planungshorizont (vgl. Biesenthal et al., 2018, S. 45; Kardes et al., 2013, S. 909), binden in hohem Umfang Ressourcen menschlicher, technologischer und finanzieller Art (vgl. Biesenthal et al., 2018, S. 45), umfassen herausfordernde technologische Neuerungen (vgl. Kardes et al., 2013, S. 909) und müssen die Interessen einer großen Zahl unterschiedlicher Stakeholder berücksichtigen (vgl. Flyvbjerg, 2005, S. 2). All diese Faktoren haben zur Konsequenz, dass sich das Management solcher Projekte mit einer merklich erhöhten Komplexität konfrontiert sieht (vgl. Kardes et al., 2013, S. 606), die mit gesteigerten Risiken in Bezug auf Kosten- und Termintreue einhergeht – dieser Tatsache wird bei der Planung und Durchführung der Projekte jedoch häufig nicht die erforderliche Aufmerksamkeit eingeräumt (vgl. Flyvbjerg, 2014, S. 9). Dies hat zur Folge, dass im Zuge der Projektplanung keine ausreichenden Puffer für mögliche Zeit- und Budgetüberschreitungen berücksichtigt werden (vgl. Flyvbjerg, 2014, S. 9), wodurch Kostenüberschreitungen von mehr als 100 % keine Seltenheit darstellen (vgl. Flyvbjerg et al., 2013, S. 44).

Bei der Betrachtung von IT-Projekten verschärft sich diese Problematik noch einmal (siehe hierzu auch Flyvbjerg, 2009, S. 363: „If a major project is not already messed up, injecting a good dose of ICT will do the job [...]“). Zwar beträgt die durchschnittliche Kostenüberschreitung hier laut einer Studie von Flyvbjerg/Budzier nur knapp 30 %, allerdings befindet sich unter den IT-Projekten eine hohe Anzahl an sogenannten ‚black swans‘, also an Projekten, die Kostenüberschreitungen von mehr als 200 % aufweisen (vgl. Flyvbjerg und Budzier, 2011, S. 24).

Da trotz dieser offensichtlichen Defizite im Bereich des Kosten- und Projektmanagements ein stetiger Anstieg in der Anzahl und im Umfang solcher Großprojekte zu verzeichnen ist (vgl. Flyvbjerg, 2014, S. 12), scheint es ratsam, die Gründe für das derart häufige Verfehlen der Projektziele näher zu betrachten und daraus entsprechende Gegenmaßnahmen abzuleiten, um die enormen wirtschaftlichen Folgen eines solchen Scheiterns für die betroffenen Unternehmen künftig vermeiden zu können (vgl. Buhl und Meier, 2011, S. 59 f.).

Flyvbjerg führt in seinen Arbeiten drei zentrale Faktoren auf, die eine mangelhafte Prognose der tatsächlichen Kosten und damit unter Umständen auch das Scheitern eines Projekts zur Konsequenz haben können (vgl. Flyvbjerg, 2005, S. 8; Flyvbjerg, 2006, S. 6). Zum einen handelt es sich hierbei um ‚technologische Faktoren‘, die in erster Linie Fehler durch unzureichende Forecasting-Methoden, schlechte Datenqualität und Prognoseprobleme aufgrund mangelnder Erfahrung erfassen (vgl. Flyvbjerg, 2005, S. 8). Zum anderen werden ‚politisch-ökonomische Faktoren‘ genannt, die das Projekt-

management dazu veranlassen können, ihr Projekt betreffende Informationen bewusst positiv verzerrt darzustellen, um beispielsweise im Wettbewerb um die begrenzten Ressourcen des Unternehmens einen Vorteil gegenüber einem Konkurrenzprojekt realisieren zu können (vgl. Flyvbjerg, 2006, S. 6 f.).

Der Fokus dieser Arbeit soll jedoch auf dem dritten von Flyvbjerg aufgezeigten Aspekt, den ‚psychologischen Faktoren‘, liegen. Die psychologischen Faktoren beschreiben das Phänomen, dass das Projektmanagement in einigen Situationen aufgrund von Rationalitätsdefiziten bzw. kognitiven Verzerrungen nicht befähigt ist, die Erfolgchancen des eigenen Projekts realistisch zu beurteilen (vgl. Flyvbjerg, 2005, S. 8 f.). Das Auftreten dieser Rationalitätsdefizite ist damit zu begründen, dass der Mensch in vielerlei Hinsicht nicht dem Idealbild des Homo Oeconomicus entspricht, sondern vielmehr aufgrund von Könnens-, Wissens- und Willensdefiziten in seinen kognitiven Fähigkeiten begrenzt ist (vgl. Hirsch et al., 2008, S. 6-8; Merchant und Van der Stede, 2017, S. 12-14). Diese Begrenzung der kognitiven Fähigkeiten, deren Auftreten sich beispielsweise mithilfe der Erkenntnisse der Dual-System Theory begründen lässt (vgl. Kahneman, 2012, S. 25; Evans, 2008, S. 266), kann in komplexen Entscheidungssituationen dazu führen, dass das Projektmanagement nicht zu einer idealen Entscheidung im Sinne der klassischen Entscheidungstheorie gelangt, sondern insbesondere aufgrund von Könnensdefiziten Fehlentscheidungen trifft, die letztlich zu suboptimalen Ergebnissen führen (vgl. Zayer und Hirsch, 2006, S. 648).

Die spezifischen kognitiven Verzerrungen, die das Entscheidungsverhalten eines Projektmanagers im Einzelfall beeinflussen können, sind so vielfältig wie die im Rahmen eines Projekts zu treffenden Entscheidungen selbst (vgl. für eine exemplarische Auswahl Lovallo und Sibony, 2010, S. 44 f.). Aus diesem Grund soll im Folgenden lediglich eine Auswahl der zentralen kognitiven Faktoren dargestellt werden, die im Laufe der Projektdurchführung zu Dysfunktionalitäten im Entscheidungsverhalten führen können. Hierbei wird bewusst auf die Entscheidungen des Projektmanagers als Individuum fokussiert; eine Betrachtung von Entscheidungsfehlern in Gruppen soll mit Rücksicht auf den Umfang der Arbeit nicht erfolgen.

3. Verhaltensorientierte Erklärungsansätze für Entscheidungsfehler bei der Durchführung großer Investitionsprojekte

3.1. Übermäßiger Optimismus im Zuge der Investitionsentscheidung

Bereits vor dem Start der Projektdurchführung, also in der Phase der vorangehenden Investitionsentscheidung, können bei den beteiligten Entscheidern wesentliche kognitive Verzerrungen auftreten, die schon bei der Auswahl des durchzuführenden Projekts zu weitreichenden Entscheidungsfehlern führen können (vgl. Weber et al., 2003, S. 11 ff.).

Insbesondere im Rahmen der Datenbeschaffung, die unter anderem der Abschätzung der aus dem Projekt resultierenden Zahlungsströme dient (vgl. Weber et al., 2003, S. 12; Zayer und Hirsch, 2006, S. 649), tritt bei den Entscheidern häufig das Problem des übermäßigen Optimismus auf (vgl. Weber et al., 2003, S. 13; Zayer und Hirsch, 2006, S. 649).

Übermäßiger Optimismus, in der englischsprachigen Literatur auch als ‚optimism bias‘ bezeichnet (vgl. Flyvbjerg, 2006, S. 6; Flyvbjerg, 2005, S. 2), beschreibt das Phänomen, dass es insbesondere bei komplexen Großprojekten regelmäßig zu einer systematisch positiv verzerrten Schätzung von Kosten, Erlösen und Implementierungszeiten kommt (vgl. Flyvbjerg et al., 2009, S. 172). So werden die Kosten und der Zeitbedarf eines Projekts durch die beteiligten Personen meist deutlich zu niedrig geschätzt, während die daraus resultierenden Erlöse merklich zu hoch angesetzt werden (vgl. Flyvbjerg et al., 2009, S. 172). Diese kognitive Fehlleistung, die in mehreren empirischen Studien konsistent nachgewiesen werden konnte (vgl. Zayer, 2011, S. 227), führt dazu, dass die Entscheider zu einer rationalen Investitionsentscheidung nicht im Stande sind. Stattdessen werden die Erfolgsaussichten des betreffenden Projekts deutlich zu positiv eingeschätzt, während mögliche Aspekte, die zum Verfehlen der Projektziele führen könnten, bei der Betrachtung außen vor bleiben (vgl. Flyvbjerg et al., 2009, S. 172; Mahlendorf, 2010, S. 111).

Um in Kapitel 4.2 geeignete Gegenmaßnahmen ableiten zu können, die dem Projektmanagement helfen, dieses kognitive Begrenzung zu überwinden und sich somit vor psychologisch bedingten Fehlentscheidungen zu schützen, gilt es zunächst zu analysieren, welche Faktoren übermäßigen Optimismus auslösen bzw. begünstigen können.

In der Literatur findet sich eine Vielzahl von Ursachen, die übermäßigen Optimismus im Zuge einer Investitionsentscheidung hervorrufen können (vgl. Lovallo und Kahneman, 2003, S. 58; Weber et al., 2003, S. 15). Einer der am häufigsten dargelegten Einflussfaktoren ist die sogenannte Kontrollillusion bzw. illusion of control (vgl. Weber et al., 2003, S. 15). Als Kontrollillusion wird der Umstand bezeichnet, dass Entscheider häufig dem Glauben unterliegen, eine Situation maßgeblich beeinflussen zu können, und zwar selbst dann, wenn ihre tatsächlichen Einflussmöglichkeiten stark begrenzt sind (vgl. Mahlendorf, 2010, S. 111; Weber et al., 2003, S. 15; Zayer und Hirsch, 2006, S. 650). In solchen Situationen, in denen ein Projektmanager seinem Empfinden nach Kontrolle über die Realisierung der Projektziele hat – was bei klassischen Investitionsprojekten üblicherweise der Fall ist (vgl. Zayer, 2007, S. 144) –, zeigt sich, dass die Entscheider dazu tendieren, die Erfolgsaussichten ihres Projekts merklich zu positiv zu beurteilen (vgl. Zayer und Hirsch, 2006, S. 650) und den Einfluss von Faktoren, die außerhalb ihrer Kontrolle liegen, deutlich zu verkennen (vgl. Gehrig und Breu, 2013, S. 48; Lovallo und Kahneman, 2003, S. 59).

Ein weiterer wesentlicher Einflussfaktor auf den übermäßigen Optimismus ist der Overconfidence-Effect (vgl. Zayer, 2007, S. 143), der insbesondere diejenigen Projektmanager betrifft, die in der Vergangenheit bereits einige Erfolge bei

der Durchführung von Investitionsprojekten erzielen konnten (vgl. Kahneman et al., 2011, S. 58). Bei Vorliegen des Overconfidence-Effects beurteilt der Entscheider die Qualität seiner Vorhersagen und Prognosen als überdurchschnittlich gut, wodurch eine kritische Auseinandersetzung mit möglichen Fehlern innerhalb der eigenen Schätzung unterbleibt (vgl. Zayer, 2007, S. 148). Das Auftreten von Overconfidence kann zweierlei Ursachen haben. Zum einen tendieren Entscheider dazu, ihre eigenen Fähigkeiten zu überschätzen, da sie zurückliegende Erfolge im Wesentlichen ihren eigenen Kompetenzen zuschreiben, während Misserfolge vorwiegend auf exogene, außerhalb ihrer Verantwortung liegende Faktoren zurückgeführt werden (vgl. Lovallo und Kahneman, 2003, S. 58; Hastorf et al., 1970, S. 73). Dies kann dazu führen, dass Projektmanager sich auch solche Erfolge anrechnen lassen, die letztlich auf glückliche Zufälle zurückzuführen sind; hierdurch wird der Glaube der Entscheider in ihre eigenen Talente übermäßig gestärkt (vgl. Lovallo und Kahneman, 2003, S. 59). Zum anderen besteht das Problem, dass in der Wahrnehmung der an der Entscheidung beteiligten Personen häufig keine klare Trennung zwischen Kompetenz und Overconfidence existiert. So sieht sich ein Manager, der offen bekennt, dass er nicht ‚confident‘ ist, der Gefahr ausgesetzt, von den anderen Beteiligten als inkompetent angesehen zu werden. Aus diesem Grund strahlen Entscheider oft auch dann Zuversicht und Selbstvertrauen aus, wenn die Faktenlage im Grunde genommen dagegen spricht (vgl. Russo und Schoemaker, 1992, S. 16).

Der letzte Einflussfaktor, der an dieser Stelle als mögliche Ursache für Überoptimismus erläutert werden soll, ist der Anchoring-Effect, dessen Entstehung in der typischen Gestaltung des Auswahlprozesses für Investitionsprojekte begründet liegt (vgl. Lovallo und Kahneman, 2003, S. 60). Da die Ressourcen, die in einem Unternehmen für Investitionsprojekte zur Verfügung stehen, in aller Regel knapp bemessen sind, geraten Projektmanager im Zuge dieses Auswahlprozesses in Versuchung, ihr Projekt in einem möglichst positiven Licht zu präsentieren, um im Vergleich zu anderen Projekten die Chancen auf eine Projektgenehmigung zu erhöhen (vgl. Kahneman und Lovallo, 1993, S. 27 f.; Weber et al., 2003, S. 16 f). Dies hat wiederum zur Konsequenz, dass die vorläufigen Schätzungen in Bezug auf Kosten und Zeitbedarf oft deutlich positiv verzerrt sind. Zwar werden diese Schätzungen im Laufe des weiteren Planungsprozesses korrigiert und beispielsweise im Hinblick auf unvorhergesehene Risiken angepasst (vgl. Lovallo und Kahneman, 2003, S. 60; Flyvbjerg et al., 2009, S. 175). Problematisch ist an dieser Stelle jedoch, dass die ursprüngliche, meist viel zu optimistische Schätzung im gesamten weiteren Planungsprozess als eine Art ‚Anker‘ fungiert, der bewirkt, dass die Entscheider sich bei späteren Anpassungen nicht weit genug vom ursprünglichen Schätzwert entfernen, sodass auch nachfolgende Schätzungen einer unangemessen optimistischen Verzerrung unterliegen (vgl. Lovallo und Kahneman, 2003, S. 60; Flyvbjerg et al., 2009, S. 175). Dies hat zur Folge, dass die positiv verzerrte Perspektive auch über den Auswahlprozess hinaus beibehalten wird, sodass eine realistische Einschätzung möglicher Kos-

tenüberschreitungen unterbleibt und der Effekt des übermäßigen Optimismus weiter verstärkt wird.

3.2. Auftreten des Sunk-Cost-Effects bei der Bewertung laufender Projekte

Neben dem Effekt des übermäßigen Optimismus sind im Zuge der Projektdurchführung weitere psychologisch bedingte Entscheidungsfehler zu beobachten, die ein rationales Handeln der Akteure behindern können. Insbesondere im Rahmen der mitlaufenden Projektkontrolle kann es beispielsweise zum Auftreten des sogenannten Sunk-Cost-Effects kommen, der zu einer Fehlbewertung von in der Umsetzungsphase befindlichen Investitionsprojekten führt (vgl. Weber et al., 2003, S. 25 ff.).

Aufgabe der mitlaufenden Projektkontrolle ist es, Divergenzen zwischen den ursprünglich geplanten Kosten und Umsetzungszeiten und den tatsächlich realisierten Werten frühzeitig aufzudecken. Im Rahmen einer Analyse der erfassten Daten gilt es anschließend zu beurteilen, ob das Projekt bei Anwendung rationaler Bewertungsmaßstäbe fortgeführt oder – bei zu stark negativen Abweichungen – abgebrochen werden sollte (vgl. Weber et al., 2003, S. 25). Ein Aspekt, der einer solchen rationalen Bewertung entgegenstehen kann, ist das Vorliegen von Sunk Costs, also jenen Kosten, die bereits in vergangenen Perioden angefallen sind und deren Existenz zum Betrachtungszeitpunkt nicht mehr zu revidieren ist (vgl. Weber et al., 2003, S. 25 und 28; Drury, 2015, S. 35). In diesem Fall tendieren Entscheider häufig dazu, die angefallenen Sunk Costs in die Projektbewertung einzubeziehen – dies läuft jedoch den Vorgaben der ökonomischen Entscheidungstheorie zuwider, da Sunk Costs per definitionem als entscheidungsirrelevant anzusehen sind (vgl. Arkes und Blumer, 1985, S. 125 f.; Zayer, 2011, S. 209) und bezüglich einer Fortführungsentscheidung lediglich zukünftige Zahlungsströme Berücksichtigung finden dürfen (vgl. Arkes und Blumer, 1985, S. 125; Thaler, 1980, S. 47).

Die fälschliche Einbeziehung der Sunk Costs in diese Entscheidung führt dazu, dass der verantwortliche Manager seinem Projekt eine unverhältnismäßig hohe Erfolgswahrscheinlichkeit einräumt (vgl. Arkes und Blumer, 1985, S. 130), da der Betrag der Sunk Costs stets einen positiven Wert annimmt und somit eine optimistisch verzerrte Bewertung stattfindet (vgl. Zayer, 2007, S. 136). Dies hat wiederum zur Konsequenz, dass häufig weitere Ressourcen für Projekte freigegeben werden, deren Scheitern bei Anwendung rationaler Bewertungsmaßstäbe bereits zu erkennen gewesen wäre (vgl. Garland, 1990, S. 731; Zayer, 2007, S. 136). Dabei zeigt sich, dass die Verzerrung umso stärker ausfällt, je höher die bereits angefallenen Sunk Costs sind (vgl. Garland, 1990, S. 731).

Die Tendenz der Entscheider, Sunk Costs ungeachtet ihrer Irrelevanz in den Entscheidungsprozess einzubeziehen, wird in der Literatur im Wesentlichen mithilfe der Erkenntnisse der Prospect Theory begründet (vgl. Arkes und Blumer, 1985, S. 124; Thaler, 1980, S. 48). Die Prospect Theory, 1979 von Kahneman/Tversky als Alternative zur klassischen Nutzentheorie vorgestellt (vgl. Kahneman und Tversky, 1979, S.

263), stellt mehrere Annahmen über die Nutzenfunktion des Entscheiders auf, die in Summe dazu führen können, dass der Entscheider in einer Situation, in der bereits Sunk Costs angefallen sind, von einer rationalen Entscheidungsfindung abweicht (vgl. Weber et al., 2003, S. 25 ff.).

Die abgebildete Funktion stellt den Nutzenverlauf des Entscheiders unter den Annahmen der Prospect Theory dar, wobei auf der horizontalen Achse objektiv bewertete Gewinne bzw. Verluste angegeben sind, auf der vertikalen Achse hingegen der subjektiv wahrgenommene Nutzen des Entscheiders (vgl. Arkes und Blumer, 1985, S. 130 f.). Kahneman/Tversky treffen in ihrer Arbeit die Annahme, dass diese Nutzenfunktion im Verlustbereich einen konvexen und im Gewinnbereich einen konkaven Verlauf aufweist (vgl. Kahneman und Tversky, 1979, S. 278).

Zu Beginn einer Investition befindet sich der Entscheider im Schnittpunkt der beiden Achsen (vgl. Arkes und Blumer, 1985, S. 131), dem sogenannten Referenzpunkt (vgl. Kahneman und Tversky, 1979, S. 277; Weber et al., 2003, S. 26). Sobald jedoch unwiederbringlich Ressourcen in ein Projekt investiert wurden, bewegt er sich im unteren linken Quadranten des Koordinatensystems, also im konvexen Bereich der Funktion, da die investierten Mittel für ihn zu diesem Zeitpunkt einen Verlust darstellen (vgl. Arkes und Blumer, 1985, S. 131). Würde der Entscheider streng rational im Sinne der klassischen Finanztheorie agieren, müsste er diesen Verlust in Höhe der Sunk Costs unmittelbar realisieren, sodass diese Kosten in zukünftigen Entscheidungssituationen keine Relevanz mehr aufweisen würden (vgl. Statman und Caldwell, 1987, S. 8 f.).

In realen Entscheidungssituationen ist es jedoch häufig so, dass Entscheider dazu tendieren, erlittene Verluste nicht unmittelbar zu realisieren, sondern „mentale Konten“ (Weber et al., 2003, S. 25) zu führen, die sie nach Möglichkeit erst dann abschließen, wenn sie diese Verluste durch einen künftigen Gewinn ausgleichen konnten. Die Ursache für dieses Verhalten ist unmittelbar dem Verlauf der Nutzenfunktion zu entnehmen: Da die Steigung im Verlustbereich aufgrund der Konvexität deutlich höher ist als im konkaven Gewinnbereich, stellt der Ausgleich des Verlusts auf diesem bereits geöffneten Konto einen größeren Nutzenzuwachs für den Entscheider dar als die Realisierung eines gleichhohen Gewinns auf einem separaten mentalen Konto (vgl. Weber et al., 2003, S. 25).

Auch die Feststellung von Kahneman/Tversky, dass sich Entscheider im Verlustbereich risikoaffin, im Gewinnbereich hingegen risikoavers verhalten (vgl. Kahneman und Tversky, 1979, S. 268; Weber et al., 2003, S. 26), lässt sich auf diese Weise erklären. Befindet sich der Entscheider bereits im konvexen Verlustbereich, ist er geneigt, weitere Ressourcen in das Projekt zu investieren, da die möglicherweise daraus resultierenden zusätzlichen Verluste lediglich einen geringen zusätzlichen Disnutzen auslösen würden. Haben die weiteren Investitionen jedoch Erfolg, ist der daraus resultierende Nutzenzuwachs für den Entscheider deutlich höher, sodass eine Fortführung des Projekts auch dann attraktiv scheinen kann, wenn sie mit einem hohen Risiko verbunden ist (vgl.

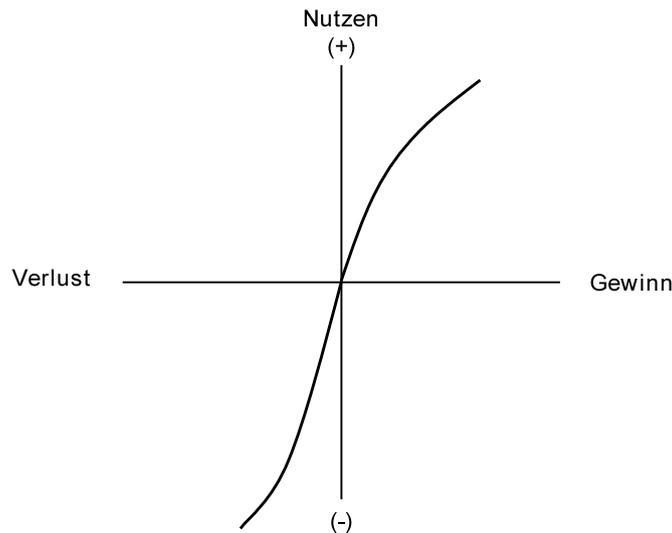


Abbildung 1: Nutzenfunktion nach Kahneman und Tversky, 1979, S. 279 und Arkes und Blumer, 1985, S. 131

Arkes und Blumer, 1985, S. 131).

Besonders bemerkenswert am Sunk-Cost-Effect ist, dass er nicht nur – wie es auf den ersten Blick vielleicht plausibel erscheinen mag – diejenigen Entscheider trifft, die mit den Regeln der ökonomischen Entscheidungstheorie nicht vertraut sind. Vielmehr konnten Arkes/Blumer diesen Effekt auch bei Studenten der Wirtschaftswissenschaften nachweisen, sodass sie davon ausgehen, dass selbst eine Ausbildung im wirtschaftswissenschaftlichen Bereich nicht davor schützt, dieser kognitiven Verzerrung anheimzufallen (vgl. Arkes und Blumer, 1985, S. 136).

Somit ist der Sunk-Cost-Effect als einer der wesentlichen Gründe dafür anzusehen, dass Entscheider aufgrund eines defizitären Vorgehens bei der Bewertung von Projekten dazu tendieren können, auch dann weiterhin an diesen Projekten festzuhalten, wenn ein Abbruch aus wirtschaftlicher Sicht zu präferieren wäre. Es handelt sich hierbei jedoch nicht um die einzige Ursache, die zu einem irrational langen Festhalten an scheiternden Projekten führen kann. Auf einen weiteren Aspekt soll im folgenden Kapitel eingegangen werden.

3.3. Übermäßiges Commitment und Self-Justification als Hindernis für Abbruchsentscheidungen

Eine weitere kognitive Verzerrung, die für das Festhalten eines Projektmanagers an einem scheiternden Projekt verantwortlich gemacht werden kann, ist die sogenannte Escalation of Commitment bzw. Eskalation des Commitments (vgl. Staw, 1981, S. 577 f.).

Grundsätzlich kann festgestellt werden, dass Personen, die in ein Projekt eingebunden sind, dazu neigen, eine gewisse Hingabe bzw. ein besonderes Engagement für das betreffende Projekt zu entwickeln, also Commitment gegenüber dem Projekt aufzubauen (vgl. Statman und Caldwell, 1987, S. 11). Dies ist zunächst nicht weiter problematisch, sondern im Grunde eine wichtige Voraussetzung dafür, dass (IT-

)Projekte sowohl während der Planungs- als auch in der Umsetzungsphase eine hinreichende Unterstützung des Managements erhalten und somit zum angestrebten Ergebnis geführt werden können (vgl. Newman und Sabherwal, 1996, S. 23). Bedenklich wird das Commitment der Projektbeteiligten jedoch, wenn es dazu führt, dass sich die Entscheider so stark an das Projekt gebunden fühlen, dass sie selbst dann, wenn sich ein Scheitern bereits abzeichnet, immer weitere Mittel investieren, statt den rational sinnvollen Weg eines Projektabbruchs einzuschlagen. In diesem Fall liegt eine Eskalation des Commitments vor (vgl. Newman und Sabherwal, 1996, S. 24; Whyte, 1986, S. 311).

Ein Einflussfaktor, der in der Literatur wiederholt als einer der wesentlichen Auslöser für eskalierendes Commitment charakterisiert wird, ist der Self-Justification-Effect (vgl. Keil, 1995, S. 66; Newman und Sabherwal, 1996, S. 25; Staw und Ross, 1987b, S. 50). Dieser Effekt besagt, dass ein Entscheider immer dann, wenn er Informationen über ein mögliches Scheitern seines Projekts wahrnimmt, das Bedürfnis entwickeln kann, das Projekt ungeachtet dieser negativen Informationen weiterzuführen, um sein Selbstbild vor negativen Einflüssen zu schützen (vgl. Mahlendorf, 2008b, S. 35 f.; Zayer, 2007, S. 115 f.). In diesem Fall können sich demnach die durch den Projektmanager verfolgten Zielgrößen zu Ungunsten des Unternehmens verschieben. Auch wenn der Projektmanager zuvor das Ziel verfolgt hat, das Projekt möglichst erfolgreich abzuschließen, um damit einen Beitrag zur Erhöhung des Unternehmenswertes leisten zu können, rückt bei Bekannt- und Bewusstwerden der negativen Informationen über das Projekt möglicherweise die Protektion seines Selbstbildes in seinen Fokus (vgl. Zayer, 2007, S. 115). Durch die Investition zusätzlicher Mittel in das scheiternde Projekt möchte er mithin den Anschein wahren, dass die ursprünglich getroffene Entscheidung zur Durchführung des Projekts gerechtfertigt gewesen sei (vgl. Mahlendorf, 2008b, S. 36;

Staw, 1976, S. 29; Staw, 1981, S. 579).

Um in Kapitel 4.4 Maßnahmen ableiten zu können, die in der Lage sind, das Auftreten des Self-Justification-Effects zu verhindern, ist es auch an dieser Stelle zunächst erforderlich, diejenigen Einflussfaktoren zu identifizieren, die eine Tendenz zur Self-Justification bei den Entscheidern auslösen können. Zunächst ist festzustellen, dass der Anreiz zur Self-Justification aus zweierlei Richtung ausgelöst werden kann (vgl. Staw, 1976, S. 41 f.). Auf der einen Seite sind Entscheider persönlich bestrebt, stets ein positives Selbstverständnis zu wahren; auf der anderen Seite soll aber auch der positive Schein gegenüber Dritten aufrechterhalten werden (vgl. Staw, 1976, S. 41 f.; Mahlendorf, 2010, S. 110). Folglich weist der Hang zur Self-Justification sowohl eine interne als auch eine externe Dimension auf (vgl. Mahlendorf, 2010, S. 110), die im Folgenden näher erläutert werden sollen.

Die interne Dimension der Self-Justification ist im Wesentlichen durch Festingers Theorie der kognitiven Dissonanzen zu erklären. Festinger postuliert im Rahmen dieser Theorie, dass sämtliche Kognitionen bzw. kognitiven Elemente eines Entscheiders – also die Gesamtheit seines Wissens, seiner Meinungen und Anschauungen (vgl. Festinger, 1957/1968, S. 3) –, die untereinander in einem relevanten Zusammenhang stehen, entweder ein konsonantes oder ein dissonantes Verhältnis zueinander aufweisen (vgl. Festinger, 1957/1968, S. 18; Weber und Riesenhuber, 2002, S. 28). Dabei liegt ein dissonantes Verhältnis bzw. eine kognitive Dissonanz immer dann vor, wenn die kognitiven Elemente nicht konsistent zueinander sind oder gar in Widerspruch zueinander stehen (vgl. Festinger, 1957/1968, S. 12 f.). Da im Allgemeinen davon ausgegangen werden kann, dass sich der Entscheider um Konsonanz innerhalb seiner kognitiven Elemente bemüht (vgl. Festinger, 1957/1968, S. 1), führt das Vorliegen einer Dissonanz dazu, dass der Entscheider deren Existenz als abträglich empfindet, sodass er bestrebt ist, die Dissonanz zu beseitigen (vgl. Festinger, 1957/1968, S. 3).

Die Übertragung dieser Überlegungen auf den Projektkontext zeigt, dass das Bewusstwerden negativer Informationen über das Projekt bei dem zuständigen Entscheider zur Entstehung einer kognitiven Dissonanz führen kann (vgl. Zayer, 2011, S. 208). Wird davon ausgegangen, dass der Entscheider über ein positives Selbstbild verfügt und sich dementsprechend für einen kompetenten Projektmanager hält (vgl. Zayer, 2007, S. 115; Mahlendorf, 2008a, S. 106), kann gefolgert werden, dass dieses Selbstbild in einer dissonanten Beziehung zu den negativen Informationen über den bisherigen Projektverlauf steht (vgl. Mahlendorf, 2008a, S. 106). In mehreren experimentellen Studien konnte gezeigt werden, dass die Stärke dieser Dissonanz von unterschiedlichen Faktoren abhängt. So wurde z. B. nachgewiesen, dass insbesondere eine hohe persönliche Verantwortung des Entscheidungsträgers zu einer Verstärkung der wahrgenommenen Dissonanz führen kann (vgl. Staw, 1976, S. 37 f.). Auch die Ausgestaltung des erhaltenen Feedbacks (vgl. Zayer, 2007, S. 127) sowie die Aussagekraft, die der Entscheider dem Fehlschlagen des Projekts in Bezug auf seine persönlichen Kompetenzen zumisst (vgl. Zayer, 2007,

S. 126), spielen für das Ausmaß der kognitiven Dissonanz eine bedeutende Rolle. Die vom Entscheider wahrgenommene Stärke der kognitiven Dissonanz hat wiederum einen wesentlichen Einfluss auf die Tendenz zur Self-Justification und somit auf die Eskalation des Commitments, da ein Entscheider mit zunehmender Stärke der Dissonanz einen vermehrten Drang zur Beseitigung der Dissonanz verspürt (vgl. Festinger, 1957/1968, S. 18).

Die externe Dimension der Self-Justification zielt im Gegensatz zur internen Dimension nicht darauf ab, das persönliche Selbstbild des Entscheiders vor Schaden zu bewahren. Bei dieser Dimension geht es vielmehr darum, dass der Entscheider Dritten gegenüber demonstrieren möchte, dass sich seine ursprüngliche Entscheidung ungeachtet des aktuell negativen Projektverlaufs langfristig als korrekt herausstellen wird (vgl. Staw, 1976, S. 42; Staw, 1981, S. 580). Er ist also bestrebt, anderen Personen gegenüber keine Fehler eingestehen zu müssen, um in den Augen Dritter nicht an Kompetenz einzubüßen (vgl. Staw und Ross, 1987a, S. 70).

Diese Tendenz, anderen gegenüber das „Gesicht [...] wahren“ zu wollen (Mahlendorf, 2008b, S. 36), ist in der Regel umso stärker ausgeprägt, je unsicherer die politische und soziale Position des Entscheiders bei einem tatsächlichen Scheitern des Projekts ist (vgl. Staw, 1976, S. 42) und je eher er in diesem Fall negative Folgen für seinen weiteren Karriereweg oder sogar einen Jobverlust zu befürchten hat (vgl. Zayer, 2007, S. 129; Staw und Ross, 1987a, S. 70).

Insgesamt haben sowohl die interne als auch die externe Dimension der Self-Justification zur Konsequenz, dass der Entscheider dazu tendiert, weitere Mittel in das scheiternde Projekt zu investieren und das Projekt fortzuführen, um dadurch seine in der Vergangenheit getroffenen Entscheidungen als gerechtfertigt erscheinen zu lassen (vgl. Mahlendorf, 2008b, S. 35 f.; Staw, 1976, S. 29) und den Anschein zu erwecken, stets rationale Entscheidungsmaßstäbe angelegt zu haben (vgl. Staw, 1976, S. 29; Staw, 1981, S. 579). Dies ermöglicht einerseits die Auflösung der kognitiven Dissonanz, da sich der Entscheider durch das Fortführen des Projekts nicht eingestehen muss, bei seiner ursprünglichen Entscheidung einen Fehler begangen zu haben (vgl. Newman und Sabherwal, 1996, S. 25 f.); andererseits hofft er, auf diese Weise negative Konsequenzen durch das Bekanntwerden des Scheiterns bei Dritten vermeiden zu können (vgl. Staw und Ross, 1987a, S. 70).

Dieses Verhalten kann den Projektmanager jedoch in eine Art Teufelskreis einschließen: Erhält er nach der erneuten Investition weiterhin negatives Feedback über den Projektverlauf, erhöht sich das Bedürfnis nach Self-Justification immer weiter und die Eskalation des Commitments wird immer stärker vorangetrieben (vgl. Staw, 1976, S. 29). Somit sollte ein Unternehmen bestrebt sein, geeignete Maßnahmen zu ergreifen, um diesen selbstverstärkenden Prozess rechtzeitig zu unterbrechen und die Investition immer weiterer Mittel in ein fehlschlagendes Projekt zu unterbinden.

4. Gegenmaßnahmen des verhaltensorientierten Kostenmanagements

4.1. Verantwortung des Controllings bei der Integration verhaltenswissenschaftlicher Erkenntnisse in das Kostenmanagement

Die soeben erläuterten verhaltensorientierten Ansätze zur Erklärung von Entscheidungsfehlern bei der Durchführung von Investitionsprojekten machen deutlich, dass in einer Vielzahl von Situationen nicht von einem vollständig rationalen Handeln des Projektmanagers bzw. Entscheiders ausgegangen werden kann. Da ein irrationales Verhalten und die daraus resultierenden Fehlentscheidungen bzw. Managementfehler jedoch im schlimmsten Fall existenzbedrohende Folgen für das Unternehmen haben können (vgl. unter anderem Flyvbjerg und Budzier, 2011, S. 23 f.), stellt sich im Folgenden die Frage, welche Maßnahmen ergriffen werden können, um ein solches irrationales Verhalten zu verhindern und wer für die Implementierung und Umsetzung dieser Maßnahmen in der Verantwortung steht.

Wie bereits zu Beginn dieser Arbeit erwähnt, ist das Projektcontrolling mit der Aufgabe betraut, das Projektmanagement bei der Erreichung der Projektziele umfassend zu unterstützen (vgl. Kapitel 1). Der Controlling-Definition nach Weber folgend gehört zu dieser Unterstützung insbesondere auch die Sicherung der Rationalität des zuständigen Managers (vgl. Weber und Schäffer, 2016, S. 47; Fiedler, 2016, S. 12). Laut dieser Definition wird dem Controlling die Aufgabe zuteil, „Fehler und andere Rationalitätsdefizite schon vor ihrem Wirksamwerden zu vermeiden“ (Weber und Schäffer, 2016, S. 47). Zu diesen Rationalitätsdefiziten gehören unter anderem auch die Könnensdefizite der Manager (vgl. Kapitel 2), die zur Folge haben, dass die zuvor beschriebenen kognitiven Fehlleistungen und Entscheidungsfehler bei den zuständigen Managern auftreten können. Diese Controlling-Definition greift somit unmissverständlich die Verantwortung des (Projekt-)Controllings auf, eine verhaltensorientierte Sichtweise einzunehmen und den Könnensdefiziten der Manager in geeigneter Weise entgegenzuwirken (vgl. Weber, 2013, S. 220). Aus dieser Definition und aus der Tatsache, dass das Kostenmanagement grundsätzlich in den Verantwortungsbereich des Controllings fällt (vgl. Weber und Hirsch, 2005, S. 11), kann gefolgert werden, dass insbesondere das Projektcontrolling in der Pflicht steht, die in Kapitel 2 dargestellten Kostenüberschreitungen, die aus ebendiesen Rationalitätsdefiziten resultieren, mit angemessenen Gegenmaßnahmen zu verhindern.

Im Folgenden sollen sowohl in Bezug auf den übermäßigen Optimismus und den Sunk-Cost-Effect als auch auf die Tendenz zur Self-Justification ausgewählte Maßnahmen vorgestellt werden, die das Projektcontrolling ergreifen kann, um diese verhaltensorientierten Probleme zu eliminieren oder zumindest abzuschwächen, sodass trotz der Könnensdefizite der Manager ein effektives Kostenmanagement im Rahmen großer Investitionsprojekte gewährleistet werden kann.

4.2. Adaption des Outside View und Durchführung einer Pre-Mortem-Analyse zur Beseitigung von übermäßigem Optimismus

Die Umsetzung von Maßnahmen zur Beseitigung von übermäßigem Optimismus stellt sich in der Praxis als vergleichsweise herausfordernd dar (vgl. Kahneman et al., 2011, S. 58) und auch die Forschung hat dieser Problemstellung bisher nur geringe Aufmerksamkeit gewidmet (vgl. Zayer und Hirsch, 2006, S. 650). Nichtsdestotrotz eröffnen sich dem Controller einige Möglichkeiten, den Manager bei der Überwindung dieser kognitiven Verzerrung zu unterstützen (vgl. Zayer und Hirsch, 2006, S. 650). Zwei dieser Möglichkeiten, zum einen die Adaption des Outside View und zum anderen die Durchführung einer Pre-Mortem-Analyse, sollen im Folgenden näher beleuchtet werden.

Der Outside View, häufig auch als Reference Class Forecasting bezeichnet (vgl. Kahneman, 2012, S. 251; Lovallo und Kahneman, 2003, S. 61), setzt zur Beseitigung von Überoptimismus an der während des Planungsprozesses angewandten Forecasting-Methode an. Wie bereits in Kapitel 3.1 erläutert, ergibt sich ein wesentlicher Teil der optimistischen Verzerrung im Rahmen der Projektplanung dadurch, dass die betroffenen Entscheider zum einen einer Kontrollillusion unterliegen und zum anderen einen Hang zur Overconfidence haben. Das Auftreten dieser Effekte wird unter anderem dadurch bedingt, dass klassische Forecasting-Methoden stark auf den individuellen Erfahrungen, Fähigkeiten und Erwartungen der Entscheider aufbauen (vgl. Lovallo und Kahneman, 2003, S. 58) und so das Auftreten optimistischer Verzerrungen begünstigen (vgl. Lovallo und Kahneman, 2003, S. 61). Diese Herangehensweise wird von Kahneman/Lovallo als Inside View bzw. Schätzung aus der Innenperspektive bezeichnet (vgl. Lovallo und Kahneman, 2003, S. 61). Bei dieser für die meisten Entscheider intuitiven Form der Planung bzw. des Forecastings konzentriert der Entscheider seine Prognosebemühungen auf die speziellen Eigenschaften und Umstände des Projekts und fokussiert seine Betrachtungen auf diejenigen Aspekte, die das Projekt einzigartig machen und es von anderen Projekten abgrenzen (vgl. Lovallo und Kahneman, 2003, S. 61; Flyvbjerg, 2006, S. 8), da er es als seine Pflicht ansieht, eine möglichst spezifische und detaillierte Schätzung vorzunehmen, um dem Projekt mit all seinen Facetten und Besonderheiten gerecht zu werden (vgl. Lovallo und Kahneman, 2003, S. 63). Trotz dieser Bemühungen, sämtliche verfügbaren Informationen über das Projekt und sein Umfeld in einer möglichst idealen Prognose zu integrieren, konnte experimentell nachgewiesen werden, dass dieser – eigentlich um höchstmögliche Genauigkeit bemühte – Ansatz zu weitaus schlechteren Prognosen führt als das sogenannte Reference Class Forecasting, das seinerseits eine Schätzung aus der Außenperspektive darstellt (vgl. Flyvbjerg, 2006, S. 8 f.).

Im Zuge des Reference Class Forecastings werden die individuellen Eigenschaften des Projekts weitgehend außer Acht gelassen; ein Versuch, den genauen zukünftigen Verlauf des Projekts zu antizipieren, wird nicht unternommen (vgl. Lovallo und Kahneman, 2003, S. 61; Flyvbjerg, 2005, S. 17). Stattdessen werden Informationen über andere, in der

Vergangenheit durchgeführte Projekte genutzt, um zu einer ersten Basisprognose für das eigentlich betrachtete Projekt zu gelangen (vgl. Flyvbjerg, 2009, S. 354). Insgesamt erstreckt sich dieser Prognoseprozess über drei Schritte. Zunächst ist es erforderlich, eine geeignete Referenzklasse für das betrachtete Projekt zu identifizieren. Eine solche Referenzklasse besteht aus in der Vergangenheit durchgeführten Projekten, die in Art und Umfang mit dem betrachteten Projekt vergleichbar sind. Nach der Beschaffung der relevanten Daten zu diesen Referenzprojekten wird im nächsten Schritt eine Wahrscheinlichkeitsverteilung erstellt, anhand derer eine Basisprognose für die gewünschten Prognosewerte des eigenen Projekts (z. B. Zeit- und Kostenbedarf) erstellt wird. Abschließend erfolgt anhand spezifischer Informationen über das betrachtete Projekt eine individuelle Anpassung der Basisprognose, um der antizipierten Stellung des Projekts innerhalb der Wahrscheinlichkeitsverteilung Rechnung zu tragen (vgl. Flyvbjerg, 2006, S. 8; Kahneman, 2012, S. 251 f.).

Das Reference Class Forecasting lässt demnach einen Großteil der bekannten Informationen über das vorliegende Projekt außer Acht und baut die Prognose im Kern auf statistischen Erkenntnissen auf, die in der Vergangenheit zu vergleichbaren Projekten erlangt werden konnten (vgl. Flyvbjerg et al., 2009, S. 183). Der wesentliche Vorteil einer solchen Vorgehensweise liegt darin, dass vom zuständigen Projektmanager nicht länger verlangt wird, umfangreiche Szenarien über den möglichen Projektverlauf zu entwickeln und seine eigenen Kompetenzen sowie den Umfang seiner Einflussmöglichkeiten in realistischer Art und Weise einzuschätzen, wodurch die kognitiven Begrenzungen, die diese Einschätzungen verzerren könnten, nicht mehr zum Tragen kommen können (vgl. Lovallo und Kahneman, 2003, S. 63).

Ein weiterer möglicher Ansatz zur Beseitigung von Überoptimismus ist die von Gary Klein entwickelte Pre-Mortem-Analyse. Diese Analyse setzt im Gegensatz zum Reference Class Forecasting nicht unmittelbar an der Planerstellung bzw. Prognose an, sondern kommt im darauffolgenden Schritt der Planevaluation zur Anwendung (vgl. Veinott et al., S. 2). In diesem Prozessschritt führen in aller Regel mehrere Faktoren dazu, dass übermäßiger Optimismus eine realistische Analyse des ausgearbeiteten Plans behindert. Zum einen fördert das hohe Maß an Arbeit, das bereits in die Planentwicklung investiert wurde, eine optimistisch verzerrte Einschätzung der Erfolgsaussichten des Plans, sodass mögliche Risiken, die einen Erfolg verhindern könnten, keine Berücksichtigung finden (vgl. Serrat, 2017, S. 224 f.). Zum anderen besteht in dieser Phase häufig eine gewisse Scheu, Bedenken gegenüber dem Plan offen anzusprechen (vgl. Klein, 2007, S. 18), da das Aufdecken von Schwächen im Plan implizit die Leistung der Planer kritisieren bzw. in Frage stellen würde (vgl. Veinott et al., S. 3). Genau diesen Problemen soll die Pre-Mortem-Analyse entgegenwirken.

Gedanklich stellt eine Pre-Mortem-Analyse das Gegenstück zu einer Post-Mortem-Analyse dar (vgl. Serrat, 2017, S. 226). Während eine solche Post-Mortem-Analyse meist im Anschluss an ein fehlgeschlagenes Projekt durchgeführt wird,

um die Gründe für das Scheitern zu analysieren und deren Eintreten in zukünftigen Projekten zu vermeiden, wird diese Betrachtung bei der Durchführung einer Pre-Mortem-Analyse an den Beginn eines Projekts gesetzt (vgl. Kahneman und Klein, 2010, S. 64; Veinott et al., S. 2).

Der Ablauf dieser Analyse ist vergleichsweise einfach gehalten: Im Rahmen eines Kick-Off-Meetings vor dem eigentlichen Projektstart erfolgt zunächst eine Beschreibung des Plans durch die Projektmitglieder (vgl. Kahneman und Klein, 2009, S. 524; Veinott et al., S. 2). Anschließend werden die Anwesenden gebeten, sich gedanklich in einen bestimmten Zeitpunkt in der Zukunft – beispielsweise ein Jahr nach Start des Projekts – zu versetzen. Zu diesem Zeitpunkt wird klar, dass der ursprünglich erarbeitete Plan nicht aufgegangen und das Projekt somit gescheitert ist (vgl. Kahneman und Klein, 2009, S. 524; Serrat, 2017, S. 227). Den Projektmitgliedern wird daraufhin die Aufgabe übertragen, sämtliche Gründe zu notieren, die zum Versagen des Plans geführt haben (vgl. Klein, 2007, S. 19). Diese Gründe werden anschließend reihum vorgetragen, bis alle Projektmitglieder sämtliche von ihnen erdachten Ursachen für das Scheitern des Projekts genannt haben (vgl. Klein, 2007, S. 19). Nachfolgend erhalten die beteiligten Personen die Anweisung, mögliche Maßnahmen aufzuzählen, die den zuvor genannten Ursachen für das Fehlschlagen des Plans entgegenwirken können (vgl. Veinott et al., S. 2). In einem letztem Schritt werden die im Rahmen der Pre-Mortem-Analyse erarbeiteten Schwachstellen und die möglichen Wege, diese zu beseitigen, in den ursprünglichen Plan eingearbeitet (vgl. Klein, 2007, S. 19; Veinott et al., S. 2).

Auch wenn dieses Vorgehen vergleichsweise simpel erscheint, bietet es doch einige nicht zu unterschätzende Vorteile und eine gute Möglichkeit, die Projektplanung von übermäßigem Optimismus zu befreien. So wird dem zumeist vorherrschenden Optimismus bewusst und explizit ein kritischer und pessimistischer Denkansatz entgegengesetzt, indem unzweifelhaft davon ausgegangen wird, dass das Projekt bereits gescheitert ist. Hierdurch wird der Projektmanager nicht in die Situation gebracht, seinen Plan verteidigen zu müssen, wie es beispielsweise der Fall wäre, wenn die Fragestellung lauten würde, warum das Projekt scheitern könnte (vgl. Serrat, 2017, S. 226; Veinott et al., S. 3). Außerdem erhalten die Projektmitglieder und insbesondere der Projektmanager bei Durchführung der Pre-Mortem-Analyse die Anerkennung anderer Beteiligter nicht länger dafür, dass sie möglichst große Zuversicht bezüglich des Projekterfolgs ausstrahlen, sondern vielmehr für die Identifikation möglicher Probleme und Risiken, die den anderen Personen entgangen wären (vgl. Klein, 2007, S. 19; Veinott et al., S. 3).

Unter Berücksichtigung des geringen Aufwandes für die Durchführung der Pre-Mortem-Analyse (vgl. Serrat, 2017, S. 227) und der im Vergleich zu anderen Maßnahmen äußerst erfolgreichen Reduktion von Overconfidence (vgl. Veinott et al., S. 1) ist es nicht verwunderlich, dass diese Methode insbesondere in der Unternehmenspraxis höchst positiv aufgenommen wird (vgl. Kahneman und Klein, 2010, S. 67).

4.3. Reduktion des Sunk-Cost-Effects durch Implementierung geeigneter Bewertungsverfahren und Ausweis von Opportunitätskosten

Auch zur Vermeidung des Sunk-Cost-Effects bei der Bewertung laufender Projekte werden in der Literatur mehrere Maßnahmen bzw. Herangehensweisen vorgeschlagen (vgl. für eine umfangreiche Literaturübersicht und Systematisierung Zayer, 2007, S. 201-207), von denen einige im Folgenden näher analysiert werden sollen. Hierbei soll zunächst auf die Bereitstellung zusätzlicher Informationen bzw. geeigneter Bewertungsverfahren eingegangen werden, anschließend wird der Ausweis von Opportunitätskosten zur Reduktion des Sunk-Cost-Effects thematisiert.

Ein erster Ansatz, dem Sunk-Cost-Effect bei der Bewertung laufender Investitionsprojekte zu begegnen, ergibt sich aus den Erkenntnissen von Conlon/Leatherwood, die bei der Analyse zweier experimenteller Studien feststellen konnten, dass eine Tendenz zum Festhalten an scheiternden Projekten aufgrund von Fehlbewertungen stärker ausgeprägt ist, wenn den Entscheidern das Risiko eines solchen Bewertungsfehlers nicht bewusst ist oder sie nicht über genügend Informationen verfügen, um eine angemessene Entscheidung treffen zu können (vgl. Conlon und Leatherwood, 1989, S. 47 f.). Somit liegt es nahe, zur Vermeidung von Bewertungsfehlern eine unmissverständliche Kennzeichnung der Sunk Costs vorzunehmen und den Projektmanager eindeutig auf die Entscheidungsirrelevanz dieser Kostenpositionen hinzuweisen (vgl. Zayer, 2007, S. 202).

Im Rahmen einer Studie von Ohlert/Weißberger konnte jedoch gezeigt werden, dass dieses Vorgehen in einer experimentellen Situation nicht die gewünschten Ergebnisse liefert, sofern sich die zusätzlich durch den Controller bereitgestellten Informationen tatsächlich rein auf die Kennzeichnung der Sunk Costs und auf eine Warnung bezüglich ihrer potenziell entscheidungsverzerrenden Wirkung beschränken (vgl. Ohlert und Weißberger, 2018, S. 95). Mahlendorf kam in einer weiteren Studie gar zu dem Ergebnis, dass eine explizite Kennzeichnung der Sunk Costs auch eine adverse Wirkung haben und den Sunk-Cost-Effect somit verschärfen kann (vgl. Mahlendorf, 2008b, S. 173).

Erfolgsversprechender scheint es indes, den Entscheidern unmittelbar geeignete Bewertungsmodelle zur Verfügung zu stellen, mit deren Hilfe sie zu einer sachgerechten Bewertung des laufenden Projekts gelangen können (vgl. Zayer, 2007, S. 203 f.; Ohlert und Weißberger, 2018, S. 94 f.). Ein Beispiel für ein solches Bewertungsverfahren, das unter anderem von Newman/Sabherwal als Maßnahme zur Abschwächung des Sunk-Cost-Effects empfohlen wird, ist das Zero-Base Budgeting (vgl. Newman und Sabherwal, 1996, S. 48). Bei dieser Form der Budgetierung bzw. Kostenschätzung erfolgt die Ermittlung der geschätzten Budgets nicht auf Basis der Werte der vorherigen Betrachtungsperiode bzw. Projektphase; stattdessen ist der zuständige Entscheider aufgefordert, den Budgetierungsprozess in jeder Betrachtungsperiode vollständig neu aufzurollen, wobei jede Budgetposition argumentativ zu legitimieren ist (vgl. Newman und Sabherwal,

1996, S. 48). Außerdem hat der Projektmanager in sämtlichen Phasen erneut darzulegen, welche Ziele mit der Durchführung des Projekts verfolgt werden, welche Erträge daraus zu erwarten sind und welche Konsequenzen ein Abbruch des Projekts nach sich ziehen würde (vgl. Pyhrr, 1970, S. 113). Die Tatsache, dass der Budgetierungsprozess im Rahmen des Zero-Base Budgeting in allen Perioden von Grund auf neu durchgeführt wird, hat automatisch zur Konsequenz, dass lediglich zukünftige Kosten- und Erlöspositionen in die Bewertung des Projekts einbezogen werden und Kostenpositionen aus vergangenen Phasen keine Berücksichtigung finden, sodass ein Auftreten des Sunk-Cost-Effects durch Anwendung dieser Methode verlässlich vermieden werden kann (vgl. Newman und Sabherwal, 1996, S. 48 f.).

Ein weiteres mögliches Bewertungsverfahren, das in der Literatur zur Vermeidung des Sunk-Cost-Effects vorgeschlagen wird, ist die Kapitalwertmethode (vgl. Mahlendorf, 2008b, S. 39; Statman und Caldwell, 1987, S. 7 und S. 11 f.). Hierbei handelt es sich um ein recht weit verbreitetes Instrument der dynamischen Investitionsrechnung, das die Bewertung eines Projekts anhand seines Barwertes vornimmt. Dieser Barwert ergibt sich aus der Summe der mit dem Projekt verbundenen (zukünftigen) Ein- und Auszahlungen, die jeweils mit einem zuvor definierten Kalkulationszinsatz auf den Betrachtungszeitpunkt abzuzinsen sind (vgl. Weber und Schäffer, 2016, S. 354; Busse von Colbe et al., 2015, S. 55 f.). Sunk Costs werden bei dieser Berechnung nicht berücksichtigt (vgl. Statman und Caldwell, 1987, S. 7). Eine Fortführung des Projekts ist stets nur dann anzustreben, wenn der sich aus den diskontierten Ein- und Auszahlungen ergebende Kapitalwert größer ist als derjenige Kapitalwert, der sich bei einem Abbruch des Projekts ergeben würde (vgl. Statman und Caldwell, 1987, S. 7). Die eindeutige Vorgabe des Bewertungsverfahrens und des dafür erforderlichen Vorgehens kann somit sicherstellen, dass Sunk Costs nicht in die Bewertung einbezogen werden.

Neben der Implementierung geeigneter Bewertungsverfahren kann es sich außerdem als hilfreich erweisen, dem Entscheider im Zuge der Bewertung nicht nur Informationen über die tatsächlichen Kosten der Projektfortführung zur Verfügung zu stellen, sondern auch die Opportunitätskosten darzulegen, die eine Fortführung des Projekts nach sich ziehen würde (vgl. Zayer, 2007, S. 205). Diese Maßnahme, die beispielsweise von Northcraft/Neale untersucht wurde, setzt unmittelbar an der Ausgestaltung bzw. dem Framing der Entscheidungssituation an (vgl. Northcraft und Neale, 1986, S. 348). Liegen zum Zeitpunkt der Entscheidungsfindung bereits Sunk Costs vor, werden diese beim Abbruch des Projekts von den Entscheidern als sichere Verluste interpretiert (vgl. Garland, 1990, S. 731). Somit steht der Projektmanager seiner Wahrnehmung nach vor der Entscheidung, einen sicheren Verlust hinzunehmen, sofern er das Projekt sofort abbricht, oder aber das Risiko eines noch höheren Verlustes einzugehen, wenn er weiterhin in das Projekt investiert – diesem Risiko steht jedoch die Chance gegenüber, das Projekt letztlich doch erfolgreich abzuschließen und einen Verlust zu vermeiden bzw. zumindest zu reduzieren (vgl. Nor-

Northcraft und Neale, 1986, S. 350). Da der Entscheider, wie in Kapitel 3.2 ausgeführt, im Verlustbereich zur Risikoaffinität neigt, tendiert er bei einem derartigen Framing der Entscheidungssituation dazu, das Projekt fortzuführen und das Risiko eines höheren Verlustes in Kauf zu nehmen (vgl. Kahneman und Tversky, 1979, S. 287).

Was der Projektmanager in diesem Falle jedoch außer Acht lässt, ist die Tatsache, dass er im Rahmen seiner Entscheidung nicht nur die tatsächlichen Kosten für die Fortführung des Projekts, sondern auch die Opportunitätskosten zu berücksichtigen hat, die daraus resultieren, dass diejenigen Mittel, die er in das scheiternde Projekt investiert, nicht mehr für alternative Investitionsmaßnahmen zur Verfügung stehen (vgl. Northcraft und Neale, 1986, S. 350). Macht der Controller diese Fehleinschätzung des Entscheiders deutlich, indem er die Opportunitätskosten beispielsweise in einem Projektbericht ausweist (vgl. Zayer, 2007, S. 205), ändert sich das Framing der Entscheidungssituation insofern, als dass der Entscheider in diesem Fall nicht mehr zwischen einem sicheren Verlust und einer unsicheren Situation, in der sowohl ein Gewinn als auch ein Verlust möglich ist, wählen muss. Stattdessen könnte auch der Abbruch des Projekts einen Gewinn nach sich ziehen – und zwar dann, wenn die freiwerdenden Mittel in ein erfolgsversprechendes neues Projekt investiert werden können (vgl. Northcraft und Neale, 1986, S. 250). Northcraft/Neale konnten die Wirksamkeit dieses Vorgehens im Rahmen einer experimentellen Studie nachweisen (vgl. Northcraft und Neale, 1986, S. 353). In der praktischen Umsetzung können sich jedoch Schwierigkeiten ergeben, da nicht in jeder Entscheidungssituation ein spezifisches Alternativprojekt zur Verfügung steht, anhand dessen mögliche Opportunitätskosten abgeleitet werden können (vgl. Zayer, 2007, S. 206 f.). Somit handelt es sich bei dieser Lösungsmöglichkeit um einen theoretisch vielversprechenden Ansatz, der jedoch in der Praxis auf einige Hindernisse stoßen könnte.

4.4. Vermeidung von Self-Justification durch Reduktion der wahrgenommenen Bedrohung und des individuellen Entscheidungsfreiraums

Auch für mögliche Maßnahmen zur Reduktion des Self-Justification-Effects bietet Zayer einen umfangreichen Literaturüberblick (vgl. Zayer, 2007, S. 185-199). Zwei der dort aufgezeigten Methoden – die Reduktion der durch den Entscheider wahrgenommenen Bedrohung und die Verringerung seines individuellen Entscheidungsfreiraums – sollen im folgenden Abschnitt Gegenstand einer näheren Betrachtung sein.

Aus der internen und der externen Dimension der Self-Justification und den jeweiligen Faktoren, die innerhalb dieser Dimensionen einen verstärkten Self-Justification-Druck auslösen können, ergibt sich eine Vielzahl unterschiedlicher Lösungsansätze, die zur Reduktion der wahrgenommenen Bedrohung des Selbstbilds des Entscheiders beitragen und somit deeskalierend auf sein Commitment einwirken können. Ein strukturierter Ansatz, der sowohl bezüglich der internen als auch der externen Dimension der Self-Justification eine deeskalierende Wirkung entfaltet, wurde

durch Heng/Tan/Wei entwickelt. Gedanklicher Ausgangspunkt dieses Ansatzes ist, dass Entscheider stets in ein soziales System, bestehend aus Vorgesetzten und Kollegen, eingliedert sind (vgl. Heng et al., 2003, S. 101). Diese beiden Parteien haben durch ihr Verhalten im Falle des Scheiterns eines Projekts einen nicht zu unterschätzenden Einfluss darauf, ob bei dem verantwortlichen Entscheider eine Tendenz zur Self-Justification entsteht oder er das Projekt im Falle einer negativen Entwicklung abbricht (vgl. Heng et al., 2003, S. 100 und S. 107).

Wie bereits in Kapitel 3.3 erläutert, ist die Höhe der persönlichen Verantwortung des Entscheiders ein wesentlicher Einflussfaktor auf das Ausmaß der kognitiven Dissonanz und somit auf die interne Dimension der Self-Justification. Diese subjektiv wahrgenommene Verantwortung kann sowohl durch das Verhalten der Vorgesetzten als auch durch das Verhalten der Kollegen entscheidend beeinflusst werden (vgl. Zayer, 2007, S. 186; Heng et al., 2003, S. 101). So stellen Heng/Tan/Wei fest, dass ein teilweiser Transfer der Verantwortung bzw. Schuld vom einzelnen Projektmanager auf seine Vorgesetzten und Kollegen dazu beitragen kann, den Verantwortungsdruck für die Einzelperson zu reduzieren (vgl. Heng et al., 2003, S. 101). Sie schlagen deshalb vor, dass die Vorgesetzten eine sogenannte „shelter strategy“ (Heng et al., 2003, S. 101) einschlagen, bei der sie dem zuständigen Entscheider einen gewissen Schutz gewähren, indem sie deutlich machen, dass sie stets eine Mitverantwortung für die Entscheidungen ihrer Mitarbeiter tragen (vgl. Heng et al., 2003, S. 101). Außerdem kann es hilfreich sein, die Unternehmenskultur dahingehend zu beeinflussen, dass die Kollegen des Projektmanagers den Erfolg oder Misserfolg eines Projekts regelmäßig als Erfolg oder Misserfolg des Teams und nicht des einzelnen Entscheiders interpretieren und mit hin eine „sharing strategy“ (Heng et al., 2003, S. 101) im Unternehmen implementieren. Insbesondere diese Strategie erwies sich in der Studie von Heng/Tan/Wei als effektives Mittel zur Deeskalation des Commitments (vgl. Heng et al., 2003, S. 107).

Diese beiden Vorgehensweisen zur Reduktion des internen Drucks zur Self-Justification werden von Zayer unter dem Begriff der „Shelter&Share-Strategy“ zusammengefasst (vgl. Zayer, 2007, S. 186).

Auch auf die externe Dimension der Self-Justification kann das soziale Umfeld des Entscheiders einen wichtigen Einfluss nehmen. Da Entscheider im Fall des Scheiterns ihres Projekts insbesondere dann zur Self-Justification und zu einer Eskalation des Commitments neigen, wenn sie einen Verlust ihres Ansehens innerhalb des Unternehmens, negative Konsequenzen für ihre Karriere oder gar den Verlust ihres Arbeitsplatzes zu fürchten haben (vgl. Kapitel 3.3), kann der Aufbau von Sicherheit und Vertrauen durch Vorgesetzte und Kollegen eine deeskalierende Wirkung entfalten (vgl. Heng et al., 2003, S. 101). Um dies zu erreichen, sollten Vorgesetzte eine „support strategy“ (Heng et al., 2003, S. 101) verfolgen, mit der sie dem Entscheider glaubhaft vermitteln, dass sein Karriereweg durch den Abbruch eines Projekts keinen Schaden nehmen wird. Mit dieser Sicherheit im Rücken

steigt die Wahrscheinlichkeit, dass der Entscheider ein Projekt nicht nur deshalb fortführt, um sein Ansehen gegenüber den Vorgesetzten zu schützen (vgl. Heng et al., 2003, S. 101). Ähnliches gilt für das Verhalten der Kollegen des Entscheiders. Durch die Anwendung einer „sympathy strategy“ (Heng et al., 2003, S. 101) können diese dazu beitragen, dass der Projektmanager keine Notwendigkeit sieht, gegenüber seinen Kollegen das Gesicht wahren zu müssen, da die Kollegen ihm im Zuge dieser Strategie glaubhaft vermitteln, dass die gegenseitige Anerkennung auch bei Abbruch des Projekts Bestand hat (vgl. Heng et al., 2003, S. 101).

Diese Wege zur Reduktion des externen Self-Justification-Drucks fasst Zayer unter dem Begriff der „Support&Sympathy-Strategy“ (Zayer, 2007, S. 191) zusammen.

Eine vielversprechende Möglichkeit, diese Ansätze in die Unternehmenskultur zu integrieren und somit systematisch für eine Reduktion der Eskalationstendenz zu sorgen, liegt in der Etablierung einer angemessenen Fehlerkultur innerhalb des Unternehmens (vgl. Heng et al., 2003, S. 107; Keil et al., 1995, S. 440). Auch eine Anpassung der Entlohnungssysteme dahingehend, dass neben der persönlichen Leistung zusätzlich die Gesamtleistung einer Gruppe in die Entlohnung einbezogen wird, könnte insbesondere die „sharing strategy“ stärken und so den Self-Justification-Druck des einzelnen Entscheiders reduzieren (vgl. Heng et al., 2003, S. 107).

Sollten diese Maßnahmen zur Reduktion der wahrgenommenen Bedrohung des Selbstbilds des Entscheiders jedoch nicht die gewünschte Wirkung erzielen – gelingt es also nicht, den Drang zur Self-Justification zu reduzieren –, muss der Entscheider dennoch daran gehindert werden, das Projekt fortzusetzen, wenn eine Fortführung den Interessen des Unternehmens entgegensteht. Um dieses Ziel zu erreichen, gilt es, den persönlichen Handlungs- und Entscheidungsfreiraum des Projektmanagers zu begrenzen (vgl. Zayer, 2007, S. 195). Eine solche Begrenzung kann beispielsweise durch die Definition eindeutiger Abbruchregeln erfolgen. Diese Regeln legen klare Kriterien bzw. Grenzen fest, deren Überschreitung zu einem Abbruch des Projekts führt. So wird der Entscheider daran gehindert, sein Commitment zu einer Fortsetzung des Projekts über diese Grenze hinaus führen zu lassen (vgl. Zayer, 2007, S. 195). Hierbei ist allerdings zu beachten, dass die Festlegung solcher Regeln einige Schwierigkeiten nach sich ziehen kann. So kann beispielsweise die Definition fester Budgetlimits ebenfalls irrationale Entscheidungen, zum Beispiel in Form eines zu frühen Projektabbruchs, zur Folge haben (vgl. Zayer, 2007, S. 196). Darüber hinaus stellt Lange fest, dass die idealen Abbruchkriterien je nach betrachteter Projektphase und Industrie variieren können (vgl. Lange, 1993, S. 100 und S. 120), sodass es bei der praktischen Umsetzung dieses Ansatzes durchaus zu Problemen kommen kann.

Eine simplere und in der Praxis recht gängige Methode zur Begrenzung des Entscheidungsfreiraums ist die Definition von Meilensteinplänen für das betreffende Projekt (vgl. Mahlendorf, 2008a, S. 109). Werden während der Projektdurchführung die Zielvorgaben an den jeweiligen Meilen-

steinen nicht erreicht, ist der Projektmanager nicht länger befugt, eigenständig über Fortgang oder Abbruch des Projekts zu entscheiden. Diese Entscheidung wird stattdessen im übergeordneten Lenkungsausschuss des Projekts getroffen (vgl. Weber et al., 2003, S. 29). Dem Projektleiter wird in diesem Fall also die Möglichkeit entzogen, das Projekt auch bei Verfehlung der gesetzten Ziele fortzuführen, um sein persönliches Ziel der Wahrung eines positiven Selbstbildes zu verfolgen.

5. Betrachtung der verhaltensorientierten Problemstellungen und möglicher Gegenmaßnahmen am Beispiel von eLWIS

Im theoretischen Teil dieser Arbeit konnte durch die Betrachtung der spezifischen Charakteristika von Großprojekten und die Analyse ausgewählter Erkenntnisse der Verhaltenswissenschaften aufgezeigt werden, dass kognitive Begrenzungen der Entscheidungsträger maßgeblich auf die Entscheidungsqualität und somit den Erfolg großer Investitionsprojekte einwirken können. Außerdem konnte mit Rückgriff auf die Controlling-Definition nach Weber argumentiert werden, dass insbesondere das (Projekt)Controlling in der Verantwortung steht, beispielsweise im Rahmen des Kostenmanagements geeignete Maßnahmen zu ergreifen, die diesen Defiziten entgegenwirken – einige potenzielle Maßnahmen wurden im vorherigen Kapitel exemplarisch dargestellt.

Im Hinblick auf die Frage, ob eine Pre-Mortem-Analyse den Tod des eingangs vorgestellten Projekts eLWIS hätte verhindern können, soll nun abschließend betrachtet werden, inwieweit es realistisch erscheint, dass die aufgezeigten kognitiven Begrenzungen auch bei den Entscheidern dieses Projekts vorgelegen haben und welche der vorgestellten Gegenmaßnahmen bei Auswertung aller verfügbaren Informationen zu diesem Projekt am erfolgversprechendsten erscheinen.

Zunächst kann festgehalten werden, dass eLWIS die klassischen Charakteristika eines Großprojekts durchweg erfüllt. So lag die Investitionssumme bereits bei Abbruch des Projekts bei rund 500 Millionen € (vgl. Schüler, 2018) und demnach deutlich über der kolportierten Grenze von 100 Millionen \$. Auch ein äußerst langfristiger Planungshorizont sowie eine hohe Ressourcenbindung lagen bei diesem Projekt vor – zwischenzeitlich sollen bis zu 1.000 Personen an eLWIS mitgewirkt haben (vgl. Kolf und Kerkmann, 2018). Durch die Verwendung der damals vollkommen neuen Datenbanktechnologie SAP HANA (vgl. Schüler, 2018) kann außerdem das Kriterium der technologischen Neuartigkeit als erfüllt angesehen werden. Gleiches gilt für die Involvierung einer großen Zahl von Stakeholdern. Berichten zufolge soll eine große Anzahl einzelner Abteilungen und Bereiche in den Entwicklungsprozess involviert gewesen sein (vgl. ?). In Summe führte dies auch bei eLWIS zu einem enorm hohen Grad an Komplexität, der charakteristisch für Großprojekte ist.

Hieraus lässt sich schließen, dass auch bei diesem Projekt die von Flyvbjerg genannten technologischen und politisch-

ökonomischen, insbesondere aber auch psychologische Faktoren (vgl. Kapitel 2) zu den wesentlichen Auslösern für das Scheitern des Projekts zählen dürften. Aufgrund der vergleichsweise geringen Menge an öffentlich zugänglichen Informationen zu diesem spezifischen Projekt ist es jedoch insbesondere im Hinblick auf den Sunk-Cost-Effect kaum möglich, zu beurteilen, inwieweit diese kognitive Verzerrung die Entscheidungsqualität innerhalb des Projekts negativ beeinflusst hat und inwiefern die Anwendung der korrespondierenden Gegenmaßnahmen ein Scheitern des Projekts hätte verhindern bzw. den Projektverlauf positiv hätte beeinflussen können.

Bezogen auf den Self-Justification-Effect ergibt sich aus den veröffentlichten Pressemeldungen hingegen ein wesentlicher Anhaltspunkt, der auf ein Wirksamwerden dieses Effekts innerhalb des eLWIS-Projekts hindeutet. So gingen der Beendigung des Projekts einige wesentliche personelle Wechsel innerhalb des Unternehmens voraus – im Jahr 2017 verließen sowohl der Head of IT als auch der Vorstandsvorsitzende die Lidl Stiftung & Co. KG (vgl. Kolf und Kerkmann, 2018). Die öffentliche Meldung über den Abbruch des Projekts folgte nur etwa ein Jahr später. Wird eine solche Abbruchentscheidung innerhalb einer so kurzen Zeitspanne nach der Neubestellung wesentlicher mit dem Projekt vertrauter Führungspositionen getroffen, kann dies als wichtiger Indikator für das Vorliegen einer Eskalation des Commitments und mithin des Self-Justification-Effects gedeutet werden (vgl. Ross und Staw, 1993, S. 724; Newman und Sabherwal, 1996, S. 43). Da die nachfolgenden Führungskräfte für die vorherigen Entscheidungen zur Durchführung des Projekts im Gegensatz zu ihren Vorgängern keine persönliche Verantwortung tragen, verspüren sie im Regelfall keinen Drang zur Self-Justification und sind aus diesem Grund deutlich eher bereit, ein zum Scheitern verurteiltes Projekt auch tatsächlich abzubrechen – hierzu waren die Vorgänger aufgrund ihres übermäßigen Commitments dem Anschein nach auch im Falle von eLWIS nicht mehr in der Lage. Somit ist zu erwarten, dass ein vorheriges Ergreifen von Maßnahmen gegen eskalierendes Commitment bzw. gegen die Tendenz zur Self-Justification bereits früher zu einer (aus ökonomischer Sicht sinnvollen) Beendigung des Projekts hätte führen können, wodurch das Unternehmen möglicherweise einen Teil der bis zum Abbruch investierten 500 Millionen € hätte einsparen können.

Noch deutlichere Anzeichen liefert die Berichterstattung dafür, dass innerhalb des eLWIS-Projekts eine Tendenz zu übermäßigem Optimismus vorgelegen haben könnte. So lassen die öffentlich verfügbaren Informationen über das Projekt die Vermutung zu, dass eLWIS selbst nach den Maßstäben eines großen IT-Investitionsprojekts eine besonders hohe Komplexität aufwies. Eine der zentralen Schwierigkeiten des Projekts bestand darin, dass eine Vielzahl der Prozesse und Abläufe innerhalb des Unternehmens mit der Standard-Ausführung der Software inkompatibel war (vgl. Kolf und Kerkmann, 2018). Da Lidl jedoch nicht bereit war, die entsprechenden Prozesse an die Gegebenheiten des Software-Systems anzupassen, waren umfangreiche Änderungen in der Programmierung erforderlich, die zum einen

grundsätzlich kosten- und zeitintensiv sind, zum anderen aber auch einen deutlich negativen Einfluss auf die Performance des Systems haben können (vgl. Kolf und Kerkmann, 2018; Lehmann, 2018, S. 4). Ein weiterer komplexitätssteigernder Faktor lag in der Tatsache begründet, dass das zuvor genutzte Warenwirtschaftssystem bereits zu Projektbeginn merklich veraltet bzw. rückständig war und aus einer Vielzahl unterschiedlicher, selbstprogrammierter Module bestand, wodurch die Migration auf ein einzelnes neues System erheblich erschwert wurde (vgl. Kolf und Kerkmann, 2018; Schüler, 2018). Diesen beiden wesentlichen Aspekten, die eLWIS im Vergleich zu anderen IT-Großprojekten besonders komplex und aufwendig machten, wurde von den zuständigen Entscheidern allem Anschein nach nicht die erforderliche Bedeutung zugemessen – der zusätzliche Aufwand scheint schlicht unterschätzt worden zu sein. Diese Vermutung wird auch durch die Aussagen eines in das Projekt involvierten Beraters gestützt, wonach die geplanten Umsetzungszeiten trotz der hohen Komplexität sogar deutlich knapper bemessen wurden als bei vergleichbaren anderen Projekten (vgl. Kolf und Kerkmann, 2018).

Eine vielversprechende Möglichkeit, derartige Probleme bzw. Herausforderungen bereits vor Beginn des Projektstarts aufzudecken und infolgedessen in angemessener Art und Weise in der Planung zu berücksichtigen, wäre sicherlich die Durchführung einer Pre-Mortem-Analyse unter Beteiligung sämtlicher Projektteilnehmer gewesen. Es ist anzunehmen, dass erfahrene Softwareexperten Bedenken bezüglich der Anpassung des Systems an die Prozesse des Unternehmens geäußert hätten, da es allgemein als erfolgsversprechender gilt, die Prozesse stattdessen an der Software auszurichten (vgl. Kolf und Kerkmann, 2018; Lehmann, 2018, S. 4). Gleiches gilt für das Problem der Migration veralteter Software auf ein neues System – auch hier ist zu vermuten, dass IT-Experten dieses Problem bei der Durchführung einer Pre-Mortem-Analyse erkannt und kommuniziert hätten.

Ein weiteres hilfreiches Instrument zur Beseitigung der optimistischen Verzerrung wäre auch die Nutzung des Reference Class Forecasting und somit die Adaption des Outside View gewesen. Das Reference Class Forecasting hätte seine Stärken in diesem Fall insbesondere deshalb ausspielen können, weil die Einführung eines neuen Warenwirtschaftssystems zwar für das einzelne Unternehmen einen einmaligen Vorgang darstellt, derartige Projekte aber in Bezug auf die gesamte Wirtschaft bei weitem keinen außergewöhnlichen Charakter aufweisen, sodass es möglich gewesen sein dürfte, eine angemessene Referenzklasse für eLWIS zu finden (vgl. Flyvbjerg, 2006, S. 9). Eine Betrachtung dieser Referenzklasse hätte schnell gezeigt, dass Lidl – wie bereits zu Beginn der Arbeit erwähnt – in den vergangenen Jahren nicht das einzige Unternehmen war, das sich mit Schwierigkeiten bei der Implementierung eines neuen SAP-Systems konfrontiert sah. Hätten die Entscheider in der Planungsphase also nicht versucht, das Forecasting anhand einer möglichst detaillierten Erfassung der spezifischen Projektinformationen zu erstellen, und hätten sie stattdessen auf die vielfach verfügbaren Informationen zu vergleichbaren Projekten zurückgegriffen (vgl.

z. B. Kroker, 2018), wären sie aller Wahrscheinlichkeit nach zu einer realistischeren Prognose der zu erwartenden Kosten und Implementierungszeiten gekommen; auch die Möglichkeit des Scheiterns des Projekts wäre dann vermutlich bewusster wahrgenommen worden.

Insgesamt ist davon auszugehen, dass insbesondere die Kombination von Pre-Mortem-Analyse und Reference Class Forecasting zu deutlich realistischeren und somit weniger optimistisch verzerrten Prognosen geführt hätte. Die Frage, ob durch die Anwendung dieser oder anderer Instrumente des verhaltensorientierten Kostenmanagements der Tod von eLWIS tatsächlich hätte verhindert werden können, kann im Rahmen einer retrospektiven Betrachtung auch aufgrund der spärlichen Informationslage nicht final beantwortet werden. Sicher ist jedoch, dass ein solches Vorgehen zu einer ausgewogeneren Einschätzung der tatsächlichen Lage des Projekts hätte führen können, wodurch wesentlich besser informierte Entscheidungen und letztlich ein zielgerichteteres Handeln der beteiligten Akteure möglich gewesen wäre.

6. Fazit und Ausblick

Abschließend bleibt festzuhalten, dass Großprojekte im Allgemeinen einige Besonderheiten aufweisen, die dazu führen, dass sie deutlich schwieriger zu managen sind als ihre kleineren Gegenstücke und dass aus diesem Grund die ursprünglich definierten Projektziele in Großprojekten überdurchschnittlich oft verfehlt werden.

Eine der wesentlichen Ursachen für dieses Verfehlen der Projektziele ist auf psychologische Faktoren zurückzuführen. Auch wenn der Manager in der klassischen Entscheidungstheorie als vollständig rational handelnder Homo Oeconomicus dargestellt wird, weist er in der Realität einige kognitive Begrenzungen auf, die eine rationale Entscheidungsfindung in zahlreichen Situationen verhindern. Drei mögliche kognitive Verzerrungen, die in unterschiedlichen Phasen der Projektdurchführung zu Entscheidungsfehlern des Managements führen können, wurden ausführlicher dargelegt. Hierbei gilt es noch einmal zu betonen, dass die hier erläuterten Verzerrungen – also der übermäßige Optimismus, der Sunk-Cost-Effect und der Self-Justification-Effect – zwar eine besonders hohe Bedeutung für die Erklärung von Fehlentscheidungen in Projekten haben, dass es insgesamt jedoch noch eine Vielzahl weiterer kognitiver Verzerrungen gibt, die in unterschiedlichsten Zusammenhängen Auswirkungen auf das Entscheidungsverhalten der Manager haben können und an dieser Stelle aufgrund von Platzrestriktionen nicht aufgenommen wurden.

Da diese kognitiven Verzerrungen erhebliche negative Folgen für die Entscheidungsqualität innerhalb des Projekts und letztlich auch für das betroffene Unternehmen haben können, ist es eine wesentliche Aufgabe des Controllings in seiner Funktion als Rationalitätssicherer, geeignete Maßnahmen zu ergreifen, die den kognitiven Begrenzungen entgegenwirken. Für diese Maßnahmen wurden in der vorliegenden Arbeit ebenfalls einige Beispiele erläutert – es gilt

jedoch auch hier, dass nur ein kleiner Ausschnitt der gesamten Vielfalt an möglichen Maßnahmen aufgezeigt werden konnte.

Zu guter Letzt sollte die Frage geklärt werden, ob eine der zuvor dargestellten Maßnahmen – die Pre-Mortem-Analyse nach Gary Klein – den Tod des Projekts eLWIS hätte verhindern können. Hierbei wurde festgestellt, dass eine solche Pre-Mortem-Analyse insbesondere in Kombination mit einem geeigneten Forecasting-Verfahren sicherlich zu einer optimierten Entscheidungsfindung innerhalb des Projekts hätte beitragen können. Ob alleine hierdurch bzw. alleine durch die Anwendung von Maßnahmen des verhaltensorientierten Controllings bzw. Kostenmanagements jedoch das Scheitern des Projekts tatsächlich hätte verhindert werden können, ist durchaus fraglich. Zwar ist eine Öffnung des Controllings für psychologische und verhaltenswissenschaftliche Erkenntnisse sicherlich empfehlenswert; es handelt sich hierbei jedoch nicht um ein Patentrezept, um Projekte vollständig vor Fehlentwicklungen und einem letztlichen Scheitern zu schützen. Auch die beiden anderen von Flyvbjerg genannten Ursachen des Scheiterns – also die technologischen und die politisch-ökonomischen Faktoren – gilt es weiterhin unbedingt zu berücksichtigen. Sowohl die Qualität der Daten, auf denen der Manager seine Entscheidungen aufbaut, als auch die Anreizstrukturen, an denen das Management sein Handeln ausrichtet, sind wichtige Eckpfeiler, um eine Entscheidungssituation zu schaffen, in der ein Entscheider auf Basis hochwertiger Informationen und gelenkt durch strategisch sinnvolle Anreize möglichst rationale Entscheidungen zum Wohle des Projekts und letztlich auch zum Wohle des Unternehmens treffen kann. Nur so ist es möglich, sowohl Wissens- und Wollens- als auch Könnensdefizite nachhaltig zu beseitigen.

7. Literaturverzeichnis

- Arkes, H. R. und Blumer, C. The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, 35(1):124–140, 1985.
- Biesenthal, C., Clegg, S., Mahalingam, A., und Sankaran, S. Applying institutional theories to managing megaprojects. *International Journal of Project Management*, 36(1):43–54, 2018.
- Buhl, H.-U. und Meier, M. C. Die Verantwortung der Wirtschaftsinformatik bei IT-Großprojekten: Symptome, Diagnose und Therapie. *Wirtschaftsinformatik*, 53(2):59–62, 2011.
- Busse von Colbe, W., Laßmann, G., und Witte, F. *Investitionstheorie und Investitionsrechnung*. Berlin, Springer-Verlag, 4 edition, 2015.
- Conlon, E. J. und Leatherwood, M. L. Sunk costs and financial decision making: Integration and implications. *Advances in Financial Planning and Forecasting*, 3:37–61, 1989.
- Deutsches Institut für Normung. DIN 69901-5: 2009-01, Projektmanagement - Projektmanagementsysteme - Teil 5: Begriffe, 2009.
- Drury, C. *Management And Cost Accounting*. Andover, Cengage Learning, 9 edition, 2015.
- Evans, J. S. B. T. Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59(1):255–278, 2008.
- Festinger, L. *A Theory of Cognitive Dissonance*. Palo Alto, Stanford University Press (Nachdruck), 1 edition, 1957/1968.
- Fiedler, R. *Controlling von Projekten: Mit konkreten Beispielen aus der Unternehmenspraxis – Alle Aspekte der Projektplanung, Projektsteuerung und Projektkontrolle*. Wiesbaden, Springer Vieweg, 7 edition, 2016.
- Flyvbjerg, B. Policy and planning for large infrastructure projects: Problems, causes, cures. 2005.
- Flyvbjerg, B. From nobel prize to project management: Getting risks right. *Project Management Journal*, 37(3):5–15, 2006.
- Flyvbjerg, B. Survival of the unfittest: Why the worst infrastructure gets built – and what we can do about it. *Oxford Review of Economic Policy*, 25(3): 344–367, 2009.
- Flyvbjerg, B. What you should know about megaprojects and why: An overview. *Project Management Journal*, 45(2):6–19, 2014.
- Flyvbjerg, B. und Budzier, A. Why your it project may be riskier than you think. *Harvard Business Review*, 89(9):23–25, 2011.
- Flyvbjerg, B., Garbuio, M., und Lovallo, D. Delusion and deception in large infrastructure projects: Two models for explaining and preventing executive disaster. *California Management Review*, 51(2):170–193, 2009.
- Flyvbjerg, B., Bruzelius, N., und Rothengatter, W. *Megaprojects and Risk: An Anatomy of Ambition*. Cambridge, Cambridge University Press, 13 edition, 2013.
- Garland, H. Throwing good money after bad: The effect of sunk costs on the decision to escalate commitment to an ongoing project. *Journal of Applied Psychology*, 75(6):728–731, 1990.
- Gehrig, M. und Brey, M. Controlling hilft, strategische denkfehler zu vermeiden. *Controlling & Management Review*, 57(3):46–53, 2013.
- Hastorf, A. H., Schneider, D. J., und Polefka, J. *Person Perception*. Reading, Addison-Wesley, 1 edition, 1970.
- Heng, C.-S., Tan, B. C. Y., und Wei, K.-K. De-Escalation of Commitment in Software Projects: Who matters? What matters? *Information & Management*, 41(1):99–110, 2003.
- Hirsch, B., Schäffer, U., und Weber, J. Zur grundkonzeption eines verhaltenorientierten controllings. *Zeitschrift für Controlling & Management*, 52 (Sonderheft 1):5–11, 2008.
- International, T. S. G. Chaos report 2015. 2015.
- Kahneman, D. *Thinking, Fast and Slow*. London, Penguin Books, 1 edition, 2012.
- Kahneman, D. und Klein, G. A. Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, 64(6):515–526, 2009.
- Kahneman, D. und Klein, G. A. Strategic decisions: When can you trust your gut? *McKinsey Quarterly*, 2010(2):58–67, 2010.
- Kahneman, D. und Lovallo, D. Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39(1):17–31, 1993.
- Kahneman, D. und Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, 1979.
- Kahneman, D., Lovallo, D., und Sibony, O. Before you make that big decision. . . . *Harvard Business Review*, 89(6):51–60, 2011.
- Kannenberg, A. "Elwis" ist tot: Lidl stoppt millionenschweres Projekt mit SAP heise online, 2018. URL <http://www.heise.de/-4111245>. Abfrage: 06.05.2019.
- Kardes, I., Ozturk, A., Cavusgil, S. T., und Cavusgil, E. Managing global megaprojects: Complexity and risk management. *International Business Review*, 22(6):905–917, 2013.
- Keil, M. Pulling the plug: Software project management and the problem of project escalation. *Management Information Systems Quarterly*, 19(4): 421–447, 1995.
- Keil, M., Mixon, R., Saarinen, T., und Tuunainen, V. Understanding runaway information technology projects: Results from an international research program based on escalation theory. *Journal of Management Information Systems*, 11(3):65–85, 1995.
- Klein, G. A. Performing a project premortem. *Harvard Business Review*, 85 (9):18–19, 2007.
- Kolf, F. und Kerkmann, C. Programmed for disaster: Lidl software disaster another example of Germany's digital failure, Handelsblatt, 2018. URL <https://www.handelsblatt.com/today/companies/programmed-for-disaster-lidl-software-disaster-another-example-of-germanys-digital-failure/23582902.html?ticket=ST-1746227-G0q1WRyHV9rUcSRpq7pd-ap4>. Abfrage: 06.05.2019.
- Kroker, M. Haribo, Lidl, Deutsche Post & Co.: Die lange Liste schwieriger und gefloppter SAP-Projekte, WirtschaftsWoche online, 2018. URL <https://www.wiwo.de/unternehmen/it/haribo-lidl-deutsche-e-post-und-co-die-lange-liste-schwieriger-und-gefloppter-sap-projekte/23771296.html>. Abfrage: 06.05.2019.
- Lange, E. C. *Abbruchentscheidung bei F&E-Projekten*. Wiesbaden, Deutscher Universitäts-Verlag, 1 edition, 1993.
- Lehmann, O. F. Mission Failure at LIDL – But Actually, What was the Mission? *PM World Journal*, 7(6):1–14, 2018.
- Lovallo, D. und Kahneman, D. Delusions of success: How optimism undermines executives' decisions. *Harvard Business Review*, 81(7):56–63, 2003.
- Lovallo, D. und Sibony, O. The case for behavioral strategy. *McKinsey Quarterly*, 2010(2):30–45, 2010.
- Madauss, B.-J. *Projektmanagement: Theorie und Praxis aus einer Hand*. Berlin, Springer Vieweg, 7 edition, 2017.
- Mahlendorf, M. D. *Eskalation des Commitments bei scheiternden Projekten: Eine empirische Untersuchung kognitiver Eskalationsfaktoren und verhaltenswissenschaftlich basierter Controllingmaßnahmen*. Vallendar, Inst. für Management und Controlling, 1 edition, 2008a.
- Mahlendorf, M. D. Verhaltensorientiertes Controlling in der Praxis: Eine am Beispiel verspäteter Projektabbrüche illustrierte Systematik zur Identifikation, Bewertung und Auswahl von Controllingmaßnahmen. *Zeitschrift für Controlling & Management*, 52(Sonderheft 1):104–112, 2008b.
- Mahlendorf, M. D. Controlling bei eskalierenden projekten. *Controlling*, 22 (2):107–112, 2010.
- Merchant, K. A. und Van der Stede, W. A. *Management Control Systems: Performance Measurement, Evaluation, and Incentives*. Harlow, Pearson, 4 edition, 2017.
- Mörsdorf, M. *Konzeption und Aufgaben des Projektcontrolling*. Wiesbaden, Gabler Verlag, 1 edition, 1998.
- Mumme, T. Wie Lidl und SAP die Digitalisierung zum 500-Millionen-Grab machten, 2018. URL <https://ngin-food.com/artikel/lidl-sap-e-lwis-beendet/>. Abfrage: 06.05.2019.
- Newman, M. und Sabherwal, R. Determinants of commitment to information systems development: A longitudinal investigation. *Management Information Systems Quarterly*, 20(1):23–54, 1996.
- Nickel, O. ELWIS: SAP-Projekt mit Lidl "rockt und rollt" nicht mehr, Golem, 2018. URL <https://www.golem.de/news/elwis-sap-projekt-mit-lidl-rockt-und-rollt-nicht-mehr-1807-135514.html>. Abfrage: 06.05.2019.
- Northcraft, G. B. und Neale, M. A. Opportunity costs and the framing of resource allocation decisions. *Organizational Behavior and Human Decision Processes*, 37(3):348–356, 1986.
- Ohlert, C. R. und Weißenberger, B. E. *Debiasing Escalation of Commitment: The Effectiveness of Decision Aids to Enhance De-escalation*, pages 72–116. Heinrich-Heine-Universität, Düsseldorf, 1 edition, 2018.
- Oliomogbe, G. O. und Smith, N. J. Value in megaprojects. *Organization, Technology and Management in Construction*, 4(3):617–624, 2012.

- Pyhrr, P. A. Zero-base budgeting. *Harvard Business Review*, 48(6):111–121, 1970.
- Rode, J. Lidl steuert erste Region mit SAP, in: *Lebensmittel Zeitung*, Nr. 28/2015, S. 37, 2015.
- Ross, J. und Staw, B. M. Organizational escalation and exit: Lessons from the shoreham nuclear power plant. *Academy of Management Journal*, 36(4):701–732, 1993.
- Russo, E. J. und Schoemaker, P. J. H. Managing overconfidence. *Sloan Management Review*, 33(2):7–17, 1992.
- Schüler, H.-P. So starb „Elwis“: Hintergründe zu Lidl's SAP-Rückzug, heise online, 2018. URL <http://www.heise.de/-4113285>. Abfrage: 06.05.2019.
- Serrat, O. *Knowledge Solutions: Tools, Methods, and Approaches to Drive Organizational Performance*. Singapore, Springer Open, 1 edition, 2017.
- Statman, M. und Caldwell, D. Applying behavioral finance to capital budgeting: Project terminations. *Financial Management*, 16(4):7–15, 1987.
- Staw, B. M. Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance*, 16(1):27–44, 1976.
- Staw, B. M. The escalation of commitment to a course of action. *Academy of Management Review*, 6(4):577–587, 1981.
- Staw, B. M. und Ross, J. Behavior in escalation situations: Antecedents, prototypes and solutions. *Research in Organizational Behavior*, 9(1):39–78, 1987a.
- Staw, B. M. und Ross, J. Knowing when to pull the plug. *Harvard Business Review*, 65(2):68–74, 1987b.
- Stockburger, M. und Fritze, H. Lidl beendet Warenwirtschaftsprojekt, STIMME.de, 2018. URL <https://www.stimme.de/heilbronn/wirtschaft/2018/Lidl-beendet-Warenwirtschaftsprojekt;art140955,4054349>. Abfrage: 06.05.2019.
- Thaler, R. Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, 1(1):39–60, 1980.
- 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*.
- Weber, J. Verhaltensorientiertes Controlling: Plädoyer für eine (nicht ganz) neue Sicht auf das Controlling. *Controlling*, 25(4/5):217–222, 2013.
- Weber, J. und Hirsch, B. Kostensenkungsmaßnahmen und die Rolle des Controllings – empirische Erkenntnisse aus deutschen Unternehmen. *Zeitschrift für Controlling & Management*, 49(Sonderheft 1):11–15, 2005.
- Weber, J. und Riesenhuber, M. *Controlling & Psychologie*. Vallendar, WHU, Lehrstuhl für Betriebswirtschaftslehre, insbesondere Controlling und Telekommunikation, 1 edition, 2002.
- Weber, J. und Schäffer, U. *Einführung in das Controlling*. Stuttgart, Schäffer-Poeschel Verlag, 15 edition, 2016.
- Weber, J., Hirsch, B., Linder, S., und Zayer, E. *Verhaltensorientiertes Controlling: Der Mensch im Mittelpunkt*, volume 1. Vallendar, WHU, Lehrstuhl für Betriebswirtschaftslehre, insbesondere Controlling und Telekommunikation, 1 edition, 2003.
- Whyte, G. Escalating commitment to a course of action: A reinterpretation. *The Academy of Management Review*, 11(2):311–321, 1986.
- Zayer, E. *Verspätete Projektabbrüche in F&E: Eine verhaltensorientierte Analyse*. Wiesbaden, Deutscher Universitäts-Verlag, 1 edition, 2007.
- Zayer, E. *Nutzung der Methode der abnehmenden Abstraktion im Rahmen der Erklärung verspäteter Projektabbrüche in der F&E*, pages 201–236. Gabler Verlag, Wiesbaden, 1 edition, 2011.
- Zayer, E. und Hirsch, B. Fehlentscheidungen bei Investitionsprojekten: Fehlerquellen und Gegenmaßnahmen. *Controlling*, 18(12):647–658, 2006.



Giving in Unilaterally Risky Dictator Games: A Model of Allocation Decisions Under Existential Threat

Sarah Franziska Kovatsch

WHU – Otto Beisheim School of Management

Abstract

This study reports experimental results from variations of the standard dictator game that capture different variants of unilaterally risky allocation decisions where only the dictator's payoff is subject to risk. Thereby, it addresses the question of whether decisions under existential threat, modeled as a risk to the dictator's payoff, encourage or discourage generosity in individual decision making. It aims at bridging the gap between experimental economics and psychological research on the behavioral impact of mortality salience. Results show that giving in unilaterally risky dictator games increases with the risk imposed on the dictator's payoff. Risk aversion falls short of explaining the increase in generosity. Instead, the observed behavior is most likely motivated by a preference for efficient capital employment. Moreover, dictator games prove to be an apt model for decisions under existential threat.

Keywords: dictator games; risky decisions; mortality salience; generosity; existential threat.

1. Introduction

How do we decide when death is looming? Whenever we talk about scarcity, we also talk about value. Scarcity does not only cause the fundamental economic resource allocation problem but typically also raises the value of goods. Recently, the infamous street artist Banksy hit the news because he equipped one of his paintings with a self-destructive mechanism which was triggered after the painting was sold during a Sotheby's auction. Although the painting was cut to pieces, its market value increased within hours after the auction. Scarcity spurs value.

This principle also extends to immaterial goods. When we are busy, hence time-constrained, we value our time higher. In fact, the association between scarcity and value is so overlearned that some researchers argue that its converse also holds: highly valued goods are often perceived as rare, referred to as the scarcity heuristic. Extending this heuristic, King et al. (2009) examine the bidirectional link between scarcity and value applied to the meaning of death. For that purpose, they define death as the scarcity of life.

But what happens to our allocation decisions when not money, but lifetime is the limited variable? Does the endowment effect increase and drive a valuation gap between our own utility expectations from a set amount of money and the expected value after discounting for a limited time horizon

(or: limited usability)? Eventually, do we become more selfish or more generous when facing death? These questions shall guide the scientific endeavor behind this experimental study. At the intersection between economics and psychology, it will investigate allocation decisions under existential threat, with special focus on generosity.

One of the most popular publications addressing the behavioral consequences of a doomsday scenario is *Death and the Afterlife* (Scheffler and Kolodny, 2013) in which the authors discuss a scenario where the inevitable demise of the earth is known to occur in 30 days and leads to a grave erosion of human values. Substantiating evidence for a higher tendency towards anti-social behavior close to doomsday can be found in a gaming study analyzing the data of more than 80.000 subjects playing a virtual apocalypse scenario (Kang et al., 2017). Results indicate that the homicide rate increases close to the end.

The present study will neither follow Scheffler and Kolodny, 2013 path of philosophical inquiry, nor study behavior in computer simulations, but instead, take an approach based on an economic game. More importantly, it will not focus on doomsday scenarios where all life on earth comes to an end, but rather consider a scenario where only one individual, the decision-maker, faces an elevated risk of dying in the close future. The experiment conducted as part of this

study is based on a dictator game modified to the effect that the dictator's payoff is subject to risk while the conditions of the recipient's payoff are equal to those in the standard dictator game.

From research on dictator games under ambiguity, it is known that dictators choose unfair splits more frequently when the recipient's payment allocation depends on an ambiguous lottery (Haisley and Weber, 2010). Put differently, dictators' estimates of the expected value of the recipient's allocation are inflated under ambiguity. Experiments with bargaining games, however, have shown that payoff uncertainty on the part of the seller is likely to influence the terms of the settlement in favor of the seller because the buyer is willing to make concessions (see, e.g., Church and Zhang, 1999). Analogously, this study will observe whether a dictator in a dictator game setting is willing to allocate more money to the recipient whose expected payoff from the money is higher relative to the dictator's. Such behavior would be indicative of strong other-regarding preferences triggered by an existential threat. Deception games, for instance, provide a strong example for such preferences: One third of the subjects in Erat and Gneezy (2012) willingly engage in lying that financially harms the sender but benefits the receiver, so-called altruistic white lying. It will be the purpose of this study to examine whether mortality awareness encourages non-selfish motives in allocation decisions.

One might ask why an elevated risk of dying should impact allocation decisions in the first place. The answer is two-fold: Firstly, many factors typically influencing decision-making become obsolete under anticipated death. Secondly, new determinants may be added.

First addressing those factors which turn obsolete, one example that comes into mind is hyperbolic discounting. The tendency to choose a smaller-sooner reward over a larger-later reward (Laibson, 1997) naturally does not play a role for death-doomed individuals. Another example would be the so-called forward consumption effect, defined by Loewenstein and Elster (1992) as deriving utility from the future through anticipatory savoring. Upon looming death, contemplating the future will most likely not affect the utility considerations behind the type of decisions this study is concerned with.

Turning to those factors which may newly emerge as decision variables under existential threat, we must first introduce the term of mortality salience. Research on this topic works with artificially activated death awareness and argues that decisions under this condition are partially dependent on self-defense mechanisms against anxiety, which usually results in selfish or socially endorsed behavior (Kasser and Sheldon, 2000). In the context of allocation decisions, one might also imagine that mortality salience and the concomitant limited time horizon to derive utility from a monetary endowment may either encourage excessive spending or large transfers to others driven by generosity and the realization that the own relative utility is lower. To test which effect is dominant, the experiment presented in this study attempts to align economic research with the tradition of mortality salience exper-

iments.

"A common method of operationalising mortality salience, however, takes the form of open-ended questions asking individuals to express the feelings and thoughts they experience when thinking about their own death," (Cozzolino et al., 2004b, p.279).

Since priming and open-ended question do not match the format of an economic experiment, the present study will attempt to model an allocation decision under existential threat by means of a modified dictator game.

Henceforth, the present study is organized as follows: Section 2 provides a review of the related economic and psychological literature. In section 3, the design of the experiment underlying this study is outlined in detail. Section 4 explains the theoretical model behind the experiment as well as the resulting hypotheses, followed by a comprehensive report of the experimental results in section 5. Finally, Section 6 discusses the findings and their implications, the limitations of this study, as well as proposals for future research.

2. Related Literature

2.1. Dictator Games under Risk

Early research on behavior in dictator game setting goes back 25 years (Forsythe et al., 1994). In the standard scenario, the game is played between two randomly paired players, one in the role of the dictator and one in the role of the recipient. The dictator may then distribute an unearned pie of \$10 between the recipient and him-/herself. According to the predictions of classic game theory, the strictly rational decision in this setting is to allocate zero funds to the recipient. In these early versions of the dictator game, average allocations are typically 10-25% (Camerer, 2003). Attempts to explain these non-trivial splits are manifold, but most of them anchor at a concept of fairness. Since allocations in dictator games are significantly different from those in simple ultimatum games where the recipient has the choice to accept or reject the dictator's proposal, fairness considerations alone cannot entirely explain behavior in either of the games, but instead strategic considerations also have an impact (Forsythe et al., 1994). In subsequent studies, the dictator's action space was anonymized (with so-called double-blind treatments) to test rule out the effect of social influence (Hoffman et al., 1994). Generally, the dictator games are sensitive to changes in the instructional setting: if the dictator earns his/her role through a contest, giving is significantly lower compared to treatments based on a random role allocation (Hoffman et al., 1994).

In the standard dictator game, the total allocation amount is fixed. Modified versions where the joint payoff can vary, however, have studied the trade-off between efficiency, i.e., choosing an allocation which maximizes the joint output, and equity, i.e., choosing a more equal allocation (Charness and

Rabin, 2002; Engelmann and Strobel, 2006). The equity-efficiency tradeoff will be helpful to the understanding of the experimental data gained in this study. Aside from concerns for efficiency, inequality aversion is another motivation prominently used to explain non-selfish allocations in dictator games (e.g., Bolton and Ockenfels, 2000).

Brock et al. (2013) extend the discussion around non-selfish giving to risky dictator treatments, which is an especially valuable research direction for the purpose of this study. Their experiment involves treatments with one-sided risk on the part of the recipient and such with two-sided risk. The discussion of the results centers around two manifestations of inequality aversion, namely *ex post* comparison and *ex ante* comparison. A fairness concept based on *ex post* comparison values equality with regard to outcomes after the resolution of uncertainty whereas *ex ante* comparison values equality with regard to chances to gain a certain income, i.e., uses expected values as benchmark variable. Brock et al. (2013) adapted these definitions from a study exploring consequentialist (*ex post*) and procedural (*ex ante*) fairness in dictator games where the outcome either depends on mutually exclusive or on independent probabilistic events, i.e., comparing competitive with non-competitive treatments (Krawczyk and Le Lec, 2010). Both papers find that a combination of the concepts may be warranted. Freundt and Lange (2017) present a paper extending the study of risky dictator games to self- versus other-risk treatments. They find that on average, giving decreases when the payoff to the recipient is risky. Risk imposed on the dictator's own payoff even leads to a relatively larger decline. The observed behavior is explained by a crowding out effect of social preferences if in conflict with risk preferences as defined by G uth et al. (2008).

Although the present study also presents an experiment based on unilaterally risky dictator treatments, the design will deviate from that of Freundt and Lange (2017) to the extent that our dictators will have safe transfer options, i.e., their own payoff risk does not increase with the amount given to the recipient. Additionally, another gap in the literature will be addressed through a variation of the risk level imposed to the dictator across all five treatments.

2.2. Mortality Salience

Progressing from the analysis of dictator games, the experiment presented in this study also takes inspiration from the psychological literature on behavior under mortality salience. Experiments working with mortality salience aim to activate subjects' awareness of their own transience and then test for changes in behavior.

A critical pillar for this stream of research is terror management theory (TMT). TMT hypothesizes that cultural beliefs are a coping mechanism against the anxiety resulting from the awareness of life's inevitable demise (Solomon et al., 1991). For capitalist cultures, or people who endorse extrinsic values, TMT predicts that mortality salience activation will promote greedy behavior. The meaning maintenance model (MMM), a related theory, interprets death as

a threat to meaning, effectively de-valuing life, and views the effect of mortality salience as an attempt to reinstate meaning (Heine et al., 2006). Both theories coincide in the assumption that death is a psychological threat necessitating defense mechanisms. This is the angle point for mortality salience research.

One economic approach used to test the behavioral effects of mortality salience is the resource dilemma. As Kortenkamp and Moore (2006) elaborate, this type of game provides two interesting properties: a social dimension, expressed in the trade-off between individual and group benefit, and a temporal dimension contingent upon the fact that resource exploitation typically has immediate positive but delayed adverse effects. Harnessing such an environment, Kasser and Sheldon (2000) designed an experiment in which subjects were instructed to represent companies bidding against each other to harvest timber in a forest. The researchers found that the mortality salience condition induces higher consumption of resources and enhances greed. However, the experiment relied on an imaginary scenario limiting its external validity. Nevertheless, other studies have corroborated the link between mortality salience and greed (see, e.g., Cozzolino et al., 2004a). Two similar studies found that death-related stimuli increase consumption quantities of luxury goods (Mandel and Heine, 1999) as well as food and beverages (Mandel and Smeesters, 2008).

Jonas et al. (2002), by contrast, provided evidence that mortality salience elicits more favorable attitudes towards charities, and increases the amount of money donated (a form of prosocial behavior), which they referred to as Scrooge effect. The observed behavior, however, is limited to charities that match the subject's worldview. Thus, the results substantiate that mortality salience triggers a desire to express culturally endorsed, in this case prosocial, behavior. Supporting evidence for a correlation between mortality awareness and prosocial behavior can be found in a study focusing on intergenerational decisions (Wade-Benzoni et al., 2012). Usually, intertemporal and interpersonal distance complicate generosity towards future generations. In this study, subjects exhibited higher intergenerational beneficence under death-priming, which the researchers attribute to a reversal of the intertemporal distance effect counteracting intergenerational discounting.

Overall, experimental evidence on the effect of mortality salience is ambivalent. Some researchers have attempted to broaden the spectrum of death priming to increase the clarity of the results. Referring back to the definition of Kasser and Ryan (1993), a study by Cozzolino et al. (2004a) postulates that individuals scoring high on extrinsic value orientation (EVO) have a higher inclination towards greed under a mortality salience condition, but not under a death reflection condition. The latter was designed to render the association of death more tangible and complement it with an element of life review. Whenever the present study refers to mortality salience, especially in the data analysis section, the term shall be defined as "mortality awareness" and thus cover a broad spectrum of death-related cognition.

3. The Experiment

3.1. Experimental Design

The experiment conducted for the purpose of this study consisted of a series of modified dictator games. An online questionnaire was used to test responses to five dictator game scenarios with different levels of risk. In each scenario, the dictator is endowed with a fund of € 10 million and must decide how to split this amount between him-/herself and an anonymous receiver. The counterpart has no decision power, i.e., he/she cannot reject the dictator's allocation. Treatments differed concerning the payoff consequences for players. One treatment replicates the standard dictator game whereas the other four treatments present the dictator with different levels of risk imposed on his/her payoff.

A total of 178 subjects, all students, and predominantly from the WHU – Otto Beisheim School of Management campus in Vallendar, responded to the invitation for the online experiment within one week in April 2019. All respondents have an academic background in either business administration or economics. Subjects were invited to participate in the experiment via e-mail. After clicking on the link to the experiment, subjects were presented a set of general instructions including an indication of the amount of time necessary to complete the experiment. Prior to the dictator tasks, the subjects were asked to state their willingness to pay for a lottery with a 50% chance of winning € 200 and an equal chance of being left with zero. In alignment with standard measurements for risk aversion (Hartog et al., 2000), this pre-test was designed to identify the individual risk attitude of each respondent. A complete overview of the instructions can be found in the Appendix (see Table A1 and Appendix 7).

Since all subjects participated in all five decision tasks, the results presented in this study are within-subject comparisons. Dictators accessed the decision form from their private devices and submitted all allocation choices online. They were not informed of the outcomes of their decisions between rounds. Lacking any decision power, the receiver was treated as a dummy variable in the setting of this experiment.

After the deadline for participation had passed, a random draw determined which subject would receive the cash payoff corresponding to the cumulative outcome of his/her allocation decisions. The winner of the draw received € 1 in cash for each € 1 million kept to him-/herself in the experiment. For the risky choices, i.e., tasks two to five, the outcome was determined according to a dice roll reflecting the respective risk attributed to the dictator's own share. The winner of the draw was contacted and paid in private. Subjects were not informed of the outcomes of their choices during or after the experiment.

3.2. Treatments

In each treatment, the dictator was asked to allocate € 10 million (experimental currency) between him-/herself and an anonymous counterpart. Thus, allocation options allowed

him/her to keep $10 - x$ [all in m€]¹ to him-/herself and transfer $x \in [0, 1, \dots, 10]$ to his/her counterpart. The choice was constrained to discrete steps of full millions.

Treatment 1 (T_C) is the baseline treatment and replicates the standard dictator game with a stake size of € 10 million. Player's payoffs can be denoted as $(\pi_1, \pi_2) = F(10 - x, x)$ where π_1 is the dictator's income and π_2 is the counterpart's income, both a function of the initial allocation decided upon by the dictator. This treatment serves as a benchmark for the other treatments and aims to position this study within the existing literature on dictator games.

In treatments 2, 3 and 4 (T_25 , T_50 , and T_75), the dictator allocates his/her endowment as in treatment 1. While the amount transferred to the recipient is not subject to any additional conditions, the amount kept by the dictator is subject to a 25% (T_25), 50% (T_50), and 75% (T_75), risk of value loss, respectively. Put differently, in each of the treatments, a lottery determines if the dictator realizes a gain equivalent to the share kept to him-/herself or realizes zero income. Chances of losing the money amount to 25%, 50%, and 75%, respectively. More formally, for all three treatments, payoffs can be denoted as

$$\pi_1 = \begin{cases} 0, & \text{with } p_k^1 \\ 10 - x, & \text{with } 1 - p_k^1 \end{cases}, \quad (1)$$

where $p_k^1 = 0,25$ for T_25 , $p_k^1 = 0,5$ for T_50 and $p_k^1 = 0,75$ for T_75 , and $\pi_2 = x$. This notation clearly shows that only the dictator him-/herself faces a risk in these treatments. Across all three allocation tasks, it holds that $E(x) < E(10 - x)$ for any unit of x . A dictator interested in maximizing the joint payoff would thus transfer the money to his/her counterpart. The purpose of these treatments is to resemble a situation as described in the introduction where the dictator has an elevated risk of sudden death. In the event of sudden death, the money will be worth nothing to the dictator. For any other person, however, the money will not use its value, as it is modeled by π_2 that has the same expected value as in the baseline treatment.

Treatment 5 (T_H) is similar to T_50 , but the role of the recipient changes from an unrelated counterpart to that of an heir. The purpose of T_H is to challenge the robustness of the experimental design and more specifically, to test whether the introduction on an heir who can, in the event of the dictator's death (in the model: realization of zero income determined by the lottery), still derive utility from the dictator's share of the endowment, has an impact on the dictator's allocation decision. Payoffs are structured similarly to T_50 . For the dictator, it can be denoted as

$$\pi_1 = \begin{cases} 0, & \text{with } p_k^1 \\ 10 - x, & \text{with } 1 - p_k^1 \end{cases}, \quad (2)$$

where $p_k^1 = 0.5$. Meanwhile, the heir's payoff function can be

¹All monetary amounts referred to in this paper are indicated in units of m€

described as follows:

$$\pi_2 = \begin{cases} x + \left(\frac{10-x}{1+0.1} \right), & \text{with } p_k^1 \\ x, & \text{with } 1 - p_k^1 \end{cases}, \quad (3)$$

It is important to highlight that the heir's payoff is conditional not only on the dictator's allocation choice but also on the lottery outcome. If the lottery determines that the dictator can keep his/her share, the heir will, as in the ordinary dictator game scenario, realize a payoff of $\pi_2 = x$. Yet, if the dictator loses his share to fate, the heir automatically gains the amount equivalent to the dictator's loss discounted by a factor of $d = 0.1$. Thus, in this treatment, the heir receives a minimum endowment of x , with an additional chance to gain the dictator's share discounted by 10%. In terms of social welfare, this treatment forecloses the possibility of money loss through the lottery. In fact, the lottery, here, serves as a second allocation mechanism. Although the dictator can still decide on the initial allocation of the endowment, the lottery decides whether the split will be implemented or not. If the outcome is at odds with the dictator's luck, the lottery prescribes a re-allocation of to the endowment to the counterpart with only a small deduction. Chances for both outcomes are equal. *T_H* was designed to render the model of allocation decisions under existential threat more realistic concerning two aspects: Firstly, it may be assumed that most people facing a 50% risk of dying within the next few days seek necessary precautions such as instructing a will which settles the allocation of their wealth after their death. Secondly, even if an individual freely decided whether or not he/she should transfer his/her endowment to somebody else prior to his/her death, the recipient of the money would most likely be a close friend or family member. Since a common attribute to legal successions is that they only comprise individuals closely related to the deceased, the designation of the recipient as an heir within the design of a dictator game equips the anonymous counterpart with a natural familiarity. The discount rate applied to the heir's additional gain – implemented only if the dictator's income is zero after the lottery draw – models a delayed payment. Since the legal proceeding of wills is typically lengthy, the heir must expect to receive the endowment significantly later compared to a situation where the ancestor decides to transfer his/her wealth prior to his/her death.

The five dictator treatments are complemented by one additional task designed to provide an indication of the experiment's external validity. Subjects were instructed to imagine a situation in which they just left the hospital after they had been diagnosed with a lethal disease predicting a 75% probability of terminating their life within the next couple of days. Given this situation, subjects were asked to indicate what they would do with their money if they had € 10 million in the bank account. They were explicitly invited to consider different consumption and transfer options. Prescribing a 75% chance of sudden death, this task was designed as a qualitative benchmark to the *T_75* dictator treatment. The framing of the question aimed at priming the subjects with a

certain degree of mortality salience. Answers could be submitted in key points or full sentences.

4. Related Theory and Hypotheses

4.1. Theoretical Model

For the development of the hypotheses to be tested in the experimental part of this study, a model that helps to test for allocation preferences needs to be formalized. [Fehr and Schmidt \(1999\)](#) and [Bolton and Ockenfels, 2000](#), for instance, propose a model of social preferences which measures payoff differences to study fairness concepts based on inequality aversion. In the present study, the focus will be on social welfare implications of initial allocations. As established in the introduction, the discussion of fairness under risk must capture two dimensions: ex ante comparison and ex post comparison. In alignment with existing literature ([Brock et al., 2012](#)), social preferences will be examined over a joint payoff function $S(\pi_1, \pi_2)$ accounting for individual utilities derived from the initial money allocation, denoted as $u(m_1)$ for the dictator and $u(m_2)$ for his/her counterpart. It will be assumed that $E[u(m_1)]$ is equal to $E(\pi_1)$. Modeling the joint payoff function over individual utilities instead of expected payoffs only, however, allows for the possibility to drop this assumption in future extensions of the model.

Tailored to our five dictator treatments, the function of joint expected payoff is described by

$$E[S(\pi_1, \pi_2)] = E[u(m_1)] + E[u(m_2)], \quad (4)$$

which can be rewritten as

$$E[S(\pi_1, \pi_2)] = u(m_1)p_k^1 + u(m_2)p_k^2 \quad (5)$$

where p is the probability with which a player will eventually realize the income corresponding to the dictator's allocation decision. For the baseline treatment, it holds that $p_k^1 = p_k^2 = 1$. In the risky treatments 2, (3; 4), $p_k^2 = 1$ remains unchanged whereas p_k^1 is equal to 0.75 (0.5; 0.25). Values for treatment 5 are equal to those of treatment 2. Starting from there, it can be established that the social welfare maximizing allocation $[m_1, m_2]$ is $[10 - x; x]$ with $x \in [0, 1, \dots, 10]$ for the baseline treatment and $[0; 10]$ for all other treatments. Yet, it must be assumed that social welfare maximization is not the most common decision rule among dictators, especially if it coincides with zero payoff for themselves (treatments 2-5). Only purely altruistic players will favor such an allocation.

Going forward, the implications of fair splits based on ex ante and ex post comparison, two alternative decision rules, shall be discussed with regard to our experimental treatments. According to the principles of ex ante comparison, only an equal split of the total endowment may be considered fair. It follows that for all our treatments, the fair allocation would be $[5; 5]$. In the baseline treatment, this split still allows for the maximization of joint payoff. For the risky treatments, social welfare implications must be assessed based

on expected individual utilities. Since $E[u(m_1)]$ decreases with a smaller p_k^1 , not only the dictator's expected payoff but also the joint payoff is negatively impacted for all allocations $\neq [0; 10]$. In treatment 2 (3; 4), a fair allocation complying with the principles of ex ante comparison yields

$$E[S(\pi_1, \pi_2)] = (10 - x)p_k^1 + xp_k^2 \tag{6}$$

with $x = 5$, $p_k^1 = 0.75$ (0.5; 0.25) and $p_k^2 = 1$, which equals an expected joint payoff $S(\pi_1, \pi_2)$ of 8.75 (7.5; 6.25). For treatment 5, the case is more complex. It can be calculated as follows:

$$E[S(\pi_1, \pi_2)] = E[u(m_1)] + E[u(m_2)] \tag{7}$$

where $E[u(m_1)] = (10 - x)0.5$ and $E[u(m_2)] = x + 0.5 \left(\frac{10 - x}{1.1} \right)$.

Substituting x with 0.5 yields $E[S(\pi_1, \pi_2)] = 9.78$ which is close to the maximum of 10. Table 1 summarizes individual expected utilities and joint payoffs for equal splits in all five treatments.

Turning to decisions based on ex post comparison, the definition of fair allocations as equal splits between the dictator and his/her counterparts becomes obsolete. Fair splits, under this principle, require equal expected utilities for both players. More formally, the requirement for a fair allocation according to the principle of ex post comparison can be denoted as:

$$E[u(m_1)] = E[u(m_2)] \tag{8}$$

For our baseline treatment, this still implies an equal split $[0; 10]$ between the players. In the risky treatments, however, the one-sided risk on the part of the dictator needs to be compensated for with a higher endowment.

This can be easily derived if we rewrite our requirement in the following way:

$$(10 - x)p_k^1 = xp_k^2 \tag{9}$$

For $p_k^1 < p_k^2$ and $x \in [0, 1, \dots, 10]$, which applies to all risky treatments of the experiment, this equation can only hold if $x < 5$. In the following, the process will be exemplified for treatment 5 (T_H). From the previous analysis, it is known that for T_H

$$E[u(m_1)] = 0.5(10 - x) \text{ and } E[u(m_2)] = x + 0.5 \left(\frac{10 - x}{1.1} \right) \tag{10}$$

Equating the two expected utilities yields $x = \frac{10}{23}$, corresponding to an allocation of $\left[\frac{220}{23}; \frac{10}{23} \right] \cong [9.57; 0.43]$ and an expected social welfare of $E[S(\pi_1, \pi_2)] = 9.57$. Table 3 summarizes individual expected utilities and joint payoffs for fair allocations according to the requirements of ex post comparison.

Notably, perfectly fair splits according to ex post comparison can only be realized in the baseline treatment and treatment four because allocation decisions in the experiment are subject to the constraint that $x \in [0, 1, \dots, 10]$. To accommodate this constraint in the experimental design, it will be assumed that a discrete split of the initial endowment is compliant with the concept of fairness based on ex post comparison if x takes on the value closest to the corresponding $x \in \mathbb{R}$. In treatment 2, for instance, a fair allocation as defined by ex post comparison would be $[6; 4]$, corresponding to an expected social welfare of $E[S(\pi_1, \pi_2)] = 8.5$. Table 2 summarizes individual expected utilities and joint payoffs for fair allocations under the constraint that the definition range of x encompasses discrete value only and compliant with the rules of ex post comparison.

It may be highlighted that the fair allocation in the heir treatment (5) is for the dictator to keep the entire endowment to him-/herself and give nothing to his/her heir. With a 50% chance of receiving a slightly discounted amount of the dictator's share, the heir's expected payoff over the initial split, even if he/she is awarded nothing, is sufficiently high to justify an allocation of $[10; 0]$ according to the principle of ex post comparison.

Fair splits compliant with the requirements of ex ante versus ex post comparison will be considered in the analysis of the experimental data. More precisely, it will be evaluated to which degree individuals are willing to compromise on both individual utility maximization and social welfare maximization in order to implement an allocation that is considered fair. Generally, it may be noteworthy that fair splits compliant with ex post comparison yield higher expected utilities for the dictator as well as a higher expected joint payoff.

Aside from inequality aversion, concerns for efficiency will also be taken into account. Hence, an additional factor considered in the analysis will be the efficiency of capital employment, defined as

$$e = \frac{E[S(\pi_1, \pi_2)]}{\max(S)} \tag{11}$$

The higher e , the closer is the expected joint payoff corresponding to the dictator's allocation choice to the maximum possible joint payoff.

4.2. Theoretical Predictions

Prediction 1: Giving in the standard dictator game is not fully consistent with allocation decisions unilaterally risky dictator games.

This assumption extends to two levels: Firstly, dictators' allocation decisions under personal risk will deviate from decisions without risk in absolute terms. Secondly, a difference will also be observed in relative terms, implying that the share of dictators complying with the concept of ex ante or ex post fairness, as well as those making purely altruistic decisions will vary across treatments. Prediction 1 allows for two alternative hypotheses as explained in the following.

H_1 : Facing personal risk, dictators will allocate more money to their counterpart.

Table 1: Expected utilities for equal splits

Exp. utilities with a fair split according to ex ante comp.					
	max (S)	Allocation	$E[u(m_1)]$	$E[u(m_2)]$	$E[S(\pi_1, \pi_2)]$
T_C	10	[5; 5]	5	5	10
T_25	10	[5; 5]	3,75	5	8,75
T_50	10	[5; 5]	2,5	5	7,5
T_75	10	[5; 5]	1,25	5	6,25
T_H	10	[5; 5]	2,5	7,28	9,78

Table 2: Expected utilities for fair splits according to ex post comparison (discrete)

Exp. utilities with a fair split according to ex post comp.					
	max (S)	Allocation (discrete)	$E[u(m_1)]$	$E[u(m_2)]$	$E[S(\pi_1, \pi_2)]$
T_C	10	[5; 5]	5	5	10
T_25	10	[6; 4]	4,5	4	8,5
T_50	10	[7; 3]	3,5	3	6,5
T_75	10	[8; 2]	2	2	4
T_H	10	[10; 0]	5	4,55	9,55

Table 3: Expected utilities for fair splits according to ex post comparison (stationary)

Exp. utilities with a fair split according to ex post comp.					
	max (S)	Allocation (stationary)	$E[u(m_1)]$	$E[u(m_2)]$	$E[S(\pi_1, \pi_2)]$
T_C	10	[5; 5]	5	5	10
T_25	10	[5,71; 4,29]	4,44	4,44	8,88
T_50	10	[6,67; 3,33]	3,33	3,33	6,66
T_75	10	[8; 2]	2	2	4
T_H	10	[9,57; 0,43]	4,785	4,785	9,57

H_1 is consistent with the idea that dictators care about their decisions' impact on joint payoff (social welfare). If their own share is subject to risk, but the counterpart's share is unaffected, dictators who honor the social welfare dimension of their decisions will give higher amounts (in relative and absolute terms) to their counterparts than non-risky conditions. On a higher level, H_1 would imply that individuals facing existential threat favor a transfer of their endowment over excessive self-consumption because others might derive more sustainable utility from the money. In other words, mortality salience will reinforce other-regarding preferences.

H_2 : Facing personal risk, dictators will keep more money to themselves.

H_2 assumes that dictators will try to equalize expected payoffs for both players, which is consistent with the concept of fairness based on ex post comparison. Compared to the standard dictator game, decisions involving personal risk will incentivize dictators to increase the share kept to themselves (in absolute terms). In the context of decisions made under an elevated risk of sudden death, this hypothesis would imply that mortality salience will reinforce selfish motives. Individuals who are aware of their limited time and consumption

horizon will try to maximize their own utility without regard for the social consequences of their consumption decisions.

Although both hypotheses can be sufficiently justified, H_1 is considered to be more likely than H_2 and will serve as the basis for predictions 3 and 4.

Prediction 2: Allocation decisions in risky dictator games are not exclusively driven by the dictator's degree of risk aversion.

Modeling existential threat through a risky lottery comes at the disadvantage that upward deviations from giving in the standard dictator game may be induced by sheer risk avoidance. Put differently, dictator's allocating more to their counterpart in the risky treatment might just be risk-averse. By contrast, prediction 2 anticipates a weak correlation between risk aversion and higher giving in risky dictator games. Instead, risk aversion should be treated as a confounding variable in the regression run on giving in the unilaterally risky dictator games dependent on giving in the baseline treatment.

Prediction 3: Neither ex post nor ex ante comparisons can explain giving in unilaterally risky dictator games. Instead, non-selfish allocations reflect a preference for efficient capi-

tal employment.

Instead of reflecting the dictator's degree of risk aversion, higher giving in risky dictator games may be explained by the dictator's preference for efficient capital allocations. This prediction is construed *ex negativo*: From the review of the existing literature on risky dictator games, it is hypothesized that neither altruism nor *ex ante* or *ex post* comparison can fully explain allocation decisions in unilaterally risky dictator games. Instead, a preference for efficient capital employment aimed at a constrained optimization of the joint payoff may be a driver behind non-zero giving in the risky treatments. An interest in social welfare, in this case, is equal to the understanding that other players may derive higher utility from the endowment.

Prediction 4: Generous giving in dictator games increases with the level of risk assigned to the dictator's own payoff. As stated earlier, prediction 3 is conditional on higher giving in the unilaterally risky dictator games relative to the baseline treatment. In extension to the preceding conjecture, it may be expected that the predicted behavior is amplified by a higher degree of risk attributed to the dictator's share. Notably, this is not equivalent to the assumption that the dictator's concern about the joint payoff increases with the level of harm he/she can cause to social welfare. Any theory based exclusively on the correlation between giving in risky dictator games and the level of unilateral risk will be not consistent with varying results across treatments T_{50} and T_H . Both treatments assign a 50% chance of value loss to the dictator's share but with significantly less downside for social welfare in the heir treatment (5).

Prediction 5: One-sided risk in dictator games can effectively simulate allocation decisions under existential threat. Lastly, the experimental design presented in this study is based on the assumption that unilateral exposure to risk in a dictator game scenario resembles a simulation where an individual faces a life-endangering hazard. Prediction 4 will be tested by comparing the subjects' responses to QMS_{75} to those of the dictator treatment T_{75} . If it holds, the relative number of selfish versus other-regarding use cases listed in QMS_{75} will mirror the quantitative split between self-use and transfer to the counterpart in the fourth dictator treatment.

5. Experimental Results (Data Analysis)

A summary of the dictator's allocation decisions is provided in Table 4.

The table average choices and the proportion of players choosing $x > 0$, $x \geq 5$ and $x = 10$ for all five treatments. Average giving in the standard dictator game, for instance, is $x = 2.63$ ($\cong 26\%$) and thus consistent with the results reported previous studies (Engel, 2011). Notably, significant positive giving can be observed in all treatments. Figure A5 shows the average contribution by task whereas Figure A6 depicts the fraction of dictator's giving non-zero amounts and those giving more than the fairness concept based on *ex post*

comparison requires (both in the Appendix). In the following, it will be explored in detail how allocation decisions differ across treatments in detail.

Result 1: Behavior in the standard dictator game is not a good predictor of giving in unilaterally risky treatments. On average, dictators behave more generously when exposed to risk.

Evidently, average contributions in the standard dictator game are significantly different (at 1% level) from giving in the risky dictator treatments (Table 5). The summary statistics of the Wilcoxon signed-rank test run on all treatments also reveals that the group means are, in fact, significantly different between any two treatments except for T_{50} and T_H . This outlier was to be expected because the heir treatment is an adaptation of T_{50} and exposes the dictator to the same level of risk.

Table 6 displays the results of a linear regression which tries to explain the allocation decisions in risky dictator games as a function of allocation choice in the standard dictator game (T_C). Regression results substantiate that giving in the standard dictator game is not always predictive of giving under personal risk: the significance of the coefficient for giving in T_C decreases with the level of risk imposed on the dictator treatments. This implies that decisions under risk are not driven by the same motives as decisions in ordinary dictator games. Nevertheless, generosity in the standard dictator game is predictive of a tendency of giving under mild risk (coefficients are significant at 1% level for T_{25} and T_H).

Result 2: The dictator's degree of risk aversion does not moderate the effect of risk on allocation decisions in dictator games.

To control for the correlation between risk profile and behavior in the dictator game treatments on the individual level, the experimental design comprised a task testing for the willingness to pay (WTP) for a risky lottery. The lottery was designed in a way that subjects would face an expected value of € 100 with a 50% chance of winning € 200. Subjects could state their upper price threshold for a ticket to participate in the lottery. Results were evaluated on both descriptive and analytical statistics.²

Firstly, subjects were assigned to a specific cluster according to the WTP stated in the lottery task. Four different clusters are used based on the definition of different levels of risk-aversion. Cluster 1 comprises strictly risk-averse subjects with a WTP between € 0 and € 49 ($n = 70$), i.e., below half of the lottery's expected value. Cluster 2 also contains risk-averse agents but with a WTP above half of the lottery's expected value, specifically between € 49 and € 98 ($n = 53$). The third cluster counts subjects ($n = 53$) with an approximately risk-neutral profile displaying a WTP between € 99 and € 100. Subjects with a risk-loving attitude, i.e., stating a WTP above the lottery's expected value, are classified as cluster 4. Yet, this cluster only comprises a very small portion of the original sample ($n = 4$). All four clusters were

²Task is referred to as QL in the data analysis

Table 4: Summary statistics of the allocation decisions

*For the individual treatments, this means: $x \geq 5$ (T_C); $x \geq 4$ (T_25), $x \geq 4$ (T_50), $x \geq 2$ (T_75); $x \geq 0$ (T_H)

	n	Mean giving	SD of giving	% of dictators giving $x > 0$	% $x \geq$ 'fair split ex post'*	% of dictators giving $x \geq 5$	% of dictators giving $x = 10$
T_C	178	2,63	2,22	68%	32%	32%	1%
T_25	178	4,65	3,42	84%	57%	47%	20%
T_50	178	5,47	3,42	90%	76%	63%	25%
T_75	178	6,22	3,75	88%	83%	67%	33%
T_H	178	5,28	3,90	79%	100%	61%	29%

Table 5: Differences in giving ($n = 178$)

Differences tested with Wilcoxon signed-rank tests. *** (**, *) indicates significance at 1% (5%, 10%) level.

	T_25	T_50	T_75	T_H
T_C	2,02 ***	2,84 ***	3,60 ***	
T_25		0,81 ***	1,57 ***	
T_50			0,76 ***	-0,19

Table 6: Linear regression of allocation decisions in a standard dictator game treatment (independent variable: x in T_C) on giving under risk (clustering at the individual level)

Differences tested with Wilcoxon signed-rank tests. *** (**, *) indicates significance at 1% (5%, 10%) level.

	Dependant variable			
	T_25	T_50	T_75	T_H
Amount given in T_C (coeff.)	0,4009 ***	0,2727 **	0,1848	0,3593 ***
Constant	3,60 ***	4,75 ***	5,74 ***	4,34 ***
Pseudo R-squared	0,07	0,02	0,01	0,04
F-statistic	12,84 ***	5,7 **	2,14	7,37 ***

evaluated with regard to deviations in average giving from the sample mean per treatment. The descriptive analysis reveals three main observations: Firstly, cluster deviations occur across all treatments including the non-risky baseline game, implying that individual risk attitudes may be influential to dictator games in general, not only in risky treatments. Secondly, average deviations in giving are positive for cluster 1 and become increasingly negative from there. Only cluster 4 seems to show an average deviation above 5% from the sample mean (close to the 20% level), but due to its small sample size, the effect may be distorted. Thirdly, there seems to be no consistent pattern as to which treatment elicits the most substantial deviations on cluster level. This may serve as an indication that risk attitudes are not predictive of giving in risky dictator games. Table A2 provides the summary statistics of differences in giving by cluster whereas Table A3 displays a consolidated overview of the standard deviation by cluster on treatment level.

Secondly, in order to verify the assumed weak correlation between individual risk attitude and the dictator game results, a regression was run on the WTP in the lottery de-

pendent on giving in the different dictator game treatments. Table 7 summarizes the results on treatment level. Notably, coefficients are negative for all treatments.

In extension to the previous analysis, the regression reveals for which treatments groups risk attitudes are most influential. The coefficient for the risk level is only significant for T_75 (at 5% level) and T_H (at 10% level). Thus, in contrast to the intuition derived from the descriptive analysis, individual risk attitudes seem to gain in decision weight with the level of risk attached to the dictator's payoff consequences. Nevertheless, the effect is negligibly weak since the coefficient size is considerably small for both treatments.

Overall, it may be concluded that allocation decisions in risky dictator games are not primarily attributable to the dictator's risk attitude, even though a small effect can be observed for the high-risk treatment.

Result 3: Neither ex post nor ex ante comparisons can explain giving in unilaterally risky dictator games. Allocation decisions may be influenced by a preference for efficient capital employment rooted in the dictator's concern for the joint payoff.

Table 7: Linear regression of risk aversion (independent variable: P in QL) on allocation decisions in dictator game treatments (clustering at the individual level)

	Dependant variable				
	<i>T_C</i>	<i>T_25</i>	<i>T_50</i>	<i>T_75</i>	<i>T_H</i>
WTP in QL (coeff.)	-0,0035	-0,0061	-0,0054	-0,0104 *	-0,0074 *
Constant	2,83 ***	5,01 ***	5,79 ***	6,84 ***	5,72 ***
Pseudo R-squared	0,01	0,01	0,01	0,03	0,01
F-statistic	1,94	3,41	2,03	6,35 **	2,94 *

While the joint payoff, or social welfare, can be maximized by any choice of x in the baseline treatment, the risky dictator games presented in this study only allow for the maximum joint payoff if the purely altruistic choice ($x = 10$) is realized. Previous studies have shown that only a minor fraction (approx. 5%) of dictators exhibits purely altruistic behavior (Engel, 2011). Consequently, if the dictator's interest in the joint payoff shall recoup attention in this study, it needs to be in form of a constrained maximization. Constrained, in this case, means that the dictator tries to employ the initial endowment of 10m€ efficiently without forgoing his own payoff entirely. More generally, it refers to all allocations which are attentive towards the social welfare implications but do not leave the dictator with an expected payoff of zero.

Prior to evaluating the relevance of social welfare concerns going along with efficient capital employment, the new experimental data will be screened for evidence supporting the application of four other decision rules emphasized in the existing literature: aversion against a purely selfish maximization of payoff chances, ex post comparison, ex ante comparison, and altruistic maximization of joint payoff (listed in order of decreasing x required to comply with the rule). Table 4, presented at the beginning of this section, shows the proportion of dictators exceeding the minimum requirements for these rules as specified in Table 8.

Overall, the experimental results show that all four types of non-selfish allocations occur across all five treatments, although altruistic maximization of the joint payoff is almost non-existent in the standard dictator game. The fraction of subjects choosing allocations which exceed one of the four thresholds for compliance seems to increase with the risk imposed to the dictator's endowment. Growth rates are generally the highest between the baseline treatment and T_{25} and maintain a rate for ex ante comparison and altruism, while the incremental fraction of subjects complying with the other two decision rules seem to abate with an increasing level of risk.

Remarkably, a closer look at the fraction of dictators implanting an allocation choice which exactly equals one of the four minimum requirements (cf. Table 9) reveals that especially for the risky treatments, the portion of subjects who do not exactly comply with any of those four strategies is considerably high (51% for T_{25} , 40% for T_{50} , and 46% for T_{75}). Since these treatments equally induce a significantly higher percentage of non-zero giving than the stan-

dard dictator game which, hence, cannot be explained by fairness concepts based on ex post or ex ante comparison, there is an explanation gap as to which other drivers motivate non-zero giving in unilaterally risky dictator games. It stands to reason that the observed behavior reflects the dictator's preference for efficient capital employment allowing for constrained joint payoff maximization.

Result 4: Generosity in risky dictator games increases with the dictator's own risk, not the risk attached to the joint payoff.

As a next step, the analysis needs to address the role of risk imposed on the dictator's own share versus the risk affecting the joint payoff. As reported earlier, unilaterally risky dictator games seem to elicit more generous giving than the baseline scenario, and the effect seems to increase with the level of risk assigned to the dictator's payoff. To test for the robustness of this result against the hypothesis that higher giving correlates with the degree to which the dictator's decision affects the joint payoff, the heir treatment (T_H) was introduced. In treatments T_{25} , T_{50} and T_{75} , a higher level of risk for the dictator comes along with a higher level of risk to the joint payoff. In T_H , however, the risk assigned to the dictator's own payoff is equal to T_{50} whereas the risk to the joint payoff is significantly cut. The minimum payoff for the dictator is 0€ at a 50% probability whereas the minimum social payoff is 9.09m€ at a probability of $\leq 50\%$. For both values, 10m€ marks the upper margin. Hence, of dictators behaved significantly more generous in T_{50} compared to T_H , this would imply that higher giving in unilaterally risky dictator games is not dependent on the risk imposed to the dictator's own payoff, but the risk imposed to the joint payoff. Results show that more dictators transfer the entire endowment to their counterparts in T_H ($x = 10$), but likewise, more dictators keep the money to themselves ($x = 0$) in T_H compared to T_{50} (Table 9). In aggregate, the two effects seem to cancel each other out. Average giving is $x = 5.28$ in T_H and $x = 5.47$ in T_{50} . A Wilcoxon signed-rank test run on the results indicates that the difference is not significant (Table 5). Consequently, the data suggests that more generous giving in unilaterally risky dictator games, de facto, increases with the level of risk attributed to the dictator's payoff, not the level of risk attributed to the joint payoff. With regard to the motivational drivers behind this behavior, the results encourage the assumption that the generosity exhibited in unilaterally risky dictator games is not

Table 8: Minimum requirements for four non-selfish decision rules in dictator games

Decision Rule	Minimum Requirement
Aversion against purely selfish maximization	Allocation such that $x > 0$
Ex post comparison of payoff chances	Allocation such that $E(\pi_1) = E(\pi_2)$ holds for $\min(x)$
Ex ante comparison of payoff chances	Allocation such that $x \geq 5$
Altruistic maximization of joint payoff	Allocation such that $x = 10$

Table 9: Fraction of dictators pursuing a one out of four strategies: pure selfishness, ex post comparison, ex ante comparison or pure altruism

*For the individual treatments, this means: $x = 5$ (T_C); $x = 4$ (T_{25}), $x = 4$ (T_{50}), $x = 2$ (T_{75}); $x = 0$ (T_H)

	Mean giving	% of subjects with $x=0$	% $x =$ 'fair split ex post**	% of subjects with $x=5$	% of subjects with $x = 10$	Remaining subjects
T_C	2,63	6%	31%	31%	1%	31%
T_{25}	4,65	4%	10%	15%	20%	51%
T_{50}	5,47	7%	8%	20%	25%	40%
T_{75}	6,22	6%	9%	6%	33%	46%
T_H	5,28	6%	21%	16%	29%	29%

induced by regret minimization but by a preference for allocating the money efficiently.

Result 5: Unilaterally risky dictator treatments hold considerable external validity for decisions under existential threat.

As stated earlier, the experiment presented in this study incorporated a qualitative task priming subjects with ephemerality-related thoughts while asking what subjects would use their money for if they only had a 25% chance of surviving the next days.

Remarkably, the responses accurately reflect the split pointed to in prediction 1 in which two different scenarios were anticipated: under existential threat, individuals would either become more generous considering that their own utility from a monetary endowment is subject to a high level of risk (H_1), or they would become increasingly selfish because the time horizon for consumption is suddenly very limited and social consequences carry little weight (H_2). Yet, while the answers carry clues for both hypotheses, they also provide a good indication of which scenario prevails. Evidence of H_1 is not only voiced more frequently but also more resolutely as illustrated by these two examples:

I would not spend the money in the next few days. [...] Spending the money on personal pleasure and dying right afterwards would be a waste.

I wouldn't consume anything but rather find a person that can make better use of the money.

Support for H_2 is usually restricted to only a fraction of the money, which can be exemplified referring to these two responses:

I would fly out all my friends and myself to Tasmania for a final dinner. [...] Whatever is left of the money after all this, I would give to the Tasmanian Wildlife fund. 10 million can be better employed elsewhere.

Give 6 million to charity and waste the rest like it is the end of the world.

For a comprehensive evaluation of the answers, six designated use cases were identified from the most frequently used keywords: self-use (luxury/gambling), self-use (event/experiences), long-term investments/trust funds, therapy research, (charitable) donations, and transfers to family and friends (F&F). While the first two use cases clearly resemble an allocation decision in favor of the dictator and the last two use cases resemble an allocation in favor of his/her counterpart, the use cases named third and fourth do not explicitly fall into one of those categories. Additionally, it is essential to note that these use cases are collectively exhaustive but not mutually exclusive since subjects were not restricted in the number of items when answering the question. Table 10 provides an overview of the popularity of each use case relative to the number of subjects and relative to the total number of items listed in the responses to QMS_75.

In order to verify the external validity of the experimental design, however, it is necessary to assess whether the relative number of selfish versus other-regarding use cases listed in QMS_75 mirrors the quantitative split between self-use and transfer to the counterpart in the T_{75} dictator game treatment. As a reminder, the mean in the fourth treatment amounts to $x = 6.22$, meaning that on average, dictators transferred 62.25% of their endowment to the other player.

Table 10: Common use cases for the endowment (sorted by popularity, $n = 52$)

* Number of responses applicable to Question QMS_75. Percentages within this row indicate how many subjects relative to the total number of subjects would spend money on the respective use case.

** Number of items accounting for the possibility of multiple answers. Percentages within this row indicate how many times a use case is mentioned relative to the total number of items listed

	n	Therapy Research	Self-use: Luxury/ Gambling	Long-term invest- ments/Trust funds	(Charitable) Donations	Self-use: Event/ Experiences	Transfer to F&F
Fraction by subjects	176*	10%	14%	24%	32%	49%	60%
Fraction by items	333**	5%	7%	13%	17%	26%	32%

Accounting for the number of use cases which clearly classify as self-use (luxury/gambling and event/experiences, to be specific) relative to the total number of items listed in the responses to QMS_75 and balancing the result against the share of other-regarding use cases (donations and transfers to F&F) while subtracting out those use cases that do not explicitly match either category yields a 33% share of selfish use-cases. This result is approximately consistent with the 37.75% share predicted by T_{75} . However, this comparison is imprecise to the extent that it only considers the relative number of times certain use cases are listed, not the amount of money allocated towards them.

For the purpose of refining the analysis of the qualitative data collected in QMS_75, only those answers which indicated a clearly quantifiable allocation of the endowment to one or more use cases were extracted from the sample ($n = 52$). Table 11 summarizes the total as well as the relative amount of money [in m€ and %, respectively] across the applicable responses allocated to any of the six designated use cases.

Coloring indicates whether the use case counts as self-use (orange), transfer to others (green), or matches neither of those categories (white). It can be discerned that only 7% of the money is used for self-consumption, 18% is allocated towards long-term investments and trust funds, and 75% is transferred to family and friends. Even if it is argued that long-term investment and trust funds can be classified as self-use and thus, on average, subjects keep 25% of their endowment to themselves, an average of 75% transferred to others still does not exactly match the average of 62.25% predicted in the corresponding dictator treatment (T_{75}). One explanation for this incongruence may be that the QMS_75 task allowed subjects to specify to whom exactly they would transfer the money and a large fraction chose to endow their family and friends, which is also the reason why one of the six high-level use cases is called “transfer to F&F” instead of “transfer to others”. Transfers to close friends or relatives can be assumed to earn subjects a higher social utility than transfers to the anonymous counterpart they were paired with in the dictator treatments. The effect of social distance is undis-

puted in the dictator game literature (Leider, Möbius, Rosenblat, and Quoc-Anh, 2009; Goeree et al., 2010). Unsurprisingly, also in the present experiment, only four subjects in the entire sample explicitly stated that they would like to transfer a portion of the money to strangers.

I would give away cash to random people on the street as well as to people on my day-to-day ways (e.g. the salesperson at REWE, the bakery etc.)

With this in mind, it may be concluded the data collected from the QMS_75 task and the T_{75} dictator game is consistent in the result that a vast majority of subjects prefers to grant a portion well above 50% of their endowment to another person. This substantiates the claim made in H_1 and opens it to a broader context: Facing personal risk, e.g., under existential threat, not only dictators in a dictator game setting but individuals, in general, become more generous and transfer a higher fraction of their endowment to others. Overall, the results suggest that the experimental design features a sufficient level of external validity for decisions under existential threat. This result is diluted only by the fact that baseline frequencies under artificial dictator games contexts, in general, are inflated relative to those exhibited in a natural setting (Winking and Mizer, 2013).

6. Discussion and Conclusion

6.1. Contribution to the Academia

In extension to the existing literature on both risky dictator games and mortality salience, the present study contributes an account of allocation decisions under existential threat. Numerous studies on giving in dictator (and ultimatum) games attempt to explain behavior inconsistent with the predictions of classic game theory by models of social preferences that allow for non-selfish motives such as inequality aversion or concerns for efficiency. It remained to be assessed how such social behavior changes with different levels of risk attached to the dictator’s payoff, or, equivalently, under existential threat. Hence, this study does not only contribute an entirely new variation of the dictator game comparing

Table 11: Fraction of money allotted to most popular use cases (sorted by popularity, $n = 52$)

* Number of subjects who clearly quantified their allocation of money in QMS_75.

	n	Therapy Research	Self-use: Luxury/ Gambling	Self-use: Event/ Experiences	Long-term invest- ments/Trust funds	(Charitable) Donations	Transfer to F&F	Total
Money allocated [m€]	52*	4	12	24	92	148	238	518
Money allocated [%]	52*	1%	2%	5%	18%	29%	46%	100%

behavior under different degrees of self-risk but simultaneously provides a first attempt to synthesize research in the field of experimental economics with (mainly psychological) research dedicated to understanding behavior under the condition of mortality salience.

The study finds that allocation decisions in dictator's games tend to become more generous when only the dictator's payoff is at risk. This is inconsistent with the findings of Freundt and Lange (2017) and is also at odds with evidence suggesting that mortality awareness spurs greed (Kasser and Sheldon, 2000; Cozzolino et al., 2004a). However, the new results can be aligned with another stream of mortality salience research arguing that existential threat actually increases pro-social behavior (Jonas et al., 2000; Wade-Benzoni et al., 2012). Similarly, consistency with the existing literature on dictator games can be stabilized under the assumption that giving in unilaterally risky dictator games is motivated by a preference for efficient capital employment (Charness and Rabin, 2002). Freundt and Lange (2017) did provide their dictators with a safe transfer option, which made giving, even in the self-risk treatments, disadvantageous to the social welfare. A concern for social welfare which motivates non-selfish allocation decisions is also consistent with findings from experiments on social lotteries indicating that subjects are also more risk-averse in social lotteries compared to a situation where only their own payoff is exposed to risk (Gaudeul, 2013).

Moreover, the presented results reveal that generosity in risky dictator games increases with the dictator's own risk, not the risk attached to the joint payoff, and show that risk aversion falls short of explaining dictators' behavior in unilaterally risky games. Therewith, this study provides a first attempt towards a better understanding of allocation decisions under unilateral risk and simultaneously bridges the gap to research on allocation decisions under existential threat. Eventually, it needs to be acknowledged that the study does not claim to portray an exhaustive model of dictator giving under elevated mortality awareness and the omission of decisive determinants cannot be precluded.

6.2. Limitations

The presented experiment is subject to a set of limitations, especially from a methodological perspective. Chiefly, the subject pool is rather homogeneously composed of students with a background in economics. Additionally, the gender split is not equal, which might skew the results downwards, since allocation decisions in all treatments show significant differences depending on the dictator's gender (Table 12; for the gender-specific breakdown of the means per treatment, see Table A4 in the Appendix). Another issue is that the within-subject design naturally bears the risk of carry-over and position effects distorting the latter treatments.

When it comes to the design of the treatments, it must be acknowledged that the standard dictator game is played with an endowment of € 10 instead of 10m€ and research has shown that a larger stake size has a small but significant adverse effect on giving in dictator games (Larney et al., 2019). Moreover, the qualitative response task (QMS_75) could be optimized for comparison with the dictator treatment T_{75} by requesting quantifiable allocations of the 10m€ to predefined use-cases. Yet, the current design gives a better account of the experiment's external validity. Overall, however, the equivalences between allocation decisions in unilaterally risky dictator treatments and such under existential threat must be proven by additional side-treatments controlling for the experiment's validity.

6.3. Concluding Remarks

Unilateral risk on the part of the dictator's payoff encourages more generous allocation decisions. The observed effect is largely independent of the dictator's risk profile and is stronger for higher-risk treatments. On a higher level, the witnessed behavior indicates an increase in other-regarding preferences under existential threat. Since psychological experiments on mortality salience are inconclusive about this effect, further research should be dedicated to modeling decisions under existential threat with different types of economic games in order to improve the robustness of the results. Another interesting angle for further research on economic decisions under existential threat would be a compar-

Table 12: Average deviation from sample mean giving by gender

	<i>n</i>	<i>T_C</i>	<i>T_25</i>	<i>T_50</i>	<i>T_75</i>	<i>T_H</i>
Dev. (F)	64	0,68	0,52	0,56	0,95	1,05
Dev. (M)	114	-0,38	-0,29	-0,32	-0,53	-0,59

ative analysis of the relevance ascribed to different consumption options for a given amount of money. Similar to the retrospective use case analysis applied to *QMS_75*, it would be intriguing to present subjects with a set of consumption options from different categories, for instance spending money on travel, luxury goods, psychedelic drugs or an R&D investment for the own family business. Such trade-off decisions would be especially interesting in combination with research suggesting the affective deconstruction of the probability weighting function. [Rottenstreich and Hsee \(2001\)](#), for instance, provide evidence that individuals' sensitivity towards changes in probabilities near certainty and impossibility increases with the level of affect attached to a lottery. On the background of these findings, it may be possible that individuals favor affect-rich consumption choices under existential threat. Finally, another research proposal would be to take an economic view on actual doomsday decisions. Doomsday decisions, as shortly referred to in the introductory section, require a situation where not only one affected party but all parties of a decision are confronted with an elevated risk of dying within the next few days. Economically speaking, not only the payoff uncertainty is affected, but the time horizon for consuming utility from money is limited for all affected stakeholders. In synopsis, it may be concluded that the academic potential in the synthesis of economic and psychological research on decisions under existential threat is far from being exhausted. Nevertheless, the present study sets the first impulse towards a better understanding of the economic preferences of terminally ill individuals. Simultaneously, artificially induced mortality salience may be leveraged to nudge people into more generous giving or to improve long-term strategic planning in (family) businesses.

References

- Bolton, G. E. and Ockenfels, A. Erc: A theory of equity, reciprocity, and competition. *American economic review*, 90(1):166–193, 2000.
- Brock, J. M., Lange, A., and Ozbay, E. Y. Dictating the risk: Experimental evidence on giving in risky environments. *American Economic Review*, 103(1):415–37, 2013.
- Camerer, C. F. *Behavioral game theory: Experiments in strategic interaction*. New York, Sage, 2003.
- Charness, G. and Rabin, M. Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117(3):817–869, 2002.
- Church, B. K. and Zhang, P. Bargaining behavior and payoff uncertainty: Experimental evidence. *Journal of economic psychology*, 20(4):407–429, 1999.
- Cozzolino, P. J., Staples, A. D., Meyers, L. S., and Samboceti, J. Greed, death, and values: From terror management to transcendence management theory. *Personality and Social Psychology Bulletin*, 30(3):278–292, 2004a.
- Cozzolino, P. J., Staples, A. D., Meyers, L. S., and Samboceti, J. Greed, death, and values: From terror management to transcendence management theory. *Personality and Social Psychology Bulletin*, 30(3):278–292, 2004b.
- Engel, C. Dictator games: A meta study. *Experimental Economics*, 14(4):583–610, 2011.
- Engelmann, D. and Strobel, M. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: Reply. *American Economic Review*, 96(5):1918–1923, 2006.
- Erat, S. and Gneezy, U. White lies. *Management Science*, 58(4):723–733, 2012.
- Fehr, E. and Schmidt, K. M. A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, 114(3):817–868, 1999.
- Forsythe, R., Horowitz, J. L., Savin, N. E., and Sefton, M. Fairness in simple bargaining experiments. *Games and Economic behavior*, 6(3):347–369, 1994.
- Freundt, J. and Lange, A. On the determinants of giving under risk. *Journal of Economic Behavior & Organization*, 142:24–31, 2017.
- Gaudeul, A. Social preferences under uncertainty. Technical report, Jena Economic Research Papers, 2013.
- Goeree, J. K., McConnell, M. A., Mitchell, T., Tromp, T., and Yariv, L. The 1/d law of giving. *American Economic Journal: Microeconomics*, 2(1):183–203, 2010.
- Güth, W., Levati, M. V., and Ploner, M. On the social dimension of time and risk preferences: An experimental study. *Economic inquiry*, 46(2):261–272, 2008.
- Haisley, E. C. and Weber, R. A. Self-serving interpretations of ambiguity in other-regarding behavior. *Games and economic behavior*, 68(2):614–625, 2010.
- Hartog, J., Ferrer-i Carbonell, A., and Jonker, N. On a simple survey measure of individual risk aversion. 2000.
- Heine, S. J., Proulx, T., and Vohs, K. D. The meaning maintenance model: On the coherence of social motivations. *Personality and social psychology review*, 10(2):88–110, 2006.
- Hoffman, E., McCabe, K., Shachat, K., and Smith, V. Preferences, property rights, and anonymity in bargaining games. *Games and Economic behavior*, 7(3):346–380, 1994.
- Jonas, E., Schimel, J., Greenberg, J., and Pyszczynski, T. The scrooge effect: Evidence that mortality salience increases prosocial attitudes and behavior. *Personality and Social Psychology Bulletin*, 28(10):1342–1353, 2002.
- Kang, A. R., Blackburn, J., Kwak, H., and Kim, H. K. I would not plant apple trees if the world will be wiped: Analyzing hundreds of millions of behavioral records of players during an mmorpg beta test. In *Proceedings of the 26th International Conference on World Wide Web Companion*, pages 435–444, 2017.
- Kasser, T. and Ryan, R. M. A dark side of the american dream: Correlates of financial success as a central life aspiration. *Journal of personality and social psychology*, 65(2):410, 1993.
- Kasser, T. and Sheldon, K. M. Of wealth and death: Materialism, mortality salience, and consumption behavior. *Psychological science*, 11(4):348–351, 2000.
- King, L. A., Hicks, J. A., and Abdelkhalik, J. Death, life, scarcity, and value: An alternative perspective on the meaning of death. *Psychological Science*, 20(12):1459–1462, 2009.
- Kortenkamp, K. V. and Moore, C. F. Time, uncertainty, and individual differences in decisions to cooperate in resource dilemmas. *Personality and Social Psychology Bulletin*, 32(5):603–615, 2006.
- Krawczyk, M. and Le Lec, F. 'give me a chance!' an experiment in social decision under risk. *Experimental economics*, 13(4):500–511, 2010.
- Laibson, D. Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2):443–478, 1997.
- Larney, A., Rotella, A., and Barclay, P. Stake size effects in ultimatum game and dictator game offers: A meta-analysis. *Organizational Behavior and Human Decision Processes*, 151:61–72, 2019.
- Loewenstein, G. and Elster, J. Utility from memory and anticipation. *Choice over time*, pages 213–234, 1992.
- Mandel, N. and Heine, S. J. Terror management and marketing: He who dies with the most toys wins. *ACR North American Advances*, 1999.
- Mandel, N. and Smeesters, D. The sweet escape: Effects of mortality salience on consumption quantities for high-and low-self-esteem consumers. *Journal of Consumer Research*, 35(2):309–323, 2008.
- Rottenstreich, Y. and Hsee, C. K. Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological science*, 12(3):185–190, 2001.
- Scheffler, S. and Kolodny, N. *Death and the Afterlife*. Oxford University Press, 2013.
- Solomon, S., Greenberg, J., and Pyszczynski, T. A terror management theory of social behavior: The psychological functions of self-esteem and cultural worldviews. In *Advances in experimental social psychology*, volume 24, pages 93–159. Elsevier, 1991.
- Wade-Benzoni, K. A., Tost, L. P., Hernandez, M., and Larrick, R. P. It's only a matter of time: Death, legacies, and intergenerational decisions. *Psychological Science*, 23(7):704–709, 2012.
- Winking, J. and Mizer, N. Natural-field dictator game shows no altruistic giving. *Evolution and Human Behavior*, 34(4):288–293, 2013.



Stochastic Optimization of Bioreactor Control Policies Using a Markov Decision Process Model

Quirin Stockinger

Technische Universität München

Abstract

Biopharmaceuticals are the fastest-growing segment of the pharmaceutical industry. Their manufacture is complicated by the uncertainty exhibited therein. Scholars have studied the planning and operation of such production systems under some uncertainties, but the simultaneous consideration of fermentation and resin yield uncertainty is lacking so far. To study the optimal operation of biopharmaceutical production and purification systems under these uncertainties, a stochastic, dynamic approach is necessary. This thesis provides such a model by extending an existing discrete state-space, infinite horizon Markov decision process model of upstream fermentation.

Tissue Plasminogen Activator fermentation and chromatography was implemented. This example was used to discuss the optimal policy for operating different fermentation setups. The average per-cycle operating profit of a serial setup was 1,272 \$; the parallel setup produced negative average rewards. Managerial insights were derived from a comparison to a basic, titer maximizing policy and process sensitivities. In conclusion, the integrated stochastic optimization of biopharma production and purification control aids decision making. However, the model assumptions pose room for further studies.

Keywords: Markov decision process; biopharmaceuticals production; fermentation uncertainty; chromatography resin; stochastic performance decay.

1. Introduction

There are currently 316 active biopharmaceutical ingredients available on the market (Walsh, 2018). Within the pharmaceutical industry, so-called biopharmaceuticals represent the fastest-growing segment, generating more than 160 bn€ in annual revenue (Otto et al., 2014). These biopharmaceuticals are not the result of chemical synthesizing, but their manufacture involves the fermentation of bacterial or mammalian cells. To capture the produced proteins, the fermentation is followed by a process of purification. However, the control of biopharmaceutical production and purification systems is highly complex because of the inherent technical uncertainties (Farid et al., 2005). This contribution studies the optimal, simultaneous control of biopharmaceutical fermentation and purification processes.

Biopharma production, on a high level, is a two-stage process: during the upstream process (USP), a cell culture is grown in sequential, volume-increasing bioreactor media wherein the fermentation environment is most commonly controlled by way of feeding substrates such as nutrients into the reactors (fed-batch fermentation) and changing the tem-

perature, pressure, and pH-value. The culture's protein production is induced in the final bioreactor of the sequence, i.e., the "production reactor," by changing the nutrient concentration in the medium accordingly. When enough product has been fermented, the protein of interest must be separated from the medium during the downstream process (DSP). The most common technique of purification is chromatography (Liu et al., 2014), but it is also the "most expensive part of the downstream process" (Nweke et al., 2018, p. 992). Because chromatography resins' capacity to bind proteins is uncertain (Farid et al., 2005) and deteriorates over time (Jiang et al., 2009), the exchange of spent chromatography resin constitutes operational complexity in the DSP. While there may be additional filtration steps and intermediate storage steps, the upstream and downstream processes are highly interdependent, and their simultaneous control poses significant room for research.

The need for optimization of biopharmaceutical production is an ongoing topic of discussion in practice. A recent report by consultancy McKinsey & Company, for example, cites "finding ways to improve the performance of the production

process” as one of the critical operational considerations that biopharmaceutical manufacturers must ponder to succeed in the face of adverse market realities such as changing payor behavior and competition from biosimilars (Otto et al., 2014, p. 6). Even though small scheduling errors can already lead to huge financial downside from lost batches due to the high value of proteins (Schmidt, 1996), operations research techniques are still only sparsely adopted in the industry (Martagan et al., 2018).

Nevertheless, academic interest in the study of controlling biopharmaceutical production has been highly active. The planning of up- and downstream capacities and processes, for example, has been discussed in literature, including chromatography column sizing (Allmendinger et al., 2014), maintenance of spent resins (Liu et al., 2014), production under chromatography yield (Liu et al., 2016) and product titer uncertainties (Lakhdar et al., 2006), and purification capacities (Siganporia et al., 2014). The control of upstream fermentation has largely focused on maximizing product concentration in a single bioreactor (Banga et al., 1997; Pandian and Noel, 2018; Peroni et al., 2005; Rocha et al., 2014; Saucedo and Karim, 1997). While these endeavors model cell-level kinetics, they often do not consider the financial trade-offs of system-level decision making, i.e., integrated decisions across all involved process steps. However, as early as 1996 (Schmidt, 1996), and more recently (Martagan et al., Accepted/In press; Martagan et al., 2016; Martagan et al., 2018), the literature on the system-level control of production and purification equipment has aimed at filling this void. Some scholars argue that purely maximizing the protein concentration during fermentation may not yield economically optimal results when considering the associated operating and purification costs (Martagan et al., 2018).

While Martagan et al. (Accepted/In press) studied simultaneous system-level control of protein production and purification processes, they abstracted upstream decision making to a single decision of how much protein to produce. Furthermore, they did not consider the issue of chromatography resin performance decay, as introduced by Liu et al. (2014). Nevertheless, its consideration is relevant for practice because resin material is a major driver of downstream operating costs (Farid, 2007). To the best of my knowledge, no existing paper has simultaneously considered the biopharmaceutical fermentation and optimal resin exchange schedule sub-problems under uncertain fermentation and stochastic performance decay. Existing literature, therefore, doesn't conclusively answer some outstanding questions about the influence of the aforementioned uncertainties. What is the optimal, simultaneous control policy for the USP and DSP? When is chromatography resin exchanged under stochastic decay? When under different minimum allowed resin capacities and different resin costs? How does the consideration of two parallel production reactors change the optimal policy?

This work provides a dynamic stochastic model spanning both upstream and downstream operations which is used to answer these questions. Using this dynamic program, this

contribution analyzes a practice-representative production and purification process using the example of the recombinant protein Tissue Plasminogen Activator (TPA), a prominent product of the early biopharma industry (Datar et al., 1993) which is still relevant today (Johnston, 2010). On a system-level, decisions about the production of TPA and its purification must be made simultaneously. During the USP decisions about how long to grow the culture before converting it into its protein-producing state and when to harvest the TPA from the medium are considered. A linear accumulation of TPA in the production medium and possible batch failure due to contamination are assumed. The states of the observed production reactors are assumed to represent physiological states during the culture's lifecycle, e.g., growth, production, and decay, but cell-level kinetics are not modeled. Because there is no intermediate storage, and both process steps are highly interdependent, harvesting of TPA constitutes two necessarily simultaneous decisions: to harvest the production medium and to accept the medium into the first chromatography step. For the purification of a single batch, five purification cycles in the first chromatography column are assumed to take place within one decision epoch. After each purified batch, the performance of the resin, i.e., what fraction of the TPA in the medium it can bind, deteriorates stochastically. This gives rise to the need for maintenance actions related to the exchange of spent resin. Maintenance activities are assumed to take one decision epoch due to their short duration. Because the first chromatography column can be regarded as the bottleneck of a multi-step chromatography process, only the first chromatography step is considered in the presented model. Analogously, only the production reactor of a seed-train is considered as its bottleneck. To test hypotheses about the parallelization of production, the provided model is extended by a second, parallel production reactor in the same seed-train and scenarios are analyzed. Sets of states for which the same control actions are optimal are discussed because of the approach's demonstrated value in previous research (Martagan et al., 2018).

By building on the theoretical foundations of Schmidt (1996), this work contributes a framework for the system-level study of simultaneous decision making in the USP and DSP under uncertainty. Contrary to existing models, it allows for the study of the optimal operation of parallel production reactors and chromatography resin maintenance. Furthermore, it explicitly models protein accumulation during production. This thesis also contributes optimal control policies for the production and purification of TPA. By studying resin exchange policies under varying process conditions, this thesis builds on Liu et al. (2014) understanding of what influences the carrying out of costly maintenance activities. Furthermore, this thesis argues for the business case of the stochastic optimization of integrated production and purification control compared to simple upstream titer maximization. The derived decision spaces aid managerial and operational decision making under the uncertain environment of biopharmaceutical production.

The remainder of this thesis is organized as follows:

Chapter 2 reviews the problem context and surveys existing literature on biopharmaceutical production. Chapter 3 details the problem characteristics and this work's research questions. Chapter 4 outlines Markov decision processes (MDPs) as the solution approach and introduces the case study of TPA production. Chapter 5 describes the developed model and Chapter 6 the results of the numerical case study. Chapter 7 concludes this work and provides an outlook for future research endeavors.

2. Review of Literature and Research

To the best of my knowledge, no existing paper studies the simultaneous, system-level control of upstream fed-batch fermentation and downstream resin exchange. In this chapter, existing literature and research on the optimization of biopharmaceutical production schedules in the light of uncertainty are reviewed. First, an introduction to the biopharmaceutical production process and their academic and economic relevance are provided and the literature review methodology summarized (Section 2.1). Next, the control of fed-batch fermentation processes based on cell-level kinetics is reviewed (Section 2.2). This is extended by a survey of the issue of planning biopharma production capacities and activities, with a focus on the inherent stochasticity of the process (Section 2.3). In Section 2.4, the middle ground between the two prior abstraction levels, optimal control policies of integrated production systems under biological uncertainties, is reviewed. Concluding this literature overview, the concrete research gaps in existing research are identified and this thesis is motivated (Section 2.5).

2.1. Context of literature research

This section provides a brief overview of the biopharmaceutical market and defines common production processes. Additionally, the methodology of the literature review is described.

2.1.1. Biopharmaceutical production

Since the first biopharmaceutical drug being commercialized in 1982, the market has continuously grown (Figure 1). Between 2014 and 2018 alone, 129 distinct biopharmaceuticals have been commercialized across the United States and European Union. Taking the 58 withdrawn active ingredients into account, currently, 316 biopharmaceutical active ingredients are available (Walsh, 2018).

Regardless of the topic's academic and commercial relevance, the lack of a uniform definition of terms has been lamented (Rader, 2005). Within the pharmaceutical industry, two terms need to be distinctly defined to allow for a concise discussion of the topic. Ordinary "drugs" are manufactured by "chemical (non-biological) means and involving small molecules", whereas "biopharmaceuticals" are "manufactured by biotechnology methods and involving complex biological molecules" (Rader, 2005). Biopharmaceuticals may be "produced from cultures of eukaryotic or prokaryotic

cells, isolated from natural sources, or made by synthetic methods" (Jagschies et al., 2018, p. 59).

The production process of biopharmaceuticals, on a high level, is separated into two sequential, highly interdependent phases: an upstream process (USP) and a downstream process (DSP). During the USP, the active pharmaceutical ingredient (API) is synthesized by cultivating the living cells under controlled nutrient conditions, i.e., fermentation. Upon reaching the required API quantity, recovery and filtration of the API must take place in the DSP before packaging and shipping of final products (Jagschies et al., 2018, p. 76). This is necessary because of impurities produced alongside the API during fermentation. The purification can directly follow the fermentation process or be postponed (Siganporia et al., 2014). One of the most common yet highly cost-intensive process steps during product recovery is chromatography (Jiang et al., 2009; Liu et al., 2014). During chromatography, the product of interest is separated from the medium and impurities based on their physiological differences. For this task, so-called chromatography columns hold resins which either bind the product of interest or the impurities (Martagan et al., Accepted/In press). Additional steps between USP and DSP, like intermediate storage and filtration, are possible. However, these are not in-scope for this contribution.

Each phase may consist of a series of bioreactors and chromatography steps, respectively. During the USP, bioreactors of increasing volumes may be used to grow an initially small amount of cell culture to commercial production scale. Such a "train" of sequentially interconnected bioreactors may be referred to as a "seed train" (Jagschies et al., 2018, p. 632). During the DSP, between two and six sequential chromatography steps may be required to meet the purity demands of the desired product (Martagan et al., 2018). A schematic of a biopharmaceutical production process using a single seed-train of three bioreactors of increasing volumes and three sequential chromatography steps is provided in Figure 2. To improve the utilization of the purification equipment, different process set-ups are possible. Appendix 1, for example, shows two serial upstream seed-trains feeding into one downstream process. It may, however, be more capacity- and cost-feasible to inoculate multiple parallel production reactors from one seed-train (Jagschies et al., 2018, 653f), visualized in Appendix 2.

Bioreactors can be operated in two modes: batch processing and continuous processing. Batch processing allows distinct production periods and lot sizes while continuous production allows continuous harvesting (Siganporia et al., 2014). Within the batch process category, two prominent operating modes exist: fermenters-batch and fed-batch. Of these, fed-batch fermentation has been the most popular mode historically. Under this production paradigm, batches of high-value products such as APIs are fermented in bioreactors under a controlled environment while being continuously fed nutrients (Banga et al., 1997). Contrasting continuous production, the product is harvested at the end of the fed-batch production (Siganporia et al., 2014)(Siganporia

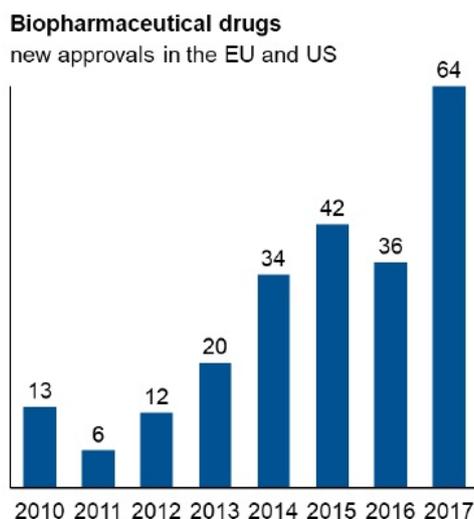


Figure 1: The number of newly approved biopharmaceutical drugs, including those with identical ingredients, in the EU and the US shows a positive, historical growth trend (Walsh, 2014; Walsh, 2018)

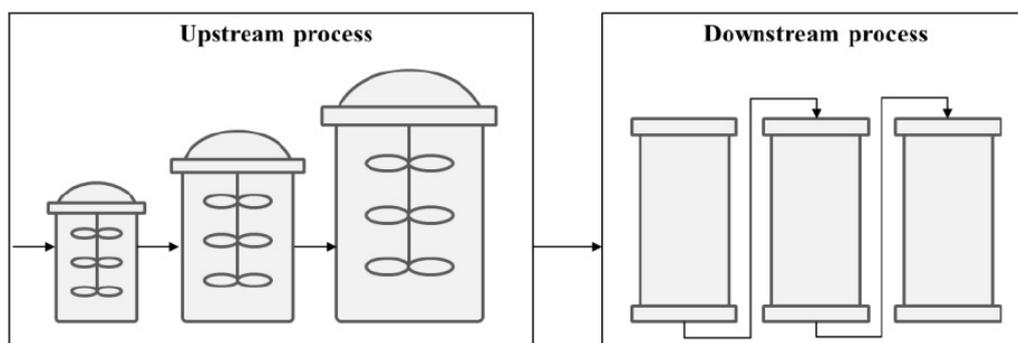


Figure 2: Schematic view of a biopharmaceutical production process with a singular serial seed-train and three sequential chromatography columns

et al., 2014). While continuous fermentation processes are regarded as more economical in some cases, the pharmaceutical, biotechnology, and food industries predominantly use batch processes. Especially for the production of Monoclonal Antibodies (MABs), one of the fastest growing product category in biopharma (Liu et al., 2016), fed-batch has been established as the most prolific production paradigm (Sigantoria et al., 2014). Its popularity is mainly due to its lower process complexity, recent increases in production titer sizes, avoidance of over-feeding and high levels of sterility (Freitas et al., 2017; Rani and Rao, 1999; Sigantoria et al., 2014).

2.1.2. Methodology of the literature survey

Determining optimal control of fed-batch fermentation has been the subject of academic research for at least twenty years (Lee et al., 1999; Rani and Rao, 1999). For the purpose of this contribution's merit, only literature published in or accepted to be published in peer-reviewed journals and conference proceedings with at least 2nd quartile Scimago or A Jourqual3 rankings were reviewed. Primary sources

for collecting relevant literature were references in seminal papers (such as Liu et al. (2014); Martagan et al. (2016); and Schmidt (1996)), Web of Science, EBSCO Host Business Source Complete, and Google Scholar. Databases were searched based on keywords from the most relevant papers. Based on the context described above (Chapter 1), existing literature was categorized along the following five characteristics:

- **Planning level:** The level of abstraction with which the biopharmaceutical production process is studied.
- **Decision space:** How the locus of decision-making is defined, i.e., what the set of necessary decisions is.
- **Uncertainties:** The process-specific uncertainties and their implications which are considered.
- **Solution method:** The mathematical approach to modeling and, finally, solving the identified problem.

- Financial trade-offs: Whether the financial value of decision making along the production process is considered.

These criteria allow the precise discussion of the existing academic literature and, later on, the proposed approach of this work. Among existing literature, three abstraction levels were identified on which fed-batch biopharmaceutical production is generally studied: (1) biomass kinetics within one bioreactor, (2) scheduling and capacity planning on facility-level, and (3) operating an integrated system of bioreactors and chromatography steps. Following this motivation of the biopharmaceutical production control problem and its study, the next three sections will review the existing literature along the three introduced planning levels: Bioreactor control (Section 2.2), Production and capacity planning (Section 2.3), Production system operations (Section 2.4). Across these levels, problem characteristics will be discussed and, finally, a worthwhile gap in the academic conversation will be concluded.

2.2. Biorreactor control

A large amount of uncertainty during the production of biopharmaceutical products stems from the unpredictability of living organisms' behavior. Cell growth, production rate, and contamination are highly non-linear in nature (Pandian and Noel, 2018). It is therefore paramount to control bioreactors effectively by taking the optimal decision in each situation, i.e., adhering to an optimal policy. In this section, existing literature on fed-batch bioreactor modeling and determination of control policies is surveyed. The non-linear nature of biopharmaceutical production processes rules out conservative linear models (Pandian and Noel, 2018). Control models used for this task are consequently based on Dynamic Programming (DP) (Rocha et al., 2014). DP, as defined by Bellman (1957b), models a dynamic problem as a set of consecutive transitions from state to state. Therefore, bioreactor control researchers have described fed-batch fermentation processes predominantly as control problems, e.g., open loop control (Rani and Rao, 1999). These control problems are often constrained by differential equations modeling the physiological kinetics of the biomass within a reactor (Banga et al., 1997). Models of biomass kinetics can be either structured, i.e., "explicitly describe intracellular processes," or unstructured, i.e., rely on "concentrations of nutrients and metabolites" (Xing et al., 2010, p. 208). Unstructured models often incorporate Monod-type bacterial growth models (Xing et al., 2010). Under the lens of process control, the most common control objective is to find the nutrient feed rate policy which maximizes product concentration at the terminal time or overall reactor productivity. Consequentially, the decision space is often limited to the feed rate of nutrients over the course of fermentation making it a singular control problem. Banga et al. (1997) and Banga et al. (2005), for example, formulate open-loop control problems, modeling the kinetics of one fed-batch fermenter with the goal to maximize the

yield of penicillin and ethanol, respectively, by way of controlling the substrate feed-rate. Ponte et al. (2018) describe fed-batch production of the recombinant fungus ROL and obtain optimal feeding trajectories. Skolpap et al. (2004) and Skolpap et al. (2008) model the control of the Monod-type kinetics of α -amylase in a fed-batch fermenter, and include the switching time from batch-mode to fed-batch mode as a decision variable. Peroni et al. (2005) propose an approximate dynamic program to determine the feed-rate profile and the fermentation end time in invertase production of *Saccharomyces Cerevisiae*.

Due to the computational challenge of determining the control parameters in higher order differential equations, stochastic optimization routines have gained popularity for solving these control problems. While Xing et al. (2010); Skolpap et al. (2004); and Skolpap et al. (2008) use Markov chain Monte Carlo simulation, Banga et al. (2005) implement control vector parameterization, stochastic approaches, such as random search, genetic algorithms, and differential evolution, as well as dynamic hill climbing, to estimate kinetic model parameters. Other stochastic heuristics include model predictive control using evolutionary computation (Ashoori et al., 2009; Freitas et al., 2017), particle swarm optimization (Liu et al., 2013), and, more recently, reinforcement learning (Pandian and Noel, 2018). Peroni et al. (2005) use a neural network for the implementation of an initial approximation and iterative improvements thereof. Earlier, a feedforward neural network was used by Chaudhuri and Modak (1998) to model the same problem.

In general, most surveyed models are deterministic and don't consider process uncertainties such as variable production rates and spontaneous cell death due to contamination. Delvigne et al. (2006), however, propose a Markov chain model of the concentration gradients to complement a deterministic kinetic model. Attempts to include stochastic cell behavior in dynamic control problems have also been made, for example, by Saucedo and Karim (1998); Saucedo and Karim (1997), who propose an MDP for modeling the concentration of ethanol produced by way of an optimal feed-rate policy.

Process engineering works towards the optimal utilization of existing process capabilities (Rocha et al., 2014). Therefore, maximum productivity of the reactor, rather than economically-optimal operation is a common objective (Peroni et al., 2005). However, some scholars even at the bioreactor control level, tie optimization to a financial tradeoff between the value of a maximized product concentration and the cost of nutrients fed into the reactor (Ponte et al., 2018; Saucedo and Karim, 1998).

Process control of batch and fed-batch biopharmaceutical production to achieve maximum productivity has been an area of active academic discussion. However, maximizing production titers during the upstream process is only one way of optimizing production. Biopharmaceutical production, as discussed above, is an interconnected system of upstream fermentation and downstream purification. Given the high costs of all involved process steps, maximizing protein production in the USP may not be an optimal policy for an economically

incentivized decision maker. The next section deals with capacity and production planning literature on the facility-level which allows the consideration of downstream processes in decision making.

2.3. Production and capacity planning

The novelty and complexity of the biopharmaceutical industry, combined with the incumbent focus on spreadsheet-based planning, led to unrealized savings from more sophisticated planning models (Lakhdar and Papageorgiou, 2008). Recently, mathematical optimization of production and capacity planning has been subject to rigorous academic research spanning both up- and downstream processes (Liu et al., 2016). This section reviews past research on this abstraction level that deals with biopharmaceutical production, it is summarized in Table 1.

Due to the high level of abstraction in medium-term planning, earlier work on biopharmaceutical production planning often considered the production process as a black box with decisions focusing on production schedules irrespective of downstream operations. Gatica et al. (2003) and Lakhdar et al. (2006), for example, determined production quantities and campaign durations for a not closer specified biopharma production process. More recently, decision making has expanded to include purification operations as well. Sigantoria et al. (2014) modeled capacity decisions in sequential upstream production and downstream purification. Liu et al. (2014), building on Lakhdar et al. (2005), studied production and maintenance planning covering both the USP and DSP, with a special focus on the issue of downstream maintenance work, i.e., when decayed chromatography resins are scheduled to be replaced. Therein, chromatography resins deteriorate after each purified batch, leading to the resin binding less of the available product from the medium in following batches. To restore the performance of the chromatography process, maintenance activities related to the exchange of the used resin are necessary. Liu et al. (2016) expanded their prior work to include sizing and sequencing decisions in both process steps for MAB production. Focusing on a DSP decision space, chromatography capacities and their operations were studied by Allmendinger et al. (2014). Furthermore, Allmendinger et al. (2014); Liu et al. (2014), and Liu et al. (2016) simulated scenarios with parallel production setups being harvested into one or more purification suites.

Due to the inherently uncertain nature of biopharmaceutical production, a stream of literature has specifically focused on dealing with this. Gatica et al. (2003), for example, modeled clinical trial success scenarios in their Mixed Integer Linear Program (MILP) determining capacity plans. Lakhdar et al. (2006) and Lakhdar and Papageorgiou (2008) expanded previous work on multi-period planning and scheduling (Lakhdar et al., 2005) by considering uncertain production titers. Liu et al. (2016), in a modification of their previously deterministic MILP (Liu et al., 2014), considered uncertainties in both the production titer and the resin

yield during chromatography purification by means of triangularly distributed stochastic parameters when determining production plans. The impact of uncertain production titer on chromatography decisions was also considered by Allmendinger et al. (2014).

For planning problems under uncertainty, literature has mostly focused on stochastic programming approaches such as 2-stage programming (Lakhdar and Papageorgiou, 2008), Chance-Constrained Programming (CCP) (Lakhdar et al., 2006, Liu et al., 2016), and scenario-based programs (Gatica et al., 2003). Allmendinger et al. (2014) formulated the closed-loop control of a process economic model which was optimized using evolutionary algorithms. Deterministic planning problems are generally modeled as MILPs and solved using standard approaches (Lakhdar et al., 2005; Liu et al., 2014; Sigantoria et al., 2014).

As optimal decision making regarding the use of financial and capital assets is one of the principal goals in production and capacity planning, it's no surprise that this is also true in biopharmaceuticals planning. In this stream of planning research, the financial trade-offs consider (1) maximizing sales revenue while incurring minimum capacity investments and operating costs or (2) minimizing costs per sold product. Gatica et al. (2003) and Sigantoria et al. (2014), for example, studied a capacity planning problem's trade-off between capacity investments and operating costs and revenues generated from selling the produced amount. Lakhdar et al. (2005); Lakhdar et al. (2006), and Lakhdar and Papageorgiou (2008) studied the classic production planning conundrum of satisfying demand under operating profit maximization. Liu et al. (2014), on the other hand, studied the financial trade-offs of production and maintenance simultaneously. They specifically focus on the trade-off between reduced purification yield from reusing the same resin and incurring costly resin maintenance. Liu et al. (2016) focused on the trade-off between sales revenues and operating costs under consideration of uncertain titers and purification yields, minimizing the total costs of goods. In the latter category of financial trade-offs, Allmendinger et al. (2014) aimed at minimizing the cost of goods per gram of sold product.

In conclusion, biopharmaceutical production planning literature seems to be playing catch-up regarding consideration of operational uncertainty. Due to the high economic impact of resin maintenance, performance decay and the related scheduling of resin exchanges constitute a further need for academic attention. Furthermore, a priori planning and scheduling policies, are not able to guide ad-hoc operations of systems of bioreactors and chromatography columns especially under the influence of biopharmaceutical process uncertainties. The next two sections focus on existing approaches which aim at closing this gap and conclude with the still remaining gap as the focus of this work.

2.4. Production system operations

Recently, academia's widespread focus on concentration maximization within singular bioreactors and its disregard of uncertainties and downstream process implications have

Table 1: Biopharma planning literature

Reference	Decision space	Uncertainties	Model
Lakhdar et al. (2005)	Campaign sequence and duration, production quantities	n/a	MILP
Siganporia et al. (2014)	Capacity plans, out-sourcing decisions	n/a	MILP
Liu et al. (2014)	Maintenance and production plans	n/a	MILP
Gatica et al. (2003)	Product portfolio, capacity planning	Clinical trial out-come	Scenario-based MILP
Lakhdar et al. (2006)	Production plans	Production titer	CCP MILP
Lakhdar and Papageorgiou (2008)	Production plans, sales and backlog profiles	Production titer	2-Stage MILP
Allmendinger et al. (2014)	Capacity plans, chromatography operations	Production titer	Closed-loop control
Liu et al. (2016)	Sequencing and sizing, production plans	Production titer, resin yield	CCP MILP

been questioned by management scholars, such as Martagan et al. (2016). They argue for the need for harvesting policies maximizing discounted financial profit under the consideration of system-level operations. This section surveys existing literature on system-level approaches to the optimization of biopharmaceuticals production.

The decision space in system-level biopharmaceutical production optimization has expanded. Originally, decision making was exclusively concerned with the operations of an upstream seed-train under a known average downstream protein yield (Schmidt, 1996). Schmidt studied the optimal scale-up and harvest policy on a seed-train consisting of sequential, in volume increasing bioreactors. Therefore, the model allowed actions involved in the inoculation of a prepared bioreactor, the feeding of nutrients to grow the culture, the transfer of the medium from a smaller reactor to the next larger one, and the facilitation of protein production and harvesting. 20 years later, Martagan et al. (2016) researched upstream harvesting decisions under explicit purification dynamics. In their model, a decision maker chooses between continuing fermentation and harvesting the protein produced thus far. Later, they studied decisions about chromatography pooling windows under explicit knowledge about upstream product yield (Martagan et al., 2018). Therein, they studied optimal chromatography pooling. Their model aimed at determining which chromatography “lanes”, i.e., the amount of protein and impurity flowing through the column per time unit, to “pool”, i.e., which lanes to capture from the chromatographic separation, under the consideration of the trade-off between impurity levels and product yield. Most recently, Martagan et al. (Accepted/In press) integrated the simultaneous determination of the upstream production quantity and downstream technology and pooling window. Herein, they modeled a decision maker’s interdependent choice about which chromatography technique to use at a given time, which chromatography lanes

to pool, and how much protein to ferment. It is interesting to note the level of abstraction to which decisions in the different process steps are studied. While Schmidt (1996) modeled detailed production process operations, excluding downstream decision making, Martagan et al. (2016) only modeled two different actions at each time interval (continue or harvest) and Martagan et al. (Accepted/In press) reduced upstream decisions to the selection of what amount of protein to produce.

Due to the stochastic nature of the dynamic control biopharmaceutical manufacturing, literature at the system abstraction level has exclusively modeled decision making using MDPs. However, the considered uncertainties differ. Schmidt (1996) accounted for the stochasticity of protein production through living organisms. Continuation of culture growth and protein production was assumed not deterministic but subject to random upsets such as contamination. Martagan et al. (2016) considered the uncertainty in the accumulation of MABs and impurities, as well as the arrival of random shocks, such as sudden failure or increased impurity accumulation, during production. Martagan et al. (2018) and Martagan et al. (Accepted/In press) studied the uncertainty of how much protein and impurity remain in the medium at the beginning of each chromatography step. Each piece of literature on system-level decision making in biopharmaceutical production considers the financial trade-off of the decision maker’s actions. Schmidt (1996) model, for example, was aimed at aiding process change decisions. Therefore, he modeled the financial impact of harvesting and continuing fermentation. On the one hand, operating revenues were based on an average protein concentration per harvested liter of medium and an average purification yield. On the other hand, each undertaken action incurred a volume-dependent cost for the fed nutrients and a fixed cost. By studying average operating profits, Schmidt (1996) was able to deduct the financial impact of process parameter

changes, such as a reduction of the contamination risk or an increase of the productive phase of the culture, for example, due to investment in process improvements. Similarly, Martagan et al. (2016) considered the financial trade-off between incurring operating costs for continuing fermentation to increase the potential payoff from selling the accumulated MABs and realizing the potential profits by harvesting before batch failure occurs. Martagan et al. (2018) studied protein yield and product purity. Inherent therein is the financial trade-off between the fixed revenue earned from satisfying customer orders and the costs incurred at each chromatography step. In their model, fixed revenue is earned if at least as much product as the customer's yield requirement is delivered, otherwise a yield penalty cost is incurred, given that purity requirements are met. In their most recent work, Martagan et al. (Accepted/In press) integrate both considerations (costs of continued fermentation and profits associated with chosen chromatography operations) into a simultaneously considered financial trade-off.

Research on the system-operations abstraction level and on all other discussed levels is summarized in the following section. It is also there, where the gap is identified which this contribution aims at closing.

2.5. Concluding remarks and identified research gap

After an introduction to the biopharmaceutical production industry in Section 2.1, Section 2.2 reviewed control theory literature on fermentation reactor control constrained by culture kinetics. Existing research on this topic has often focused on the maximization of fermentation productivity with little regard for the financial trade-offs between costly fermentation and associated purification operations. Section 2.3 introduced planning level considerations and the notion of chromatography resin performance decay in the work of Liu et al. (2014). This linear, deterministic planning model, among other decisions, determined maintenance schedules for the replacement of spent chromatography resins. While Liu et al. (2016) model did not consider resin decay, it modeled the uncertainty of the resin's yield. Uncertain, decaying resin yield, however, has not yet been considered.

Section 2.4 reviewed existing literature on the control of fermentation and purification systems. Although research has explored the control problem of interconnected fermentation and purification systems (Martagan et al., Accepted/In press), a paucity of further study of the problem appears to exist. Existing academic work on the topic seems to either overly simplify fermentation operations (Martagan et al., Accepted/In press; Martagan et al., 2016) or exclude control of purification processes (Schmidt, 1996). Furthermore, when to incur chromatography raw material costs, i.e., when to exchange spent resin, has not yet been considered on the system abstraction-level.

Due to the interdependence of up- and downstream decision making in the complex production of highly valuable products, optimal policies for simultaneous operation of protein production and chromatography maintenance under fermentation and resin decay uncertainties constitute a worth-

while endeavor for academia and practice. This contribution, therefore, aims at closing this gap by providing optimal decision policies for the control of an integrated biopharmaceutical production system. The next chapter summarizes the concrete characteristics of this identified problem and introduces the research questions of this contribution.

3. Problem Statement & Research Questions

Following a summary of the setting of the problem studied in this contribution (Section 3.1), this chapter introduces the research questions this contribution aims at answering (Section 3.2).

3.1. Problem context

As per the identified gap in biopharmaceutical production control literature, this contribution aims at the simultaneous optimization of two interconnected sub-problems, i.e., the upstream fermentation policy and the downstream resin exchange schedule. Building on existing literature (see Sections 2.2-2.4), the following problem setting is considered:

Uncertainties: Both, product fermentation and purification are stochastic in nature. These uncertainties are hypothesized to play a critical role in the optimal control of a biopharmaceutical production system. Due to the inherently non-linear nature of the living organisms used in biopharmaceutical production, the physiological states of the observed fermentation culture are highly uncertain. Although fed-batch fermentation is argued to be highly sterile, a risk of batch failure, e.g., due to contamination, persists and is hypothesized to influence optimal decision making (Martagan et al., 2016; Schmidt, 1996). As the performance decay of chromatography resins has thus far only been assumed to be deterministic (Liu et al., 2014) but resin yields constitute a relevant uncertainty (Farid et al., 2005), maintenance decisions under stochastic performance decay are studied.

Planning level: The identified problem is examined on a system-level. The optimal control of the production reactor(s) depends on the state of the fermentation process as observed by a hypothetical production operator. While these states are informed by culture physiologies (Martagan et al., 2016; Schmidt, 1996); explicit culture kinetics are not in-scope for this work. Such approximations are considered "sufficient to describe the evolution of the [fermentation] system through time" (Schmidt, 1996, p. 607). Analogously, the operation of the product purification, while depending on the physiological details of the chromatography technique, are not considered herein.

Decision space: The decision space of interest is the simultaneous control of the highly coupled USP and DSP. The control of the production reactors in a predetermined upstream seed-train configuration to produce a predetermined protein, including decisions regarding the preparation of the reactors, the culture's growth, continuation of production or harvesting the accumulated product are studied in this contribution as these realistically constitute the object of daily

decision making for production operators. This contribution considers only the production reactor of a predetermined USP. For one, it's mainly interested in the production phase and harvesting decisions, and these take place in the production reactor, rather than during the volume-scale up from inoculum flasks to the production volume across the preceding seed-train. As a second argument for this decision, the production reactors of a seed-train can be considered as the bottleneck of that seed-train. For multiple parallel production reactors, the USP decision space consequently spans all bottleneck reactors. The parallel production reactors are considered the bottleneck of parallel production because their operation limits the batch throughput; furthermore, parallel production reactors inoculated by a single seed-train are considered more economical than entire parallel seed-trains (Jagschies et al., 2018, p. 653). The control of the purification process, on the other hand, focuses on the first chromatography step for simplicity, in line with previous work (Liu et al., 2014). Consequently, policies for the first chromatography column with a predetermined resin and capacity must be determined. Decisions include accepting the harvest from the upstream process and exchanging used resin to restore chromatography performance.

Financial trade-offs: The financial trade-off of both sub-problems are integrated and considered simultaneously. During the USP, the financial trade-off between continuing fermentation by incurring operational costs and "locking in" value by harvesting before the whole batch is lost due to sudden failure from contamination has been identified as the most relevant (Martagan et al., 2016; Schmidt, 1996). Downstream decisions about resin maintenance are made in the face of the financial trade-off between scheduling the costly chromatography resin exchange and postponing the exchange, thus accepting a lower financial payoff from the next harvest. On the system-level, a controller, therefore, must integrate the decisions of both sub-problems. This integration considers the financial trade-off of when to "lock in" the revenue from already produced protein before losing the whole batch and the implied cost of reducing the resin performance by purifying a harvested batch. On the one hand, accepting a harvest for purification secures the revenue from the fermented protein but, on the other hand, also accelerates the necessity of costly resin exchange. Meaning, whenever a batch is purified, the need for exchanging the resin is moved closer to the present.

These problem characteristics, compared to existing literature, are summarized in Appendix 3. A formal statement of the problem is provided in Appendix 4.

3.2. Research questions

This contribution poses several questions which existing literature has yet to address but which could contribute to academia's and practice's understanding of this complex environment. This section derives these questions from the identified gaps.

Schmidt (1996) modeled the simultaneous control of sequential upstream bioreactors and Martagan et al. (Ac-

cepted/In press) considered both up- and downstream operations but reduced protein production to a single decision. Because of the benefit of simultaneously optimizing up- and downstream operations demonstrated by Martagan et al. (Accepted/In press), there exists paucity to answer the following research question.

RQ1: What is the optimal, simultaneous control policy for the USP and DSP?

Furthermore, while Liu et al. (2014) determined optimal resin maintenance schedules, only deterministic decay was considered, even though resin yields constitute a source of uncertainty (Allmendinger et al., 2014; Farid et al., 2005; Liu et al., 2016). However, maintenance decisions under uncertain remaining resin capacities, haven't been considered. They can constitute a relevant endeavor for practical decision making because of the high financial impact of chromatography material (Farid, 2007).

RQ2: When is the chromatography resin exchanged under stochastic decay?

RQ3: What is the influence of changes to process parameters such as the resin cost and the minimum viable capacity on the optimal policy?

Lastly, because parallel fermentation can increase the batch throughput of a biopharmaceutical production process, seed-trains may be set up to inoculate multiple parallel production reactors to increase utilization of purification equipment (Jagschies et al., 2018, p. 652). While Allmendinger et al. (2014) and Liu et al. (2016) modeled different ratio setups between upstream seed-trains and downstream purification, the operations-level implications of multiple parallel production reactors have yet to be studied, giving the following research question interest and relevance.

RQ4: How does the consideration of two parallel production reactors change the optimal policy from RQ1?

The related, non-directional hypotheses are summarized in Appendix 5. How this contribution aims at answering these questions is described in the following chapter (Chapter 3). Afterward, the respective analyses are carried out (Chapter 4).

4. Solution Approach

In this thesis, a stochastic dynamic programming approach is developed which allows the study of simultaneous decision making regarding the fed-batch fermentation process, the timing of harvesting, and the maintenance of chromatography resin.

This chapter argues for the use of stochastic dynamic programming over purely stochastic approaches to study the bioreactor system control problem (Section 4.1), reviews the theoretical background of Markov decision processes (Section 4.2), and introduces the numerical case study used to study the identified problem (Section 4.3).

4.1. Dynamic vs. stochastic programming

After reviewing the problem characteristics in the previous chapter (Section 3.1), the curious reader may ponder the question: “under which mathematical framework could one study such a complex environment?” This section argues for MDP as the most fitting modeling technique in this context.

Before discussing possible approaches, the model requirements of the proposed context are reviewed. From the problem context, one can derive the following three requirements of the model approach:

- **Regular observation:** The model must allow regular observation of a complex system under uncertainties. A hypothetical production system controller must be able to observe the production and purification processes and the cell cultures, periodically.
- **Decisions based on state:** To study optimal decision making in the control of biopharmaceutical production and purification systems, the model must allow a hypothetical system controller to form a decision based on their observed information. In other words, one observation must provide enough information for the choice of the next decision.
- **Required optimality:** Due to the significant financial trade-off underlying decisions during biopharmaceutical production and purification, the model should provide an optimal prescription on what actions to take.

Two mathematical modeling techniques may constitute candidates for studying the research questions this contribution poses: stochastic programming and dynamic programming.

Stochastic programming deals with decision problems which can be formulated as follows: “Some decisions must be made today, but important information will not be available until after the decision is made” (King and Wallace, 2012, p. 2). In principle, the problem studied in this thesis can be formulated to fit this description. However, two arguments discourage the use of stochastic programming in this context. Firstly, as Powell (2014) notes, stochastic programs are practically restricted to two decisions instances, so-called two-stage approximations. In two-stage programs, a decision x_t in time t under uncertainty is made, followed by scenario information becoming known and a second decision instance in which decisions for all following epochs are made $x_{t+1} \cdots x_T$ (Powell, 2014). Due to the high dimensionality of scenario trees, which must capture all historical information up to the current decision epoch, stochastic programs are virtually only applicable for problems which can be approximated as two-stage problems. Because biopharmaceutical production systems are monitored and operated in discrete time intervals (Martagan et al., 2016) over sometimes week-long production campaigns (Schmidt, 1996), it seems evident that a two-stage approximation is not feasible. Secondly, as multi-stage stochastic programs (two-stage programs included) are approximations, Powell (2014, p.

111) further postulates that the optimal solution to such an approximation is “(with rare exceptions) not an optimal policy.” As this work is interested in finding an optimal policy to study its research questions, stochastic programming can be eliminated as a viable candidate for modeling the identified problem.

Dynamic programs, on the other hand, can be argued to fit the requirements of the presented problem context. Firstly, dynamic programs deal with sequential decision problems and, therefore, allow the study of multi-stage decision making. Due to their dynamic nature, these problems don’t approximate multi-stage problems using scenarios but allow modeling them stage-by-stage. Secondly, properly modeled dynamic programs are Markovian (Powell, 2014) and, therefore, aren’t history dependent. These first two characteristics allow the modeling of a decision problem, in which an observer sequentially takes in information about the controlled system and can inform a control decision solely based on the last available state description. Thirdly, for subclasses of dynamic programs, it can be shown that an optimal policy exists. For example, using the Banach fixed-point theorem, it can be shown that an optimal policy exists for the total discounted reward problem of an infinite horizon MDPs (Saucedo and Karim, 1997) and relative value iteration converges to the optimal value function of the average cost problem of an infinite horizon MDP (Gupta et al., 2015). Markov decision processes are dynamic programs with stochastic state transitions. Section 4.2 provides more detail on their theoretical background.

This section concludes that dynamic programming meets the requirements of the problem identified in this thesis and, furthermore, MDPs, as a special case of dynamic programs, provide optimal policies for the control of stochastic systems.

4.2. Markov decision processes

This section introduces the notion of Markov decision processes, stochastic dynamic programs which are commonly used to model the control of a complex system which evolves according to a controlled Markov process.

MDPs are tools for analyzing dynamic systems in which state transitions are stochastic and can be influenced by the actions taken by a controller. As such, MDPs combine characteristics of DP and Markov chains (Tijms, 2003, p. 233). The theoretical foundations of MDPs lie in Bellman (1957a) and Howard (1960) for the interested reader.

An MDP is a partly controlled Markov process, i.e., a process following the Markov property (cf. Markov (1954)), in which the transition probability to the next state is solely dependent on the process’s current state and the controller’s chosen action. Puterman (2014, p. 17–20) mathematically formulates such a process in terms of the 5-tuple $(\mathcal{T}, \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R})$:

- a set of N decision epochs \mathcal{T} , over which the system evolves. If N is a finite number, the MDP is referred to as a finite horizon MDP, otherwise as an infinite horizon MDP.

- a set of possible states \mathcal{S} , which the system may occupy at any epoch $t \in \mathcal{T}$
- a set of actions \mathcal{A} , which the controller of the system can choose. In any given state $s \in \mathcal{S}$, only a subset $\mathcal{A}_s \subseteq \mathcal{A}$ may be permissible. Therefore, $\mathcal{A} = \bigcup_{s \in \mathcal{S}} \mathcal{A}_s$
- a set of transition probabilities P of dimension $|\mathcal{S}| \times |\mathcal{S}| \times |\mathcal{A}|$, which determines the probabilities $p(s, s', a)$ of the system progressing to state s' after choosing action $a \in \mathcal{A}_s$ in state s
- a rewards function R which maps $\mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$. In addition to progressing the system into its next state, taking an action $a \in \mathcal{A}_s$ in state s results in an immediate reward $r(s, a)$ for the controller. Depending on the sign of the reward, this may be interpreted as an income (if positive) or a cost (if negative) (Saucedo and Karim, 1998).

Generally, the output of an MDP is a policy π , i.e., the contingency plan determining the action to take in every state of the system in a specified decision epoch. A policy π is therefore a sequence of decision rules $(d_1, d_2, \dots, d_{N-1})$ where $d_t \in D, \forall t \in (1, 2, \dots, N-1)$ for $N \leq \infty$. Such a policy is called stationary, if $d_t = d, \forall t \in \mathcal{T}$, if all decisions are independent of the current decision epoch. Stationary policies are integral to interpreting infinite-horizon MDPs (Puterman, 2014). An optimal action optimizes a predetermined performance index. Mainly two optimality criteria exist for the solution of MDPs: maximum total expected discounted rewards and average expected rewards. In the case of the total discounted rewards problem of an infinite horizon MDP, the function $v_\lambda^\pi(s)$, which maps a value to any given state s under policy π and the discount factor λ , takes the form of the Bellman's equation

$$v_\lambda^\pi(s) \equiv \sup_{a \in \mathcal{A}_s} \left\{ r(s, a) + \sum_{s' \in \mathcal{S}} \lambda(s, s', a) v(s') \right\} \quad (1)$$

which recursively determines the discounted, expected reward of choosing the optimal action for any given state s . Therefore, the optimal policy π^* for a given discount rate λ can be determined by solving the following equation:

$$v_\lambda^{\pi^*}(s) \equiv \sup_{\pi \in \Pi} v_\lambda^\pi(s) \quad (2)$$

An optimal policy regarding the total expected reward criterion consequently maximizes the total expected reward, whereas an optimal policy under the average reward criterion maximizes the average reward. Extending Bellman's work in which these MDPs were to be solved backward iteratively by value iteration, Howard (1960) proposed the policy iteration algorithm for solving infinite horizon MDPs. MDPs allow modeling the control of complex systems which evolve over time following a Markov process. As such, Boucherie and van Dijk (2017) provide examples of MDPs' application in highly complex industries, such as healthcare, transportation, and financial modeling.

A prerequisite of the applicability of MDPs to a given problem, however, is the definition of a state space such that all relevant information is captured in each state and the stochastic process over them becomes Markovian. A Markovian process's state in the epoch epoch $t + 1$ solely depends on its state in epoch t , i.e. the process is path-independent or memory-less. As the state of a bioreactor can be described by the biological and thermodynamic characteristics of its content, and these characteristics are path-independent, it has been argued that the process of a bioreactor is Markovian (Schmidt, 1996). A respective assumption is reasonable for the dynamics of chromatography resins performance decay (Liu et al., 2014). Therefore, the requirements of a Markov decision process in this context are met.

4.3. TPA production and purification

To study the research questions introduced in Section 3.2, this contribution analyses the case study of recombinant Tissue Plasminogen Activator (TPA) production and purification. This section argues for this decision and introduces the numerical case study. To draw managerial conclusions on the determined optimal policies, their sensitivity to process changes is analyzed and they are compared to a basic, upstream titer maximizing policy.

TPA is a recombinant protein with medical applications in the dissolution of blood clots. It was considered a "flagship product of the young biotechnology production industry" (Datar et al., 1993, p. 349), but remains an important protein today (Johnston, 2010). TPA can be produced by mammalian, i.e., Chinese hamster ovary, cells or bacterial, i.e., E.coli, cells (Datar et al., 1993).

The case study of TPA was chosen due to two reasons. Firstly, its importance led to wide coverage in academic literature. Because reliability and availability of specific numerical information are scarce in a secretive industry such as biopharmaceuticals (Schmidt, 1996), being able to rely on previously peer-reviewed publications provides a level of data validity. Secondly, TPA is an economically and curatively important product of the biopharma industry to this day. It generates yearly revenues of c.22 m€ in the US and its use adds 0.75 quality-adjusted life years by reducing disability after ischemic strokes (Johnston, 2010).

Schmidt (1996) summarizes the production of TPA as follows: mammalian cells are attached to inexpensive microcarriers (0 \$/g) in a medium which includes the costly nutrient Fetal Calf Serum (FCS), because TPA exhibits anchorage-dependent growth. Furthermore, TPA production only starts after the culture has finished growing, because TPA exhibits non-growth dependent production. Therefore, a protein producing phase is induced following a growth phase. The medium required for the continuation of the growth phase (12.8 \$/L) is more expensive than that which is used during protein production (2 \$/L) because of the higher FCS concentration of the former. The medium is exchanged every 12 hours and a growing culture is switched into its TPA producing state by exchanging growth for production medium. Schmidt (1996) assumed that exchanging the medium is

equivalent to harvesting all TPA from the medium. By doing this, however, his model is constrained to considering an average TPA concentration of 15 mg/L instead of the time-dependent accumulation of up to 33.5 mg/L of TPA reviewed by Datar et al. (1993).

This thesis models fed-batch dynamics in which the medium is only harvested at the end of the batch. After each cycle, the spent medium in the bioreactor is exchanged with new medium without reducing the TPA concentration, i.e., new medium is added to the reactor while the TPA concentration is growing continuously. In practice, this can be accomplished, for example, feeding new FCS into the spent medium or by separating the spent medium from the TPA without harvesting. The production reactor of a three-reactor seed-train can have a volume of c.160 L and the value of TPA is c.24,000 \$/g (Schmidt (1996)). As the weekly contamination risk in TPA production is between 5 and 10 % and a week has 14 12-hour production cycles, Schmidt (1996) assumes the probability of a successful state transition of a bioreactor to be 99.3 % for all non-zero and non-one probabilities.

After reaching an economical concentration of TPA in the production medium, the protein is separated from the medium without intermediate storage. Lin et al (1993), for example, described a process of affinity purification during which an anti-TPA MAB is used to ease the binding of TPA during chromatography. In Liu et al. (2014) numerical case study, the cost of chromatography maintenance was assumed to be equivalent to three-quarters of the revenue from one sold production batch. Because the value of a full batch of TPA, i.e., 5.36 g of pure TPA, is c.128,640 \$, and resin material is the main driver of these maintenance costs, chromatography resin is assumed to cost c.96,480 \$. This aligns with the magnitude of reported per resin L-prices and necessary resin volumes for chromatography resins by Allmendinger et al. (2014) and Liu et al. (2016). Every action is assumed to incur 100 \$ of fixed costs, mainly due to labor costs.

5. Mathematical Model Formulation

This chapter presents a stochastic dynamic program to study the system-level control of biopharmaceutical production and purification under fermentation and chromatography performance decay uncertainty. The presented model aims at finding simultaneous optimal policies for fermentation control and purification to maximize average expected operating profit. Section 5.1 provides an overview of the used nomenclature. Section 5.2 introduces the proposed discrete state space, infinite time horizon MDP. Section 5.3 and 5.4 consider two peculiarities of the proposed model: the accumulation of protein in the medium and the decay of the chromatography resin. Section 5.5 introduces the objective function of the model.

5.1. Nomenclature

5.2. MDP formulation

The model proposed in this contribution extends Schmidt (1996) upstream model by a downstream component. Furthermore, it allows for multiple parallel production reactors being harvested into a single downstream process to test the presented hypotheses about the implications of the purification process on optimal policies in the USP seed-train and vice versa (see Section 3.2).

The presented framework consists of the two sub-problems described above and is mathematically formulated in terms of the 5-tuple $(\mathcal{T}, \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R})$ using the notation introduced in Section 4.2:

Decision epochs: In accordance with Schmidt (1996), this contribution proposes a discrete-time, infinite time horizon Markov decision process. Let $\mathcal{T} = \{0, 1, 2, \dots, N\}$ be the set of discrete, equidistant decision epochs with an infinite N . Each decision epoch $t \in \mathcal{T}$ represents an observation of the system. A discrete time representation was chosen as it best represents industry practice. Industry practice for biopharmaceutical production is seldom real time monitoring. While real-time, on-line, monitoring of culture physiologies is the topic of ongoing academic discussion (e.g., Abu-Absi et al. (2011), Wechselberger et al. (2013)), on-line monitoring is often high-priced or limited in practicality and functionality (Lourenço et al., 2012). In practice, therefore, monitoring of biomass physiologies is often carried out off-line (Lourenço et al., 2012). Thereby, samples are either taken automatically or manually from the bioreactor and transferred to a laboratory for analysis. Because this is a time-intensive process, a discrete-time representation of the modeled process makes sense. Martagan et al. (2016), for example, observe measurement intervals of two to three days in practice.

An infinite horizon is modeled to allow for the study of stationary policies.

State space: During upstream production, multiple sequential bioreactors may be used to scale-up a cell culture to production volume. As described above, this work focuses only on the production reactors as the bottleneck of the process. The state space of the upstream fermentation policy problem for one production reactor consists of approximations of the culture's thermodynamic properties and was defined by Schmidt (1996). Let

$$\mathcal{S}^U = \{\text{empty}, \text{ready}, \text{growth}_i, \text{production}_j, \text{upset} : \forall i, j \in \mathbb{N}, \forall i, j \leq n_g, n_p\} \quad (3)$$

define the set of feasible states of a production reactor. Let n_g, n_p be the predetermined, finite number of periods in which the culture exhibits cell growth and protein production, respectively. The indices i and j , therefore, represent the number of cycles since the start of the respective phase. For example, production_{30} is the beginning of the 30th production cycle, i.e., the culture has traversed 29 cycles since it entered the protein production phase. The state $s^U \in \mathcal{S}^U$,

Table 2: Nomenclature used in the presented Markov decision process model

Superscripts	
k	Process component, $k \in \mathcal{K}$
Indices	
i	Growth cycle, $1 \leq i \leq n_g$
j	Production cycle, $1 \leq j \leq n_p$
l	Parallel upstream production reactor, $1 \leq l \leq n_U$
m	Number of the next purified batch, $1 \leq m \leq \text{MaxBatch}$
Sets	
\mathcal{T}	Decision epochs
\mathcal{S}	State space of the MDP
\mathcal{S}^k	State space of process component k
\mathcal{A}	Action space of the MDP
\mathcal{A}^k	Action space of process component k
\mathcal{A}_s	Feasible actions in state s
\mathcal{A}_{s^k}	Feasible actions in the k -th process component in state s^k
\mathcal{P}	Transition probabilities of the MDP
\mathcal{P}^k	Transition probabilities of process component k
\mathcal{K}	Process components $\mathcal{K} = \mathcal{U} \cup D$, where \mathcal{U} are the parallel upstream reactors and D is the first downstream chromatography step
\mathcal{U}	Set of n_U parallel upstream production reactors, $\mathcal{U} = \{U_1, \dots, U_l, \dots, U_{n_U}\}$
\mathcal{G}_p	Production-competent growth states which can be converted into production states within one decision epoch
Functions	
\mathcal{R}	Rewards function, $\mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$
π	Policy function, $\mathcal{S} \rightarrow \mathcal{A}$
π^*	Optimal policy function, which maximizes the objective criterion
$\mathcal{V}^\pi(s)$	Value function for policy π , $\mathcal{S} \rightarrow \mathbb{R}$
Parameters	
N	Number of decision epochs
t	Decision epoch, element of T
n_U	Number of parallel upstream production reactors, $n_U = U $
n_g, n_p	Number of growth and production cycles, respectively
Δ	Resin state, in which no further harvests can take place and the resin must be exchanged
capacity_m	Remaining capacity of the resin to bind protein in the m -th harvested batch, %
MaxBatch	Maximum number of batches purifiable using the same chromatography resin, before it must be exchanged
$p(s^k, s'^k, a^k)$	Probability of the k -th process component transitioning from its state s^k in a decision epoch to s'^k in the following decision epoch, if action a^k is taken in it
$p(s, s', a)$	Probability of the system transitioning from state s in a decision epoch to s' in the following decision epoch, if action a is taken
$r(s^k, a^k)$	Reward obtained from choosing action a^k in state s^k in the k -th process component
$r(s, a)$	Reward obtained from choosing action a in state s

(Continued)

Table 2—continued

Parameters	
V	Volume of a production reactor, L
c_g, c_p	Cost of a liter of growth and production medium, respectively, \$/L
c_b	Cost of a gram of microcarriers, \$/g
fixed(a^k)	Fixed cost of carrying out action a^k in the k -th process component, \$
v_g	Value of a liter of growth medium, \$/L
v_p	Value of a gram of protein, \$/g
x_j	Concentration of protein in the production medium in the j -th production cycle, g/L
x_{max}	Maximum possible concentration of protein in the production medium by the end of the production phase, g/L
c_{resin}	Material cost of exchanging spent resin, \$

therefore, describes a bioreactor in terms of its readiness to be used and the physiology of the contained culture. Accordingly, a reactor may be empty, sanitized and ready for inoculation, its culture may traverse states of growth and states of protein production. Out-of-the-ordinary states such as contamination are summarized as upset. In growth-independent production cultures, an ordered set of production-competent growth states \mathcal{G}_p exists. In these states, the culture exhibits kinetics which allow its conversion into a protein producing state within one cycle. For n_U parallel seed-trains, the state of the upstream sub-problem is represented by the vector $s^U = (s^{U_1}, \dots, s^{U_1}, \dots, s^{U_{n_U}}) \in \mathcal{S}^{U_1} \times \dots \times \mathcal{S}^{U_1} \times \dots \times \mathcal{S}^{U_{n_U}}$. For simplicity, the assertion

$$\mathcal{S}^{\mathcal{U}} = \mathcal{S}^{\mathcal{U}_l}, \forall \mathcal{U}_l \in \mathcal{U} \quad (4)$$

is made, where $\mathcal{S}^{\mathcal{U}_l}$ is the state space of the l -th in the set \mathcal{U} of all $n_{\mathcal{U}}$ parallel production reactors.

The state space of the downstream resin exchange schedule problem consists of the remaining performance of the resin in the chromatography column, i.e., what percentage of the protein in the medium can successfully be bound by the resin during purification. The performance of the resin used during chromatography decays after each purified batch. Based on research on resin performance for the industry-standard Protein A chromatography by Jiang et al. (2009), research by Liu et al. (2014) assumed a small number of batches purified at full performance, followed by a linear performance decay to a minimum viable resin performance. This work formulates Liu et al. (2014) resin decay pattern as the state space of the downstream sub-problem. Let

$$\mathcal{S}^D = \{\text{capacity}_m : m \in \mathbb{N}, 1 \leq m \leq \text{MaxBatch}\} \cup \Delta \quad (5)$$

define the set of feasible states of a downstream chromatography column. Therein, MaxBatch is the finite, maximum number of batches which can be purified using the same resin before it must be exchanged. For MaxBatch = 10, the capacity of the resin may start at full capacity before the first batch

is purified, i.e., capacity₁=100%, and deteriorate over pre-defined, intermediate steps to its minimum viable capacity, i.e., capacity₁₀, before purification of the tenth batch. If the resin is used at its minimum capacity, capacity_{MaxBatch}, it deteriorates to a state in which it must be replaced, Δ . The state $s^D \in \mathcal{S}^D$, therefore, is a rational, non-negative number in $[100\%, \text{capacity}_{\text{MaxBatch}}]$, describing the remaining protein binding capacity of the resin at the start of each decision epoch, or Δ .

The state space of a production system with n_U parallel bioreactors and one chromatography column is, therefore, defined as $\mathcal{S} = \mathcal{S}^{U_1} \times \dots \times \mathcal{S}^{U_{n_U}} \times \mathcal{S}^D$, with the vector $s(t) = (s^{U_1}, \dots, s^{U_{n_U}}, s^D) \in \mathcal{S}$ identifying the state of the system at the beginning of epoch t . The vector (empty,100%), for example, describes a system with a single production reactor in a state, wherein the reactor is empty, and the chromatography resin is at its full performance.

Assumption 1: The state space is finite. Therefore, $|\mathcal{S}| < \infty$.

Action space: Let \mathcal{A}^U be the set of feasible actions in a production reactor as part of the upstream fermentation sub-problem. The system operator chooses an action $a^{U_l}(s^{U_l})$ from the set $\mathcal{A}_{s^{U_l}}^{U_l} \subseteq \mathcal{A}^{U_l}$ of feasible actions to take for the l -th parallel upstream reactor in state s^{U_l} . The upstream action space, \mathcal{A}^U , is then the union $\cup_{s,l} \mathcal{A}_{s^{U_l}}^{U_l}$ of all feasible actions in all possible upstream reactor states. Schmidt (1996) action space is adapted to fit the proposed model. As this contribution only considers the production reactors and, therefore, abstracts from the scaling-up of the culture medium from small-scale reactors to the production reactor, actions related to these activities are omitted. The goal of the upstream fermentation control is the production of the optimal amount of protein by way of scheduling the culture growth and production phases within a production reactor. This leads to an action set available for the control of the production reactor which modifies the model of the growth-independent production of TPA by Schmidt (1996). The operator controls a reactor to grow a cell culture and facilitate protein production. To this end, they sanitize an empty reactor for inocula-

tion (prep), feed nutrients into the reactor to grow the culture (addgm), feed a different mix of nutrients into the reactor to convert the growing culture into producing culture and continue this production (addpm), or extract the medium from the bioreactor for harvesting (harvest). In a state of non-production, e.g., upset or growth, harvest is equivalent to dumping the contents of the reactor out. In line with the flexibility allowed by Schmidt (1996), the presented model also includes the combination of harvesting/ dumping and preparation as one action (hprep). The action set is summarized in Table 3. However, not all actions are feasible in every state of a production reactor. Adding production medium, for example, is only feasible for production-competent growth states (thus, inducing production) and all but the final production state (thus, continuing production). Adding growth medium is only feasible in a ready state (thus, inducing growth) and during the growth phase of the culture except for the final growth state.

Let $a^U = (a^{U_1}, \dots, a^{U_l}, \dots, a^{U_{n_U}}) \in \mathcal{A}^{U_1} \times \dots \times \mathcal{A}^{U_l} \times \dots \times \mathcal{A}^{U_{n_U}}$ be the vector of actions taken in an upstream process of n_U parallel production reactors. Again, for simplicity, the following assertion is made for the action spaces of each production reactor:

$$\mathcal{A}^U = \mathcal{A}^{U_l}, \forall U_l \in U \quad (6)$$

Consequently, let A^D be the set of feasible actions in the downstream resin exchange schedule sub-problem. This contribution incorporates Liu et al. (2014) notion of the maintenance activity required to replace deteriorated resin into the action space of a biopharmaceutical production system operator. At any given state of the chromatography resin, the actions $\mathcal{A}_{s^D}^D \subseteq \mathcal{A}^D$ are feasible and $a^D(s^D) \in \mathcal{A}_{s^D}^D$ denotes the action chosen in the downstream sub-problem. Equivalently to the upstream sub-problem, the action space of the downstream sub-problem, \mathcal{A}^D , is the union $\cup_s^D \mathcal{A}_{s^D}^D$ of the feasible actions in all states of the chromatography resin. An operator must choose between doing nothing (none), accepting the harvest from the upstream process (accept) or exchanging the resin to restore full chromatography performance (exresin). The action space of this sub-problem is summarized in Table 4. Whenever the resin capacity deteriorates past the predetermined minimum viable threshold, $\text{capacity}_{\text{MaxBatch}}$, it must be exchanged, therefore, only action 3 is feasible in the state $s^D = \Delta$.

The action space of the simultaneous USP and DSP control problem presented herein is defined as the set of actions $A = A^{U_1} \times \dots \times A^{U_{n_U}} \times A^D$ from which the controller of the production system can choose at the beginning of each decision epoch. Let $a(s) \in A_s$ denote the vector of actions taken in system state s representing the joint decision made by the operator. Therein, A_s is the set of feasible actions in states s . The action vector (addpm,exresin) or (3,3), for example, indicates that, in a system of a single production reactor and one DSP production medium is added to the production reactor and the resin of the chromatography column is exchanged in the same decision epoch.

As introduced earlier, the simultaneous control of the entire production system is complicated by the inter-dependencies between the up- and the downstream part of the process. Because both processes are assumed to be coupled, i.e., there is no intermediate storage of harvested medium, harvesting a production reactor (harvest or hprep) is only feasible if the chromatography step chooses to accept the medium for purification (accept) in the same decision epoch. Furthermore, in a system of multiple parallel reactors, only one can be harvested at a time.

Transitions: The observed production and purification system evolves over time from state to state, i.e., along the path $\{s(t_0), s(t_1), \dots, s(t_N)\}$, with $s(t) \in \mathcal{S}, \forall t \in \mathcal{T}$. As described above, the states which a reactor in the upstream fermentation sub-problem occupies over time are defined by approximations of the culture's cell concentrations, thermodynamic quantities and time passed since the start of either growth or production (Schmidt, 1996). A production reactor, therefore, evolves over time according to the physiological rules which govern cell cultures (as captured by cell-level models in Section 2.2) and the actions taken by the controller. In the highly uncertain environment of living cell cultures, however, this evolution is not deterministic.

Let

$$\begin{aligned} p(s^{U_l}, s'^{U_l}, a^{U_l}) = \\ \mathbb{P}[s^{U_l}(t+1) = s'^{U_l} | s^{U_l}(t) = s^{U_l}, a^{U_l}(t) = a^{U_l}], \\ \forall s^{U_l}, s'^{U_l} \in \mathcal{S}^{U_l}, \forall a^{U_l} \in \mathcal{A}_{s^{U_l}}^{U_l}, \\ \forall U_l \in U, \forall t \in T \end{aligned} \quad (7)$$

determine the probability of the l -th production reactor to evolve from state s^{U_l} in one decision epoch t to state s'^{U_l} in the next, $t+1$, if action a^{U_l} is chosen. While this model allows for transition probabilities in the bioreactor to be set one for one, a simplification is made to reduce the number of necessary process parameters: every non-zero, non-one probability is identical for every production reactor U_l . However, in line with Schmidt (1996), the cell culture decline towards the end of its lifecycle manifests by transition probabilities decreasing over the last sixth of the production phase. Using the state and action schemas defined above, Table 5 summarizes the transition probabilities of the upstream fermentation sub-problem.

The transitions of the chromatography step introduce uncertainties to the downstream resin exchange schedule sub-problem. The state of the chromatography step evolves over the predefined set of remaining capacities the resin can have. As studied by Jiang et al. (2009), chromatography resins decay with increased usage and their capacity to bind proteins of interest deteriorates. Accepting a harvest from the upstream part has two consequences: Firstly, an amount of protein is yielded depending on the remaining capacity of the resin and, secondly, the resin's capacity deteriorates to the next capacity step in the following epoch. As the purification of a batch in the first chromatography step is assumed to take five cycles (Liu et al., 2014) and within six to

Table 3: Action space of the control of a production reactor

Action	Abbreviation	code
Do nothing	none	1
Add growth medium	addgm	2
Add production medium	addpm	3
Harvest or dump	harvest	4
Prepare	prep	5
Harvest or dump and prepare	hprep	6

Table 4: Action space for the control of the first chromatography step

Action	Abbreviation	code
Do nothing none	None	1
Accept a harvest for purification	accept	2
Exchange resin	exresin	3

Table 5: Non-zero transition probabilities of an upstream sub-problem, adapted from Schmidt (1996)

$s^{U_i^1}$	$\text{index}^2(s^{U_i}, s'^{U_i}, a^{U_i})$		
e		none	$p(e, e, 1)$
		prep	$p(e, r, 5), p(e, u, 5) = 1 - p(e, r, 1)$
r		none	$p(r, u, 5) = 1$
		harvest	$p(r, e, 4) = 1$
		hprep	$p(r, r, 6) = p(e, r, 5), p(r, u, 6) = 1 - p(e, r, 5)$
g_i	$1..n_g$	none	$p(g_i, u, 1) = 1$
	$1..n_g - 1$	addgm	$p(g_i, g(i + 1), 2), p(g_i, u, 2) = 1 - p(g_i, g(i + 1), 2)$
	$i \in G_p$	addpm	$p(g_i, p_1, 3), p(g_i, u, 3) = 1 - p(g_i, p_1, 3)$
	$1..n_g$	harvest	$p(g_i, e, 4) = 1$
p_j	$1..n_g$	hprep	$p(g_i, r, 6) = p(e, r, 5), p(g_i, u, 6) = 1 - p(e, r, 5)$
	$1..n_p$	none	$p(p_j, u, 1) = 1$
	$1..n_p - 1$	addpm	$p(p_j, p(j + 1), 3), p(p_j, u, 3) = 1 - p(p_j, p(j + 1), 3)$
	$1..n_p$	harvest	$p(p_j, e, 4), p(p_j, u, 4) = 1 - p(p_j, e, 4)$
u	$1..n_p$	hprep	$p(p_j, r, 6) = p(e, r, 5), p(p_j, u, 6) = p(e, u, 5)$
		none	$p(u, u, 1) = 1$
		harvest	$p(u, e, 4) = 1$
hprep		$p(u, r, 6) = p(e, r, 5), p(u, u, 6) = p(e, u, 5)$	

eight hours (Martagan et al., 2018), harvests aren't blocked by ongoing purification. Yields from chromatography, however, have been considered a source of uncertainty, e.g., by Liu et al. (2016) Therefore, the model proposed in this contribution allows for stochasticity in the performance decay of the chromatography step. Let $p(s^D, s'^D, a^D)$ be the probability of the resin in state s^D decaying to its new state s'^D , if action a^D is taken. This allows modeling of uncertain deterioration: after a purification cycle, resin performance may, for example, reduce to the next lower step, remain at the same performance level or deteriorate by more than one step. Whenever the resin is exchanged, the performance is deterministically restored to the resin's original value, i.e., $p(s^D, \text{cap}_1, 3) = 1, \forall s^D \in S^D$. Maintenance activities are not time-intensive (Liu et al., 2014) and are, therefore, assumed to be feasible in one decision epoch. Furthermore, the resin

must be exchanged when it deteriorates past its minimum allowed capacity. These dynamics are detailed in Section 5.4. Table 6 details the non-zero transition probabilities.

Transition probabilities within bioreactors (Schmidt, 1996) and within chromatography resins are assumed to be independent. Therefore, let

$$p(s, s', a) = \prod_{k \in \mathcal{K}} p(s^k, s'^k, a^k) \tag{8}$$

define the probability of the system evolving from state s in decision epoch t to state s' in $t+1$, where \mathcal{K} is the set of the

¹State descriptions have been shortened for readability, i.e., empty to e , ready to r , growth to g , production to p , and upset to u

²Index range, in which action a^{U_i} is feasible

biopharmaceutical process “components”, i.e., all production reactors and the chromatography column.

Assumption 2: Transition probabilities are stationary, i.e., do not vary with time. Therefore, $p(s, s', a) = p(s(t), s(t+1), a(t))$, $\forall t \in \mathcal{T}$.

Let \mathcal{P} constitute the set of all transition probabilities $p(s, s', a)$.

Rewards: Production and purification system operations deal with the financial trade-off between operating costs, caused by continued fermentation and chromatography resin exchange, and revenue from purified protein. The rewards obtained from a USP reactor in the state s^{U_i} by choosing the action a^{U_i} are denoted $r(s^{U_i}, a^{U_i})$. In line with Schmidt (1996), actions incur small fixed costs and volume dependent variable costs or revenues. The cost of adding new growth medium (addgm) is, therefore, the sum of the cost of the growth medium and a fixed cost for carrying out the action. The former is dependent on the two parameters: the volume of the reactor (V in L) and the cost of the growth medium per volume (c_g in \$/L): $r(s^{U_i}, \text{addgm}) = -c_g V - \text{fixed}(\text{addgm})$. Therein, c_g depends on the concentration of FCS in the growth medium and its cost. Consequently, $r(s^{U_i}, \text{addpm}) = -c_p V - \text{fixed}(\text{addpm})$ is the cost of inducing or continuing the production by adding nutrients into the spent medium up to a production medium concentration, where c_p is the per-liter-cost of the production medium, depending on its FCS concentration. Preparing an empty reactor (prep) with a growth medium incurs costs for the used growth medium and microcarriers allowing for anchorage-dependent culture growth. Because both costs are volume-dependent, $r(s^{U_i}, \text{prep}) = -(c_g + c_b) V - \text{fixed}(\text{prep})$ defines the cost of this action, where c_b is the cost of the microcarrier (in \$/L). If the medium is dumped during the growth phase (harvest, hprep), salvage revenue from the spent growth medium is earned. Let, therefore,

$$r(s^{U_i}, a^{U_i}) = \begin{cases} v_g V - \text{fixed}(\text{harvest}) & \text{if } a^{U_i} = \text{harvest} \\ (v_g - (c_g + c_b)) V - \text{fixed}(\text{hprep}) & \text{if } a^{U_i} = \text{hprep} \end{cases} \quad (9)$$

where $s^{U_i} \in \{\text{growth}_i : 1 \leq i \leq n_g\}$. If a batch is lost due to doing nothing, negative reward in the amount of the opportunity costs are earned. Let

$$r(s^{U_i}, \text{none}) = \begin{cases} -v_g V & \text{if } s^U \in \{\text{growth}_i : 1 \leq i \leq n_g\} \\ -v_p x_j V & \text{if } s^U \in \{\text{production}_i : 1 \leq i \leq n_p\} \end{cases} \quad (10)$$

determine the opportunity costs of losing a batch due to doing nothing, where v_g is the value of the growth medium in \$/L, v_p is the value of TPA in \$/g, x_j the concentration of TPA in the production medium at the time of losing the batch in g/L, and V the volume of the reactor in L. Rewards from

harvesting production medium (harvest, hprep) are only obtained after purification, therefore, these are included in the rewards function of the DSP sub-problem.

Equivalently, the rewards obtained from operating the chromatography step in state s^D by choosing action a^D are denoted $r(s^D, a^D)$, which is defined for all $a^D \in \mathcal{A}^D$ and all $s^D \in \mathcal{S}^D$. The material cost of chromatography resin is significant (Allmendinger et al., 2014; Farid, 2007), therefore the cost of maintenance activities mainly depends on the material costs. Let $r(s^D, \text{exresin}) = -c_{\text{resin}} - \text{fixed}(\text{exresin})$ describe the cost of exchanging the chromatography resin. The revenue obtained from purified product depends on the volume of TPA in the medium at the point of harvesting and the remaining binding capacity of the resin. Let

$$r(s^{U_i}, \text{accept}) = \begin{cases} -v_p x_j V \text{capacity}_m - \text{fixed}(\text{harvest}) & \text{if } \exists! U_l \in \mathcal{U} : a^{U_l} = \text{harvest} \\ -v_p x_j V \text{capacity}_m - \text{fixed}(\text{hprep}) & \text{if } \exists! U_l \in \mathcal{U} : a^{U_l} = \text{hprep} \end{cases} \quad (11)$$

define the operating profit obtained from accepting a harvest from exactly one production reactor which depends on the value of the production medium at the time of harvesting and what percentage of that value is captured during purification (capacity_m in %). While Schmidt (1996) implies harvesting every time spent production medium is exchanged for new medium in the reactor (addpm), the presented model only allows harvesting as an action distinct from continuing fermentation which returns the bioreactor to an empty state, to better model the practical realities of fed-batch fermentation as presented by Martagan et al. (2016).

Let the aggregate rewards function

$$\mathcal{R}(s, a) = \sum_{k \in \mathcal{K}} r(s^k, a^k) \quad (12)$$

denote the sum of costs/ revenues of carrying out each action in the action vector a in system state s on their respective reactor/ chromatography column.

Assumption 3: Like transition probabilities, rewards are stationary and bounded. Therefore, $r(s, a) = r(s(t), a(t))$, $\forall t \in \mathcal{T}$ and $|r(s, a)| \leq M < \infty$, $\forall s \in \mathcal{S}, \forall a \in \mathcal{A}(s)$.

By assumption, the reward obtained from harvesting the medium depends strongly on the amount of protein in the medium and the yield performance of the chromatography resin. As these variables constitute important aspects of the presented model, special consideration should be given to their modeling. The following two sections are aimed at this.

5.3. TPA accumulation during production

Protein production in living cell organisms most prominently follows one of two patterns: growth related or non-growth related (Schmidt, 1996). In growth related protein

Table 6: Non-zero transition probabilities of the downstream sub-problem

s^D ³	$index$ ⁴	a^D	$p(s^D, s'^D, a^D)$
cap_m	1..MaxBatch – 1	none	$p(cap_m, cap_m, 1) = 1$
	1..MaxBatch – 1	accept	$\sum_{0 \leq s \leq 1} p(cap_m, cap_{m+s}, 2) +$ $p(cap_m, \Delta, 2) = 1$
	MaxBatch	accept	$p(cap_m, cap_m, 2)$ $p(cap_m, cap_m, 2), p(cap_m, \Delta, 2) =$ $1 - p(cap_m, cap_m, 2)$
	1..MaxBatch	exresin	$p(cap_m, cap_1, 3) = 1$

formation, the product is produced as a by-product of culture growth, i.e., during growth states. In non-growth-related protein accumulation, however, proteins are produced only after the cell culture has finished growing and started producing, i.e., during production states. Figure 3 illustrates this difference. Towards the end of their life, cell cultures exhibit decline (Schmidt, 1996)(Schmidt, 1996). This is modelled by the probabilities of a successful transition decreasing over the last sixth of the production phase: $p(p_{30}, p_{31}, 3), \dots, p(p_{35}, p_{36}, 3)$ have the values 0.84, 0.67, 0.5, 0.34, 0.17, 0.05, which were adapted from Schmidt (1996). TPA, the subject of this work's numerical study, is produced in a non-growth-related pattern.

Assumption 4: The protein concentration x_j monotonically increases in j . Therefore, $x_j \leq x_{(j+1)}, \forall j: 1 \leq j \leq n_p$.

Although stochastic formulations of protein accumulation exist (Martagan et al. (2016)), one may sensibly approximate protein accumulation as a linear relationship as illustrated in Figure 3. Therefore, the proposed model describes a linear growth of TPA in the production medium starting in the first production state and ending with a maximum concentration x_{max} in the final production state. Let the concentration at the beginning the production cycle j be defined by the equation

$$x_j = \frac{x_{max}}{n_p - 1} \times (j - 1), \forall j: 1 \leq j \leq n_p, \forall n_p > 1 \quad (13)$$

The product of concentration (in g/L) and bioreactor volume (in L), therefore, describes the amount of protein in the reactor (in g).

5.4. Resin performance decay

Chromatography resins decay due to being repeatedly exposed to process conditions; their decay is most commonly monitored based on product yield (Nweke et al., 2018). The implied financial trade-off between reduced revenues from harvesting fermented product and the scheduling of costly resin exchange activities is studied in this contribution.

Therefore, modeling resin decay is given special consideration in this section.

The notion of maintenance actions related to resin performance decay is based on Liu et al. (2014), who expanded on prior work with a focus on downstream planning (Allmendinger et al., 2014). Liu et al. (2014) further built on Jiang et al. (2009) study of resin performance decay. They assume that each chromatography step, i.e., the purification of a batch in one chromatography column, takes five cycles. Therefore, they aggregate Jiang et al. (2009) experimental results of resin decay over multiple chromatography cycles to a per-batch decay pattern in their first numerical example. Similar to Liu et al. (2014), this work assumes a resin deterioration over 11 batches, with each batch taking five chromatography cycles. After the eleventh purified batch, the resin must be exchanged. Therefore, after four batches purified at full performance, the resin starts to deteriorate until its minimum capacity is reached before batch 11 (see Figure 4). This decay, however, was previously only deterministic in nature. In practice, this may not be the case and constitute an additional source of uncertainty and operational complexity. Because resin yield is one of the uncertainties which have an impact on production cost and delivery (Farid et al., 2005), Liu et al. (2016) considered the notion of uncertain resin in capacity planning. They modeled resin yield deviations following a triangular distribution. In their numerical case, they assume that resin yields deviate at most 5 % from their standard value depending on the type of resin.

However, because the model in this contribution has a finite, discrete state space, assumptions about the evolution of the resin capacity must be made. The performance decay in this model follows a similar pattern as in the Liu et al. (2014) model but introduces an element of uncertainty by assuming the decay steps to be stochastic rather than deterministic.

Assumption 5: After a batch is purified, the resin either doesn't decay, decays as expected after one purification, or decays as if two batches were purified. Therefore, $\sum_{0 \leq s \leq 2} p(cap_m, cap_{m+s}, 2) = 1, \forall m < MaxBatch$ and $p(cap_m, cap_{m+1}, 2) + p(cap_m, \Delta, 2) = 1$, if $m = MaxBatch$

The fifth assumption introduces stochasticity in the resin yields. As indicated in Table 6, after the purification of one batch, the yield capacity of the resin may stay the same, dete-

³capacity_m has been shortened to cap_m for readability

⁴Index range, in which action a^D is feasible

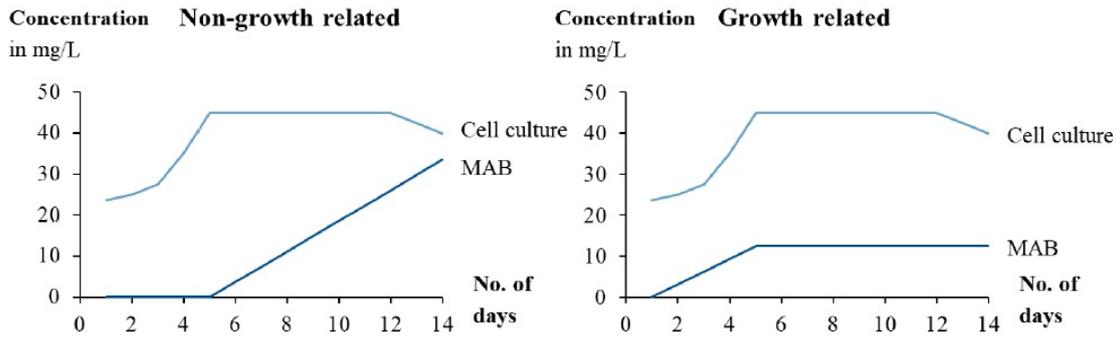


Figure 3: Comparison of non-growth related and growth-related production patterns, adopted from Schmidt (1996)

riorate by one step, or deteriorate by two steps. For example, a resin that yielded 95 % of the product in batch five, may also yield 95 % of batch six, or, more likely, only yield 90 % or even only 85 % of batch six (see some scenarios in Figure 4). In a worst-case scenario (double deterioration after every batch), this resin could reach its minimum capacity of 65% after only two more batches rather than six. For the edge case of the resin being one decay step away from its minimum capacity where a decay by two steps is not feasible, this work assumes that only one decay step is possible. Because of their short duration (Liu et al., 2014), maintenance activities related to restoring the resin capacity take only one decision epoch.

5.5. Optimality objective and optimization

The Markov decision process defined herein aims at determining the optimal stationary policy for operating upstream fermentation, harvesting, and downstream resin exchange. As argued by Martagan et al. (2016), discounting future rewards too strongly on this operations-planning horizon could bias the results. Therefore, the presented problem is studied under average rewards. To achieve this, the model's objective is to maximize the average reward over an infinite time horizon. Contrasting Equation 1, let the average expected reward criterion

$$v(s, \pi) \equiv \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}_s^\pi \left\{ \sum_{t=1}^N r(s(t), a^\pi(t)) \right\} \quad (14)$$

be defined for all starting states $s \in \mathcal{S}$ and all $\pi \in \Pi$ where Π is the set of all possible policies. Therein, \mathbb{E}_s^π is the expected value under policy π of starting in state s , dependent on the reward $r(s(t), a^\pi(t))$ of choosing the action $a^\pi(t)$ in the state $s(t)$ as prescribed by policy π . Therein, the objective is to find $V^*(s, \pi^*) \equiv \sup_{\pi \in \Pi} v(s, \pi)$. The average reward criterion problem is solved using the relative value iteration algorithm as implemented by Chades et al. (2014) in Matlab (The MathWorks Inc., 2018). While the standard value iteration algorithm does not converge to the optimal value function for average reward problems, the relative value iteration does so (Gupta et al., 2015). For robustness, the total

expected discounted reward problem with a discount factor λ close to one is also solved to optimality. In line with Equation 1, let the total expected discounted reward criterion

$$v_\lambda^\pi(s) \equiv \sup_{a \in A_s} \left\{ r(s, a) + \sum_{s' \in \mathcal{S}} \lambda p(s, s', a) v_\lambda^\pi(s') \right\} \quad (15)$$

be defined for all $s \in \mathcal{S}$. The optimal policy π^* optimizes the objective $v_\lambda^*(s, \pi^*) \equiv \sup_{\pi \in \Pi} v_\lambda^\pi(s, \pi)$. For this, the policy iteration algorithm, also implemented by Chades et al. (2014), is used. All model code is provided in Appendix 7 - Appendix 9.

6. Case Study Results

To investigate the posed research questions, a numerical case study of TPA production and purification, as introduced in Section 4.3, was implemented. To test specific hypotheses, sensitivity analyses were run; the resulting policies were critically analyzed and compared.

This task is complicated by the presented research questions pertaining to two different process setups, or USP to DSP ratios, i.e., how many parallel production reactors are harvested into one downstream process. A basic "1:1" model with one USP production reactor and one DSP, where $\mathcal{K} = \{U, D\}$, and a "2:1" model with two parallel production reactors and one DSP, where $\mathcal{K} = \{U_1, U_2, D\}$ were implemented based on the formulation in Section 5.1. Section 6.1 discusses the optimal control of the 1:1 process setup. Section 6.2 extends Section 6.1 by adding a second parallel production reactor in a 2:1 process setup and the implications thereof are discussed. All case study data is provided in Appendix 12 - Appendix 16.

6.1. Optimal, simultaneous control of a serial USP and the DSP

In this section, the 1:1 setup of one production reactor and one chromatography step is studied. In answering RQ1 and RQ2, the average per-cycle operating profit maximizing policy is presented and discussed. Furthermore, managerial insights are derived from the optimal reward's sensitivity to process parameter changes. Conclusions are also

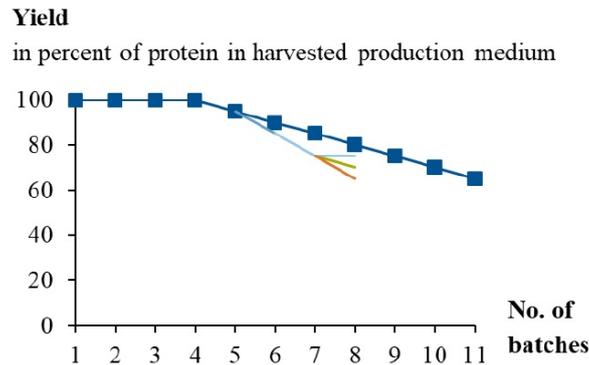


Figure 4: Deterministic performance decay pattern of a candidate chromatography resin over 11 purified batches, with selected stochastic decay scenarios starting after the fifth batch, adapted from Liu et al. (2014)

checked for robustness using a total expected discounted reward problem formulation and a formulation with longer decision epochs (36 instead of 12 hours). First, the resulting optimal policy is described in sub-section 6.1.1 before its theoretical and managerial implications are discussed (sub-section 6.1.2) and further insights are derived from analyzing sensitivities (sub-section 6.1.3).

6.1.1. Optimal harvesting and resin exchange control

The average reward-maximizing policy for the simultaneous control of a single USP production reactor and the first chromatography step is depicted in Table 7.

The visualization of the optimal policies in this work is based on the format used by Schmidt (1996). As such, the rows in Table 7 represent the state of the production reactor; the columns represent the state of the chromatography column. Areas represent “decision zones” of states in which the optimal action vector is identical. The number pairs within each decision zone of the table represent the optimal action vector in those states of the system, in accordance with the coding scheme introduced earlier (Table 3 and Table 4). Thereby, the first digit of the ordered tuple is the optimal action to be taken in the production reactor, and the second digit is the optimal action in the chromatography column. For example, if the production reactor is within the last seven cycles of the production phase, and the resin has at least 75 % remaining capacity, it is optimal to harvest the production medium, return the production reactor to a ready state, and purify the harvested TPA, i.e., action vector (6,2) or (hprep, accept). The average reward of the optimal policy is $\mathcal{V}^{\pi^*} = 1,272\$$ of operating profit per 12-hour cycle, leading to the following observation.

Observation 1: The maximum average per cycle operating profit is positive. Therefore, it may be economically feasible to operate the serial 1:1 production system under the assumed process parameters.

Should the system start with a chromatography resin at full performance, e.g., (empty,100 %), the optimal contin-

gency plan (given no contamination) is to prepare the reactor for inoculation (prep), to grow the culture over five cycles, or 2.5 days, (addgm), to convert the growing culture into a TPA-producing culture by lowering its FCS-concentration in its first production-competent state, g6, (addpm) and to continue TPA-production for 29 cycles, or 14.5 days, (addpm) before harvesting in the 30th cycle and preparing the bioreactor for the next batch (hprep).

For almost all USP states, it is optimal to exchange the chromatography resin only when it is completely depleted (state Δ), i.e., only sometimes is it optimal to prematurely incur resin exchange costs. Between the 27th and 29th production cycle of the USP reactor, for example, resin exchange is optimal even though it’s still at as much as 80 % of its full performance. In these cycles, the value of the TPA batch lies between 95,561 \$ and 102,912 \$, but if only 80 % of the TPA were bound, it would be worth only between 76,449 \$ and 82,330 \$. At this point, it is, therefore, optimal to exchange the resin (thus, incurring 96,480 \$ material net of fixed costs) and continue production until the 30th production cycle, where 100 % of the then c.4.4 g of TPA can be purified, realizing 106,587 \$ of operating revenue.

Observation 2: Premature resin exchange seems to be optimal only sometimes. Robustly, it is optimal for resin performances under 85 %, before the culture reaches the optimal harvesting decision zone.

Policy-related insights remained robust under longer decision epochs and total expected discounted reward maximization. For the total expected discounted reward maximizing policy, policy iteration with a discount factor close to one ($\lambda = 0.99$) was used. The optimal total discounted reward function in the initial state $V^*(s^U = \text{empty}, s^D = 100\%) = 117,395\$$ was obtained. Observations 1 and 2 remained robust with two exceptions: (1) given remaining resin performance of 65% (70%), cultures in their first to fifth (first and second) growth cycle were dumped and the reactor prepared to start a new batch, and (2) resin exchange seemed more risk-seeking, i.e., premature resin exchange was postponed

Table 7: Average reward maximizing policy for the operation of one TPA production reactor

	100%	100%	100%	100%	95%	90%	85%	80%	75%	70%	65%	Δ
e	51 Prepare reactor for inoculation											53
r	21 Add growth medium to continue culture growth phase											23
g_1												
\dots												
g_5												
g_6												
g_7	31 Add production medium to convert production-competent growth states, or to continue TPA production											
g_8												
p_1												
\dots												
p_4												
p_5												
p_6												
\dots												
p_{17}												
p_{18}												
p_{19}	33 Add growth medium in reactor, exchange chromatography resin											
p_{26}	62 Harvest production medium and prepare reactor for next batch											13
p_{27}												
p_{28}												
p_{29}												
p_{30}	61 Dump upset medium out											63
p_{31}												
\dots												
p_{36}												
u												

to the 65% remaining performance and production cycles 29 and 30. The optimal policies of the respective robustness checks are visualized in Appendix 10 and Appendix 11.

Following this description of the case study results, their theoretical and managerial implications are discussed in the next sub-section.

6.1.2. Theoretical and practical contributions

Fed-batch fermentation and purification of TPA under the assumed process parameters is economically feasible, yielding a positive average operating profit (see Observation 1). This is in line with the findings of Schmidt (1996). The average reward of the presented model, however, is c.88 % smaller than Schmidt's result. This may seem counterintuitive because of the presented model's support for increasing the TPA concentrations up to 33.5 mg/L and higher purification yields. However, because the presented model incorporates batch-ending harvesting dynamics and substantial resin material costs, the lower average reward is reasonable. To offset continuously incurred operating costs, revenues can only be realized at two points in a batch's lifecycle: scrapping it during growth (at its salvage value) or terminating fermentation by harvesting the TPA produced up to that point. Furthermore, harvesting seems to be optimal

in the states just before the culture enters decline. The production medium is harvested in the 30th cycle already, given at least 75% resin capacity (Table 7). If the batch was not harvested in this state, the probability of successfully continuing would be reduced to 84%. This observation holds robust under total expected discounted reward maximization. Furthermore, it corroborates the findings of Martagan et al. (2016). In their study of IgG₁ harvesting decisions under uncertain fermentation dynamics, they observe the counterintuitive optimality of harvesting before the culture decline phase at the end of the fermentation time. However, they attribute this to the costs associated with the additionally produced by-products. In the presented study, impurities are not explicitly modeled, yet we observe a similar pattern. During advanced cycles of the production phase, resin exchanges are optimal before the minimum allowed capacity is reached whether resin decay is deterministic or uncertain. The earliest resin exchange is carried out after seven batches have been purified and the resin's performance has deteriorated to 80 %, i.e., in the state (p_{18} , 80 %) and the latest after the eleventh purified batch (Table 7). This is even earlier than previously reported under deterministic resin decay. Liu et al. (2014) found in their numerical study of fermentation and purification schedules that at most ten batches are

purified (leaving capacity at 65 %) before the chromatography resin is exchanged if the maintenance costs are equivalent to three-quarters of a full production batch. Comparing the scenario of stochastic decay to that of deterministic decay, the resin is exchanged after the same number of purified batches in both cases. Under deterministic resin decay⁵ i.e., $p(\text{cap}_m, \text{cap}_{m+1}, 2) = 1, \forall m : 1 \leq m \leq \text{MaxBatch} - 1$ and $p(\text{cap}_{\text{MaxBatch}}, \Delta, 2) = 1$, we find that the resin is exchanged the earliest after the seventh batch has been purified, i.e., in the state $(p_{29}, 80\%)$.

While the effect of stochasticity on the timing of exchanges was limited, it had an effect on the optimal average operating profit obtained under stochastic decay compared to under deterministic decay. The potential downside of a decay by two steps seems to outweigh the potential upside of the resin not decaying, in terms of average operating profits. The sensitivity of the optimal average reward, \mathcal{V}^* , and the total expected discounted reward in the initial state, V^* (e, 100%), are compared to the proposed stochastic resin decay in Table 8. Moving from the proposed stochastic decay to a deterministic decay, ceteris paribus, yielded a 3% reduction in \mathcal{V}^* , whereas \mathcal{V}^* (e, 100%) remained virtually constant. In addition to its theoretical contributions, this case study also provides relevant managerial insights.

Biopharmaceuticals producers seem to benefit from the stochastic optimization of the integrated control of up- and downstream processes. We can compare the optimal policy to a primitive, production titer maximizing policy, π_{MaxTiter} , which might be used in biopharmaceutical manufacturing without decision support systems integrating system-level trade-offs in a worst-case scenario. From this comparison, the value of the presented optimal policy can be quantified. π_{MaxTiter} is provided in Appendix 18. Compared to π_{MaxTiter} , the optimal policy obtained from the presented model outperforms by a wide margin. Titer maximizing production under the assumed process parameters leads to a negative optimal total discounted rewards value function in the initial state, $V^{\pi^*}(s^U = \text{empty}, s^D = 100\%) = -76,311\$$. Therefore, while titer maximizing production may sound reasonable under consideration of only the upstream process, it may not be economically feasible on a system-level. The presented optimal policies, on the other hand, provide a business case for profitable production on a system-level.

Furthermore, by studying the sensitivity of average operating profits to changes in the fermentation reliability, managerial implications about the business case of reliability investments can be derived. While a reduction of the per-cycle probability of a successful transition has a negative effect on average and total discounted rewards, an increase in reliability by 0.2 percentage points (from 99.3% to 99.5%) has the potential to increase average and total discounted rewards by 6% and 7%, respectively (Table 9). The present value of this average operating profit increase in perpetuity is at least

c.1.4 m€ (10% p.a. discount rate, 730 yearly cycles). Therefore, investments aimed at increasing process reliability, e.g., in newer fermentation equipment, are economically feasible up to this amount. The effect of further probability increases, however, showed opposite effects between the two measures. Their managerial interpretation should, therefore, be subject to further validation.

This work's answers to RQ1 and RQ2 extend academia's and practice's understanding of fed-batch fermentation control and harvesting decisions under simultaneous consideration of downstream purification and maintenance operations. The presented Markov decision process provides a dynamic framework for the stochastic optimization of biopharmaceutical operations. Furthermore, it corroborates existing research on the optimal harvesting decisions under fermentation uncertainties (Martagan et al., 2016; Schmidt, 1996), extends the literature on the timing of resin exchange activities (Liu et al., 2014), and aids practical decision making on an operational- and financial-level.

6.1.3. Sensitivity to varying resin exchange-related parameters

This sub-section answers RQ3 by discussing the optimal policy's sensitivity to process parameter variations relating to the chromatography resin economics. To this end, the 1:1 model analyzed above was run with different parameter sets and the effects were studied.

Because of its high economic impact, the timing of resin exchange, i.e., how long this costly activity is at least postponed, is particularly interesting for academia and practice. Decisions regarding the timing are mainly driven by the trade-off between the expected higher payoff of future harvests and the incurring of high material costs. This sub-section, therefore, analyses the sensitivity of the timing of the earliest exchange activity to changes in related parameters, such as resin exchange costs and minimum viable resin capacities. The "earliness" of resin exchanges in a policy is operationalized as the least number of batches purified before any resin exchange. Visually, this is the upper-left-most occurrence of the downstream sub-problem action 3 in the optimal policy as pictured in Table 7. In this specific case, that occurrence is in the state $(p_{18}, 80\%)$, meaning four batches have been purified at 100% and one batch each at 95, 90, and 85%. It can be found by traversing the optimal policy column-wise, starting from the upper left state (e, 100%). The optimal average reward was analyzed as a second important Key Performance Indicator (KPI).

Observation 3: Higher resin material cost negatively influenced both monitored KPIs; higher minimum resin capacities led to higher average rewards.

Resin material cost had a negative impact on both the optimal average rewards and earliness of the resin exchanges. The rest of this thesis considers resin material costs, c_{resin} , equivalent to 0.75 times the value of a full TPA batch, i.e.,

⁵The average reward maximizing policy under deterministic decay is provided in Appendix 17

Table 8: Sensitivity of the optimal average reward and total expected discounted reward functions to changes in the resin decay probabilities in the DSP

Probability scenarios				\mathcal{V}^*	$\mathcal{V}^*(e, 100\%)$
	No decay	by one step	by two steps		
Base case	5%	90%	5%	1.00	1.00
Deterministic case	0%	100%	0%	0.97	1.00

Table 9: Sensitivity of the optimal average reward and total expected discounted reward functions to changes in the probability of a successful transition in the USP

Probability scenarios	\mathcal{V}^{*6}	$\mathcal{V}^*(e, 100\%)^7$
Base case $p(s_t^U, s_{t+1-U}, a_{t+1}^U) = 99.3\%$	1.00	1.00
90%	0.40	0.40
99%	0.90	0.89
99.5%	1.06	1.07
99.8%	1.01	1.19
99.9%	0.76	1.23

96,480\$. For this analysis, c_{resin} is varied in the range [192,960\$, 189,0220\$..0\$], where 192,960\$ is equivalent to one and a half times the value of a full TPA batch. Figure 5 visualizes the negative correlation of more expensive resin exchanges and earliness of the exchange. For increasing resin costs, it becomes optimal to postpone the earliest exchange further. If the resin material were free, it can be optimal to exchange the chromatography resin as early as after four batches, i.e., when 100% of the produced TPA could be purified for the last time.

On the other extreme end of the range, for material costs of at least one and a half times the value of a full TPA batch, resin exchange is postponed until nine batches have been purified and the performance of the resin has decayed to 75%. Furthermore, the negative correlation of higher resin costs and the maximum achievable average operating profit is apparent in Figure 5. This relationship should not surprise. As the cost of replacing spent resin increases, resin exchanges are postponed further and further. This puts negative pressure on the average operating profit from three directions: (1) when resin exchange is carried out, higher one-time costs are incurred, (2) because the exchange is postponed, more batches are harvested at lower relative purification yields, and (3) because more batches are harvested before the resin is exchanged, operating costs from ongoing fermentation are incurred for longer. While the minimum allowed resin capacity had a strong negative impact on the earliness of the resin exchanges, its effect on the optimal average reward was modestly positive. Values for $\text{capacity}_{\text{MaxBatch}}$ were varied in the interval [100%,98% ..0%]. However, the maximum number of batches purified before the resin must be exchanged, MaxBatch , was not altered. Therefore, lower val-

ues of $\text{capacity}_{\text{MaxBatch}}$ imply a steeper decay profile after the fourth batch. Increasing minimum resin capacities showed a strong negative correlation with the earliness of resin exchange activities, as visualized in Figure 6. For higher minimum allowed purification yield and, therefore, a flatter resin decay pattern, more batches are purified before the earliest optimal resin exchange. For $\text{capacity}_{\text{MaxBatch}} \geq 90\%$, a premature exchange is no longer optimal. In these cases, the resin is always fully depleted and only exchanged in its final unusable state. In contrast to the effect which increasing resin costs had on the optimal average reward, higher minimum resin capacities led to slightly higher average operating profits (Figure 6). This positive correlation could be driven by the comparatively higher number of batches which can be purified at close to full yield due to the flatter slope of the resin's performance decline.

In providing an answer to RQ3, this contribution extends the existing literature on the scheduling of chromatography resin maintenance activities by a consideration of its system-level operational implications. Additional theoretical contributions are made by considering stochastic decay and different resin costs. The analyses in this work provide an understanding of the sensitivity of optimal maintenance timing and average profitability to these process parameter changes. The discussed sensitivities also aid managerial decision making under the uncertainties of biopharmaceutical production. The next section discusses the 2:1 case of two parallel production reactors being harvested into one chromatography step.

6.2. Optimal control of two parallel USPs and chromatography

In this section, the scenario of two parallel production reactors being harvested into one chromatography step in a 2:1 ratio is studied to answer RQ4. After describing the

⁶Average reward maximizing value function (normalized)

⁷Total expected discounted reward value function in the initial system state (normalized)

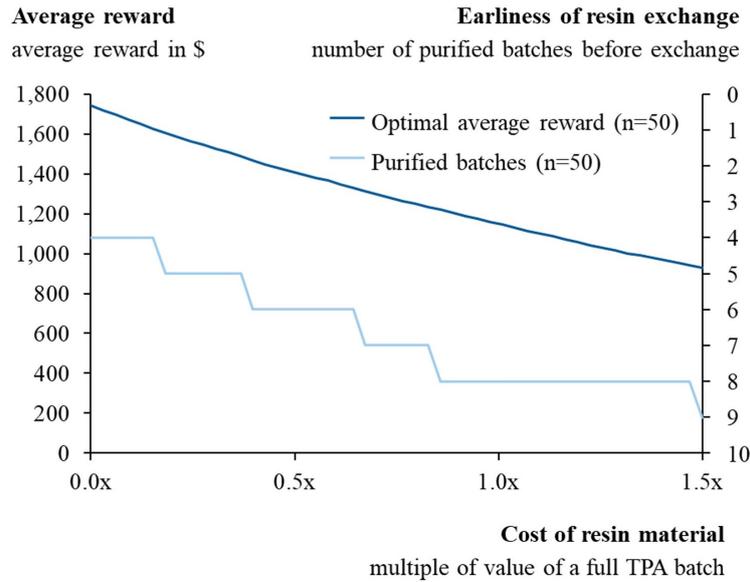


Figure 5: Sensitivity of the optimal average reward and the lateness of the first scheduled exchange to changes of the chromatography resin's material cost

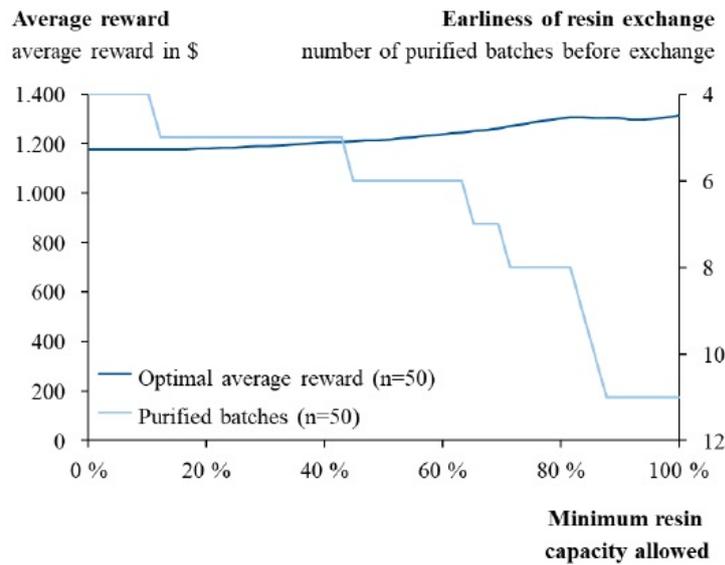


Figure 6: Sensitivity of the optimal average reward and the lateness of the first scheduled exchange to changes of the minimum allowed chromatography resin capacity

necessary adjustments to the model (sub-section 6.2.1), the optimal control policy is discussed (sub-section 6.2.2) and its sensitivity to process parameter changes analyzed (sub-section 6.2.3).

6.2.1. Model adjustment for the 2:1 case

This sub-section introduces the adjustments to the proposed model which were necessary to study parallel production reactors. Both production reactors are inoculated by a single seed-train (see Appendix 2). For this to be feasible, the

following equation must hold

$$N_{PR} = \frac{46 \times 1}{\max(T)} \geq 2 \tag{16}$$

Therein, N_{PR} is the number of parallel production reactors which can be inoculated by a single seed-train, $1+1+8+36 = 46$ is the number of 12-hour cycles in a production reactor from an empty reactor to the culture's first growth state until the latest possible time of harvesting the batch (see optimal policy in Table 7), 1 is the number of seed-trains feeding into

the production reactors, and $\max(T)$ is the maximum seed-train occupancy time (Jagschies et al., 2018, p. 654). For $\max(T) \leq 23$, i.e., at most 23 12-hour cycles or 11.5 days, it is feasible for a single seed-train to inoculate two production reactors. Because seed-train dynamics are not explicitly discussed in this contribution, this is assumed to be given.

Due to the exponential growth of the state space from adding a second reactor, the original model had to be scaled down to make computation feasible. This process is described in the following paragraph before the analysis of RQ4 commences.

Both parallel production reactors and the first chromatography step must be represented to study RQ4. However, the addition of a second, identical production reactor to the model drastically increased the state space from $|\mathcal{S}| = (1 + 1 + 8 + 36 + 1) \times 12 = 564$ states to $|\mathcal{S}| = (1 + 1 + 8 + 36 + 1)^2 \times 12 = 26,508$

The resulting transition matrix had c.12.6 billion entries, which made the calculation of optimal policies infeasible on available equipment. Therefore, to reduce the state space of the MDP and escape the curse of dimensionality, the number of decision epochs was scaled down linearly by a factor of three. Each decision epoch, therefore, represented 36 hours, compared to 12 hours in the basic model. This reduced the growth phase from eight epochs to three, and the production phase from 36 to twelve epochs. Additional repercussions on some process dynamics include: the culture decline phase at the end of the production phase was reduced to two transitions (a sixth of the now 12 production states) with probabilities 84% and 34%, respectively, and G_p only included the last growth state. The size of the resulting state space was $|\mathcal{S}| = 3888$ and a transition matrix with c.272 million entries. The reduced scale made optimization of the 2:1 model feasible while holding the model robust (see robustness check for longer decision periods in the single reactor case above). Furthermore, the probability of a successful state transition had to be adjusted for the reduced number of cycles per week. Following Schmidt (1996) argumentation of a weekly contamination probability of 5% to 10% and the number of cycles per week scaled-down to from 14 to c.4.67, the probability of a successful cycle transition, $p(s^{U_i}, s'^{U_i}, a^{U_i})$, was assumed to be 97.8%.

6.2.2. Optimal control of the parallel reactors and the DSP

This sub-section discusses the optimal policy obtained during this variation of the TPA case study.

The optimal policy for the two production reactors if the chromatography resin is yet unused is presented in Table 10. Following the visualization scheme introduced in Section 6.1, the rows and columns represent the states of the two production reactors, respectively, while the remaining performance of the chromatography resin is held constant. The ordered triplets within each section prescribe the optimal action to take in the first reactor (represented by the rows), the second reactor (represented by the columns), and in the chromatography column. Average reward maximizing policies for

all other resin capacities are provided in Appendix 19 - Appendix 23. Policy-related insights remain robust under total expected discounted reward maximization, although slight individual differences in the optimal policies exist.

We can't robustly conclude the economic feasibility of producing TPA in two parallel production reactors under the assumed process parameters. The average reward maximizing policy for the control of two parallel production reactors and one chromatography column is $V^* = -24,122\$$ per decision epoch. However, under the total expected discounted reward criterion, the optimal value function in the initial state is $V^*(e, e, 100\%) = 827,959\$$, and, therefore, strongly positive.

The decision zones of both production reactors are fairly symmetric, independent of the remaining resin performance. We can deduce this visually. For example, regard the state pair (p_{10}, p_{12}) in Table 10, i.e., reactor one is in its tenth production cycle and reactor two is in its twelfth. The optimal action is $(3,6,2)$, i.e., to add production medium in the first reactor and harvest the second reactor. Conversely, in the state (p_{12}, p_{10}) , the optimal actions are mirrored: $(6,3,2)$. This symmetry holds for all USP reactor-state pairs except those in which both reactors are in the same state within the last third of their production phase, i.e., $(s^{U_i}, s^{U_i}) \in \{(p_j, p_j) : 9 \leq j \leq 12\}$. In these states, both reactors have a TPA concentration of at least 24.4 mg/L (batch value of at least 93,556\$).

This symmetry should not surprise the attentive reader. Contrary to a serial set-up, such as the one modeled by Schmidt (1996) where intra-reactor transfers between sequentially operating bioreactors are possible, interdependencies between parallel reactors only come into play in harvesting decisions. As described earlier (Section 5.2), both reactors can't be harvested simultaneously. Additionally, because purification only takes one decision epoch (by assumption, as discussed in Section 5.4), purification of one reactor's medium never blocks the other reactor.

Observation 4: The decision zones of two parallel production reactors exhibit strong symmetry due to limited interdependency between the reactors them.

For the optimal average financial reward, both reactors are operated in a "staggered" fashion. Following the first harvested batch, one reactor is delayed from the other by one cycle. Regard the optimal path of a system with two parallel bioreactors and one DSP, starting from its initial state $(e, e, 100\%)$, in Table 11. Because simultaneous harvesting is not feasible, reactor two is harvested prematurely by one production cycle (in p_9) and a new batch started, followed by harvesting reactor one in the next decision epoch. This leads to reactor two being one cycle further in its process than reactor one because the chromatography step only takes one decision epoch. This allows harvesting future batches of both reactors in the optimal tenth production cycle. For all subsequent batches until the tenth, it is optimal to harvest in their tenth production cycle, i.e., p_{10} , or on the 13th production phase day. These yield revenues of 105,251\$, 105,251\$,

Table 10: Optimal policy for the operation of two parallel bioreactors into a single chromatography column, before the first and second purified batch

	<i>e</i>	<i>r</i>	<i>g</i> ₁	<i>g</i> ₂	<i>p</i> ₁₀	<i>p</i> ₁	...	<i>p</i> ₈	<i>p</i> ₉	<i>p</i> ₁₀	<i>p</i> ₁₁	<i>p</i> ₁₂	<i>u</i>
<i>e</i>	551		521			531				562			561
<i>r</i>													
<i>g</i> ₁	251		221			231				262			261
<i>g</i> ₂													
<i>p</i> ₁₀										362			
<i>p</i> ₁	351		321			331			Continue production in reactor one, harvest reactor two			361	
...													
<i>p</i> ₈													
<i>p</i> ₉													
<i>p</i> ₁₀	652		622			632							662
<i>p</i> ₁₁													
<i>p</i> ₁₂												162	
<i>u</i>	651		621			631				662			661

105,251\$, 99,988\$, 95,726\$, 89,463\$, 84,201\$, 78,938\$, 73,676\$, respectively. The eleventh batch is then purified in its eleventh production cycle, generating 76,015 \$ of operating revenue. Afterward, the resin is exchanged for the first time.

Furthermore, in states where the resin is already depleted to its minimum capacity, one batch is repeatedly dumped while the other is continued. Regard the sequence (*r*, *g*₁, 65%), (*g*₁, *g*₂, 65%), (*r*, *g*₃, 65%), (*g*₁, *p*₁, 65%) in Table 11. Because the resin is at its minimum capacity, operating costs are spent on producing only one batch (in reactor two) while the other (in reactor one) is repeatedly dumped (realizing its salvage value) and restarted in the first growth cycle.

6.2.3. Sensitivity to process parameter changes

Finally, this sub-section analyzes the optimal policy’s sensitivity to process parameter changes.

Observation 5: Both optimal average rewards and earliness of resin exchange were strongly sensitive to changes in the cost of resin material and the minimum allowed resin capacity.

Decreases in the cost of resin material, on the one hand, led to decreasingly increasing optimal average rewards. Below resin costs equivalent to half of the value of a full TPA batch, they led to almost linearly earlier resin exchanges. Firstly, cheaper resins are necessary for the profitable production and purification of TPA under the proposed process parameters. The average operating profit was negative for resin material costs of 0.75 times the value of a full TPA batch. While the single production reactor setup was profitable for all resin costs, the system of two parallel reactors only breaks even for reductions at least past 51,193\$ (c.0.24x) (Figure 7). However, reductions past 31,503\$(c.0.3x), show only small marginal profit improvements. Marginal gains per additional 4,000\$ cost reductions diminish to under 10%. Therefore, even if resin material were free, only a maximum

of 937\$ of average operating profit per 36 hour-cycle would be earned. Additionally, the cheaper the resin material is, the earlier is its earliest exchange optimal. Figure 7 visualizes this correlation. Below the break-even point, the number of batches before the earliest resin exchange decreases almost linearly in resin cost decreases. On the other hand, both optimal average rewards and earliness of resin exchanges were decreasing in minimum allowed resin capacities. This contrasts with the single production reactor setup.

While delaying the earliest optimal resin exchange does not surprise, the decrease of the maximum average operating profit seems counterintuitive. Increasing the minimum allowed resin capacity flattens the decay profile and a higher number of batches can be harvested at relatively higher resin yields. Therefore, resin exchange becomes necessary later and later. Intuitively, harvesting more batches at almost full resin yield should also increase average rewards. This is not what we observe.

Per-cycle average fermentation operating costs surge in adding a second, parallel production reactor. It seems logical to conclude that the comparatively higher revenues from harvested batches can’t fully compensate for this downward pressure on average profits. With the flatter decay curve, average revenues from harvested batches increase (for same protein concentrations at the time of harvesting). This leads to postponing the earliest resin exchange. Because more batches are purified before the resin is exchanged, fermentation operating costs are incurred longer – an effect which the higher average revenues observably don’t fully counter.

Nevertheless, if minimum allowed resin performances are reduced below c.16%, thereby reducing the number of batches before the earliest resin exchange, operating the parallel production reactor setup turns profitable. At best, however, an average operating profit of only c.294 \$ is achieved. In answering RQ4, this section discussed the model adjustments necessary to study the simultaneous control of two parallel production reactors, discussed the optimal policies, and analyzed their sensitivity to process parameter changes.

Table 11: Optimal path of a 2:1 system, given no upsets

Path of system state left to right, starting at $(e, e, 100\%)$				
$(e, e, 100\%),$	$(r, r, 100\%),$	$(g_1, g_1, 100\%),$...	$(p_9, p_9, 100\%),$
$(p_{10}, r, 100\%),$	$(r, g_1, 100\%),$...	$(g_2, g_3, 100\%),$	$(g_3, p_1, 100\%),$
...	$(p_9, p_{10}, 100\%),$	$(p_{10}, r, 100\%),$	$(r, g_1, 100\%),$...
$(g_2, g_3, 100\%),$	$(g_3, p_1, 100\%),$	$(p_9, p_{10}, 100\%),$	$(p_{10}, r, 95\%),$	$(r, g_1, 95\%),$
...	$(p_9, p_{10}, 95\%),$	$(p_{10}, r, 90\%),$	$(r, g_1, 85\%),$...
$(p_9, p_{10}, 85\%),$	$(p_{10}, r, 80\%),$	$(r, g_1, 75\%),$...	$(p_9, p_{10}, 75\%),$
$(p_{10}, r, 70\%),$	$(r, g_1, 65\%),$	$(g_1, g_2, 65\%),$	$(r, g_3, 65\%),$	$(g_1, p_1, 65\%),$
$(r, p_2, 65\%),$	$(g_1, p_3, 65\%),$...	$(p_6, p_{11}, 65\%),$	$(p_7, r, \Delta),$
$(p_8, g_1, 100\%),$	$(p_9, g_2, 100\%),$	$(r, g_3, 100\%),$...	

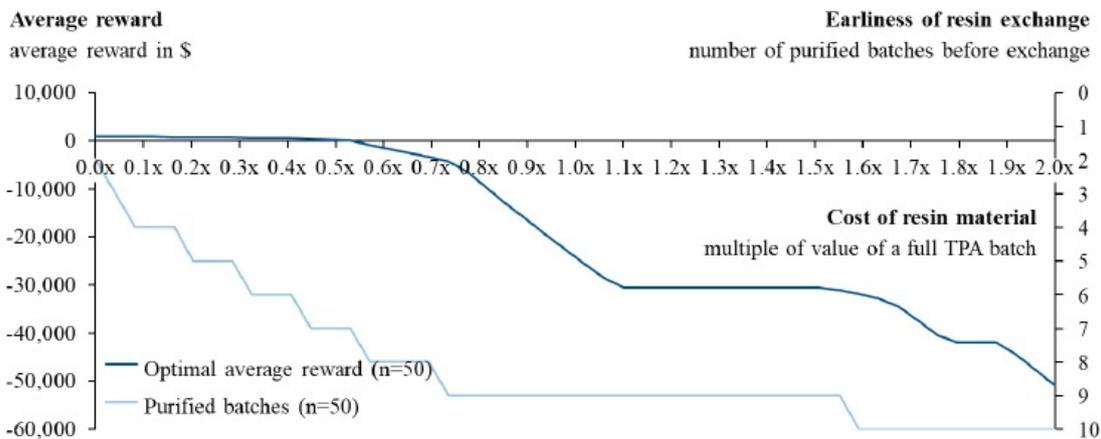


Figure 7: Sensitivity of the optimal average reward to changes of the chromatography resin’s material cost

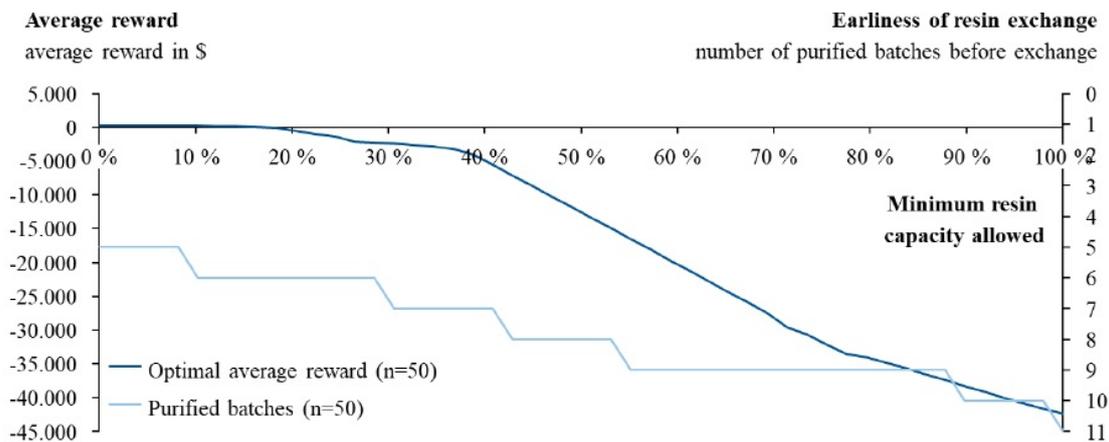


Figure 8: Sensitivity of the optimal average reward to changes of the minimum allowed chromatography resin capacity

In conclusion, the production and fermentation of TPA may not be economically feasible under the proposed setup, however, we were able to generate insights related to the staggered operation of the reactors and necessary parameter changes to make the setup profitable. The next chapter concludes this thesis and provides directions for future research endeavors on the topic.

7. Conclusions & Outlook

Producing biopharmaceutical products is a complex endeavor due to the high stochasticity of the living organisms used in fermentation and uncertainties about purification yields. Due to the high economic and curative value of biopharmaceutical APIs, optimal decision making during

their manufacture is especially important. Currently, operations research methods are under-appreciated in practice, even though, the academic discussion on the topic has been growing since the advent of biopharmaceutical production processes.

Academia has produced contributions on three different abstraction levels: single bioreactor control, capacity and operations planning, and the system-level control of integrated production processes. Literature on the bioreactor-level has dealt mostly with the maximization of protein concentrations constrained by dynamic models of cell-level kinetics. The capacity planning-level considers scheduling interdependencies between up- and downstream operations and the notion of uncertain or decaying chromatography resin yield. Furthermore, it introduced maintenance activities related to the exchange of spent chromatography resin. System-level literature, on the other hand, has modeled the dynamic control of entire production systems. However, it lacks work which considers integrated fermentation and purification control under relevant process uncertainties.

This work, therefore, was motivated by the paucity of literature examining the system-level control of interdependent up- and downstream processes under stochastic fermentation and resin decay. Thus far, no paper has simultaneously considered the upstream fermentation and downstream resin exchange control-problems. In the former sub-problem, decisions about growing the culture, its conversion into the protein-producing phase, and when to harvest the protein from one or multiple parallel production reactors must be made under the risk of costly batch failure due to contamination. After each purified batch, the capacity of the used chromatography resin to bind the protein of future harvests deteriorates stochastically. In the resin exchange sub-problem, therefore, decisions must be made under the trade-off between potentially reduced yields from future harvests and costly maintenance activities to restore the resin to its full performance.

This thesis aimed at answering the following four research questions. What is the optimal, simultaneous control policy for the USP and DSP? When is chromatography resin exchanged under stochastic decay? When under different minimum allowed resin capacities and different resin costs? How does the consideration of two parallel production reactors change the optimal policy?

To this end, an infinite horizon MDP with a discrete state space was proposed. The model allows for a single production reactor in the USP coupled with a single DSP, or, to depict common production realities, parallel USP production reactors coupled with a single DSP. It models the evolution of the production reactors over states, representing the cell culture lifecycle (growth, production, decline). Every decision epoch, the system controller decides whether to do nothing, prepare the empty bioreactor, add growth medium to facilitate culture growth, add production medium to start or continue protein production, or harvest the produced protein. When harvested, the production medium must be accepted into the chromatography column because intermediate stor-

age is not assumed. At some point, the spent chromatography resin is exchanged for fresh resin to restore full chromatography performance.

Furthermore, the case study of TPA fermentation and purification was analyzed and optimal policies interpreted. Both production setups with one production reactor and two parallel reactors were implemented and the sensitivity of critical process parameters was analyzed. The case study results were already summarized in their respective sections in Chapter 6. Here they are synthesized to answer the research questions of this work.

In answering RQ1, we thusly postulate three takeaways: Firstly, the state of the downstream process influences optimal upstream decision making and vice versa. This extends the existing research on system level control in which only one process part was considered in detail while the other was strongly abstracted or omitted. Secondly, batches should optimally be harvested before the culture enters decline. This corroborates prior findings (Martagan et al., 2016) and aids operative decision making. Lastly, a business case for the simultaneous stochastic optimization of up- and downstream decision making exists compared to a rudimentary policy focusing only on the output of the upstream process.

For RQ2, we conclude that premature resin exchanges become optimal under stochastic decay. The timings are even earlier than existing findings under deterministic decay (Liu et al., 2014).

The findings support the hypotheses related to RQ3. Thusly, resin-related parameters affect the profitability and resin exchange timing as expected, aiding managerial decision making and capacity planners.

By considering two parallel production reactors in studying RQ4, we gained the following four insights: For one, the simultaneous operation of two parallel reactors was only interdependent at the point of harvesting, therefore exhibiting strong symmetry in their respective decision zones. Secondly, the financial results cause doubt over this production setup's financial viability. Because the parallelization of production reactors is a common practice, managers must critically examine their production lines for profitability. This corroborates findings on increasing costs of goods in the case of parallel USPs (Liu et al., 2016). Thirdly, parallel production reactors should be operated in a staggered fashion to compensate for the time it takes to harvest one batch. Lastly, contrasting the findings under RQ3, the results fail to support the hypothesis of how flatter resin decay affects financial results of parallel operation.

Nevertheless, generalizability of the presented research has some limitations and warrants future research on the topic.

The presented model incorporates a linear protein accumulation pattern, i.e., the amount of the API of interest in the production medium is linearly dependent on the number of passed production phase cycles. Martagan et al. (2016), however, argue for stochastic accumulation of protein during fermentation based on probability distributions from fermentation experiments. Future research, therefore, could extend

the presented discrete state space formulation into a continuous state space, allowing the modeling of stochastic accumulation based on experimentally derived probability distributions.

The model presented in this contribution makes some significant assumptions about the purification process. This contribution abstracts the purification decision space to whether to accept a harvest for chromatography. In reality, purification decisions are more intricate. [Martagan et al. \(2018\)](#), for example, model the practically relevant decision about which chromatography lanes to pool. Future research on the simultaneous control of fermentation and purification operations should include this level of complexity to better aid practical decision making.

This contribution studied the TPA production case study, which [Schmidt \(1996\)](#) proposed. While the argument for choosing this numerical exercise has been made, corroboration of the presented academic and managerial insights requires the study of additional cases. Future research could, for example, model the fermentation of an MAB, a class of biopharmaceuticals of high commercial prominence.

Furthermore, due to the lack of high-grade literature on the topic of stochastic chromatography performance decay, model assumptions regarding the decay probabilities should be validated in future experimental research.

Lastly, future research on the topic should draw a more concrete picture of the fermentation process by including additional process variables in the state description, such as nutrient concentration, pressure, pH, and oxygen, instead of merely state approximations. Furthermore, future models could consider more concrete control actions such as the substrate feed-rate to even better aid operations.

Ultimately, academia's progress in modeling bioprocesses continues to be hindered by our limited understanding of living organisms' complexities ([Jagschies et al., 2018](#), p. 98). However, this contribution aids practice and academia in increasing their understanding of controlling production systems considering the presented complexities and argues for the simultaneous optimization of up- and downstream operations.

References

- Abu-Absi, N. R., Kenty, B. M., Cuellar, M. E., Borys, M. C., Sakhamuri, S., Strachan, D. J., Hausladen, M. C., and Li, Z. J. Real time monitoring of multiple parameters in mammalian cell culture bioreactors using an in-line Raman spectroscopy probe. *Biotechnology and Bioengineering*, 108 (5):1215–1221, 2011. ISSN 00063592.
- Allmendinger, R., Simaria, A. S., Turner, R., and Farid, S. S. Closed-loop optimization of chromatography column sizing strategies in biopharmaceutical manufacture. *Journal of Chemical Technology & Biotechnology*, 89 (10):1481–1490, 2014.
- Ashoori, A., Moshiri, B., Khaki-Sedigh, A., and Bakhtiari, M. R. Optimal control of a nonlinear fed-batch fermentation process using model predictive approach. *Journal of Process Control*, 19(7):1162–1173, 2009.
- Banga, J. R., Alonso, A. A., and Singh, R. P. Stochastic dynamic optimization of batch and semicontinuous bioprocesses. *Biotechnology Progress*, 13(3): 326–335, 1997.
- Banga, J. R., Balsa-Canto, E., Moles, C. G., and Alonso, A. A. Dynamic optimization of bioprocesses: efficient and robust numerical strategies. *Journal of Biotechnology*, 117(4):407–419, 2005.
- Bellman, R. Dynamic programming. *Princeton University Press*, 89:92, 1957a.
- Bellman, R. A Markovian decision process. *Journal of Mathematics and Mechanics*, pages 679–684, 1957b.
- Boucherie, R. J. and van Dijk, N. M., editors. *Markov decision processes in practice*, volume 248 of *International series in operations research & management science*. Berlin, Springer, Cham, Switzerland, 1st edition edition, 2017. ISBN 978-3-319-47764-0.
- Chades, I., Chapron, G., Cros, M.-J., Garcia, F., and Sabbadin, R. MDPtoolbox: a multi-platform toolbox to solve stochastic dynamic programming problems. *Ecography*, 37(9):916–920 (v. 1.6), 2014. ISSN 0906-7590. doi: 10.1111/ecog.00888.
- Chaudhuri, B. and Modak, J. M. Optimization of fed-batch bioreactor using neural network model. *Bioprocess Engineering*, 19(1):71–79, 1998. ISSN 1615-7605.
- Datar, R. V., Cartwright, T., and Rosen, C.-G. Process economics of animal cell and bacterial fermentations: a case study analysis of tissue plasminogen activator. *Bio/Technology*, 11(3):349–357, 1993. doi: 10.1038/nbt0393-349.
- Delvigne, F., Lejeune, A., Destain, J., and Thonart, P. Stochastic models to study the impact of mixing on a fed-batch culture of *Saccharomyces cerevisiae*. *Biotechnology Progress*, 22(1):259–269, 2006.
- Farid, S. S. Process economics of industrial monoclonal antibody manufacture. *Journal of Chromatography B*, 848(1):8–18, 2007.
- Farid, S. S., Washbrook, J., and Titchener-Hooker, N. J. Decision-support tool for assessing biomanufacturing strategies under uncertainty: stainless steel versus disposable equipment for clinical trial material preparation. *Biotechnology Progress*, 21(2):486–497, 2005.
- Freitas, H., Olivo, J., and Andrade, C. Optimization of bioethanol in silico production process in a fed-batch bioreactor using non-linear model predictive control and evolutionary computation techniques. *Energies*, 10 (11):1–23, 2017.
- Gatica, G., Papageorgiou, L. G., and Shah, N. An aggregation approach for capacity planning under uncertainty for the pharmaceutical industry. *Foundations of Computer-Aided Process Operations*, 4:245–248, 2003.
- Gupta, A., Jain, R., and Glynn, P. W. An empirical algorithm for relative value iteration for average-cost mdps. In *2015 54th IEEE Conference on Decision and Control*, pages 5079–5084. IEEE, 2015. ISBN 1479978868. doi: 10.1109/CDC.2015.7403014. URL <https://doi.org/10.1109/CDC.2015.7403014>.
- Howard, R. A. *Dynamic programming and Markov processes*. Cambridge, MIT Press, Cambridge, MA, 1960.
- Jagschies, G., Lindskog, E., Lacki, K., and Galliher, P. M. *Biopharmaceutical processing: Development, design, and implementation of manufacturing processes*. Amsterdam, Elsevier, Saint Louis, UK, 2018. ISBN 9780128125526. URL <http://ebookcentral.proquest.com/lib/munichentech/detail.action?docID=5154046>.
- Jiang, C., Liu, J., Rubacha, M., and Shukla, A. A mechanistic study of Protein A chromatography resin lifetime. *Journal of Chromatography A*, 1216(31):5849–5855, 2009. ISSN 0021-9673.
- Johnston, S. C. The economic case for new stroke thrombolytics. *Stroke*, 41 (10 Suppl):S59–S62, 2010.
- King, A. J. and Wallace, S. W. Modeling with stochastic programming. In Mikosch, T., Resnick, S. I., and Robinson, S. M., editors, *Springer Series in Operations Research and Financial Engineering*. Berlin, Springer Science & Business Media, New York, NY, 2012.
- Lakhdar, K. and Papageorgiou, L. G. An iterative mixed integer optimisation approach for medium term planning of biopharmaceutical manufacture under uncertainty. *Chemical Engineering Research and Design*, 3(86):259–267, 2008. ISSN 0263-8762.
- Lakhdar, K., Zhou, Y., Savery, J., Titchener-Hooker, N. J., and Papageorgiou, L. G. Medium term planning of biopharmaceutical manufacture using mathematical programming. *Biotechnology Progress*, 21(5):1478–1489, 2005.
- Lakhdar, K., Farid, S. S., Titchener-Hooker, N. J., and Papageorgiou, L. G. Medium term planning of biopharmaceutical manufacture with uncertain fermentation titers. *Biotechnology Progress*, 22(6):1630–1636, 2006.
- Lee, J., Lee, S. Y., Park, S., and Middelberg, A. P. J. Control of fed-batch fermentations. *Biotechnology Advances*, 17(1):29–48, 1999.
- Liu, C., Gong, Z., Shen, B., and Feng, E. Modelling and optimal control for a fed-batch fermentation process. *Applied Mathematical Modelling*, 37(3): 695–706, 2013.
- Liu, S., Yahia, A., and Papageorgiou, L. G. Optimal production and maintenance planning of biopharmaceutical manufacturing under performance decay. *Industrial & Engineering Chemistry Research*, 53(44):17075–17091, 2014.
- Liu, S., Farid, S. S., and Papageorgiou, L. G. Integrated optimization of upstream and downstream processing in biopharmaceutical manufacturing under uncertainty: a chance constrained programming approach. *Industrial & Engineering Chemistry Research*, 55(16):4599–4612, 2016.
- Lourenço, N. D., Lopes, J. A., Almeida, C. F., Sarraguça, M. C., and Pinheiro, H. M. Bioreactor monitoring with spectroscopy and chemometrics: a review. *Analytical and Bioanalytical Chemistry*, 404(4):1211–1237, 2012. ISSN 1618-2650. doi: 10.1007/s00216-012-6073-9.
- Markov, A. A. The theory of algorithms. *Trudy Matematicheskogo Instituta Imeni VA Steklova*, 42:3–375, 1954.
- Martagan, T., Krishnamurthy, A., and Maravelias, C. T. Optimal condition-based harvesting policies for biomanufacturing operations with failure risks. *IIE Transactions*, 48(5):440–461, 2016.
- Martagan, T., Krishnamurthy, A., Leland, P. A., and Maravelias, C. T. Performance guarantees and optimal purification decisions for engineered proteins. *Operations Research*, 66(1):18–41, 2018.
- Martagan, T., Krishnamurthy, A., and Leland, P. A. Managing trade-offs in protein manufacturing: how much to waste? *Manufacturing & Service Operations Management*, Accepted/In press. ISSN 1523-4614.
- Nweke, M. C., Rathore, A. S., and Bracewell, D. G. Lifetime and aging of chromatography resins during biopharmaceutical manufacture. *Trends in Biotechnology*, 36(10):992–995, 2018. ISSN 0167-7799.
- Otto, R., Santagostino, A., and Schrader, U. Rapid growth in biopharma: challenges and opportunities, 2014. URL <https://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/rapid-growth-in-biopharma>.
- Pandian, B. J. and Noel, M. M. Control of a bioreactor using a new partially supervised reinforcement learning algorithm. *Journal of Process Control*, 69:16–29, 2018.
- Peroni, C. V., Kaisare, N. S., and Lee, J. H. Optimal control of a fed-batch bioreactor using simulation-based approximate dynamic programming. *IEEE Transactions on Control Systems Technology*, 13(5):786–790, 2005.
- Ponte, X., Barrigón, J. M., Maurer, M., Mattanovich, D., Valero, F., and Montesinos-Seguí, J. L. Towards optimal substrate feeding for heterologous protein production in *Pichia pastoris* (Komagataella spp) fed-batch processes under P AOX1 control: a modeling aided approach. *Journal of Chemical Technology & Biotechnology*, 93(11):3208–3218, 2018.
- Powell, W. B. Clearing the jungle of stochastic optimization. In *Bridging data and decisions*, INFORMS TutORials in Operations Research, pages 109–137. INFORMS, 2014.
- Puterman, M. L. *Markov decision processes: discrete stochastic dynamic programming*. Hoboken, John Wiley & Sons, 2014.
- Rader, R. A. What is a biopharmaceutical? (bio)technology-based definitions. *BioExecutive International*, 1:60–65, 2005.
- Rani, K. Y. and Rao, V. S. R. Control of fermenters – a review. *Bioprocess Engineering*, 21(1):77–88, 1999. ISSN 1615-7605.
- Rocha, M., Mendes, R., Rocha, O., Rocha, I., and Ferreira, E. C. Optimization

- of fed-batch fermentation processes with bio-inspired algorithms. *Expert Systems with Applications*, 41(5):2186–2195, 2014.
- Saucedo, V. M. and Karim, M. N. Experimental optimization of a real time fed-batch fermentation process using Markov decision process. *Biotechnology and Bioengineering*, 55(2):317–327, 1997. ISSN 00063592. doi: 10.1002/(SICI)1097-0290(19970720)55:2<textles s>317::AID-BIT9<textgreater>3.0.CO;2-L.
- Saucedo, V. M. and Karim, M. N. Real time optimal feeding in a fermentor using a Markov decision algorithm. *Computers & Chemical Engineering*, 22(4-5):547–558, 1998.
- Schmidt, R. L. A stochastic optimization model to improve production planning and R&D resource allocation in biopharmaceutical production processes. *Management Science*, 42(4):603–617, 1996.
- Siganporia, C. C., Ghosh, S., Daszkowski, T., Papageorgiou, L. G., and Farid, S. S. Capacity planning for batch and perfusion bioprocesses across multiple biopharmaceutical facilities. *Biotechnology Progress*, 30(3):594–606, 2014.
- Skolpap, W., Scharer, J. M., Douglas, P. L., and Moo-Young, M. Fed-batch optimization of alpha-amylase and protease-producing *Bacillus subtilis* using Markov chain methods. *Biotechnology and Bioengineering*, 86(6): 706–717, 2004. ISSN 00063592.
- Skolpap, W., Nuchprayoon, S., Scharer, J. M., Grisdanurak, N., Douglas, P. L., and Moo-Young, M. Fed-batch optimization of alpha-amylase and protease-producing *Bacillus subtilis* using genetic algorithm and particle swarm optimization. *Chemical Engineering Science*, 63(16):4090–4099, 2008.
- The MathWorks Inc. Matlab and statistics toolbox r2018b, 2018. URL <https://www.mathworks.com/products/matlab.html>.
- Tijms, H. C. *A first course in stochastic models*. Hoboken, John Wiley & Sons, Chichester, UK, 2003. ISBN 9780470013632. doi: 10.1002/047001363X.
- Walsh, G. Biopharmaceutical benchmarks 2014. *Nature Biotechnology*, 32(10):992–1000, 2014.
- Walsh, G. Biopharmaceutical benchmarks 2018. *Nature Biotechnology*, 36(12):1137–1145, 2018.
- Wechselberger, P., Sagmeister, P., and Herwig, C. Real-time estimation of biomass and specific growth rate in physiologically variable recombinant fed-batch processes. *Bioprocess Engineering*, 36(9):1205–1218, 2013. ISSN 1615-7605. doi: 10.1007/s00449-012-0848-4.
- Xing, Z., Bishop, N., Leister, K., and Li, Z. J. Modeling kinetics of a large-scale fed-batch CHO cell culture by Markov chain Monte Carlo method. *Biotechnology Progress*, 26(1):208–219, 2010.



Value Co-Creation: Eine Analyse von Treibern und Gestaltungsansätzen für B2B- vs. B2C-Kunden

Felix Konstantin Steinhardt

Freie Universität Berlin

Abstract

Value Co-Creation (CC) wird eine hohe Bedeutung zugesprochen, da sie ermöglicht, Kunden aktiv in die Wertschöpfung von Unternehmen einzubeziehen und so neue Produkte und Dienstleistungen zu kreieren, die optimal auf Bedürfnisse abgestimmt sind. In dieser Arbeit werden Beweggründe bzw. Treiber für die CC-Teilnahme aus Kundensicht identifiziert. Dies geschieht anhand exemplarischer Szenarien mit quantitativen Auswertungen. Hierbei wird erstmalig auf Gemeinsamkeiten und Unterschiede der Bedeutung der Treiber von CC bei business-to-business-Kunden (Geschäftskunden) vs. business-to-consumer-Kunden (Privatkunden) im direkten Vergleich eingegangen. Was bewegt die Kunden dazu, sich an anbieterinitiierten CC-Aktivitäten zu beteiligen? Wie unterscheidet sich die Bedeutung dieser Treiber für beide Kundentypen B2B vs. B2C im Detail? Bei B2B-Kunden werden hierfür zwei Betrachtungsebenen unterschieden: die organisationale, strategische Ebene des gesamten Unternehmens („firm-level“) und die individuelle Perspektive der Mitarbeiter eines Unternehmens, welches Geschäftskunde ist („individual-level“). Zusätzlich werden praktische Gestaltungsansätze für CC diskutiert. Die Ergebnisse tragen zum tieferen CC-Verständnis bei und geben dabei Anhaltspunkte für die Initiierung von CC-Projekten.

Keywords: Co-Creation; Customer Engagement; B2B-Treiber; B2C-Treiber; Kundenmanagement.

1. Einleitung

1.1. Relevanz der Thematik

„Der Kunde ist König!“¹ Dieser Slogan ist seit langem verbreitet und hat bis heute nichts an Aktualität verloren. Im Rahmen der Globalisierung stehen Unternehmen immer mehr im verstärkten Wettbewerb zueinander und haben sich gegen Konkurrenten durchzusetzen (Cunat und Guadalupe, 2009, S. 180). Um hierbei bestehen zu können, müssen die Produkte, Dienstleistungen und angebotenen Lösungen der Unternehmen die Bedürfnisse der Kunden besser erfüllen, als es die Angebote der Konkurrenz tun (Narver et al., 2004, S. 343). Die zunehmende Digitalisierung bietet neue Möglichkeiten, kundenorientiert Prozesse der Wertschöpfung zu gestalten (Pralhad und Ramaswamy, 2000, S. 80). So ist es bspw. möglich, über Onlineplattformen schnell eine große Reichweite zu schaffen, direkt potentielle Kunden anzusprechen und dabei Meinungen sowie neue Ideen einzuholen

(Breidbach und Maglio, 2016, S. 73; Dellarocas, 2003, S. 1408 f.).

Das Konzept der Value Co-Creation (VCC) hat in den letzten Jahren eine starke Aufmerksamkeit sowohl in der Forschung als auch in der Managementpraxis erhalten (Gemser und Perks, 2015, S. 660; Ranjan und Read, 2016, S. 290). Auch in Blogs werden immer wieder Ansätze dazu thematisiert. Co-Creation (CC)² wird oftmals eine hohe Bedeutung zugesprochen, da sie ermöglicht, Kunden aktiv einzubeziehen: „Die Stimme des Kunden wird zum entscheidenden Wettbewerbsvorteil im digitalen Zeitalter“ (Peuster, 2016). Es gibt unterschiedliche Definitionsansätze, auf die in Kapitel 2.1 näher eingegangen wird. Grundlegend beschreibt VCC die Zusammenarbeit verschiedener Stakeholder im Rahmen der Wertschöpfung von Unternehmen, das heißt z. B. bei der Entwicklung von Produkten oder Dienstleistungen (Ramaswamy, 2009, S. 11; Storbacka et al., 2016, S. 3008). Dabei kann CC mit unterschiedlichen Stellen intern, das heißt innerhalb einer Organisation, oder mit externen Par-

¹Für die bessere Lesbarkeit wird in dieser Arbeit auf eine geschlechterspezifische Sprache verzichtet. Wenn im Folgenden die männliche Begriffsform verwendet wird (z. B. Mitarbeiter), sollen damit immer ausdrücklich Personen aller Geschlechter gemeint sein.

²Die Begriffe Value Co-Creation und Co-Creation (als Kurzform) sind als synonym anzusehen und werden daher im Folgenden gleichbedeutend verwendet.

teien (z.B. Zulieferer, Geschäftspartner, Kunden) stattfinden. Ein besonderer Fokus der Marketingliteratur liegt auf der VCC gemeinsam mit Kunden (Pralhad und Ramaswamy, 2004a, S. 8). Dies stellt auch den Betrachtungsschwerpunkt dieser Arbeit dar.

Anbieter bzw. Unternehmen nutzen CC mit Kunden, um neue Produkte und Dienstleistungen zu kreieren, die bestmöglich auf die jeweiligen Kundenbedürfnisse abgestimmt sind. Die Aktivitäten dieser Kundenintegration sind dabei vielfältig (Füller, 2010, S. 98; Nambisan und Baron, 2007, S. 43). Ein prominentes Beispiel der CC im Konsumenten Umfeld ist das von McDonald's initiierte „MeinBurger“-Programm, bei dem sich Privatkunden online individuell Burger mit verschiedenen Zutaten zusammenstellen können. Die Nutzer bzw. Kunden können anschließend über die Kreationen abstimmen. Gewinnerburger werden zeitweise in das Angebot landesweit aufgenommen (Handelsblatt, 2018). Auch im Geschäftskunden Umfeld findet häufig CC statt (Kohtamäki und Rajala, 2016, S. 5). Beispielsweise starten IT-Anbieter Programme, bei denen Geschäftskunden in Workshops oder über Onlineplattformen aktuelle Herausforderungen und Problematiken schildern. In Zusammenarbeit mit ihren Kunden können die Anbieter dann Lösungen entwickeln, die auf die spezifischen Geschäftsanforderungen abgestimmt sind (Nambisan und Baron, 2007, S. 44; Petri und Jacob, 2016, S. 65 f.; Siepmann 2018).

Es gibt zahlreiche Studien zum Themenfeld der VCC mit verschiedenen Untersuchungsschwerpunkten. Die Arbeiten reichen von qualitativen Analysen der Offline-Co-Creation in Servicesystemen des öffentlichen Personennahverkehrs (Jaakkola und Alexander, 2014) über Online-Co-Creation in der Produktentwicklung (Füller, 2010) bis hin zu Metastudien, die versuchen, konzeptionelle Modelle publizierter Forschungsergebnisse zusammenzubringen (Ranjan und Read, 2016, S. 290; Storbacka et al., 2016). Die Begrifflichkeiten sind in der Literatur jedoch nicht einheitlich verwendet. Dies spiegelt sich auch in der Praxis wider. Neben Value Co-Creation werden häufig auch Bezeichnungen wie Customer Co-Creation, Customer Participation, Customer Integration, Customer Innovation, Customer Engagement, Co-Innovation oder Co-Production angewandt. Auch der Begriff Customization wird in diesem Zusammenhang immer wieder kontrovers diskutiert (Gemser und Perks, 2015, S. 661). Allein die Vielzahl dieser Begriffe macht deutlich, wie präsent die Thematik in der Marketing- und Managementliteratur bereits ist. Die unterschiedlichen Konzepte haben teilweise große inhaltliche Überschneidungen, setzen jedoch häufig einen anderen Fokus bei der Begriffsherleitung. Auch die Ergebnisse unterschiedlicher Arbeiten zur Thematik sind divers und zeigen – je nach konkret betrachtetem Szenario – meist einen positiven, seltener aber auch einen negativen Einfluss oder gar keinen Einfluss auf die Kundenzufriedenheit, Kaufbereitschaft und/oder Mund-zu-Mund-Propaganda (Dong und Sivakumar, 2017, S. 945; Jaakkola und Alexander, 2014, S. 257).

1.2. Zielsetzung, Methodik und Aufbau der Arbeit

Diese Masterarbeit befasst sich mit der VCC von Unternehmen mit Kunden. Für Unternehmen stellt sich immer wieder die Frage, wie sie Privat- und/oder Geschäftskunden zu einer Teilnahme an CC-Programmen bewegen können (O'Hern und Rindfleisch, 2008, S. 97). Hauptziel dieser Arbeit ist es, solche Beweggründe bzw. Treiber für die CC-Teilnahme aus Kundensicht zu identifizieren. Dabei soll auf die Gemeinsamkeiten und Unterschiede der Bedeutung der Treiber von CC bei business-to-business-Kunden (B2B, Geschäftskunden) vs. business-to-consumer-Kunden (B2C, Privatkunden) eingegangen werden. Was bewegt die Kunden dazu, sich an anbieterinitiierten CC-Aktivitäten zu beteiligen? Wie unterscheiden sich diese Treiber und deren Bedeutung für die beiden Kundentypen B2B vs. B2C im Detail? Bei B2B-Kunden werden hierfür zwei Betrachtungsebenen³ unterschieden: zum einen die organisationale, strategische Ebene des gesamten Unternehmens als Geschäftskunde („firm-level“, B2B-Organisation) und zum anderen die individuelle Perspektive der Mitarbeiter eines Unternehmens, welches Geschäftskunde ist („individual-level“, B2B-Mitarbeiter). Zusätzlich sollen mögliche Gestaltungsansätze für CC diskutiert werden. Durch welche Aktivitäten lässt sich CC, mit Blick auf die Kundentreiber, optimal für B2B- bzw. B2C-Kunden gestalten?

Die bisherige Literatur im Themenbereich der CC analysiert unterschiedliche Fragestellungen meist entweder mit Blick auf B2B- oder B2C-Kunden, ohne dabei näher auf einen direkten Vergleich dieser Kundentypen einzugehen. Beispielsweise werden in Studien Treiber von B2C-Kunden analysiert, jedoch wird kein Bezug zur B2B-Kundenmotivation hergestellt und umgekehrt. Ein quantitativer, direkter Vergleich der Treiber von B2C- vs. B2B-Kunden zur VCC, im Rahmen eines vergleichbaren Szenarios, bleibt bisher aus. Die direkte Gegenüberstellung ist aus theoretischer Sicht bedeutend, da sie in dieser Form erstmalig erfolgt. Dabei werden unterschiedliche Konzepte und Erkenntnisse zur VCC beider Forschungsbereiche (B2B und B2C) zusammengebracht. Ein solcher Ansatz kann zu neuen theoretischen Erkenntnissen führen (Mencarelli und Rivière, 2015, S. 201). Für die Praxis kann außerdem das Bewusstsein für die Wirkung der Treiber unterschiedlicher Kundentypen gestärkt werden. Dies gilt insbesondere für Anbieter, die sowohl Kontakte zu B2B- als auch zu B2C-Kunden pflegen. So schafft diese Arbeit erste Anhaltspunkte für die Gestaltung von CC-Aktivitäten in beiden Bereichen der Kundenbeziehung, da die Treiber der Kundentypen direkt vergleichend gegenübergestellt werden. Die Arbeit folgt so dem Aufruf von Jaakkola und Alexander, die zu mehr Forschung hinsichtlich VCC in Verbindung mit B2C- und B2B-Prozessen aufrufen (2014, S. 259). Auch Dong und Sivakumar äußern sich hierzu und stellen fest, dass durchaus Unterschiede der CC im B2B- und B2C-Umfeld bestehen, ohne dabei auf Details einzugehen:

³B2B-Ebenen bzw. B2B-Kundengruppenperspektiven in Anlehnung an Cater und Cater, 2009, S. 1162

„a significant body of innovation research resides in the business-to-business contexts and covers various types of innovation collaboration (e.g. business customers collaborating with suppliers to create upstream innovations). Customer-firm interactions in business-to-business contexts have different dynamics than those in business-to-customer [specifically consumer] contexts“ (2017, S. 953)

Mencarelli und Rivière (2015, S. 212) fordern in ihrer konzeptionellen Untersuchung zu empfundenen Werten in B2B- und B2C-Kontexten direkt einen „increased scope for value co-creation“, bei dem Erkenntnisse zur CC aus B2B- und B2C-Kundensicht zusammengebracht werden sollten.

Die zentrale Fragestellung der Arbeit lautet folglich: Was sind die Treiber, Motivationsfaktoren bzw. Erklärungsansätze aus Kundensicht für eine CC-Teilnahme und wie unterscheiden sich diese Treiber in der Bedeutung im Vergleich von B2B- und B2C-Kunden? Hier schließt sich die Frage nach optimalen Gestaltungsansätzen für Anbieter als Managementimplikation an.

Um die Fragestellung zu beantworten, finden unterschiedliche Methoden Anwendung. Zunächst werden bereits veröffentlichte Studien und existierende Literatur zum Themenfeld der VCC (und verwandten Begriffen) sowie von B2B- und B2C-Kundenbeziehungen betrachtet. Hierbei wird sukzessiv-iterativ vorgegangen, denn zentrale Literaturquellen von Publikationen werden ebenfalls in die Analyse einbezogen. Über die im akademischen Umfeld etablierten Onlineportale EBSCOhost und ScienceDirect wird mit Hilfe der Suchbegriffe

- „Co-Creation“
- „Customer Engagement“
- „Customer Participation“
- „B2B“ bzw. „business-to-business“
- „B2C“ bzw. „business-to-consumer“

im Bereich Abstract, Title, Keywords nach relevanten Artikeln gesucht. Die Suchergebnisse werden zunächst anhand der Titel auf mögliche Relevanz für diese Arbeit geprüft (Vorauswahl), das heißt, inwieweit der Artikel aufgrund des Titels voraussichtlich Erkenntnisse für die Zielsetzung der Arbeit beiträgt (Treiber zur CC aus Kundenperspektive und/oder Gestaltungsmöglichkeiten). In einem zweiten Schritt wird der Abstract übriggebliebener Artikel betrachtet, welcher Aufschluss über den tatsächlichen Informationsgehalt für die Zielsetzung geben soll (finale Quellenauswahl).

Zusätzlich werden die Erkenntnisse der Literaturanalyse durch Primärinformationen aus zwei qualitativen, semi-strukturierten Experteninterviews ergänzt (Broß, 2018; Siep-

mann, 2018).⁴ Auf Basis dieser theoretischen Hintergründe wird die Onlinebefragung von potentiellen B2C- und B2B-Kunden durchgeführt. Anhand fiktiver Szenarien werden die Probanden gebeten, entsprechende Treiber zu beurteilen. Diese quantitative Studie dient als Primärerhebung dazu, bisherige Erkenntnisse zu Kundentreibern der CC zu vertiefen. Erstmals werden hiermit Rückschlüsse in einem direkten Vergleich der Kundengruppen B2C, B2B-Mitarbeiter und B2B-Organisation ermöglicht.

Der Aufbau der Arbeit ist wie folgt gestaltet: Zunächst werden zentrale Begriffe und Konzepte aus dem VCC-Themenbereich erläutert und in einen Zusammenhang gebracht. Anschließend findet eine Auseinandersetzung mit charakteristischen Eigenschaften von B2B- und B2C-Kundenbeziehungen statt. Zentrale Aspekte werden im Verlauf auf den Bereich der VCC übertragen. Es folgt ein kurzer Überblick zu Beweggründen bzw. Treibern von Anbietern, CC-Programme mit Kunden zu initiieren. Die Ergebnisse der Literaturanalyse zu Kundentreibern einer CC-Teilnahme sind ebenfalls als kurzer Überblick zusammenfassend präsentiert. Im Folgenden wird auf den theoretischen Hintergrund der Gestaltungsmöglichkeiten und Phasen im Rahmen von VCC eingegangen. Auf Basis dieser Erkenntnisse wird das konzeptionelle Modell vorgestellt, welches die Zusammenhänge der CC aus Kundenperspektive verdeutlichen soll. Dabei werden Hypothesen über die Bedeutung einzelner Treiber von B2B- vs. B2C-Kunden und zu CC-Aktivitäten formuliert. Nach Erläuterung der Vorgehensweise der quantitativen Studie erfolgt die Diskussion und Interpretation der Ergebnisse. Im Anschluss kommt es zur Ableitung von Implikationen für die Managementpraxis. Abschließend wird auf Restriktionen der empirischen Studie eingegangen und ein Ausblick auf mögliche zukünftige Forschungsfelder in diesem Zusammenhang gegeben.

2. Theoretische Hintergründe

2.1. Begriffsklärung zu Value Co-Creation und verwandten Konzepten

2.1.1. Hintergrund der VCC in der Service-Dominant Logic

Der Begriff VCC wurde durch Vargo und Lusch (2004, S. 10) geprägt, indem sie den Gedanken der Service-Dominant Logic (SDL) formuliert haben. Dabei geht es um den Wandel von der bis dahin vorherrschenden Perspektive der Goods-Dominant Logic (GDL), bei der primär die rein anbietergesteuerte Produktion und der Verkauf von Gütern betrachtet wird, hin zur Service-Dominant Logic. Die SDL forciert eine neuere Sichtweise, bei der Dienstleistungen die grundlegende Basis des ökonomischen Handelns darstellen. Der Fokus liegt auf Austauschprozessen von Dienstleistungen, bei denen Kunden eine bedeutendere Rolle einnehmen:

⁴Für eine höhere Aussagekraft waren ursprünglich fünf Experteninterviews geplant. Aufgrund vertraglicher Umstände der Ansprechpartner, Unternehmenswechsel und zeitlichen Engpässen konnten jedoch nur zwei Interviews durchgeführt werden.

“tangible goods serve as appliances for service provision rather than ends in themselves. [...] It positions service [...] as the core of the firm’s mission. [...] Relationship building with customers becomes intrinsic not only to marketing but to the enterprise as a whole. All employees are [...] service providers with the ultimate goal of satisfying the customer” (Vargo und Lusch, 2004, S. 13)

Begründet wird die SDL anhand der „Foundational Premises“, welche fortlaufend aktualisiert und ergänzt wurden (Vargo und Lusch, 2004; Vargo und Lusch, 2008; Vargo und Lusch, 2016), vgl. Tabelle 10 im Anhang. Tatsächlich ist dieser Wandel weltweit bei Industrienationen zu beobachten, die sich zu modernen Dienstleistungsgesellschaften entwickeln (Statista, 2018b; Bundesamt, 2018). Die Interaktion mit Kunden ist dabei ein entscheidender Aspekt: „The customer is always a co-creator of value“ (Vargo und Lusch, 2008, S. 2). Nach neuerem Verständnis haben Kundenbeziehungen oftmals einen interaktiven Charakter, das heißt, Kunden nehmen häufiger aktive Rollen ein. Obwohl dieser Gedanke ursprünglich aus der Betrachtung von Dienstleistungen stammt, wird die gemeinsame Wertschöpfung im Sinne der CC inzwischen sehr vielseitig angewandt und diskutiert. Dies gilt bspw. für den Kontext der Produktentwicklung oder die Gestaltung neuer Dienstleistungsprozesse von Anbietern mit Kunden. So argumentieren einige Autoren bereits für die Notwendigkeit der Einführung einer weiteren neuen Sichtweise, die sog. Customer-Dominant Logic (Heinonen et al., 2010).

2.1.2. Definition von VCC und verwandten Begriffen

Value Co-Creation beschreibt die interaktive Zusammenarbeit verschiedener Stakeholder bei Wertschöpfungsprozessen von Unternehmen (Ramaswamy, 2009, S. 11; Storbacka et al., 2016, S. 3008). Dabei werden unterschiedliche Akteure mit ihren Ressourcen, wie z. B. ihrem spezifischen Wissen oder ihrer Arbeitszeit, in die Unternehmensprozesse eingebunden. Durch die Weitergabe eigener Ressourcen nehmen die Akteure eine aktive Rolle ein, die durch eine direkte oder indirekte Zusammenarbeit auf einer oder mehreren Ebenen im Geschäftsprozess gekennzeichnet ist (Payne et al., 2008, S. 84). Das Engagement geht dabei deutlich über ein Mindestmaß bzw. den üblichen Austauschprozess am Markt „Geld gegen Ware bzw. Dienstleistung“ hinaus (Pralhad und Ramaswamy, 2004b, S. 14). Co-Creation ist folglich ein kollaborativer Prozess, um gemeinsam neue Werte zu schaffen – sowohl im materiellen als auch symbolischen Sinne (Galvagno und Dalli, 2014, S. 644).

Grundsätzlich umfasst VCC nicht nur die Integration von Kunden, sondern auch die Zusammenarbeit mit anderen Akteuren wie Zulieferern, Vertriebspartnern oder anderen Dienstleistern (Vargo und Lusch, 2016, S. 6 f.; Ramaswamy, 2009, S. 11). In der Marketingliteratur konzentriert man sich jedoch auf VCC von Unternehmen mit Kunden, da Kunden als direkte Zielgruppe die wichtigsten Akteure dar-

stellen (Ranjan und Read, 2016, S. 292). So hat sich auch der Begriff der Customer Co-Creation entwickelt (Gemser und Perks, 2015, S. 661; Pee, 2016; S. 217). Dieser wird häufig parallel zu VCC verwendet und stellt streng genommen eine Unterkategorie der CC mit Kunden dar. Es gibt zahlreiche Beispiele für die Integration von Kunden in Entwicklungs- bzw. Herstellungsprozesse von Unternehmen bei Produkten und Dienstleistungen. Dazu zählen die Gestaltung von Textilien, Wohnmöbeln oder Infotainmentsystemen im Auto, aber auch neue Impulse für Softwarelösungen oder Dienstleistungsprozesse im Privat- oder Geschäftskundenumfeld (Füller, 2010, S. 109; Mustak et al., 2016, S. 260 f.; Petri und Jacob, 2016, S. 63). Der Großteil akademischer Publikationen verwendet dabei weiterhin die übergreifende Terminologie der VCC (Breibach und Maglio, 2016, S. 74; Kohtamäki und Rajala, 2016, S. 7). Da der Begriff in Forschung und Praxis geläufiger ist, wird dieser Ausdruck hier beibehalten.

VCC mit Kunden kann so betrieben werden, dass nur individuelle Kunden selbst davon profitieren (z. B. gemeinsames Erschaffen von Designs/Lösungsansätzen, die nur dem einen Kunden zur Verfügung stehen). Dies ist jedoch seltener der Fall. Häufiger kommt es zu CC-Aktivitäten, bei denen am Ende zumindest indirekt verschiedene Kunden profitieren können (z. B. gemeinsames Erschaffen von Designs/Lösungsansätzen, die mehreren Kunden zur Verfügung gestellt werden). Beide Arten treten im B2C- und B2B-Kundenumfeld auf (Dong und Sivakumar, 2017, S. 952; Witell et al., 2011, S. 143). Auch wenn dabei im B2B-Kontext eher Lösungen für einen spezifischen Geschäftskunden erarbeitet werden, lernen Anbieter daraus und können Anwendungsfälle ggf. auf andere Kundensituationen übertragen. Allerdings wird dies in Studien meist nicht differenziert betrachtet. Vielmehr wird betont, dass B2C-/B2B-Kunden bei CC immer eine aktive Rolle einnehmen.

Während einige Publikationen unterschiedlich über den Stellenwert von Co-Production und Value-in-Use in Verbindung mit Value Co-Creation urteilen (Etgar, 2007, S. 97; Galvagno und Dalli, 2014, S. 644; Kohtamäki und Rajala, 2016, S. 5, 7), identifizieren Ranjan und Read diese Konstrukte als essentielle Teildimensionen von VCC. Ferner erläutern sie zentrale Elemente, die ein umfassendes Bild der VCC-Definition ermöglichen.

Co-Production, als erste Teildimension der VCC, zielt auf den Aspekt des Ressourcenaustausches der Akteure im Rahmen der Lösungsentwicklung ab, das heißt die aktive Integration der Kunden in den Wertschöpfungsprozess des Unternehmens:

„Co-production consists of direct or indirect, co-working with customers‘ [...] or participation in the product/service design process“ (Ranjan und Read, 2016, S. 292).

Zentrale Elemente der Co-Production sind 1. Wissensaustausch der Kunden hinsichtlich Kreativität und neuer Ideen („Knowledge sharing“), 2. Fairness/Gerechtigkeit in Form von Kundenfokus, Kontrollabgabe und Umgebungsbereitstellung des Anbieters („Equity“) und 3. Interaktionsprozesse

(„Interaction“) zwischen Anbieter und Kunde (ebd., S. 295).

Value-in-Use hingegen, als zweite Teildimension der VCC, betrachtet das Phänomen, dass der eigentliche Wert von Anbieterlösungen erst während der Nutzung durch und für den Kunden geschaffen wird (Vargo und Lusch, 2016, S. 6).

„While value can be derived through interaction with the firm and its offerings, it can also arise through a process of consumption, which may be mostly independent of the company’s intervention or exchange“ (Ranjan und Read, 2016, S. 293)

Die wesentlichen Elemente hier sind 4. Erlebnisse für den Kunden schaffen („Experience“), 5. Personalisierung der Nutzungsprozesse für Kunden („Personalization“) und 6. Beziehungsmanagement („Relation“) zwischen Anbieter und Kunden aufbauen (ebd., S. 295).

Im Zusammenhang mit VCC werden in der Literatur häufig auch andere Begrifflichkeiten verwendet. Dazu gehören bspw. die übergeordneten Begriffe der Customer Integration und Customer Innovation. Aber auch die Terminologien Customization, Customer Participation und Customer Engagement werden oftmals genutzt, um die Integration von Kunden in die unternehmerische Wertschöpfung zu analysieren. Es gibt zwar starke inhaltliche Überschneidungen, die Konzepte sind aber theoretisch nicht vollends deckungsgleich.

Abbildung 1 stellt verwandte Begriffe und Konzepte zu VCC in einer Übersicht dar. Dabei sind Überkategorien in Kreisen visualisiert, um Teilbereiche und Überschneidungen zu veranschaulichen. Diese Abbildung erhebt keinen Anspruch auf Vollständigkeit, das heißt, dass nur zentrale Aspekte berücksichtigt sind. Es folgt eine kurze Erläuterung der Begriffe, wobei für weitere Details diesbezüglich auf die angegebenen Literaturquellen verwiesen wird.

Customer Integration beinhaltet als Überkategorie auch Prozesse, bei denen Kunden nur passiv integriert werden, das heißt ohne eine aktive Rolle im Sinne der CC mit Interaktion und Engagement dem Anbieter gegenüber einzunehmen (Bruhn und Strauss, 2009, S. 6; Kurzmann und Reinecke, 2009, S. 199). Werden bspw. einem Kunden bei einem Friseur die Haare geschnitten, so muss sich dieser nicht zwangsläufig mit eigenen Ideen aktiv einbringen. Gleiches gilt für das Messen von Nutzungsverhalten über Datenauslesung bei technischen Anwendungen. Jeder CC-Prozess ist somit eine Form der Kundenintegration, aber nicht jede Kundenintegration führt zu VCC im engeren Sinne. Customer Innovation hingegen betrachtet Prozesse der CC im Innovationsumfeld (Reichwald/Piller 2005, S. 2, 4). Auch wenn VCC besonders häufig in diesem Kontext Anwendung findet, ist sie nicht darauf beschränkt, vgl. Kapitel 3.3. Bei Customization im engeren Sinne werden Kunden häufig nur in späten Phasen der Lösungsentwicklung integriert und nur in limitierter Art, da sie dabei aus vorgefertigten Optionen wählen (Kaplan und Haenlein, 2006, S. 170). Customer Participation hat seinen Ursprung im Dienstleistungsumfeld und beinhaltet auch Formen, die nur eine passive Einbeziehung von Kunden betrach-

ten (Dong und Sivakumar, 2017, S. 950). Customer Engagement aggregiert multiple Wege, wie Kundenverhalten über Standardtransaktionen hinaus Firmen beeinflussen kann und kommt damit dem Verständnis der VCC sehr nahe (Jaakkola und Alexander, 2014, S. 248).

Letztlich lassen sich alle diese Begriffe jedoch auf den Umstand zurückführen, dass Kunden in Wertschöpfungsprozesse von Anbietern mehr oder weniger aktiv miteinbezogen werden. Daher sind grundsätzlich Publikationen zu all diesen Terminologien in der Literaturanalyse zu VCC mit berücksichtigt. Nach Erläuterung zentraler Aspekte der VCC, sowie dem Überblick zu verwandten Begriffen und Konzepten, soll für den folgenden Teil der Arbeit die unten stehende Definition als grundlegender Bezugspunkt dienen: Der Begriff Value Co-Creation beschreibt Prozesse, in denen Produkte und/oder Dienstleistungen gemeinsam von Unternehmen mit Kunden entwickelt werden. Dies geschieht durch direkte oder indirekte Zusammenarbeit und Ressourcenaustausch über eine oder mehrere Ebenen hinweg. Dabei nehmen Kunden eine aktive Rolle ein und schaffen neue Werte in Kooperation mit Anbieterunternehmen. Außerordentliches Engagement, Interaktion und Erfahrungsaustausch sind wichtige Elemente der Co-Creation. Grundsätzlich nimmt VCC eine übergeordnete Rolle zu den Konzepten der Co-Production und Value-in-Use ein, das heißt, diese Begriffe sind als Teildimensionen von VCC anzusehen.⁶

2.2. Merkmale von B2B- und B2C-Kundenbeziehungen

Der Unterschied zwischen B2B- und B2C-Kundenbeziehungen liegt in den verschiedenen Kundentypen. Während beim business-to-business-Geschäft Unternehmen bzw. Organisationen Geschäftskunden eines Anbieters sind (Kleinaltenkamp und Saab, 2009, S. 1), stellen Endverbraucher bzw. Konsumenten die Kunden im business-to-consumer-Geschäftsmodell dar. Wertschöpfung findet in beiden Bereichen statt. Eine starke Interaktion zwischen Anbietern und Kunden zeigt sich naturgemäß im B2B-Umfeld, da Unternehmen häufiger enge Bindungen eingehen und spezielle Lösungen gemeinsam erarbeiten (Petri und Jacob, 2016, S. 63; Vargo und Lusch, 2016, S. 7). So liegt der Ursprung der akademischen Auseinandersetzung mit VCC im B2B-Kontext (Vargo und Lusch, 2011, S. 181). Die Bedeutung dieses Sektors zeigt sich auch darin, dass 70% aller Umsätze in entwickelten Volkswirtschaften über B2B-Beziehungen erzielt werden (Kleinaltenkamp und Saab, 2009, S. 3).

Die Systematik von B2C und B2B ist in Abbildung 2 dargestellt und soll mit den erläuterten Beispielen verdeutlicht werden.

Wenn bspw. IBM als Anbieter von IT-Lösungen mit einem anderen Unternehmen wie Bosch, BMW oder Bayer Geschäfte macht und Softwareanwendungen verkauft, liegt eine business-to-business-Kundenbeziehung vor. Im B2B-Bereich

⁶Eigene Definition, in Anlehnung an Erläuterungen von Jaakkola und Alexander, 2014; Payne et al., 2008; Prahalad und Ramaswamy, 2000; Ramaswamy, 2009; Ranjan und Read, 2016

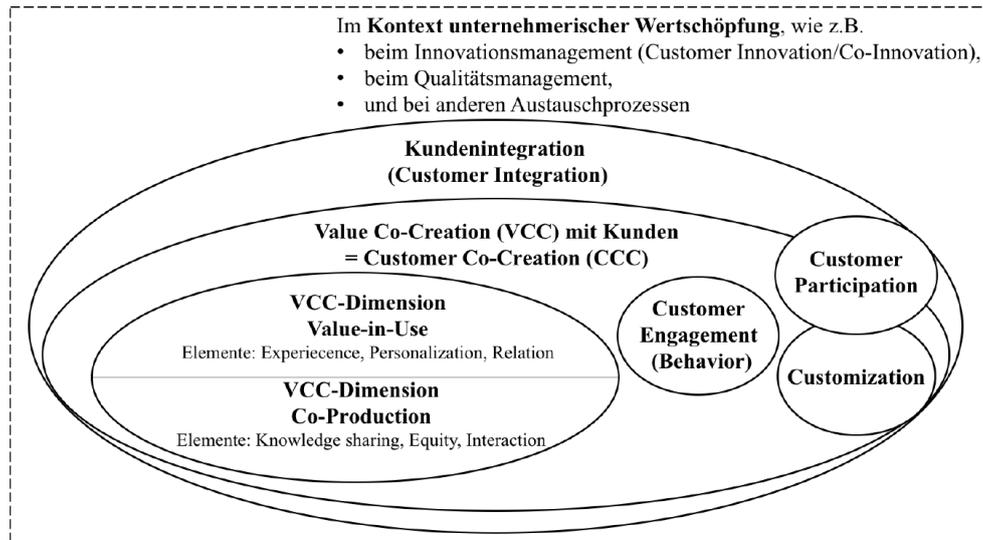


Abbildung 1: Überblick zu VCC und verwandten Begriffen⁵

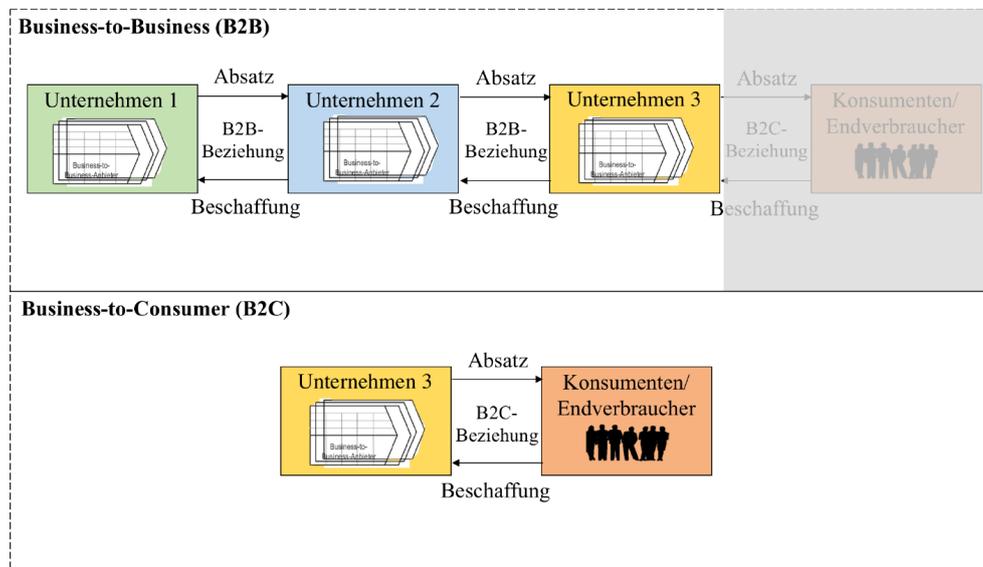


Abbildung 2: Systematik von B2B- und B2C-Kundenbeziehungen⁷

kann eine unterschiedliche Anzahl an Unternehmen beteiligt sein, ehe schließlich ein Unternehmen steht, das direkt Geschäfte mit Privatkunden bzw. einem Konsumenten macht (B2C). Wenn z. B. EDEKA als Supermarkt Geschäfte mit Einzelpersonen macht, die Lebensmittel für den individuellen Verzehr kaufen, wird von einer business-to-consumer-Beziehung gesprochen.

Die genannten Beispiele sind eindeutig. Es gibt jedoch auch Anbieterunternehmen, die sowohl business-to-business- als auch business-to-consumer-Beziehungen pflegen bzw. strategische Implikationen beider Bereiche beachten müssen. Unilever vertreibt als Anbieter viele unterschiedliche Marken, die direkt Konsumenten ansprechen und zum Kauf anregen sollen, wie z. B. Langnese Eiscreme und Axe Deodorant (Deutschland, 2017). Die einzelnen Unilever-Marken

formulieren ihre Marketingstrategie mit Blick auf Konsumenten (B2C) als Zielgruppe. Umsatz generiert Unilever jedoch mit dem Verkauf der Markenprodukte an Handelsunternehmen (B2B). Erst diese Händler erreichen ihre Umsätze mit dem Weiterverkauf an Konsumenten (B2C). Ein weiteres Beispiel ist der IT-Anbieter Microsoft. Das Unternehmen hat viele B2B-Kunden, die Business-Lizenzen für Office-Produkte (Word, PowerPoint, Excel) erwerben. Außerdem zielt Microsoft auch darauf ab, Privatanutzer anzusprechen, die Office-Produkte einsetzen. Die Trennlinie zwischen B2B- und B2C-Kundenbeziehungen ist nicht immer eindeutig zu ziehen.

⁵Eigene Darstellung

⁷Eigene Darstellung, in Anlehnung an Kleinaltenkamp und Saab, 2009, S. 2

Der Begriff Kunde („customer“) ist daher zunächst nicht spezifisch, denn sowohl Unternehmen/Organisationen (B2B) als auch Privatpersonen (B2C) können Kunde eines Anbieters sein. Die Marketingaktionen sollten an die jeweiligen Kundentypen angepasst werden. Häufig haben Unternehmen als Kunde auf organisationaler, strategischer Ebene andere Kauf- und Nutzungskriterien als private Einzelpersonen. In Kaufprozessen verfolgen Unternehmen bei der Beschaffung bestimmte Ziele, die von Anbietern zu erfüllen sind. Diese sind Teil des Zielsystems des gesamten Unternehmens und lassen sich in Teildimensionen gliedern (Günter und Kuhl, 2000, S. 377 ff.):

- Kostenziele (Kosten minimieren, die mit der Anschaffung verbunden sind)
- Qualitätsziele (Beitrag der Leistung zum Unternehmenserfolg)
- Risikoziele (Risikoreduktion durch Vertragsklauseln oder Eigenherstellung)
- Flexibilitätsziele (möglichst hohe Ressourcenflexibilität schaffen)

Ist ein Unternehmen Geschäftskunde eines Anbieters, sind verschiedene Rollen bzw. Ebenen beachtenswert. So befasst sich die Marketingliteratur mit dem Konzept des Buying-Centers (Brinkmann und Voeth, 2007, S. 1001). Strategische Entscheidungen in Unternehmen werden meist kollektiv in Gremien getroffen. Das Buying-Center ist das Einkaufsgremium einer Organisation, das heißt eine Gruppe von Personen, die am Einkaufsprozess beteiligt sind und als Mitarbeiter des Unternehmens unterschiedliche Rollen einnehmen. Ziel ist es, die beste Entscheidung für das Unternehmen aufgrund gemeinsamer Erfahrungen zu treffen. Die verschiedenen Rollen innerhalb des Buying-Centers sind von Anbieterseite bei der Kundenansprache zu beachten (Töllner et al., 2011, S. 713).

Im Buying-Center haben die Mitarbeiter unterschiedliche Rollen, u. a. als

- „Gatekeeper“, verteilen oder blocken von Informationen (z. B. Sekretär/In)
- „Influencer“, sind Fachexperten im Unternehmen (z. B. Referenten)
- „Einkäufer“, verantworten Kaufprozesse (z. B. Sachbearbeiter im Einkauf)
- „Anwender“, erkennen Bedarf und regen Kauf an (z. B. Mitarbeiter einer Fachabteilung)
- „Entscheider“, treffen endgültige Managemententscheidungen zur Geschäftsbeziehung mit Anbietern (z. B. Bereichsleiter mit Budgetverantwortung)

Je nach Rolle benötigen die Mitglieder einen engen Dialog, Produktargumente, tiefgehende Informationen, Nutzenargumente, transparente Preise und Sicherheit im unternehmerischen Kaufprozess (ebd., S. 713 f.). Buying-Center können variieren und je nach Kontext unterschiedlich zusammengesetzt sein.

Cova und Salle (2007, S. 8) beschreiben Unterschiede des B2B- und B2C-Kundenmarketings anhand verschiedener Perspektiven im Zeitverlauf. Demnach seien Einkaufs- und Servicetransaktionen im B2C-Kontext eher isolierte Events mit emotionaler Bindung, während im B2B-Umfeld von langfristigen, rationalen Beziehungen auszugehen ist. Konsumenten können passive und aktive Rollen einnehmen. Im B2B-Kontext haben Kunden grundsätzlich eine höhere Anbieterinteraktion. Der B2C-Markt ist geprägt durch viele unabhängige Konsumenten und Anbieter. B2B-Märkte hingegen bestehen aus wenigen unabhängigen, aber über Netzwerkstrukturen verbundenen Kunden und Anbietern. Auch auf führenden Social-Media-Plattformen werden Unterschiede diskutiert. So schreibt Nasr, dass die Hauptdifferenzen zwischen B2B- und B2C-Kundenbeziehungen anhand folgender Aspekte zu bestimmen sind (2017): Komplexität industrieller Produkte und Dienstleistungen, Diversität von industriellen Bedürfnissen, signifikante Unterschiede bei der Anzahl an Kunden, höhere Umsätze pro Kunde (B2B), Kundenrationalität (B2B), Verlauf der Kaufprozesse und grundsätzlich hohe Interaktion in Geschäftsprozessen zwischen Anbieter und Kunde (B2B).

Foscht et al. (2017, S. 14 ff.) stellen Eigenschaften von privatem und organisationalem Handeln gegenüber. Dabei fallen Differenzen auf, die den bereits erläuterten Aspekten ähneln. Beim privaten wirtschaftlichen Handeln besteht meist eine „relative Kurzfristigkeit des Beziehungsgefüges“, während beim organisationalen Verhalten eine „relative Langfristigkeit“ charakteristisch ist. Privat gibt es eine geringere Prozessorientierung bei der Beschaffungsentscheidung; organisationale Entscheidungen sind stark prozessorientiert. Der Formalisierungsgrad von Beschaffungs- und Nutzungsprozessen ist bei B2B-Kunden deutlich höher als bei B2C-Kunden. B2B-Entscheidungen sind zudem geprägt von „Fremddeterminiertheit“ (z. B. aus vorgegebenen Partnern wählen) und „Multipersonalität/-organisationalität“. Bei B2C-Kunden werden üblicherweise individuelle Entscheidungen getroffen. Es liegen im Vergleich zu organisationalen Entscheidungen deutlich geringere Anreiz- und Sanktionsmechanismen vor (z. B. Provision, Quotendruck). Während privat oft emotionale und impulsive Entscheidungen getroffen werden können, spielen organisational eher rationale, objektiv nachvollziehbare Entscheidungen eine Rolle. Selektive Wahrnehmungen, soziale Beeinflussbarkeit und das Streben nach Vereinfachung im Umfeld der B2C-Kunden stehen der vermeintlich objektiven Beurteilungsfähigkeit, „unermüdlichen Informationsverarbeitung“ und Einbindung in faktenorientierten Entscheidungsgremien von B2B-Kunden gegenüber. Allerdings seien die Grenzen nicht immer eindeutig zu ziehen und auch ambivalentes Verhalten beobachtbar. Solche zentralen Unterschiede im Kundenverhalten prägen

voraussichtlich auch die Bedeutung einzelner Kundentreiber für die Teilnahme an CC.

3. Aktueller Kenntnisstand zu Treibern und Gestaltungsmöglichkeiten von VCC mit Kunden

Grundsätzlich lassen sich unterschiedliche Treiber heranziehen, die Akteure dazu bewegen, sich an VCC zu beteiligen. Dabei ist zwischen der Anbieter- und der Kundenperspektive zu unterscheiden. Die Anbieterperspektive betrachtet strategische Treiber der Unternehmen, die Produkte bzw. Dienstleistungen am Markt anbieten und VCC zur Gestaltung der Angebote initiieren. Die Kundenperspektive hingegen geht auf die Gründe von Kunden ein, sich an VCC-Aktivitäten der Anbieter zu beteiligen.

3.1. Überblick zu Beweggründen von Anbietern

Der Hauptgrund von Unternehmen für die Initiierung von VCC-Aktivitäten mit Kunden ist, verbesserte Produkte und Dienstleistungen gemeinsam zu kreieren, die optimal auf Kundenbedürfnisse abgestimmt sind. Dies gilt prinzipiell sowohl für B2B-Kunden als auch für B2C-Kunden (Petri und Jacob, 2016, S. 63; Ramaswamy, 2009, S. 11). Sofern dabei tatsächlich Lösungen entstehen, die den Bedürfnissen der Kunden mehr entsprechen, als es bei den Angeboten konkurrierender Anbieter der Fall ist, kann von nachhaltig geschaffenen Wettbewerbsvorteilen gesprochen werden (Porter, 2010, S. 27). Für die strategische Ausrichtung von Anbieterunternehmen nimmt VCC mit Kunden daher zunehmend eine entscheidende Rolle ein.

Durch die Kundenintegration in Entwicklungsprozessen im Rahmen von CC kann die allgemeine Kundenzufriedenheit, die Kundenbindung (z. B. in Form von Wiederkauf) und positive Mund-zu-Mund-Propaganda deutlich gesteigert werden (Dong und Sivakumar, 2017, S. 945; Mustak et al., 2016, S. 250 f.; Jaakkola und Alexander, 2014, S. 257). Ein weiterer Grund für VCC aus Anbieterperspektive ist die mögliche Reduktion von Fehlern und Risiken bei Produkteinführungen, die sich durch eine interaktive, frühzeitige Beteiligung zentraler Kundengruppen im Wertschöpfungsprozess ergibt. So können neue Angebote, die nicht den Kern der Kundenbedürfnisse treffen, rechtzeitig angepasst werden, ehe es zu größeren Fehlinvestitionen kommt. Hinzu kommt eine mögliche Steigerung der Produktqualität und eine höhere Marktakzeptanz (Hoyer et al., 2010, S. 283). Schließlich sind auch interne Aspekte beachtenswert. Zum einen kann VCC zu Kosteneinsparungen beim Anbieter führen, da zentrale Aufgaben auf Kunden oder Partner übertragen werden und dabei kostengünstig auf deren Ressourcen (Wissen, Zeit, usw.) zugegriffen wird – anstatt nur interne Ressourcen hierfür bereitzustellen (Galvagno und Dalli, 2014, S. 646). Zum anderen können die Aktionen Abwechslung in den Alltag der Mitarbeiter des Anbieterunternehmens bringen und so die Mitarbeitermotivation zusätzlich erhöhen (Ramaswamy und Gouillart, 2010, S. 104).

Allerdings sei darauf hingewiesen, dass VCC kein „Allheilmittel“ ist. Auch das Aufsetzen von CC-Aktionen ist mit Aufwand und Kosten für Anbieter verbunden. Außerdem ist es entscheidend, die richtigen Kunden für eine Beteiligung zu gewinnen – das heißt Kunden, die über Wissen und Fähigkeiten verfügen, mit dem Anbieter zu kooperieren. Die Kundenansprüche können sehr heterogen sein. Als Anbieter sollte das gesamte Spektrum der Zielgruppen berücksichtigt werden.

3.2. Überblick zu Beweggründen von Kunden

Die Literaturanalyse zeigt, dass es zahlreiche Gründe für Kunden gibt, sich an anbietergesteuerten CC-Programmen zu beteiligen. Diese Treiber der Kunden reichen von direkten/indirekten finanziellen Anreizen über Lerneffekte und möglichen Reputationsabsichten bis hin zu bislang noch unbefriedigten Bedürfnissen mit derzeitigen Anbieterlösungen. Auch die Chance dabei Networking zu betreiben (mit dem Anbieter und/oder anderen Kunden) sowie ein grundsätzliches Interesse an Innovationen und Spaß an der Aktivität selbst können Beweggründe zur Teilnahme von Kunden sein. Schließlich kann auch ein altruistischer Gedanke (zu helfen, etwas Gutes zu tun) Grund für eine CC-Beteiligung sein; vgl. dazu auch Kapitel 4.1. Tabelle 1 zeigt eine Übersicht verschiedener Publikationen, die sich mit solchen Kundentreibern auseinandersetzen.

Die Übersicht kategorisiert die Art der Studie (quantitativ, qualitativ oder konzeptionell) und gibt den Kontext der Analyse an (B2B- oder B2C-Kundenfokus). Die Kennzeichnung X in einer Zelle bedeutet, dass der jeweilige Treiberfaktor von den Autoren in der Arbeit identifiziert und näher erläutert wird. Ein leeres Feld bedeutet hingegen, dass der Treiber in dieser Publikation nicht näher erläutert wird. Dabei gilt in Studien jeweils die Grundannahme, dass eine höher assoziierte bzw. empfundene Ausprägung der jeweiligen Kundentreiber bei VCC-Aktivitäten zu einer höheren Teilnahmebereitschaft von Kunden führt. Das heißt bspw. für den Treiberfaktor Lerneffekt: Je höher bzw. nützlicher Lerneffekte durch die konkrete VCC-Aktivität assoziiert oder empfunden werden, desto höher ist prinzipiell die Teilnahmebereitschaft. Entsprechendes gilt für die anderen genannten Kundentreiber.

Auszüge wortwörtlicher Formulierungen der Autoren zu den Treibern sind im Anhang in Tabelle 11 hinterlegt. Die letzte Zeile fasst Ergebnisse der zwei Experteninterviews zusammen, welche im Rahmen der kurzen qualitativen Voruntersuchung stattgefunden haben. Diese wurden mit Repräsentanten von Unternehmen durchgeführt, die regelmäßig CC betreiben und begleiten. Die Interviews sind nur als Ergänzung und daher nicht gleichwertig zu den wissenschaftlichen Publikationen anzusehen. Sie sind mit einem * unten in kursiver Schrift separat gekennzeichnet.

Viele der identifizierten Treiber sind potentiell sowohl für B2C-Kunden als auch für B2B-Kunden von Bedeutung. Dies wird in der folgenden quantitativen Erhebung genauer überprüft. Zudem ist auch der Zeitaufwand für Kunden in Verbindung mit VCC als negativer Einflussfaktor zu berücksichtigen.

Zwar wurde dieser Faktor in den zitierten Publikationen nicht direkt untersucht. *Etgar* (2007, S. 100) stellt jedoch fest, dass Zeit eine der Hauptressourcen ist, welche Kunden bei CC einbringen. Er bezeichnet diese als rare Ressource für alle Individuen. Dies gilt auch für Geschäftskunden (Broß, 2018; Siepmann, 2018). So ist anzunehmen, dass ein höherer Zeitaufwand einen negativen Einfluss auf die Teilnahmebereitschaft von Kunden hat. Es gilt daher, CC-Aktivitäten möglichst zeitsparend zu gestalten.

Grundvoraussetzung für die Teilnahme an VCC ist ein Mindestmaß an Involvement der Kunden. Involvement beschreibt die empfundene Relevanz und Erregung z. B. in Bezug auf Objekte, basierend auf eigenen Bedürfnissen, Werten und Interessen (von Loewenfeld, 2003, S. 3). Insofern bezieht sich das Involvement auf das gedankliche Engagement und eine damit verbundene Aktivierung, mit der sich Kunden einer Sache bzw. Aktivität zuwenden. Da die Kundenbeteiligung bei VCC per Definition über das übliche Ausmaß an Interaktion zwischen Anbieter und Kunde hinausgeht, ist hierfür ein gewisses Involvement unerlässlich. Dieses kann sich bei Privatkunden (B2C) und Geschäftskunden (B2B) jeweils auf die Effektivität aller Treiber auswirken (Nambisan und Baron, 2007, S. 46). Beispielsweise können mögliche Lerneffekte zu einer Lösung eines Anbieters nur als Treiber für die CC-Teilnahme gesehen werden, wenn der Kunde dem Produkt bzw. der Dienstleistung eine entsprechend hohe Bedeutung zuweist. Denn nur dann kann es als Anreiz verstanden werden, mehr über die Lösung zu erfahren und zusätzliches Wissen über Anwendungen und Methoden diesbezüglich aufzubauen. Ähnliches gilt für den Treiberfaktor Reputation. Nur wenn die Lösung des Anbieters ein Mindestmaß an Relevanz bzw. Interesse beim Kunden hervorruft, ist es ein Anreiz, im Zusammenhang mit der Lösung die eigene Reputation zu erhöhen. Dies wäre z. B. bei der anbieterseitigen Nennung als Produktoptimierer der Fall oder auch wenn Kunden öffentlich als Referenzkunde genannt werden. Entsprechendes gilt für die anderen Kundentreiber. Einzige Ausnahme bildet der Treiber Finanzieller Anreiz, da zumindest eine direkte Aufwandsentschädigung für eine mögliche CC-Teilnahme der Kunden selbst nicht zwingend Involvement voraussetzt. Indirekte finanzielle Anreize wie z. B. Rabatte für zukünftige Angebote oder Einsparpotential durch das Anpassen der Lösung auf individuelle Bedürfnisse mit neuen Funktionen setzen hingegen wieder ein grundsätzliches Maß an Involvement voraus.

Zusammenfassend gilt: Nur wenn ein Mindestmaß an Involvement vorliegt, ist eine interaktive Beteiligung der Kunden im Rahmen von VCC überhaupt sinnvoll bzw. möglich. Zudem sollte es Ziel des Anbieters sein, nur solche Kunden für CC-Aktivitäten zu gewinnen, die ein intensives Interesse an der Lösung, ein gewisses Trendgefühl und die Bereitschaft haben, ihre Erfahrungen mit dem Anbieterunternehmen zu teilen. In diesem Zusammenhang wird häufig von Leadusern gesprochen (Sänn, 2017, S. 9).

3.3. Gestaltungsmöglichkeiten und Phasen der Kundenintegration

Die Möglichkeiten der B2B- und B2C-Kundenintegration sind vielseitig. Nach *Vargo und Lusch* (2016, S. 7) werden dabei immer Akteure einbezogen, indem sie ihre Ressourcen integrieren. So findet ein Austausch an gegenseitig erbrachten „Services“ statt, auf Basis von Absprachen und gemeinsamen Richtlinien. Diese Annahme lässt zunächst großen Handlungsspielraum bezüglich der Gestaltung von VCC-Aktivitäten.

Payne et al. stellen ein Modell auf, dass die prozesstechnische Darstellung von CC-Aktivitäten konkretisiert (vgl. Abbildung 3). Entscheidend für Kunden ist in diesem Fall die Handlungsbeziehung („Relationship Experience“), welche sich durch die Interaktion mit dem Anbieter ergibt. Bestimmt wird diese durch Emotionen, Kognitionen und Verhalten der Kunden. Parallel dazu ergibt sich der Kundenlernprozess über alle CC-Aktivitäten hinweg. Auf der Anbieterseite stehen Prozesse der CC- und Handlungsbeziehungsgestaltung im Vordergrund, mit denen die Anbieter durch Planung, Implementierung und verschiedene Ansätze („Opportunities“) versuchen, die Beziehungen zum Kunden im Rahmen der Interaktionsprozesse positiv zu beeinflussen. Organisationales Lernen ist auch ein wesentlicher Bestandteil für das Anbieterunternehmen (Payne et al., 2008, S. 86).

Bei der CC mit B2B-Kunden interagieren Anbieter oftmals nicht nur mit ihren direkten Geschäftskunden, sondern auch mit dem Netzwerk dieser Kunden, um umfassende Lösungen kundenspezifisch zu erschaffen (Cova und Salle, 2007, S. 6 ff.). So werden in Zusammenarbeit z. B. Schnittstellen in Softwarelösungen etabliert, um systemübergreifend im Partnernetzwerk arbeiten zu können. Dies hat zur Folge, dass CC-Programme insbesondere im B2B-Umfeld so zu gestalten sind, dass mehrere Stakeholdergruppen zeitgleich angesprochen werden können.

Nambisan und Baron (2007, S.43) identifizieren und erläutern vier Hauptphasen der VCC mit Kunden. Diese beziehen sich auf unterschiedliche Bereiche der Unternehmenswertschöpfung, bei denen Kunden integriert werden können (Grönroos, 2011, S. 283; Porter, 2010, S. 66). Zwar wurden diese VCC-Phasen mit Fokus auf B2C-Kunden definiert. Ein Vergleich mit den angewandten VCC-Praktiken von *Marcos-Cuevas et al.* (2016, S. 99 ff.) aus dem B2B-Umfeld zeigt jedoch, dass diese Phasen einen eher allgemeingültigen Charakter aufweisen. Grundlegend sind die Phasen Product Ideation, Product Design & Development, Product Testing und Product Support Services zu unterscheiden.

Product Ideation umfasst das gemeinsame Finden bzw. Ausarbeiten von Ideen für grundlegend neue Produkte oder Dienstleistungen eines Anbieters. Ziel von Ideation ist es, konkrete Ansätze für bisher nicht existierende Lösungen des Anbieters zu schaffen. Ein Einbezug von Kunden in dieser frühen Phase der Wertschöpfung eines Unternehmens kann dazu führen, neue Trends und Bedürfnisse rechtzeitig zu erkennen und Lösungen auf den Markt zu bringen, die der Konkurrenz in Bezug auf Kundenbedürfnisse voraus sein können. Co-Diagnosis (Sammeln und Aufbereiten

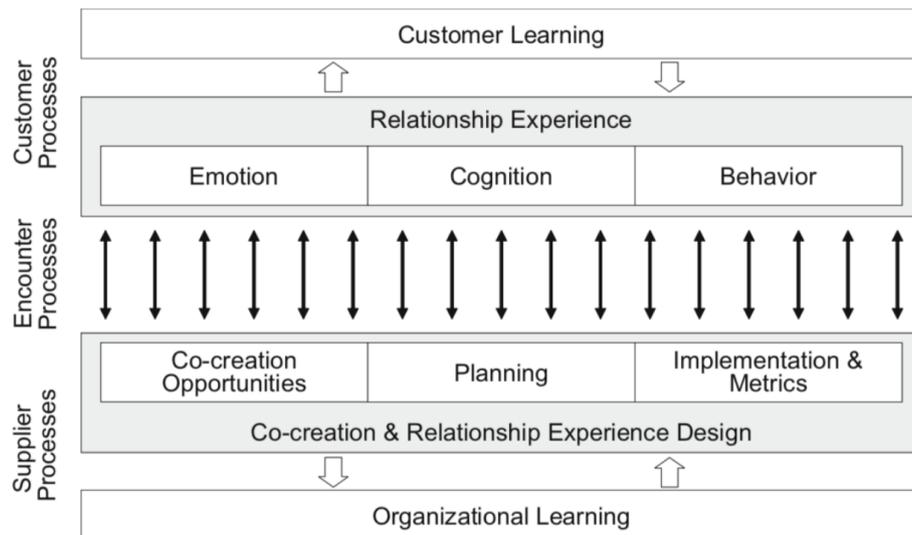


Abbildung 3: Prozesstechnische Darstellung von CC mit Kunden ⁸

von Informationen), Co-Ideation (Teilen neuer Ideen) und Co-Evaluation (Beurteilen und Priorisieren von Ideen) sind Praktiken, die im Zusammenhang mit dieser Phase stehen (Marcos-Cuevas et al., 2016, S. 99 ff.). Product Design & Development hingegen bezieht sich auf das Erarbeiten konkreter Vorschläge zu Produktfunktionen und die inkrementelle Verbesserung bereits bestehender Angebote. Welche Funktionen wünschen sich die Kunden bei existierenden Lösungen, wo genau gibt es Anwendungsprobleme und wie können diese zukünftig durch gemeinsame Entwicklungen behoben werden? Marcos-Cuevas et al. (2016, S. 99 ff.) sprechen hierbei von Co-Design, wenn gemeinsam mit Kunden Konzepte oder Prototypen erarbeitet werden. Die Grenzen von Product Ideation sowie Product Design & Development sind nicht immer eindeutig. Der Übergang der Phasen kann fließend verlaufen. Product Testing oder Co-Testing bezieht Kunden in konkrete Testszenarien ein. Das bedeutet, die Phase beschreibt die Interaktion mit Kunden in Hinblick auf Produkttests mit Prototypen zur Fehleridentifikation oder das Testen von Dienstleistungsangeboten.

Eine ausführliche Dokumentation der Nutzererfahrungen ist dabei essentiell, um später Anpassungen vornehmen zu können (Marcos-Cuevas et al., 2016, S. 99 ff.; Nambisan und Baron, 2007, S. 43). Product Support Services hingegen umfassen sämtliche Aktivitäten der Informationsverbreitung von Kunden über Produkte und „Services“ der Anbieter sowie die Hilfestellung für andere Nutzer. Im weitesten Sinne gehört auch die Mund-zu-Mund-Propaganda dazu, ebenso wie das Bereitstellen von Hilfe über Onlineforen oder eigene Erfahrungsberichte. Je nachdem, in welchen Phasen sich Kunden einbringen, können sie unterschiedliche Rollen als Akteur im Rahmen der CC einnehmen. Sie sind folglich Product Designer, Product Tester etc. (Nambisan und Nambisan, 2008, S. 55).

Abhängig von der Zielsetzung des Anbieters macht ein Einbezug von Kunden frühzeitig (Ideation, Development)

oder in späteren Phasen der Wertschöpfung Sinn (Testing, Support). Unterschiede bei der CC mit B2C- vs. B2B-Kunden können sich in der konkreten Ausgestaltung von CC-Programmen ergeben (Marcos-Cuevas et al., 2016, S. 99 ff.; Nambisan und Baron, 2007, S. 43), z. B. mit Blick auf

- einmalige vs. wiederkehrende CC-Programme
- die Dauer von CC-Programmen (kurzfristig vs. langfristig angelegt)
- die Anzahl beteiligter Personen (viele Konsumenten vs. ausgewählte (Business-)Nutzergruppen; Personaleinsatz des Anbieters)
- die Intensität der Interaktion von Kunden und Anbietern (direkter vs. indirekter Austausch; geplante vs. spontane Interaktionen; Einheitendauer)
- den Ressourceneinsatz (Kostenaufwand des Anbieters für CC-Aktionen; konkrete Einsatzbereitschaft der Kunden)
- Kundenpräferenzen für die Art der konkreten Aktivitäten (online vs. offline).

Hinsichtlich der Art von CC-Aktivitäten kann grundsätzlich zwischen Offline- und Onlineaktivitäten unterschieden werden (Daiberl et al., 2016, S. 12 f.; Gyrd-Jones und Kornum, 2013, S. 1486). Hierauf wird in Kapitel 4.3 näher eingegangen.

4. Konzeptionelles Modell zur Analyse von Kundentreibern für die Teilnahme an VCC-Programmen

Im Rahmen der Literaturanalyse wurden unterschiedliche Treiber von Kunden zur Beteiligung an VCC-Programmen

⁸Entnommen aus Payne et al., 2008, S. 86

identifiziert (vgl. Tabelle 1). Die Erhebung im Folgenden untersucht die Bedeutung bzw. Wichtigkeit dieser Treiber als Variable in Abhängigkeit von den Gruppen B2C- vs. B2B-Kunden eines Anbieters im direkten Vergleich. Dabei werden drei unterschiedliche Kundentypen bzw. -perspektiven berücksichtigt, die jeweils als eigene Kundengruppe betrachtet werden:

1. Perspektive von Konsumenten bzw. Privatkunden (B2C)
2. Perspektive von Unternehmen bzw. Organisationen als Geschäftskunden eines Anbieters (d. h. Unternehmen als Kunden, „firm-level“, B2B-Organisation)
3. Perspektive von Mitarbeitern in einem Unternehmen, welches Geschäftskunde eines Anbieters ist („individual-level“, B2B-Mitarbeiter)

Durch diese Analyse soll die Frage geklärt werden, welche Gemeinsamkeiten und Unterschiede es hinsichtlich der Bedeutung der Treiber für eine CC-Teilnahme gibt. Es wird demnach überprüft, ob die Kundenperspektive bzw. der Kundentyp (B2C vs. B2B-Mitarbeiter vs. B2B-Organisation) als unabhängige Variable einen Einfluss darauf hat, wie wichtig einzelne Treiber für die Bereitschaft einer CC-Teilnahme empfunden werden (abhängige Variablen).

Einerseits kommt es hierbei zum Vergleich der Bedeutung einzelner Treiberfaktoren zwischen den jeweiligen Kundengruppen. Das heißt, jeder Treiberfaktor wird zunächst einzeln betrachtet, wobei die Werte zur Bedeutung dieses Treiberfaktors zwischen den Kundengruppen miteinander verglichen werden – kundengruppenübergreifende Analyse.

Andererseits wird die Bedeutung aller Treiberfaktoren innerhalb einer Kundengruppe für sich selbst noch einmal priorisiert bzw. separat analysiert. Das heißt, jede Kundengruppe wird für sich einzeln betrachtet, wobei die Werte zur Bedeutung aller Treiberfaktoren innerhalb dieser Kundengruppe miteinander verglichen und priorisiert werden – kundengruppeninterne Analyse.

Zusätzlich werden exemplarisch häufig initiierte Aktivitäten von CC-Programmen betrachtet (Workshops, Interviews, Onlinetools, Onlineforen). Welche Präferenzen zeigen die Kundengruppen hinsichtlich Online- und Offlineaktivitäten und wie ist die Teilnahmebereitschaft für diese im Vergleich?

Somit ergibt sich ein formatives Modell für den Gesamtzusammenhang der quantitativen Untersuchung (Abbildung 4). Die Kundentreiber werden als latente Variablen mit jeweils zwei bzw. drei Items reflexiv gemessen (vgl. Kapitel 5.1.2).

4.1. Kategorisierung rationaler und emotionaler Kundentreiber zur VCC-Teilnahme

Die identifizierten Treiber der Literaturanalyse lassen sich auf verschiedene Arten kategorisieren. Üblicherweise werden intrinsische und extrinsische Motivationsfaktoren unterschieden (Heckhausen und Heckhausen, 2006, S. 5). Intrinsisch bedeutet, dass die Motivation aus der Tätigkeit selbst bzw. dem Vollzug der Tätigkeit herrührt. Extrinsisch sind hingegen anreizbesetzte Ereignisse oder Veränderungen, die sich einstellen, wenn diese Tätigkeit erfolgreich

erledigt ist. Dies sind z. B. materielle Belohnungen oder die Annäherung an langfristige Ziele als Folgen der Handlung (Heckhausen und Heckhausen, 2006, S. 5; Rheinberg, 2006, S. 333).

Für die betrachteten Kundenperspektiven dieser Arbeit eignet sich jedoch die Einteilung nach rationalen und emotionalen Treibern von Kunden besser (Bagozzi, 2006, S. 453 f.; Cater und Cater, 2009, S. 1151; Chen et al., 2012, S. 145; Döring, 2007, S. 363; Mencarelli und Rivière, 2015, S. 209).

Rationale Kundentreiber sind solche, die Kunden primär einen direkten ökonomischen oder funktionalen Nutzen bieten. Insofern können rationale Treiber auch als impliziter „Kosten-Nutzen-Vergleich“ angesehen werden. Rationale Treiber sind zweckorientiert und utilitaristisch (Mencarelli und Rivière, 2015, S. 209). Typischerweise sind diese Treiber mit dem Gedanken verbunden, eine direkte Gegenleistung für ein eigenes Engagement zu bekommen (Bagozzi, 2006, S. 453). Rationale Treiber sind demnach kalkulativer Natur. Hierbei geht es um die Beurteilung der Anreize nach objektiven Kriterien. Zusammenfassend lassen sich diese Treiber mit der Frage nach dem ökonomischen und funktionalen Nutzen von Aktionen identifizieren („benefit“) (Cater und Cater, 2009, S. 1152).

Emotionale Treiber sind jene, die keinen direkten ökonomischen oder funktionalen Nutzen bieten. Sie stellen den Gegenpol zu rationalen Treibern dar. Hier kann eher von selbstbezogenen bzw. sozialen Emotionen gesprochen werden (Bagozzi, 2006, S. 453). Es geht dabei um ein positives Gefühl, persönlichen Geschmack („liking“) und Identifikation der Kunden (Cater und Cater, 2009, S. 1152). In Verbindung damit kann auch von „affektivem Commitment“ gesprochen werden. Diese Treiber sind folglich nicht kalkulativer, sondern mit sonstigen Bedürfnissen verbunden (Mencarelli und Rivière, 2015, S. 209). Hierbei geht es um die Beurteilung der Anreize nach subjektiven Kriterien. Kurzum beschreiben emotionale Treiber jegliche Anreize, die keinen direkten ökonomischen oder funktionellen Nutzen haben. Sie sind emotional gesteuert und lassen sich mit der Frage nach Gefühlen bzw. Geschmack identifizieren („like“) (Cater und Cater, 2009, S. 1152).

Die Kategorisierung nach rationalen und emotionalen Kundentreibern stellt zwei gegensätzliche Pole dar. Bestimmte Treiber lassen sich den Kategorien genau zuordnen. Andere hingegen sind nicht eindeutig einer Kategorie angehörig, da sich Argumente für beide Richtungen anbringen lassen. Im Folgenden werden die vorab identifizierten Kundentreiber kurz erläutert. Dabei kommt es zur Formulierung der Hypothesen, die im Rahmen der quantitativen Studie zu überprüfen sind.

Treiberfaktor Finanzieller Anreiz: Es gibt direkte und indirekte finanzielle Anreize für Kunden, um sich an VCC zu beteiligen (Hoyer et al., 2010, S. 288). Mit direktem finanziellen Anreiz ist gemeint, dass Kunden für eine CC-Teilnahme eine direkte Bezahlung bzw. Aufwandsentschädigung vom Anbieter erhalten (Füller, 2010, S. 118). Indirekte finanzielle Anreize wären, wenn Kunden selbst Kosten einsparen können (Etgar, 2007, S. 101; Petri und Jacob, 2016, S. 66), z. B.

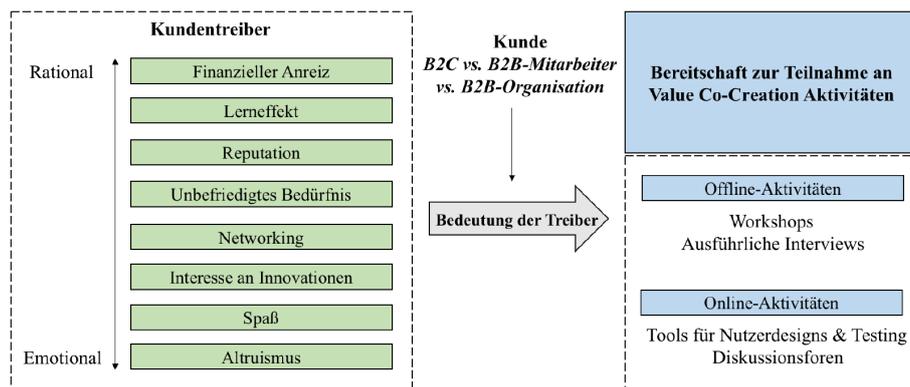


Abbildung 4: Konzeptionelles Modell der Treiber zur VCC-Teilnahmebereitschaft ⁹

durch das Anpassen der Lösung zugunsten ihrer Bedürfnisse mit neuen Funktionen, die sich aus ihrer CC-Teilnahme ergeben. Auch mögliche Rabatte für zukünftige Versionen bzw. Produkte des Anbieters für die CC-Teilnahme stellen einen finanziellen Anreiz dar. Insofern ist dieser Faktor direkt mit einem ökonomischen Nutzen für den Kunden verknüpft und lässt sich eindeutig den rationalen Treibern zuordnen.

Treiberfaktor Lerneffekt: Kunden können durch die CC-Teilnahme neues Wissen generieren, das heißt selbst daraus lernen (Etagar, 2007, S. 102; Payne et al., 2008, S. 87). Beispielsweise können Kunden bei CC mehr über das Produkt bzw. die Lösung des Anbieters und Anwendungsfälle dazu erfahren (Hoyer et al., 2010, S. 288). Ebenso kann die Interaktion dem Kunden mehr Erkenntnisse über Technologien und Methoden des Anbieters einbringen oder Lösungen für derzeitige Anwendungsprobleme generieren (Nambisan und Nambisan, 2008, S. 55; Petri und Jacob, 2016, S. 66). Lerneffekte haben somit einen direkten funktionalen Nutzen. Neue Erkenntnisse lassen sich vom Kunden direkt für die zukünftige Anwendung von Lösungen einbringen. Es kann zusätzliches Wissen für Kunden generiert werden, das sonst nur schwer zugänglich ist. Entsprechend lässt sich der Treiber der rationalen Kategorie zuordnen.

Treiberfaktor Reputation: Ein weiterer Anreiz zur Teilnahme an VCC-Aktivitäten kann ein möglicher Reputationsgewinn des Kunden in als wichtig empfundenen Communities sein (Nambisan und Nambisan, 2008, S. 55). Dies bedeutet, dass die Wahrnehmung des Kunden selbst durch andere Unternehmen, Anbieter, Kunden und Nutzer positiv beeinflusst wird (Hoyer et al., 2010, S. 288). Konkret ist dies möglich, wenn der Anbieter den Kunden offiziell als Referenzkunden oder Produktoptimierer nennt, z. B. über die Website. Außerdem können die Beiträge des Kunden selbst Aufmerksamkeit und Reputation mit sich bringen, bspw. wenn Designvorschläge öffentlich online über Tools eingereicht werden oder andere von Workshopbeiträgen profitieren (Nambisan und Baron, 2007, S. 61). Der Faktor Reputation ist somit indirekt als rationaler Treiber einzuordnen. Je nach Intensität des Effekts lassen sich dadurch z. B. Kosten für eigene Werbe-

maßnahmen von Geschäftskunden reduzieren, wenn über eine CC-Teilnahme mit Beiträgen eine hohe Reichweite erzielt wird. Eine positive Reputation hat auch einen funktionalen Nutzen, denn der Kunde wird dadurch von anderen als zuverlässiger Partner für zukünftige Beziehungen angesehen.

Treiberfaktor Unbefriedigtes Bedürfnis: Bislang noch unbefriedigte Bedürfnisse mit dem derzeitigen Angebot sind ebenfalls ein häufiger Grund, sich als Kunde an VCC zu beteiligen. Der Anreiz dabei ist, Produkte im Hinblick auf eigene Bedürfnisse zu verbessern (Füller, 2010, S. 118; Jaakkola und Alexander, 2014, S. 252; Siepmann, 2018). So benötigt der Kunde das Produkt bzw. die Lösung grundsätzlich, ist aber bspw. mit aktuellen Funktionen oder der Bedienbarkeit noch unzufrieden (Hoyer et al., 2010, S. 288). Je nachdem, welche Art von Bedürfnissen unbefriedigt ist, lässt sich dieser Kundentreiber eher der rationalen oder emotionalen Kategorie zuordnen. Rational vor allem dann, wenn es um ein Bedürfnis geht, das direkte ökonomische oder funktionale Auswirkungen hat (z. B. neue Funktionen, um Kosten im Betrieb einzusparen). Emotional eher dann, wenn es um eine angenehmere Bedienung oder andere Aspekte geht, die vornehmlich emotionalen Mehrwert bieten.

Treiberfaktor Networking: Der Wunsch als Kunde direkt einen Kontakt zum Anbieter oder zu anderen Nutzern zu haben, kann ein Anreiz sein, sich an VCC-Aktivitäten zu beteiligen (Jaakkola und Alexander, 2014, S. 252). Hierbei geht es weniger um die eben erläuterte Reputation. Fokus ist in diesem Fall das Kennenlernen, der Austausch und die Interaktion selbst mit Anbietern oder Nutzern im Rahmen von Offline- oder Onlineaktivitäten (Ranjan und Read, 2016, S. 294). Dieser Anreiz kann sowohl als rationaler Treiber wie auch als emotionaler Treiber angesehen werden – abhängig vom Kontext. Weil mit Networking jedoch nicht zwingend ein direkter ökonomischer oder funktionaler Nutzen verbunden sein muss, wird der Faktor tendenziell eher in Richtung emotionaler Treiber eingeordnet.

Treiberfaktor Interesse an Innovationen: Ein (intrinsisches) Interesse an Innovationen und neuen Dingen allgemein kann auch Beweggrund sein, sich als Kunde an VCC zu beteiligen (Etagar, 2007, S. 102; Füller, 2010, S. 105; Siepmann 2018). Grundsätzlich ist damit gemeint, dass sich

⁹Eigene Darstellung

manche Kunden gerne mit Innovationsthemen auseinandersetzen, auch wenn dabei nicht immer ein direkter Bezug zu funktionalem oder ökonomischem Nutzen herzustellen ist. Diese reine „Neugier“ ist somit als emotionaler Treiber anzusehen. Das Interesse kann aber im Nachgang langfristig auch noch einen funktionalen oder sogar ökonomischen Nutzen haben, wenn es bspw. dazu führt, dass Kunden Trends aufspüren und für sich nutzen können, noch bevor dies andere tun.

Treiberfaktor Spaß: Besonders häufig wird der Treiber Spaß diskutiert. Gemeint ist, dass Kunden an CC-Aktivitäten teilnehmen, weil Sie es als positive Bereicherung empfinden. Die Aktivität selbst macht Kunden Spaß, das heißt, die Teilnahme löst ein positives Gefühl aus (Etagar, 2007, S. 102; Füller, 2010, S. 105; Nambisan und Baron, 2007, S. 47). Häufig wird in diesem Kontext von hedonistischen Motiven der Teilnahme gesprochen, als „mentally stimulating or entertaining“ (Nambisan und Nambisan, 2008, S. 56). Spaß ist ein emotionaler Treiber, der keinen direkten ökonomischen bzw. funktionalen Nutzen bietet.

Treiberfaktor Altruismus: Das Bedürfnis als Kunde helfen zu wollen, ist für manche ein weiterer Anreiz, an VCC teilzunehmen (Hoyer et al., 2010, S. 289). Dieses Bedürfnis kann sich auf den Anbieter fokussieren, das heißt, Kunden wollen einen Anbieter bei der Produkt- bzw. Lösungsentwicklung unterstützen. Es ist jedoch auch möglich, dass Kunden mit einer Teilnahme auch anderen Kunden bzw. Nutzern helfen möchten, indem sie auf Problematiken und Wünsche bei der Nutzung eingehen und neue Ideen einbringen (Chen et al., 2012, S. 145). Altruismus gehört zur Kategorie der emotionalen Treiber, da auch hier kein direkter ökonomischer oder funktionaler Nutzen für den Kunden selbst ausgeht.

4.2. Hypothesenbildung zu Kundentreibern für eine VCC-Teilnahme

Es stellt sich die Frage, welche der erläuterten Treiber für die jeweiligen Kundengruppen bzw. -perspektiven B2C, B2B-Mitarbeiter und B2B-Organisation eine hohe Wichtigkeit haben und worin die Gemeinsamkeiten und Unterschiede bezüglich der Bedeutung liegen. Insgesamt gesehen liegt die Vermutung nahe, dass einzelne Treibern, je nach Perspektive, eine unterschiedliche Bedeutung haben. Wie beschrieben, verfolgen Unternehmen als organisationaler Kunde vor allem gewinnorientierte Ziele (Foscht et al., 2017, S. 14 ff.; Günter und Kuhl, 2000, S. 377 ff.). Mitarbeiter von Unternehmen sind im Rahmen ihrer beruflichen Tätigkeit an die Unternehmensrichtlinien gebunden, bringen jedoch auch individuelle Sichtweisen mit ein (Storbacka et al., 2016, S. 487 ff.; Töllner et al., 2011, S. 713). Privatpersonen können als Konsumenten ganz individuell und frei darüber entscheiden, ob sie sich an VCC-Aktivitäten beteiligen wollen. Daraus ergibt sich einleitend zunächst folgende allgemein übergreifend formulierte Hypothese für den Untersuchungskontext:

H1: Die Bedeutung mindestens eines Treibers für eine Co-Creation-Teilnahme der Kunden ist für

die Kundengruppen B2C, B2B-Mitarbeiter und B2B-Organisation unterschiedlich.

Wie in Kapitel 2.2 ausführlich erläutert, basiert das Handeln in Geschäftssituationen (B2B) meist auf rationalen, objektiv nachvollziehbaren Argumenten (Lippold, 2017, S. 199). In diesem Kontext kommt es zu Kosten-Nutzen-Kalkulationen, um objektiv vorteilhafte Handlungsoptionen abzuwägen, die einen direkten ökonomischen oder funktionalen Mehrwert bieten (Eggert und Ulaga, 2002, S. 109; Mencarelli und Rivière, 2015, S. 208 f.). Die Ziele von Unternehmen, der Druck in Einkaufsgremien sowie die strategische Bedeutung von genutzten Gütern und Dienstleistungen in Unternehmensprozessen führen zu einer vorherrschenden utilitaristischen Perspektive in B2B-Kundenbeziehungen (Corsaro und Snehota, 2010, S. 990 f.). In letzter Zeit wird jedoch auch in der B2B-Literatur zunehmend auf die Bedeutung eher emotionaler und symbolischer Aspekte hingewiesen – insbesondere mit Blick auf die Mitarbeiter eines Unternehmens (Leek/Christodoulides 2012, S. 106; Lynch und de Chernatony, 2004, S. 403). Dennoch sind rationale Argumente weiterhin dominierend, wenn es um Entscheidungen im organisationalen, strategischen Kontext geht (Mencarelli und Rivière, 2015, S. 209). Auch bei der gemeinsamen Entwicklung von Business Solutions im Rahmen von CC-Aktivitäten geben B2B-Kunden an, insbesondere aus finanziellen Gründen, wegen der methodischen/funktionalen Expertise der Anbieter und aufgrund einer höheren Legitimation nach innen mit externen Anbietern zusammenzuarbeiten. Durch eine VCC-Teilnahme wollen Kundenunternehmen die Lösungen des Anbieters zu eigenen Gunsten optimieren und so erfolgreicher in die eigene Wertschöpfung einbringen (Petri und Jacob, 2016, S. 66). Insofern erscheint es naheliegend, dass B2B-Kunden insgesamt gesehen eher den rationalen Treibern eine hohe Bedeutung zuordnen. Dies gilt insbesondere für die organisationale Ebene, das heißt das strategische Management eines Unternehmens (B2B-Organisation). Aber auch die einzelnen Mitarbeiter eines Unternehmens, die ein Anbieter im Kundenunternehmen adressieren muss, sind angehalten, nach organisationalen Richtlinien zu agieren.

H2a: Für B2B-Kunden (B2B-Organisation, B2B-Mitarbeiter) sind rationale Treiber relativ gesehen von höherer Bedeutung als emotionale Treiber. Das heißt, dass Finanzieller Anreiz, Lerneffekt und Reputation kundengruppenintern tendenziell als wichtiger wahrgenommen werden als Altruismus, Spaß und Interesse an Innovationen.¹⁰

Lange wurde der Konsument (B2C-Kunde) im Marketing als „homo oeconomicus“ betrachtet, also als rein rational

¹⁰Die Treiberfaktoren Unbefriedigtes Bedürfnis und Networking werden nicht in die Hypothesenbildung aufgenommen, da sie nicht eindeutig der Kategorie rationaler oder emotionaler Treiber zuzuordnen sind (vgl. Kapitel 4.1).

handelnder Mensch, der Informationen sammelt, abwägt und „objektiv vernünftig“ entscheidet.

Diese Idee ausschließlich rational handelnder Konsumenten ist inzwischen überholt (Halfmann, 2014, S. 4). Immer häufiger spielen Gefühle, emotionale Argumente und spezifische Situationen des Handlungskontextes eine entscheidende Rolle. Diese sind nicht mit rationalen Nutzenfaktoren zu erklären (Franke, 2013, S. 2). Bei VCC geht die aktive Beteiligung von Kunden über das übliche Maß der Interaktion im Austauschprozess zwischen Anbieter und Kunde hinaus. Kunden entscheiden sich bewusst dafür, gemeinsam mit Unternehmen neue Produkte und Dienstleistungen zu entwickeln oder ein bestehendes Angebot zu verbessern. Die bisherige Forschung dazu zeigt, dass eine Teilnahmeentscheidung von Privatkunden häufig durch psychologische, emotionale Treiber bestimmt wird (Chen et al., 2012, S. 145; Füller, 2010, S. 108 ff.; Hoyer et al., 2010, S. 289; Nambisan und Baron, 2007, S. 52 ff.). Rationale Treiber sind jedoch nicht völlig unbedeutend. Insbesondere finanzielle Anreize können einen zusätzlichen Anreiz zur VCC-Teilnahme darstellen. Emotionale Treiber werden von Konsumenten jedoch oftmals als besonders wichtig empfunden, wenn es um die Teilnahmebereitschaft für VCC-Aktivitäten geht (Füller, 2010, S. 108 ff.). So lassen sich, in Anlehnung an die bisherige Argumentation, folgende Hypothesen weiterführend aufstellen.

H2b: Für B2B-Kunden (B2B-Organisation, B2B-Mitarbeiter) sind rationale Treiber relativ gesehen von höherer Bedeutung als für B2C-Kunden, das heißt Finanzieller Anreiz, Lerneffekt und Reputation betrachtet im kundengruppenübergreifenden Vergleich zwischen den Gruppen.

H3a: Für B2C-Kunden sind emotionale Treiber relativ gesehen von höherer Bedeutung als rationale Treiber. Das heißt, dass Altruismus, Spaß und Interesse an Innovationen kundengruppenintern tendenziell als wichtiger wahrgenommen werden als Finanzieller Anreiz, Lerneffekt und Reputation.

H3b: Für B2C-Kunden sind emotionale Treiber relativ gesehen von höherer Bedeutung als für B2B-Kunden (B2B-Organisation, B2B-Mitarbeiter), das heißt Altruismus, Spaß und Interesse an Innovationen betrachtet im kundengruppenübergreifenden Vergleich zwischen den Gruppen.

4.3. Kategorisierung von Offline- und Onlineaktivitäten der VCC

Offlineaktivitäten sind charakterisiert durch einen persönlichen, direkten Kontakt mit Interaktion von Anbietern und Kunden. Im Gegensatz zu Onlineaktivitäten haben sie eine geringere Reichweite, das heißt, dass weniger Akteure einbezogen werden können. Allerdings ist der Austausch komplexer Informationen und impliziten Wissens durch

die Präsenz vor Ort meist besser möglich. Der direkte Austausch ermöglicht ein Nachfragen und das Deuten nonverbaler Kommunikation beteiligter Kunden. Häufig vollzogene Offlineaktivitäten der CC sind Workshops. In Workshops werden unterschiedliche Methoden genutzt (Broß 2018; Daiberl et al., 2016, S. 13). Design Thinking ist eine Methode, die häufig in Workshops zur Generierung und Umsetzung neuer Ideen Anwendung findet. Hierbei kommen verschiedene Disziplinen zusammen. Ziel ist das agile Entwickeln von Lösungen in diversen, kreativen Teams mit Fokus auf spezifische Kundenbedürfnisse (Mosely et al., 2018, S. 177). Dabei werden unterschiedliche Probleme der Praxis aufgezeigt, Beispiele diskutiert und visualisiert (z. B. in unterschiedlichen Klebezettelgruppierungen, die sich immer wieder neu anordnen und erweitern lassen). Es ist zu unterscheiden zwischen verschiedenen Design Thinking-Typen: result-focused, convention-based, situation-based, strategy-based, experience-based, developing new schema und redefining the field (Mosely et al., 2018, S. 179). Auch das Durchführen von Rollenspielen/Planspielen ist eine Methode, mit der Workshops gestaltet werden können. Teilnehmer sollen sich hierbei in Situationen hineinversetzen und ihre Beweggründe hinter getroffenen Maßnahmen schildern (Daiberl et al., 2016, S. 13). Neben Workshops bieten semi-strukturierte, ausführliche Interviews mit Nutzergruppen eine Alternative (Witell et al., 2011, S. 141). Im Gegensatz zu Workshops lassen persönliche Interviews noch mehr Raum für spontane „Ausschweifungen“ von Kunden, die bspw. im engen Zeitplan eines Workshops weniger möglich sind. Der Vorteil hierbei kann sein, dass noch tiefergehende Informationen über Handlungsmuster, Bedürfnisse und Emotionen einzelner Kunden (Einzelinterview) oder Kundengruppen (Gruppeninterview bzw. Fokusgruppen) gewonnen werden können. Zudem können es Kunden selbst als weniger aufwendig empfinden, lediglich über Erfahrungen und Wünsche zu berichten, anstatt selbst Lösungen in Workshops zu erarbeiten. Oftmals erscheint die Kombination aus Einzel- und Gruppeninterviews sinnvoll.

Onlineaktivitäten haben eine deutlich höhere Reichweite im Vergleich zu Offlineaktivitäten. Allerdings erreichen Anbieter so häufig nur Kunden mit einer hohen Onlineaffinität (Daiberl et al., 2016, S. 12). Oft finden CC-Programme über speziell dafür geschaffene Onlineplattformen des Anbieters oder von Dritten statt. Je nach Kontext gibt es unterschiedliche Toolkits zum Entwerfen, Bewerten und Testen von Nutzerdesigns und Prototypen, auf die Kunden bzw. Nutzer direkt Zugriff über die Anbieter haben können (Füller, 2010, S. 109). Auch Blogs, Wikis und eigens geschaffene Foren von Anbietern, in denen Kunden Anwendungsfälle diskutieren, sich gegenseitig Hilfestellung geben und neue Informationen verbreiten können, sind für Online-CC von hoher Relevanz (Nambisan und Nambisan, 2008, S. 43). Über die Tools und Foren lassen sich Ideenwettbewerbe starten oder virtuelle Innovationcommunities gestalten. Beides kann für Anbieter relevante Erkenntnisse über Kundenbedürfnisse und Nutzungsverhalten liefern. Zudem gibt es für Anbieter auch die Möglichkeit, CC-Aktionen über bereits etablierte Social-

Media-Plattformen wie Facebook etc. zu initiieren. Hierbei werden u. a. Kommentare ausgewertet und darauf Bezug genommen oder gepostete Designs von Nutzern im Austausch weiter verarbeitet (Ramaswamy und Ozcan, 2018, S. 198).

4.4. Hypothesenbildung zu Offline- und Onlineaktivitäten von VCC

Bei Betrachtung bisheriger Forschung zu VCC zeigt sich, dass hinsichtlich anbieterinitiiertes CC mit B2C-Kunden ein deutlicher Schwerpunkt auf Onlineaktivitäten gesetzt wird (Füller, 2010; Hoyer et al., 2010; Nambisan und Baron, 2007; Nambisan und Nambisan, 2008; Prahalad und Ramaswamy, 2000; Ramaswamy und Ozcan, 2018). Mit Hilfe onlinebasierter Methoden können Anbieter eine hohe Reichweite mit geringen Kosten erreichen. Zudem bieten Onlineaktivitäten eine Vielzahl an Möglichkeiten der Kundenbeteiligung. Der Zeitaufwand für Kunden kann geringer sein, da eine evtl. Anreise entfällt und Onlinemethoden zeitlich flexibler gestaltet sind. Aufgrund dessen und der real existierenden Häufigkeit liegt die Vermutung nahe, dass B2C-Kunden privat Onlineaktivitäten für eine VCC-Teilnahme bevorzugen.

H4: Aus privater Perspektive werden Onlineaktivitäten bevorzugt. Das heißt, dass die Gruppe B2C-Kunden CC-Aktivitäten über das Internet bevorzugt.

Für VCC mit B2B-Kunden, das heißt mit Unternehmen als organisationalen Kunden und deren Mitarbeitern, setzen Anbieter sowohl Onlinemethoden als auch Offlineaktivitäten vor Ort ein (Broß 2018; Mustak et al., 2016; Siepman 2018). Bei zentralen CC-Prozessen im Geschäftskundenumfeld kommt es jedoch überwiegend zu einem persönlichen Austausch (Petri und Jacob, 2016, S. 64). Oftmals findet dieser offline in Form von Workshops beim Anbieter oder Kundenunternehmen vor Ort statt. In diesen werden Design Thinking-Methoden zur Problemlösung angewandt oder andere Strategien der Ideenfindung (Broß 2018; Mosely et al., 2018, S. 177). Auch ausführliche Interviews mit Nutzergruppen spielen eine zentrale Rolle (Siepman 2018). Aufgrund der Häufigkeit von Workshops und Events, die im Business-Umfeld für CC-Aktivitäten stattfinden, lässt sich daraus die Annahme ableiten, dass Offlineaktivitäten im beruflichen Kontext präferiert werden könnten (B2B-Kunden).

H5: Aus beruflicher Perspektive werden Offlineaktivitäten bevorzugt. Das heißt, dass die Gruppe B2B-Mitarbeiter und die Gruppe B2B-Organisation CC-Präsenzveranstaltungen vor Ort bevorzugen.

5. Empirische Untersuchung

5.1. Vorgehensweise der Datenerhebung

5.1.1. Überblick zur allgemeinen Prozedur

Mit Hilfe einer Onlineumfrage über das Tool Unipark (ww2.unipark.de) wird die Bedeutung einzelner Treiber für

eine CC-Teilnahme der unterschiedlichen Kundengruppen B2C, B2B-Mitarbeiter und B2B-Organisation quantitativ gemessen. Hierfür sind drei unterschiedliche Teilbereiche in der Umfrage gestaltet, jeweils mit eigenem Szenario und Frageteil pro Gruppe. Jeder Proband wird nur einem dieser Teilbereiche zugewiesen. Ein Teilnehmer beantwortet die Fragen daher entweder aus der Perspektive B2C („Als Privatperson...“), B2B-Mitarbeiter („Im Rahmen meiner beruflichen Tätigkeit...“) oder der Perspektive B2B-Organisation („Das Unternehmen, bei dem ich arbeite...“), vgl. Fragebogen im Anhang A3. Zudem werden die Präferenzen und die Teilnahmebereitschaft zu unterschiedlichen VCC-Aktivitäten erhoben.

Die Zuteilung auf eines der drei Szenarios erfolgt zunächst randomisiert. Nur wenn ein Proband in ein B2B-Kundenszenario eingeteilt wird, das heißt B2B-Mitarbeiter oder B2B-Organisation, erscheint folgende Kontrollfrage vorab:

„Haben Sie mind. 1 Jahr Berufserfahrung in einem Unternehmen, indem Sie täglich IT-Anwendungen (z. B. Instant Messaging, E-Mail) für die Geschäftskommunikation im Unternehmen oder mit Kunden nutzen? Hinweis: Beide Antwortkategorien sind gleichermaßen relevant und dienen lediglich der Zuordnung auf ein Szenario.“

Diese Frage soll sicherstellen, dass sich Probanden auf die Perspektive als Mitarbeiter einlassen können bzw. in der Lage sind, ihren Arbeitgeber als Unternehmen hinsichtlich der Treiber einer CC-Teilnahme einzuschätzen. Hiervon kann nach mind. einem Jahr Unternehmenserfahrung zumindest tendenziell ausgegangen werden. Antworten Probanden hierbei mit „Nein“, werden diese direkt in das B2C-Szenario umgeleitet, damit die Personen dennoch an der Umfrage aus Sicht einer Privatperson teilnehmen können. Bei „Ja“ werden die Teilnehmer im Szenario normal weitergeführt. Dieses Vorgehen birgt das Risiko, dass mehr Befragte im B2C-Teil antworten, da zufällig zugeteilte Probanden in einem B2B-Szenario mit „Nein“ bei der Kontrollfrage umgeleitet werden. Allerdings wird durch die Zufallszuweisung ein systematischer Fehler bei der Gruppeneinteilung ausgeschlossen. Zudem ist eine vergleichbare Gruppengröße mit Randomisierung eher zu erwarten als wenn sich die Probanden selbst zuordnen. Dann wäre laut Pretestgesprächen damit zu rechnen, dass das Privatszenario bevorzugt wird, da es auf den ersten Blick einfacher erscheint.

Zu Beginn wird allen Probanden eine VCC-Definition mit Beispielen erläutert, um ein Grundverständnis sicherzustellen. Anschließend kommt es zur Darstellung eines fiktiven Szenarios für ein CC-Programm. Dieses Szenario ist im Wortlaut an den Kontext der Kundengruppen angepasst. Danach erhalten Teilnehmer die Aufforderung Aussagen zu beurteilen, die Treiber zur CC-Teilnahme messen. Im Anschluss erfolgt eine Präzisierung der Umstände des vorgestellten CC-Szenarios, woraufhin Probanden direkt Präferenzen sowie die Teilnahmebereitschaft insgesamt und für einzelne Aktivitäten einschätzen sollten. Am Ende werden demographische

Daten erhoben.

5.1.2. Angewandte Szenarien und Operationalisierung

VCC ist eine vielschichtige Thematik, die exemplarisch anhand kurzer Szenarien nur schwer zu erfassen ist. Gleichwohl sollten die Szenarien einer Onlinebefragung möglichst prägnant sein, um die Aufmerksamkeitsspanne hoch zu halten. Entscheidend ist, dass die Szenarien inhaltlich vergleichbar sind und so direkte Rückschlüsse auf die CC-Treiber von Privat- und Geschäftskunden zulassen.

Instant Messaging ist von hoher Relevanz für Privatkunden. Im Alltag nutzen viele Menschen regelmäßig solche Dienste für die private Kommunikation (Statista, 2018a). Doch auch im Businessbereich spielen diese eine zunehmend bedeutende Rolle. Immer mehr Unternehmen setzen bewusst solche Lösungen zusätzlich zur E-Mail-Kommunikation ein. Der Markt hierfür ist groß und wird weiter wachsen. Anbieter haben bereits eine Vielzahl an Lösungen für Geschäftsanwendungen entwickelt (Flynn, 2005, S. 6; Koch et al., 2007, S. 453). Durch die hohe Bedeutung und vermeintliche Berührungspunkte bei Privat- und Geschäftskunden gleichermaßen eignen sich exemplarische Szenarien zur VCC im Instant Messaging-Umfeld gut für einen direkten Vergleich der Kundengruppen. Dahinter verbirgt sich die Annahme, dass ein Instant Messaging-Anbieter sowohl Privat- als auch Geschäftskunden als Zielgruppe bedienen kann, indem unterschiedliche App-Versionen angeboten werden. Dies ist bspw. bei WhatsApp der Fall (Inc., 2018). Die formulierten Szenarien für die Onlineumfrage sind in Tabelle 2 dargestellt.

Durch Anwendung der drei Szenarien handelt es sich um ein „multiple-case design“ (Yin, 1994, S. 45). Für die Messung der Kundentreiber sind Items als Teilaussagen in Anlehnung an die Füller (2010) sowie Nambisan und Baron (2007) formuliert, welche sich diesbezüglich wiederum auf Arbeiten anderer Autoren beziehen (Franke und Shah, 2003; Hertel et al., 2003; Kollock, 1999; Wasko und Faraj, 2000). Ergänzend dazu werden selbst konstruierte Items einbezogen, um weiterführende Präferenzen zu erfassen. Auf 7er-Likert-Skalen sollen die Probanden jeweils ihre Zustimmung zu den Aussagen einschätzen (von „stimme überhaupt nicht zu“ [1] bis „stimme voll und ganz zu“ [7]). Die Treiberfaktoren werden als latente Variablen erhoben, welche für sich anhand der verschiedenen Teilaussagen (Items) reflexiv gemessen werden. Diese Items zur Messung der unterschiedlichen Treiber für die Kundengruppen sind in Tabelle 3 gegenübergestellt.

Die Bedeutung der Treiber wird anhand von Mittelwerten für die Kundengruppen operationalisiert. Zunächst sind hierfür von jedem Probanden Werte zu den Treiberfaktoren generiert, indem die Beurteilung der Teilaussagen zu den jeweiligen Faktoren zusammengefasst wird. Alle zwei bzw. drei Items eines Faktors sind dabei gleich gewichtet, um die Vergleichbarkeit über die Kundengruppen hinweg zu wahren. So setzt sich bspw. der Wert des Faktors Finanzieller Anreiz für einen Probanden wie folgt zusammen:

$$[\text{Gesamtwert des Treiberfaktors Finanzieller Anreiz}] = [1/3 * \text{Wert Bezahlt werden}] + [1/3 *$$

$$\text{Wert Kosten sparen}] + [1/3 * \text{Wert Rabatte bekommen}]$$

Anschließend werden über alle Probanden einer Kundengruppen hinweg die Mittelwerte für die Bedeutung der CC-Treiber gebildet und verglichen. Das heißt, in der Analyse wird die „mittlere Bedeutung der Treiber“ für die jeweilige Kundengruppe betrachtet. Fokus sind sowohl kundengruppenübergreifende Analysen (Beurteilung eines Treiberfaktors der Gruppen B2C vs. B2B-Organisation vs. B2B-Mitarbeiter) als auch kundengruppeninterne Gegenüberstellungen (Bedeutung bzw. Priorisierung der Treiberfaktoren innerhalb einer Gruppe, z. B. bei B2C). So wird der Stellenwert der Treiber für die Kundengruppen auf unterschiedliche Arten herausgestellt.

Im Anschluss kommt es zur Abfrage der Teilnahmebereitschaft für das CC-Programm allgemein und dessen einzelne Aktivitäten (Workshops, Interviews, Onlinetools, Onlineforen). Hierfür finden ebenfalls 7er-Likert-Skalen Anwendung (Teilnahmebereitschaft von „sehr gering“ [1] bis „sehr hoch“ [7]). Schließlich erfolgt eine Präferenzabfrage hinsichtlich der Phase, in der sich die Kundengruppe am ehesten in der Unternehmenswertschöpfung beteiligen möchte (Ideation, Design & Development, Testing oder Support Services). Analog zu den Szenarien sind auch diese Items und Fragen im Wortlaut an die jeweiligen Kundenperspektiven angepasst. Auch hinsichtlich der Teilnahmebereitschaft werden gemäß beschriebener Operationalisierung Mittelwerte gebildet und verglichen. Die Präferenzen zur CC-Phase sind anhand einer Rangordnung erfasst und daher mittels deskriptiver Statistik ausgewertet.

5.1.3. Reichweite der Umfrage und Stichprobendaten

Für die Umfrage dieser Masterarbeit wird die Convenience Sampling Methode angewandt. Es werden also Probanden herangezogen, die „am besten zu erreichen“ sind (Jackson 2016, S. 102). Zusätzlich soll das Prinzip des Snowball Sampling verwendet werden, bei dem die Probanden selbst zum Verbreiten der Umfrage angehalten sind. So kann sichergestellt werden, dass eine zufriedenstellende Stichprobengröße für die Kundengruppen erreicht wird, die eine Auswertung mittels statistischer Methoden zulässt.

Um viele Teilnehmer zu gewinnen, ist ein Anreiz in Form einer Verlosung von fünf Amazon-Gutscheinen à 25 EUR gesetzt worden. Die Umfrage startete am 28.06.2018. Sie war bis zum 11.07.2018 online erreichbar und wurde in der Nacht zum 12.07.2018 geschlossen. Für eine möglichst hohe Repräsentanz ist der Link breit gestreut worden. Über LinkedIn wurde eine direkte Reichweite von 312 Kontakten z. T. in verschiedenen Unternehmen erzielt. Bei Xing waren es 332 Kontakte, wobei einige davon in beiden Netzwerken angesprochen wurden. Zusätzlich gab es fünf Reposts von Kontakten, mit denen wiederum ihr breites Netzwerk zur Teilnahme aufgefordert wurde. Der Link wurde außerdem in den zielgruppenspezifischen Facebook-Gruppen „Students at IBM“ (1.204 Mitglieder, darunter viele Alumni, die inzwischen bei anderen Unternehmen arbeiten), „M&M Master Manage-

Tabelle 2: Szenarien der Kundengruppen für die Onlinebefragung

Kundengruppe/ perspektive	- Szenario
B2C (Privatkunde)	<p><i>Stellen Sie sich vor...</i> Im Alltag nutzen Sie sehr häufig und engagiert eine Instant Messaging App (z. B. WhatsApp oder andere Anbieter), um mit Freunden und Bekannten zu kommunizieren. Nun kontaktiert Sie ein Mitarbeiter des Anbieters der App direkt und bittet Sie, an einem Co-Creation Programm für Privatkunden teilzunehmen. Ziel ist es, die App gemeinsam mit Ihnen als Privatanutzer zu verbessern und so optimal auf Kundenbedürfnisse einzustellen. Dabei sollen u. a. neue Funktionen eingeführt und getestet werden. Außerdem soll die Nutzerfreundlichkeit erhöht werden. Das Co-Creation Programm umfasst verschiedene Aktivitäten (Workshops, Interviews, Online Tools, Online Forum), an denen Sie teilnehmen können. <u>Bitte beantworten Sie die folgenden Fragen aus Ihrer persönlichen Sichtweise, d. h. als Privatperson bzw. Verbraucher.</u></p>
B2B-M (Mitarbeiter)	<p><i>Stellen Sie sich vor...</i> Sie sind derzeit Mitarbeiter eines Unternehmens. Im Arbeit-salltag nutzen Sie sehr häufig und engagiert eine Instant Messaging App (z. B. WhatsApp oder andere Anbieter), um mit Kollegen und Kunden zu kommunizieren. Ihr Arbeitgeber hat als Unternehmen Lizenzen für die Business-Version der App und ist somit Geschäftskunde des Instant Messaging Anbieters. Nun kontaktiert Sie ein Mitarbeiter des Anbieters der App direkt und bittet Sie, an einem Co-Creation Programm für Geschäftskunden teilzunehmen. Ziel ist es, die App gemeinsam mit Ihnen als Businessnutzer zu verbessern und so optimal auf Kundenbedürfnisse einzustellen. Dabei sollen u. a. neue Funktionen eingeführt und getestet werden. Außerdem soll die Nutzerfreundlichkeit erhöht werden. Das Co-Creation Programm umfasst verschiedene Aktivitäten (Workshops, Interviews, Online Tools, Online Forum), an denen Sie im Rahmen Ihrer beruflichen Tätigkeit teilnehmen können. Ihr Arbeitgeber stimmt einer möglichen Beteiligung grundsätzlich zu. <u>Bitte übertragen Sie die Situation auf Ihre eigene Berufserfahrung. Beantworten Sie die folgenden Fragen aus Ihrer Sicht als Mitarbeiter eines Unternehmens, das Geschäftskunde des Instant Messaging Anbieters ist.</u></p>
B2B-O (Organisation)	<p><i>Stellen Sie sich vor...</i> Sie sind derzeit Mitarbeiter eines Unternehmens. Im Arbeit-salltag nutzen die Teams im Unternehmen sehr häufig und engagiert eine Instant Messaging App (z. B. WhatsApp oder andere Anbieter), um mit Kollegen und Kunden zu kommunizieren. Ihr Arbeitgeber hat als Unternehmen Lizenzen für die Business-Version der App und ist somit Geschäftskunde des Instant Messaging Anbieters. Nun kontaktiert ein Mitarbeiter des Anbieters der App Ihren Arbeitgeber und bittet das Unternehmen, an einem Co-Creation Programm für Geschäftskunden teilzunehmen. Ziel ist es, die App gemeinsam mit Businessnutzern zu verbessern und so optimal auf Kundenbedürfnisse einzustellen. Dabei sollen u. a. neue Funktionen eingeführt und getestet werden. Außerdem soll die Nutzerfreundlichkeit erhöht werden. Das Co-Creation Programm umfasst verschiedene Aktivitäten (Workshops, Interviews, Online Tools, Online Forum), an denen Ihr Arbeitgeber teilnehmen kann. <u>Bitte übertragen Sie die Situation auf Ihre eigene Berufserfahrung. Beantworten Sie die folgenden Fragen mit Blick auf das strategische Management (Leitung) Ihres Arbeitgebers. Sollten Sie verschiedene Arbeitgeber im Kopf haben, konzentrieren Sie sich auf das eine Unternehmen, bei dem das Szenario am ehesten zutreffen könnte. Bei den folgenden Fragen sollen Sie die Beweggründe Ihres Arbeitgebers einschätzen, als Geschäftskunde am Co-Creation Programm des Anbieters teilzunehmen.</u></p>

Tabelle 3: Skalennormen zur Messung der Kundentreiber für VCC

Kundengruppe	B2C „Als Privatperson könnte ich mir vorstellen, am Co.Creation Programm teilzunehmen...“	B2B-M „Im Rahmen meiner beruflichen Tätigkeit könnte ich mir vorstellen, am Co-Creation Programm teilzunehmen...“	B2B-O „Ich könnte mir vorstellen, dass das Unternehmen bei dem ich arbeite (gearbeitet habe), am Co-Creation Programm teilnimmt...“
Finanzieller Anreiz	<ul style="list-style-type: none"> • Wenn ich dafür vom Anbieter bezahlt werde (direkte Aufwandsentschädigung) • Wenn ich durch das Anpassen der App auf meine Bedürfnisse (z.B. neue Funktionen) Kosten einsparen kann • Wenn ich durch eine Teilnahme zukünftige Versionen/andere Produkte des Anbieters vergünstigt bekomme 	<ul style="list-style-type: none"> • Wenn ich dafür vom Anbieter bezahlt werde (direkte Aufwandsentschädigung) • Wenn ich durch das Anpassen der App auf meine Bedürfnisse (z.B. neue Funktionen) beruflich Kosten einsparen kann • Wenn ich durch eine Teilnahme zukünftige Versionen/andere Produkte des Anbieters beruflich vergünstigt bekomme 	<ul style="list-style-type: none"> • Wenn das Unternehmen dafür vom Anbieter bezahlt wird (direkte Aufwandsentschädigung) • Wenn das Unternehmen durch das Anpassen der App auf die Bedürfnisse (z.B. neue Funktionen) Kosten einsparen kann • Wenn das Unternehmen durch eine Teilnahme zukünftige Versionen/andere Produkte des Anbieters beruflich vergünstigt bekomme
Lerneffekt	<ul style="list-style-type: none"> • Um mein Wissen über Instant Messaging und mögliche Anwendungen zu vergrößern • Um neue Erkenntnisse über die Technologien und Methoden des Anbieters zu erlangen • Um Lösungen für spezifische Probleme bei der Anwendung von Instant Messaging zu erhalten 	<ul style="list-style-type: none"> • Um mein berufliches Wissen über Instant Messaging und mögliche Anwendungen zu vergrößern • Um neue berufliche Erkenntnisse über die Technologien und Methoden des Anbieters zu erlangen • Um Lösungen für spezifische, berufliche Probleme bei der Anwendung von Instant Messaging zu erhalten 	<ul style="list-style-type: none"> • Um das Wissen über Instant Messaging und mögliche Anwendungen im Unternehmen zu vergrößern • Um neue Erkenntnisse über die Technologien und Methoden des Anbieters zu erlangen • Um Lösungen für spezifische Probleme bei der Anwendung von Instant Messaging im Unternehmen zu erhalten
Reputation	<ul style="list-style-type: none"> • Um als Co-Inventor/Produktoptimierer bekannt zu werden • Um als Referenz-Kunde genannt zu werden • Damit ich meine Sichtbarkeit bei Anbietern und Nutzern erhöhe 	<ul style="list-style-type: none"> • Um beruflich als Co-Inventor/Produktoptimierer bekannt zu werden • Um beruflich als Referenz-Kunde genannt zu werden • Damit ich meine berufliche Sichtbarkeit bei Anbietern(n) und Business-Nutzern erhöhe 	<ul style="list-style-type: none"> • Um als Co-Inventor/Produktoptimierer bekannt zu werden • Um als Referenz-Kunde genannt zu werden • Damit das Unternehmen seine Sichtbarkeit bei Anbieter(n) und Business-Nutzern erhöht

(Continued)

Tabelle 3—continued

<p>Unbefriedigtes Bedürfnis</p>	<ul style="list-style-type: none"> • Wenn die App noch nicht alle Funktionen enthält, die ich mir wünsche • Wenn ich die Bedienung der App noch nicht optimal finde 	<ul style="list-style-type: none"> • Wenn die App noch nicht alle Funktionen enthält, die ich mir beruflich wünsche • Wenn ich die Bedienung der App noch nicht optimal finde 	<ul style="list-style-type: none"> • Wenn die App noch nicht alle Funktionen enthält, die ich mir Unternehmen wünscht • Wenn ich die Bedienung der App noch nicht optimal finde
<p>Networking</p>	<ul style="list-style-type: none"> • Um mein Kontaktnetzwerk auszubauen (Networking betreiben) • Weil ich mit dem Anbieter in Kontakt treten möchte • Weil ich mit anderen Nutzern in Kontakt treten möchte 	<ul style="list-style-type: none"> • Um mein berufliches Kontaktnetzwerk auszubauen (Networking betreiben) • Weil ich beruflich mit dem Anbieter in Kontakt treten möchte • Weil ich beruflich mit anderen Nutzern in Kontakt treten möchte 	<ul style="list-style-type: none"> • Um das Kontaktnetzwerk auszubauen (Networking betreiben) • Weil das Unternehmen mit dem Anbieter in Kontakt treten möchte • Weil das Unternehmen mit anderen Nutzern in Kontakt treten möchte
<p>Interesse an Innovationen</p>	<ul style="list-style-type: none"> • Weil ich gern mit neuen Produkten/Lösungen zu tun zu habe • Weil ich neugierig auf Neues am Markt bin 	<ul style="list-style-type: none"> • Weil ich beruflich gern mit neuen Produkten/Lösungen zu tun zu habe • Weil ich beruflich neugierig auf Neues am Markt bin 	<ul style="list-style-type: none"> • Weil das Unternehmen gern mit neuen Produkten/Lösungen zu tun zu habe • Weil das Unternehmen neugierig auf Neues am Markt bin
<p>Spaß</p>	<ul style="list-style-type: none"> • Weil mir die Aktivität selbst Spaß macht • Weil es mir Freude bereitet, neue Ideen zu generieren und Probleme zu lösen 	<ul style="list-style-type: none"> • Weil mir die Aktivität selbst Spaß macht • Weil es mir Freude bereitet, beruflich neue Ideen zu generieren und Probleme zu lösen 	<ul style="list-style-type: none"> • Weil die Aktivität Mitarbeitern im Unternehmen Spaß macht • Weil es Mitarbeitern im Unternehmen Freude bereitet, neue Ideen zu generieren und Probleme zu lösen
<p>Altruismus</p>	<ul style="list-style-type: none"> • Weil ich dem Anbieter durch meine Teilnahme helfen möchte • Weil ich anderen Nutzern durch meine Teilnahme helfen möchte 	<ul style="list-style-type: none"> • Weil ich dem Anbieter durch meine Teilnahme helfen möchte • Weil ich anderen Nutzern durch meine Teilnahme helfen möchte 	<ul style="list-style-type: none"> • Weil das Unternehmen dem Anbieter durch die Teilnahme helfen möchte • Weil ich anderen Business-Nutzern durch die meine Teilnahme helfen möchte

ment & Marketing FU Berlin“ (291 Mitglieder) und „Circle of Excellence in Marketing“ (426 Mitglieder) gepostet. Hinzu kommt das Verzeichnis von www.surveycircle.com, über das jedoch nur zwei Teilnehmer akquiriert wurden. Schließlich wurde der Link in unternehmensinternen E-Mail-Verteilern an Mitarbeiter der Unternehmen IBM Deutschland GmbH, Dachser SE & Co. KG sowie SAP SE verbreitet.

Innerhalb von 14 Tagen wurde eine Gesamtreichweite von 1.380 Ansichten des Umfragelinks erzielt (Brutto-Reichweite). Von diesen 1.380 Ansichten haben insgesamt 330 Personen die Umfrage in einem der drei Szenarien abgeschlossen (unbereinigtes Gesamtsample). Dies entspricht einer Beendigungsquote von 23,91%.

Insgesamt sind 176 dieser Befragten männlich (53,3%), 150 Probanden weiblich (45,5%) und vier wollen sich nicht zuordnen (1,2%). Der Altersdurchschnitt aller Teilnehmer, die ein Alter angeben (321 Probanden), liegt bei 37,78 Jahren. Die jüngsten Teilnehmer sind 16 Jahre, der älteste ist 75 Jahre alt. Hinsichtlich Berufstätigkeit ergibt sich folgende Aufteilung: 225 Arbeitnehmer (68%), 51 Studierende (15,5%), 20 Selbstständige (6%), 17 machen keine Angabe (5,2%), 10 Schüler (3%), 6 Beamte (1,8%) und eine Person der 330 Teilnehmer ist arbeitssuchend (0,3%). Durch die Zuteilung auf die Szenarien haben sich unbereinigt folgende Teilnehmerzahlen für die Kundengruppen ergeben: 128 für B2C, 109 für B2B-Mitarbeiter und 93 für B2B-Organisation.

5.2. Auswertung der Daten

5.2.1. Pretests zum allgemeinen Verständnis

Vor Start der Onlineumfrage wurden zwei Pretests durchgeführt. Als Erstes hat ein qualitativer Pretest offline mit zwölf Testern stattgefunden. Dabei wurden die Szenarien mit Fragen jeweils vier unterschiedlichen Personen der Zielgruppen mit entsprechender (Berufs-)Erfahrung vorgelegt (4 x B2C, 4 x B2B-Mitarbeiter und 4 x B2B-Organisation). Im persönlichen Austausch sollten die Tester offen Unklarheiten und Verständnisfragen kommunizieren. Dabei stellte sich heraus, dass bei B2B-Szenarien Fragen nach Arbeitgeber, Unternehmensgröße und -branche auf Vorbehalte stießen. Deshalb wurde danach auf ein Erfassen solcher Daten verzichtet. Der zweite Pretest fand mit elf anderen Testern statt. Diese hatten Zugang zum Umfragelink und sollten die Umfrage online ohne weitere Hilfestellung bearbeiten. Anschließend kam es erneut zum Austausch. Schließlich wurden letzte Anpassungen für die finale Version der Onlineumfrage vorgenommen.

5.2.2. Datenbereinigung und Ausschluss von Probanden

Die Datenbereinigung vor der weiteren Auswertung ergibt einen Ausschluss von insgesamt acht Probanden des 330er-Gesamtsamples (davon 3 B2C, 4 B2B-Mitarbeiter, 1 B2B-Organisation). Dieser Ausschluss erfolgt, weil die Probanden im Kommentarfeld angeben, dass das angewandte Szenario für sie nicht zutreffend sei. Zudem werden auch Teilnehmer ausgeschlossen, welche alle Treiberitems durchgehend gleich einschätzen, das heißt ohne jegliche Differenzierung die Aussagen mit auffallend kurzer Bearbeitungszeit

„durchgeklickt“ haben. Mit Ausschluss dieser Teilnehmer sollen Verzerrungen der Ergebnisse minimiert werden. Somit ergibt sich ein bereinigtes Gesamtsample von 322 Probanden.

5.2.3. Stabilität des Messmodells der Treiber für die Kundengruppen

Wie in Kapitel 5.1.2 beschrieben, werden die Treiberfaktoren anhand von je zwei bzw. drei Items pro Faktor und Kundengruppe gemessen. Um die Güte bzw. Stabilität dieser Messung für die Treiber zu überprüfen, wird Cronbachs Alpha als Gütekriterium verwendet. Diese Maßzahl bestimmt die interne Konsistenz einer Skala, das heißt die Konsistenz der Items, deren Werte zu einem Treiberfaktor zusammengefasst werden. Es ist somit das Ausmaß bezeichnet, indem die Teilaussagen (Items) miteinander in Beziehung stehen. Für die Interpretation der Werte gilt ab $> 0,9$ ein exzellenter, $> 0,8$ guter, $> 0,7$ akzeptabler, $> 0,6$ fragwürdiger und $> 0,5$ schlechter Zusammenhang der Items. Die Analyse zeigt für die Umfrage folgende Ergebnisse (vgl. Tabelle 4):

Grundsätzlich zeigen die Cron α -Werte zuverlässige Ergebnisse. Die Items für die Messung der Treiberfaktoren weisen eine relativ hohe interne Konsistenz auf, sodass sie für die Erhebung geeignet sind. Einzige Ausnahme bilden die drei rot markierten Cron α -Werte, welche unter 0,7 liegen und somit für die Erhebung bei diesen Kundengruppen weniger geeignet erscheinen. Insgesamt zeigt sich die Messung der Faktoren jedoch als stabil. Die Analyse der MW-Differenzen ist demnach wie geplant durchzuführen.

5.2.4. Bedeutung der Kundentreiber zur Teilnahme an VCC Prüfung der Voraussetzungen der ANOVA-Methode

Im Statistikprogramm JMP Pro 13 wird mittels ANOVA-Berechnungen und Post-Hoc-Tests die Bedeutung der Treiberfaktoren für die Kundengruppen überprüft. Wie bei wirtschaftswissenschaftlichen Arbeiten üblich ist für alle nachfolgenden Tests ein Signifikanzniveau von $\alpha = 0,05$ festgelegt. Somit ist ein Test auf MW-Differenzen signifikant, wenn der dazugehörige p-Wert $< 0,05$ ist. Die Voraussetzungen der ANOVA können insgesamt als erfüllt angesehen werden: Pro Gruppe sind mind. 20-30 unabhängige Beobachtungen erfasst,¹¹ die kleinste Gruppengröße hat ein n von 92. Die Gruppen sind ähnlich groß. Der maximal zulässige Faktor $n_{\max}/n_{\min} = 1,5$ wird nicht überschritten, da $125/92 = 1,36 < 1,5$. Die verwendeten 7er-Likert-Skalen für die Messung der abhängigen Variablen sind annähernd intervallskaliert

¹¹Die Unabhängigkeit der Gruppen/Beobachtungen ist bei der kundengruppenübergreifenden Analyse gegeben. Beim kundengruppeninternen Vergleich jedoch nicht, da die Treiber-MW innerhalb einer Gruppe verglichen werden. Diese Einschätzungen innerhalb der Gruppe zu den Treiberfaktoren sind nicht voneinander unabhängig, da ein Proband der Befragung alle acht Treiberfaktoren für die Kundengruppe einschätzt. Dennoch wird die ANOVA angewandt, da sie als Methode noch bestmöglich geeignet ist, um die Bedeutung der Treiberfaktoren gruppenintern zu priorisieren. Die Ergebnisse zeigen sich auch für die kundengruppeninterne Analyse als stabil, das heißt mit eindeutigen p-Werten.

Tabelle 4: Analyseergebnisse zur internen Skalenkonsistenz nach Cronbachs Alpha

Kundengruppe/ -perspektive	B2C (Privatkunde)	B2B-M (Mitarbeiter)	B2B-O (Organisation)
Sample	$n = 125$	$n = 105$	$n = 92$
Faktor <i>Finanzieller Anreiz</i> (3 Items)	Cron $\alpha = 0,76$	Cron $\alpha = 0,54$	Cron $\alpha = 0,60$
Faktor <i>Lerneffekt</i> (3 Items)	Cron $\alpha = 0,88$	Cron $\alpha = 0,79$	Cron $\alpha = 0,69$
Faktor <i>Reputation</i> (3 Items)	Cron $\alpha = 0,89$	Cron $\alpha = 0,73$	Cron $\alpha = 0,80$
Faktor <i>Unbefriedigtes Bedürfnis</i> (2 Items)	Cron $\alpha = 0,93$	Cron $\alpha = 0,94$	Cron $\alpha = 0,85$
Faktor <i>Networking</i> (3 Items)	Cron $\alpha = 0,84$	Cron $\alpha = 0,78$	Cron $\alpha = 0,80$
Faktor <i>Interesse an Innovationen</i> (2 Items)	Cron $\alpha = 0,81$	Cron $\alpha = 0,79$	Cron $\alpha = 0,85$
Faktor <i>Spaß</i> (2 Items)	Cron $\alpha = 0,78$	Cron $\alpha = 0,78$	Cron $\alpha = 0,84$
Faktor <i>Altruismus</i> (2 Items)	Cron $\alpha = 0,76$	Cron $\alpha = 0,76$	Cron $\alpha = 0,85$

(de Winter und Dodou, 2010; Winship und Mare, 1984). Die visuelle Prüfung der Verteilung der abhängigen Variablen ergibt, dass diese optisch einer Normalverteilung entsprechen. Auch wenn der Shapiro-Wilk-W-Test dies rechnerisch nicht bestätigt, liegt das n pro Kundengruppe so hoch, dass diese Voraussetzung erfüllt ist. Dies gilt auch für die Varianzhomogenität (Groß, 2018).¹² An dieser Stelle sei darauf hingewiesen, dass bei nachfolgender Ergebnisdokumentation immer der adjustierte R^2 -Wert angegeben ist.

Vergleich der Bedeutung einzelner Treiber zwischen den Kundengruppen

Die kundengruppenübergreifende Analyse der Mittelwerte zu den Kundentreibern für eine VCC-Teilnahme zeigt signifikante Unterschiede zwischen den Gruppen B2C, B2B-Mitarbeiter und B2B-Organisation. Die Ergebnisse hierzu sind in Tabelle 5 zusammengefasst.

Hypothese 1 ist bestätigt. Die Varianzanalyse belegt Unterschiede in der Bedeutung aller Treiberfaktoren für die verschiedenen Kundengruppen. Einzige Ausnahme ist der Treiber Finanzieller Anreiz, bei dem sich kein signifikanter Unterschied für die Wichtigkeit feststellen lässt. Das heißt, dieser Treiberfaktor wird von allen drei Kundengruppen

ähnlich bedeutend eingeschätzt. Die anderen Treiber werden insbesondere von der B2C-Gruppe im Vergleich zu den B2B-Kundengruppen häufig unterschiedlich bewertet, wohingegen auch Differenzen in der Einschätzung zwischen den Gruppen B2B-Organisation und B2B-Mitarbeiter zu erkennen sind.

Grundsätzlich ist festzustellen, dass die Streuung in den Kundengruppen relativ hoch ist und sich diese z. T. mit der in den anderen Kundengruppen überschneidet. Die erklärte Effektstärke der MW-Differenzen zur Bedeutung einzelner CC-Treiber durch die Kundengruppen ist daher insgesamt betrachtet vergleichsweise gering (interne Validität). Hierauf weist der adjustierte R^2 -Wert hin. Der nur durch die unterschiedlichen Kundengruppen erklärte Teil der Gesamtvarianz liegt, je nach Treiberfaktor, bei ca. 2 bis zu 21% (vgl. Tabelle 5). Die MW zeigen allgemein eine „Tendenz zur Mitte“ der 7er-Likert-Skala, das heißt, sie sind häufig um den Wert 4 gelagert. Für die signifikanten MW-Differenzen gilt also, dass diese insbesondere Aufschluss für große Kundengruppen geben können, das heißt nur insgesamt „als großes Mittel“ zu betrachten sind und weniger Einzelfallvorhersagen für die Beurteilungen eines Probanden einer Kundengruppe zulassen.

Für die Gruppen lassen sich bei der Analyse einzelner

¹²Die Ergebnisse der statistischen Tests zur Normalverteilung (Shapiro-Wilk-W-Test) und zur Varianzhomogenität (Levene-Test) sind im Anhang in Tabelle 12 zusammengefasst.

¹³MW auf Basis von 7er-Likert-Skalen. Es wird der adjustierte R^2 -Wert angegeben.

Tabelle 5: Kundengruppenübergreifender Vergleich zur Bedeutung der CC-Treiber¹³

Kundengruppe/-perspektive	B2C (Privatkunde)	B2B-M (Mitarbeiter)	B2B-O (Organisation)	Auswertung (je Treiberfaktor eine ANOVA-Berechnung; UV = Kundengruppe, AV = Bedeutung des Treiberfaktors als MW)
Sample	<i>n</i> = 125	<i>n</i> = 105	<i>n</i> = 92	Probanden je Kundengruppe; <i>n</i> gesamt = 322
Faktor Finanzieller Anreiz (3 Items)	5,315 (MW) 0,103 (Std-F)	5,095 (MW) 0,101 (Std-F)	5,217 (MW) 0,094 (Std-F)	F (2; 319) = 1,24. <i>p</i> _{gesamt} = 0,2918. <i>R</i> ² = 0,00147. Keine signifikanten MW-Unterschiede
Faktor Lerneffekt (3 Items)	4,155 (MW) 0,114 (Std-F)	5,308 (MW) 0,125 (Std-F)	4,801 (MW) 0,133 (Std-F)	F (2; 319) = 23,84. <i>p</i> _{gesamt} < 0,0001. <i>R</i> ² = 0,12458. Post-hoc-Tests: B2B-M vs. B2C (<i>p</i> < 0,0001) sowie B2B-O vs. B2C (<i>p</i> < 0,0007) und B2B-M vs. B2B-O (<i>p</i> = 0,015) unterscheiden sich signifikant.
Faktor Reputation (3 Items)	2,893 (MW) 0,124 (Std-F)	4,235 (MW) 0,136 (Std-F)	4,533 (MW) 0,145 (Std-F)	F (2; 319) = 44,45. <i>p</i> _{gesamt} < 0,0001. <i>R</i> ² = 0,21302. Post-hoc-Tests: B2B-O vs. B2C (<i>p</i> < 0,0001) sowie B2B-M vs. B2C (<i>p</i> < 0,0001) unterscheiden sich signifikant.
Faktor Unbefriedigtes Bedürfnis (2 Items)	5,188 (MW) 0,118 (Std-F)	5,571 (MW) 0,128 (Std-F)	5,119 (MW) 0,137 (Std-F)	F (2; 319) = 3,54. <i>p</i> _{gesamt} = 0,0301. <i>R</i> ² = 0,01559. Post-hoc-Tests: B2B-M vs. B2B-O unterscheiden sich signifikant (<i>p</i> = 0,0443).
Faktor Networking (3 Items)	3,520 (MW) 0,122 (Std-F)	4,730 (MW) 0,133 (Std-F)	4,511 (MW) 0,142 (Std-F)	F (2; 319) = 2,94. <i>p</i> _{gesamt} < 0,0001. <i>R</i> ² = 0,13449. Post-hoc-Tests: B2B-M vs. B2C (<i>p</i> < 0,0001) sowie B2B-O vs. B2C (<i>p</i> < 0,0001) unterscheiden sich signifikant.
Faktor Interesse an Innovationen (2 Items)	4,925 (MW) 0,110 (Std-F)	5,633 (MW) 0,121 (Std-F)	4,995 (MW) 0,129 (Std-F)	F (2; 319) = 10,36. <i>p</i> _{gesamt} < 0,0001. <i>R</i> ² = 0,05512. Post-hoc-Tests: B2B-M vs. B2C (<i>p</i> < 0,0001) sowie B2B-O vs. B2B-M (<i>p</i> = 0,0010) unterscheiden sich signifikant.

(Continued)

Tabelle 5—continued

Faktor Spaß (2 Items)	5,072 (MW) 0,112 (Std-F)	5,757 (MW) 0,124 (Std-F)	4,505 (MW) 0,132 (Std-F)	F (2; 319) = 24,55. $p_{gesamt} < 0,0001$ R ² = 0,12796. Post-hoc-Tests: B2B-M vs. B2B-O (p < 0,001) sowie B2B-M vs. B2C (p < 0,001) sowie B2C vs. B2B-O (p = 0,0033) unterscheiden sich signifikant.
Faktor Altruismus (2 Items)	4,236 (MW) 0,127 (Std-F)	4,638 (MW) 0,139 (Std-F)	3,973 (MW) 0,149 (Std-F)	F (2; 319) = 5,53. $p_{gesamt} = 0,0043$. R ² = 0,02746. Post-hoc-Tests: B2B-O vs. B2B-M (p = 0,0033) unterscheiden sich signifikant.
Gesamtscore Rationale Treiber	4,121 (MW) 0,077 (Std-F)	4,879 (MW) 0,084 (Std-F)	4,850 (MW) 0,089 (Std-F)	Gesamtscore gleich gewichtet mit je 1/3 aus den Faktoren Finanzieller Anreiz, Lerneffekt, Reputation F (2; 319) = 28,50. $p_{gesamt} < 0,0001$. R ² = 0,14629. Post-hoc-Tests: B2B-M vs. B2C (p < 0,0001) sowie B2B-O vs. B2C (p < 0,0001) unterscheiden sich signifikant.
Gesamtscore Emotionale Treiber	4,752 (MW) 0,096 (Std-F)	5,343 (MW) 0,105 (Std-F)	4,491 (MW) 0,112 (Std-F)	Gesamtscore gleich gewichtet mit je 1/3 aus den Faktoren Interesse an Innovationen, Spaß, Altruismus F (2; 319) = 16,52. $p_{gesamt} < 0,0001$. R ² = 0,08817. Post-hoc-Tests: B2B-M vs. B2B-O (p < 0,0001) sowie B2B-M vs. B2C (p = 0,0001) unterscheiden sich signifikant.

Treiberfaktoren häufig Unterschiede zwischen der B2C-Perspektive und den B2B-Kundenperspektiven feststellen (vgl. Tabelle 5):

- Lerneffekte werden von B2C-Kunden als signifikant weniger bedeutsam wahrgenommen im Vergleich zur Perspektive B2B-Mitarbeiter und B2B-Organisation. Doch auch für die B2B-Gruppen unterscheidet sich die Wichtigkeit dieses Treibers. Für Mitarbeiter eines Kundenunternehmens ist der Lerneffekt als Treiber zur VCC-Teilnahme noch wichtiger als für die strategische Sichtweise von Unternehmen, die als organisationale Geschäftskunden an CC teilnehmen können (B2B-Organisation).
- Reputation ist für B2C-Kunden signifikant weniger bedeutend als für beide B2B-Kundengruppen. Mit einem MW von 2,89 fällt die B2C-Einschätzung gering aus, sodass streng genommen nur von einem schwachen Treiber für Privatkunden gesprochen werden kann. Für die Gruppen B2B-Mitarbeiter und B2B-Organisation ist die Bedeutung insgesamt höher. Dabei besteht kein signifikanter Unterschied zwischen diesen beiden B2B-Perspektiven.
- Unbefriedigtes Bedürfnis ist als Treiber bedeutend über alle Kundengruppen hinweg. Ein signifikanter Bedeutungsunterschied ist nur zwischen den Gruppen B2B-Mitarbeiter und B2B-Organisation festzustellen. B2B-Mitarbeiter schätzen diesen Faktor noch wichtiger ein.
- Networking ist ein weiterer Treiber, bei dem signifikante Differenzen zwischen der Wichtigkeit für B2C- und B2B-Kunden festzustellen sind. Während die Bedeutung für B2B-Kunden deutlich höher ausfällt, nehmen Privatkunden den Faktor als weniger wichtigen Anreiz für eine CC-Teilnahme wahr. Die Perspektiven B2B-Mitarbeiter und -Organisation unterscheiden sich in der Einschätzung hierzu nicht signifikant.
- Interesse an Innovationen wird von allen Kundengruppen als relativ bedeutender Treiber angesehen. Die Wichtigkeit ist jedoch für B2B-Mitarbeiter signifikant

höher als für die Kundengruppen B2B-Organisation und B2C. Letztgenannte Gruppen schätzen diesen Faktor gleich hoch ein.

- Spaß wird von allen Kundengruppen signifikant unterschiedlich beurteilt. Die höchste Bedeutung hat Spaß demnach für B2B-Mitarbeiter, gefolgt von B2C-Kunden und B2B-Organisationen – jeweils mit einem eher größeren Abstand.
- Altruismus ist insbesondere für B2B-Mitarbeiter und B2C-Kunden von eher höherer Bedeutung. Lediglich zwischen den Gruppen B2B-Mitarbeiter und B2B-Organisation lassen sich signifikante Differenzen nachweisen. Demnach ist dieser Faktor für die Sicht von Unternehmen als Geschäftskunden am unbedeutendsten (B2B-Organisation).

Zusätzlich zu dieser Einzelbetrachtung der Faktoren können diese inhaltlich in Gesamtscores zusammengefasst werden. Der Gesamtscore zu rationalen Treibern lässt sich aus den drei Faktoren Finanzieller Anreiz, Lerneffekt und Reputation generieren, welche eindeutig als rationale Treiber zu kategorisieren sind. Um die Vergleichbarkeit über alle Kundengruppen hinweg zu wahren, werden die Faktoren gleich gewichtet. Gleiches gilt für den Gesamtscore zu emotionalen Treibern, welcher sich analog aus den Faktoren Interesse an Innovationen, Spaß und Altruismus zusammensetzt.

- Der Gesamtscore Rationale Treiber zeigt, dass signifikante Unterschiede zwischen B2C- und den B2B-Kundenperspektiven bestehen. Rationale Treiber sind demnach insgesamt betrachtet für die Gruppen B2B-Mitarbeiter und B2B-Organisation wichtiger als für Privatkunden (B2C).
- Für den Gesamtscore Emotionale Treiber lässt sich eine signifikante Differenz zwischen B2B-Mitarbeiter und B2B-Organisation sowie B2B-Mitarbeiter und B2C feststellen. Emotionale Treiber werden von der Gruppe B2B-Mitarbeiter als besonders bedeutend wahrgenommen, gefolgt von B2C und der Gruppe B2B-Organisation.

Hypothese 2b ist demnach bestätigt. Es zeigt sich, dass die rationalen Treiber Lerneffekt und Reputation einzeln betrachtet für die Perspektiven der B2B-Kunden (B2B-Organisation, B2B-Mitarbeiter) jeweils signifikant bedeutender sind als für B2C-Kunden. Beim Faktor Finanzieller Anreiz sind jedoch prinzipiell keine Gruppenunterschiede feststellbar. Wird der Gesamtscore rationaler Treiber betrachtet, ist hier die Bedeutung für die B2B-Kundengruppen signifikant höher als für B2C.

Hypothese 3b ist nicht bestätigt. Die Analyse zeigt, dass der emotionale Treiber Spaß für B2C-Kunden signifikant bedeutender ist als für die Gruppe B2B-Organisation. Allerdings gilt dies nicht für die anderen emotionalen Faktoren Interesse an Innovationen und Altruismus. Auffallend ist, dass die

Gruppe B2B-Mitarbeiter den Treibern Interesse an Innovationen und Spaß signifikant höhere Bedeutungen zuweist, als es B2C-Kunden tun. Bei Betrachtung des Gesamtscores zu emotionalen Treibern ergeben sich ähnliche Ergebnisse: Es gibt keine signifikanten Unterschiede zwischen B2C und B2B-Organisation. B2B-Mitarbeiter und B2C unterscheiden sich signifikant, und zwar insofern, als B2B-Mitarbeiter emotionalen Faktoren insgesamt eine noch höhere Bedeutung zuweisen, als es die Gruppe B2C tut. Andere Ergebnisse zeigt die folgende kundengruppeninterne Analyse.

Vergleich der Bedeutung der Treiber innerhalb der Kundengruppen

Um Aussagen über die Priorisierung einzelner Treiberfaktoren innerhalb der Kundengruppen treffen zu können, wird jeweils ein kundengruppeninterner Vergleich durchgeführt. Die Ergebnisse dieser ANOVA-Berechnungen pro Kundengruppe sind in Tabelle 6 und 7 zusammengefasst. Kundengruppenintern zeigen sich signifikante Unterschiede für Treiber. Das heißt, dass den Treibern innerhalb einer Kundengruppe auch unterschiedliche Bedeutungen eingeräumt werden. Analog zu den Ergebnissen des Kapitels 5.2.4.2 ergeben sich hierbei vergleichsweise niedrige adjustierte R^2 -Werte von ca. 2 bis zu 27%. Die Bedeutung aller nicht über den gleichen statistischen Buchstaben verbundenen CC-Treiber in Tabelle 6 unterscheidet sich signifikant innerhalb der entsprechenden Kundengruppe.

Hypothese 2a ist nur teilweise bestätigt, nämlich für die Gruppe B2B-Organisation. Für B2B-Organisation sind Finanzieller Anreiz und Lerneffekt sowie Unbefriedigtes Bedürfnis und Interesse an Innovationen die bedeutendsten Treiber (statistische Gruppe A). Die beiden erstgenannten Faktoren sind rationale Treiber. Unbefriedigtes Bedürfnis lässt sich nicht eindeutig zuordnen. Interesse an Innovationen ist ein eher emotionaler Faktor, dem eine relativ hohe Wichtigkeit zugewiesen wird. Doch die Bedeutung der anderen beiden emotionalen Faktoren Spaß und Altruismus (statistische Gruppe C) ist für B2B-Organisation signifikant geringer. Die Betrachtung der Gesamtscores zeigt, dass rationale Treiber insgesamt für B2B-Organisation signifikant wichtiger sind als emotionale Treiber (vgl. Tabelle 7). Für die Kundengruppe B2B-Organisation kann 2a somit als bestätigt angesehen werden. Für B2B-Mitarbeiter zeigt sich ein gänzlich anderes Bild. Spaß wird zusammen mit Interesse an Innovationen, Unbefriedigtes Bedürfnis und Lerneffekt als wichtigster Treiber wahrgenommen (statistische Gruppe A). Altruismus und Reputation hingegen werden als signifikant weniger bedeutsam erachtet (statistische Gruppe D). Insbesondere die auffallend hohe Bedeutung für Spaß und das Innovationsinteresse führen zur Ablehnung von Hypothese 2a für die Gruppe B2B-Mitarbeiter. Dies zeigt auch die Betrachtung der Gesamtscores, nach denen emotionale Faktoren für B2B-Mitarbeiter insgesamt betrachtet signifikant wichtiger sind als rationale Faktoren.

Hypothese 3a ist bestätigt. Spaß, Interesse an Innovationen, Unbefriedigtes Bedürfnis und Finanzieller Anreiz

sind die wichtigsten CC-Treiber für Privatkunden (statistische Gruppe A). Mit Ausnahme von Finanzieller Anreiz, über dessen Sonderstellung bereits berichtet wurde, zeigt sich hierbei, dass eher die emotionalen Faktoren als wichtiger erachtet werden. Reputation ist als ein rationaler Faktor am unbedeutendsten. Lerneffekte nehmen hingegen eine mittlere Position ein.

Betrachtet man die kundengruppeninternen Analysen von B2B-Mitarbeiter und B2C, lassen sich diverse Gemeinsamkeiten in der Bedeutung bzw. internen Priorisierung der Treiber erkennen. Beide Gruppen für sich betrachtet geben den Treibern Spaß, Interesse an Innovationen und Unbefriedigtes Bedürfnis die höchste Bedeutung. Sie sind folglich stärker emotional getrieben (statistische A-Gruppen). Reputation ordnen beide Gruppen die geringste Wichtigkeit zu (statistische D-Gruppen). Die Gesamtscores zeigen das gleiche Bild. Insgesamt gesehen werden die drei emotionalen Treiber als signifikant wichtiger eingeschätzt als die rationalen Treiber. Rationale Treiber sind jedoch für die Gruppe B2B-Organisation am bedeutendsten.

5.2.5. Aktivitätenpräferenz und CC-Teilnahmebereitschaft der Kundengruppen

Weitere Ergebnisse der Umfrage sind in Tabelle 8 zusammengefasst. Die nominal-skalierte Direktabfrage ergibt, dass alle drei Kundengruppen Onlineaktivitäten der VCC bevorzugen (65% bei B2C, 45% bei B2B-Mitarbeiter, 45% bei B2B-Organisation). Allerdings haben bei den Gruppen der B2B-Kunden verhältnismäßig viele Probanden angegeben, keine eindeutige Präferenz für Online- oder Offlineaktivitäten zu haben (30% bei B2B-Mitarbeiter, 27% bei B2B-Organisation). Hypothese 4 ist somit bestätigt. Hypothese 5 ist nicht bestätigt.

Vor Abfrage der Teilnahmebereitschaft für das CC-Programm allgemein und für einzelne Aktivitäten werden die Umstände der Szenarien gemäß üblicher CC-Programme weiter konkretisiert. So gäbe es im Falle der Teilnahme vom Anbieter keine direkte Entschädigung. Allerdings würden viele der Ideen umgesetzt, Teilnehmer als Co-Producer genannt sowie Rabatte für zukünftige Produkte des Anbieters gewährt. Die Auswertung zeigt, dass die allgemeine Teilnahmebereitschaft für beide B2B-Perspektiven signifikant höher ist als für B2C-Kunden. Gleiches gilt für die Teilnahmebereitschaft zu Workshops und Interviews. Bei Onlinetools hingegen gibt es keine signifikanten Unterschiede in der Teilnahmebereitschaft. Für Onlineforen ist die Teilnahmebereitschaft von B2B-Organisation höher als die von B2C.

Stellt man den Probanden anschließend doch eine direkte finanzielle Entschädigung für eine Aktivität in Aussicht, entscheidet sich der überwiegende Teil der B2B-Gruppen für die Teilnahme an Workshops (53% bei B2B-Mitarbeiter, 47% bei B2B-Organisation). Bei B2C ergibt sich dabei keine eindeutige Präferenz.

Das Ranking der Kundengruppen zu CC-Phasen ergibt ein differenziertes Bild. Es zeigt sich, dass die Gruppe B2B-Organisation insgesamt gesehen die Präferenz äußert, sich in der Ideation-Phase an der Wertschöpfung zu beteiligen.

Bei der Gruppe B2B-Mitarbeiter gibt es insgesamt ein höheres Interesse für Development, während B2C eher Testing im Rahmen von CC bevorzugt.

5.3. Darstellung und Interpretation der Umfrageergebnisse

Ein Überblick der Hypothesen und Ergebnisse hierzu ist nachfolgend zusammengestellt (vgl. Tabelle 9). Insgesamt bringt die Umfrage ambivalente Ergebnisse, die zur Bestätigung und Ablehnung zentraler Annahmen führen.

Auf der folgenden Seite sind die Bedeutungsunterschiede der CC-Treiber zunächst kundengruppenübergreifend und anschließend innerhalb der einzelnen Kundengruppen visualisiert. Zudem ist die CC-Teilnahmebereitschaft sowie die Aktivitätenpräferenz der Gruppen dargestellt (vgl. Abbildung 5).

Die Betrachtung des MW-Liniendiagramms oben zeigt, dass die Kundengruppe B2B-Mitarbeiter Treibern im Vergleich zu den beiden anderen Gruppen insgesamt eine höhere Bedeutung zuweist. Die Einschätzung der B2C-Gruppe ist hingegen oftmals am geringsten. Die höchsten Differenzen in Hinblick auf die Bedeutung der Treiber zwischen den Gruppen zeigen sich bei Reputation, Networking, Lerneffekt und Spaß. Den geringsten bzw. keinen Unterschied zwischen den Gruppen gibt es bei Finanzieller Anreiz und Unbefriedigtes Bedürfnis.

Unbefriedigte Bedürfnisse sind für alle Gruppen ähnlich wichtig, wenn es um die Entscheidung bzgl. einer CC-Teilnahme geht. Dies erscheint plausibel. Wird die Anbieterlösung grundsätzlich benötigt, aber derzeitigen Ansprüchen noch nicht gerecht, liegt es nahe, sich an einer CC zu beteiligen. Durch diese Integration können Kunden in direktem Austausch mit dem Anbieterunternehmen auf Problematiken hinweisen. Nur wenn der Anbieter erfährt, welche Bedürfnisse derzeit nicht erfüllt werden, kann darauf angemessen mit neuen Funktionalitäten reagiert werden, die den Kunden entgegenkommen. Dies gilt für den B2B- und B2C-Kontext gleichermaßen.

Finanzieller Anreiz nimmt eine Sonderrolle ein. Die höchste Bedeutung für B2C-Kunden ist überraschend und widerspricht z. T. bisherigen Forschungsergebnissen (Chen et al., 2012, S. 145; Füller, 2010, S. 108 ff.; Hoyer et al., 2010, S. 289). Geld scheint einen starken zusätzlichen Anreiz für CC-Teilnahmen darzustellen, der auch unabhängig von anderen Treibern betrachtet werden kann. Die hohe Bedeutung für B2B-Organisation ist hingegen dadurch zu erklären, dass Unternehmen als Kunden gewinnorientiert agieren. Daher sind Unternehmen insbesondere dann zu einer CC-Teilnahme bereit, wenn sie hierin einen finanziellen Mehrwert direkter oder indirekter Natur sehen. Oftmals kann dieser durch Einsparungen bei internen Prozessabläufen erreicht werden, da durch VCC die Anbieterlösungen eine kundenorientierte Anpassung erfahren. Die mittlere Bedeutung für B2B-Mitarbeiter ist ggf. dadurch zu erklären, dass Mitarbeiter im Rahmen der beruflichen Tätigkeit nur bedingt arbeitgeberkonformes Verhalten priorisieren. Zwar sind Mitarbeiter eines Unternehmens an die Ziele des Arbeitgebers gebunden, jedoch ist diese Orientierung nicht

Tabelle 6: Kundengruppeninterner Vergleich zur Bedeutung der CC-Treiber^a

Treiberfaktor	Finanzieller Anreiz (3 items)	Lerneffekt (3 items)	Reputation (3 items)	Unbefriedigtes Bedürfnis (2 items)	Networking (3 items)	Interesse an Innovationen (2 items)	Spaß (2 items)	Altruismus (2 items)
B2C (Privatkunde) <i>n</i> = 125	5,315 (MW) 0,120 (Std-f) Gruppe A	4,155 (MW) 0,120 (Std-f) Gruppe B	2,893 (MW) 0,120 (Std-f) Gruppe D	5,188 (MW) 0,120 (Std-f) Gruppe A	3,520 (MW) 0,120 (Std-f) Gruppe C	4,925 (MW) 0,120 (Std-f) Gruppe A	5,072 (MW) 0,120 (Std-f) Gruppe A	4,236 (MW) 0,120 (Std-f) Gruppe B
Innerhalb der Kundengruppe B2C ergeben sich signifikante Bedeutungsunterschiede für CC-Treiber. $F(7; 992) = 52/65$. $P_{gesamt} < 0,0001$. $R^2 = 0,26586$. Post-hoc-Tests ergeben:								
Ingesamt unterscheidet sich die Bedeutung aller Treiber für B2C signifikant, die nicht über den gleichen statistischen Gruppen-Buchstaben verbunden sind (A-B).								
Die p-Werte der jeweils signifikanten Paarvergleiche liegen zwischen 0,0001 und 0,0059. Das heißt z. B. die Bedeutung für <i>Finanzieller Anreiz</i> vs. <i>Reputation</i> unterscheidet sich signifikant ($p < 0,0001$), während sich die Bedeutung für <i>Finanzieller Anreiz</i> vs. <i>Interesse an innovation</i> nicht signifikant unterscheidet ($p < 0,3804$)								
B2B-M (Mitarbeiter) <i>n</i> = 105	5,095 (MW) 0,119 (Std-F) Gruppe B, C	5,308 (MW) 0,119 (Std-F) Gruppe A, B	4,235 (MW) 0,119 (Std-F) Gruppe D	5,571 (MW) 0,119 (Std-F) Gruppe A, B	4,730 (MW) 0,119 (Std-F) Gruppe C, D	5,633 (MW) 0,119 (Std-F) Gruppe A	5,757 (MW) 0,119 (Std-F) Gruppe A	4,638 (MW) 0,119 (Std-F) Gruppe C, D
Innerhalb der Kundengruppe B2B-Mitarbeiter ergeben sich signifikante Bedeutungsunterschiede für CC-Treiber, $F(7; 832) = 20,99$. $P_{gesamt} < 0,0001$. $R^2 = 0,14298$ Post-hoc-Tests ergeben:								
Ingesamt unterscheidet sich die Bedeutung aller Treiber für B2B-M signifikant, die nicht über den gleichen statistischen Gruppen-Buchstaben verbunden sind (A-B).								
Die p-Werte der jeweils signifikanten Paarvergleiche liegen zwischen 0,0001 und 0,0307. Das heißt z. B. die Bedeutung für <i>Spaß</i> vs. <i>Reputation</i> unterscheidet sich signifikant ($p < 0,0001$), während sich die Bedeutung für <i>Networking</i> vs. <i>Reputation</i> nicht signifikant unterscheidet ($p < 0,0651$)								
B2B-O (Organisation) <i>n</i> = 92	5,217 (MW) 0,136 (Std-F) Gruppe A	4,801 (MW) 0,136 (Std-F) Gruppe A,B	4,533 (MW) 0,136 (Std-F) Gruppe B, C	5,119 (MW) 0,136 (Std-F) Gruppe A	4,511 (MW) 0,136 (Std-F) Gruppe B, C	4,995 (MW) 0,136 (Std-F) Gruppe A, B	4,505 (MW) 0,136 (Std-F) Gruppe B, C	3,973 (MW) 0,136 (Std-F) Gruppe C
Innerhalb der Kundengruppe B2B-Organisation ergeben sich signifikante Bedeutungsunterschiede für CC-Treiber, $F(7; 728) = 9,08$. $P_{gesamt} < 0,0001$. $R^2 = 0,07145$ Post-hoc-Tests ergeben:								
Ingesamt unterscheidet sich die Bedeutung aller Treiber für B2B-O signifikant, die nicht über den gleichen statistischen Gruppen-Buchstaben verbunden sind (A-C).								
Die p-Werte der jeweils signifikanten Paarvergleiche liegen zwischen 0,0001 und 0,0474. Das heißt z. B. die Bedeutung für <i>Finanzieller Anreiz</i> vs. <i>Altruismus</i> unterscheidet sich signifikant ($p < 0,0001$), während sich die Bedeutung für <i>Reputation</i> vs. <i>Altruismus</i> nicht signifikant unterscheidet ($p=0,0708$)								

^aMW auf Basis von 7er-Likert-Skalen. Es wird der adjustierte R^2 -Wert angegeben, auch in Tabelle 7.

Tabelle 7: Gesamtscores der kundengruppeninternen Analyse

Treiber-kategorisierung	Gesamtscore <i>Rationale Treiber</i>	Gesamtscore <i>Emotionale Treiber</i>	Auswertung (je Kundengruppe eine ANOVA-Berechnung; UV = Treiberkategorisierung rational/emotional, vgl. S. 47; AV = Bedeutung der Treiberkategorien als MW ergeben die Gesamtscores)
B2C (Privatkunde) n = 125	4,121 (MW) 0,091 (Std-F)	4,752 (MW) 0,091 (Std-F)	F (1; 248) = 23,81. $p < 0,0001$. $R^2 = 0,08393$. Emotionale Treiber haben eine signifikant höhere Bedeutung für B2C im Vergleich zu rationalen Treibern.
B2B-M (Mitarbeiter) n = 105	4,879 (MW) 0,083 (Std-F)	5,343 (MW) 0,083 (Std-F)	F (1; 208) = 15,31. $p < 0,0001$. $R^2 = 0,06407$. Emotionale Treiber haben eine signifikant höhere Bedeutung für B2B-M im Vergleich zu rationalen Treibern.
B2B-O (Organisation) n = 92	4,850 (MW) 0,108 (Std-F)	4,491 (MW) 0,108 (Std-F)	F (1; 182) = 5,51. $p = 0,02$. $R^2 = 0,02404$. Rationale Treiber haben eine signifikant höhere Bedeutung für B2B-O im Vergleich zu emotionalen Treibern.

immer vom Management nachzuvollziehen. Besonders interessant ist in diesem Zusammenhang, dass B2B-Mitarbeiter Spaß als bedeutendsten Treiber für die eigene CC-Teilnahme einschätzen, gefolgt vom Interesse an Innovationen. Dies zeigt, dass B2B-Mitarbeiter eher emotionalen Treibern eine starke Bedeutung zuordnen. Zwar wurde schon in der Literatur auf zunehmend emotionale Beweggründe im B2B-Kontext hingewiesen. Dennoch ist ein solch deutliches Ergebnis überraschend. Begründet werden könnte dies damit, dass Mitarbeiter im Rahmen ihrer täglichen Arbeit durch ihr Arbeitgeberunternehmen rational gesehen abgesichert sind. Das Maß an gewünschten Finanzen, Lernprogrammen und Reputation kann im Berufsalltag oftmals im eigenen Unternehmen erlangt werden, ohne sich dafür zusätzlich an anbieterinitiierten CC-Aktivitäten als Kunde beteiligen zu müssen. Insofern ist der hohe Stellenwert an Spaß und Innovationsinteresse für B2B-Mitarbeiter nachvollziehbar. Die Gruppe sieht eine VCC-Teilnahme weniger mit funktionalem Nutzen, sondern eher als Arbeitsalltagbereicherung. Dies erklärt auch die Ablehnung der Hypothese 3b. Die Bedeutung emotionaler Treiberfaktoren wird von B2B-Mitarbeitern noch höher eingeschätzt als von der Gruppe B2C. Zwar schätzt die B2C-Gruppe diese wiederum als bedeutender ein als B2B-Organisation. Diese Differenzen sind jedoch nicht durchgehend signifikant unterschiedlich – wie die kundengruppenübergreifende Analyse gezeigt hat.

Für B2B-Organisation bestätigt sich insgesamt eine höhere Wichtigkeit rationaler CC-Treiber. Spaß von Mitarbeitern und Altruismus haben die geringste Bedeutung. Der Anreiz von Finanzen und Unbefriedigten Bedürfnissen ist am höchsten. Das Interesse an Innovationen wird als Treiber unerwar-

tet hoch eingeschätzt, was damit zusammenhängen könnte, dass Unternehmen als Kunden diesem Faktor doch einen eher funktionalen Nutzen zuordnen. Innovationszyklen spielen für organisationale Kunden eine entscheidende Rolle. Im Ziel, durch eine CC-Teilnahme auch innovationstechnisch aktiv zu sein bzw. den Markt zu beobachten, liegt ein möglicher Erklärungsansatz. Ebenso ist die Bedeutung von Lerneffekten bei Kundenunternehmen hoch. Schließlich geht es darum, Wissen der Anbieter für interne Belange zu übertragen.

27% der Gruppe B2C geben an, bereits CC-Erfahrung zu haben. Bei der Gruppe B2B-Mitarbeiter sind es hingegen 34% und bei B2B-Unternehmen sogar 57%. Dies hängt damit zusammen, dass bei letzterer Gruppe das gesamte Unternehmen einzuschätzen ist. Die Chance bisheriger CC-Teilnahmen ist daher insgesamt recht hoch.

Alle drei Kundengruppen bevorzugen Onlineaktivitäten der CC. Während dies für B2C vorhergesagt wurde, ist das Ergebnis für die B2B-Gruppen unerwartet – insbesondere deshalb, weil im B2B-Kontext häufig Workshops als Präsenzveranstaltungen zur VCC stattfinden. Es zeigt sich aber, dass auch die B2B-Einschätzungen der Teilnahmebereitschaft für Onlinetools leicht höher als für Workshops ausfällt. Dies könnte damit zusammenhängen, dass der assoziierte Arbeitsaufwand online deutlich geringer als bei Präsenzveranstaltungen eingeschätzt wird. Zudem ist ein Großteil befragter Mitarbeiter in Unternehmen von sich aus onlineaffin. Die Nutzung von Onlinetools/-foren ist außerdem zeitunabhängig möglich – eben dann, wenn es sich in den Arbeitsalltag integrieren lässt.

¹⁴MW auf Basis von 7er-Likert-Skalen. Es wird der adjustierte R^2 -Wert angegeben

Tabelle 8: Aktivitätenpräferenz und CC-Teilnahmebereitschaft der Kundengruppen¹⁴

Kundengruppe/-perspektive	B2C (Privatkunde)	B2B-M (Mitarbeiter)	B2B-O (Organisation)	Auswertung
Sample	$n = 125$	$n = 105$	$n = 92$	Probanden je Kundengruppe; n gesamt = 322
bisherige CC-Erfahrung in bzw. aus der jeweiligen Perspektive der Kundengruppe	70% Nein (87), 27% Ja (34), 3% Weiß nicht, keine Angabe (4)	63% Nein (66), 34% Ja (34), 5% Weiß nicht, keine Angabe (5)	57% Ja (52), 29% Weiß nicht, keine Angabe (27), 14% Nein (13)	Bei B2C und B2B-M gibt es überwiegend bislang noch keine CC-Erfahrung. Die Unternehmensperspektive (B2B-O) zeigt, dass Arbeitgeber häufiger an CC teilnehmen.
Präferenz zur Art der CC-Aktivitäten	65% Online (81), 19% Offline (24), 16% Keine Präferenz (20)	45% Online (47), 30% Keine Präferenz (31), 26% Offline (27)	45% Online (41), 28% Offline (26), 27% Keine Präferenz (25)	B2C zeigt eine deutliche Präferenz für Onlineaktivitäten. Für die B2B-Gruppen zeigt sich ebenfalls eine solche Präferenz.
Aktivitätenwahl bei direkter Entschädigung des Anbieters	36% Workshop (45), 34% Online Tool (43), 18% Online Forum (23), 11% Interview (14)	53% Workshop (56), 18% Online Tool (19), 15% Interview (16), 13% Online Forum (14)	47% Workshop (43), 27% Online Tool (25), 17% Online Forum (16), 9% Interview (8)	Bei Entschädigung würden sich ca. 50% der Probanden aus B2B-M und B2B-O für eine Workshopteilnahme entscheiden. Bei der Gruppe B2C gibt es hingegen keine eindeutige Präferenz.
Teilnahmebereitschaft am CC-Programm allgemein	3,896 (MW) 0,131 (Std-F)	4,838 (MW) 0,143 (Std-F)	4,5000 (MW) 0,153 (Std-F)	$F(2; 319) = 12,33$. $p_{gesamt} < 0,0001$. $R^2 = 0,06594$. Post-hoc-Tests: B2B-M vs. B2C ($p < 0,0001$) sowie B2B-O vs. B2C ($p = 0,0079$) unterscheiden sich signifikant .
Teilnahmebereitschaft Workshops	3,536 (MW) 0,146 (Std-F)	4,438 (MW) 0,159 (Std-F)	4,413 (MW) 0,170 (Std-F)	$F(2; 319) = 4,42$. $p_{gesamt} < 0,0001$. $R^2 = 0,06096$. Post-hoc-Tests: B2B-M vs. B2C ($p < 0,0001$) sowie B2B-O vs. B2C ($p = 0,0003$) unterscheiden sich signifikant .
Teilnahmebereitschaft Interviews	3,216 (MW) 0,139 (Std-F)	4,143 (MW) 0,153 (Std-F)	3,935 (MW) 0,163 (Std-F)	$F(2; 319) = 11,23$. $p_{gesamt} < 0,0001$. $R^2 = 0,05994$. Post-hoc-Tests: B2B-M vs. B2C ($p < 0,0001$) sowie B2B-O vs. B2C ($p = 0,0026$) unterscheiden sich signifikant .

(Continued)

Tabelle 8—continued

Teilnahmebereitschaft Onlinetools	4,232 (MW) 0,141 (Std-F)	4,714 (MW) 0,154 (Std-F)	4,717 (MW) 0,165 (Std-F)	F (2; 319) = 3,61. $p_{gesamt} = 0,0281$. $R^2 = 0,01601$. Keine signifikanten MW-Unterschiede (bei Post-hoc-Tests ergeben sortierte Differenzen ein $p > 0,05$ für jede Stufe).
Teilnahmebereitschaft Onlineforen	3,816 (MW) 0,146 (Std-F)	4,035 (MW) 0,160 (Std-F)	4,685 (MW) 0,171 (Std-F)	F (2; 319) = 7,67. $p_{gesamt} < 0,0006$. $R^2 = 0,03991$. Post-hoc-Tests: B2B-O vs. B2C ($p = 0,0004$) unterscheiden sich signifikant .
Präferenzwahl Ranking der Ideation	bzw. Phase 1. Präferenz 26% (32), 2. Präferenz 23% (29), 3. Präferenz 30% (38), 4. Präferenz 21% (26)	1. Präferenz 33% (35), 2. Präferenz 21% (22), 3. Präferenz 26% (27), 4. Präferenz 20% (21)	* 1. Präferenz 53% (48) , 2. Präferenz 23% (21), 3. Präferenz 10% (9), 4. Präferenz (13)	Die Präferenzen sind ähnlich verteilt. Es ist keine deutliche Tendenz erkennbar. *B2B-O hier $n = 91$, da Präferenzwahl bei einem Probanden technisch fehlerhaft war.
Präferenzwahl Ranking der Development	bzw. Phase 1. Präferenz 24% (30), 2. Präferenz 39% (49) , 3. Präferenz 23% (29), 4. Präferenz 14% (17)	1. Präferenz 28% (29), 2. Präferenz 46% (48), 3. Präferenz 21% (22), 4. Präferenz 6% (6)	* 1. Präferenz 25% (23), 2. Präferenz 55% (50) , 3. Präferenz 14% (13), 4. Präferenz 5% (5)	Fast die Hälfte der Probanden aus B2B-M geben <i>Development</i> als zweite Präferenz an.
Präferenzwahl Ranking der Testing	bzw. Phase 1. Präferenz 42% (53) , 2. Präferenz 23% (29), 3. Präferenz 16% (20), 4. Präferenz 18% (23)	1. Präferenz 34% (36), 2. Präferenz 19% (20), 3. Präferenz 29% (30), 4. Präferenz 18% (19)	* 1. Präferenz 11% (10), 2. Präferenz 13% (12), 3. Präferenz 52% (47) , 4. Präferenz 24% (22)	Der Großteil an B2C-Kunden geben <i>Testing</i> als erste Präferenz an.
Präferenzwahl Ranking der Support	bzw. Phase 1. Präferenz 8% (10), 2. Präferenz 14% (18), 3. Präferenz 30% (38), 4. Präferenz 47% (59)	1. Präferenz 5% (5), 2. Präferenz 14% (15), 3. Präferenz 25% (26), 4. Präferenz 56% (59)	* 1. Präferenz 11% (10), 2. Präferenz 9% (8), 3. Präferenz 24% (22), 4. Präferenz 56% (51)	<i>Support</i> scheint für alle Kundengruppen eher weniger reizvoll. Die meisten Probanden stellen diese Phase im Ranking an die letzte Position für eine mögliche Teilnahme.

Tabelle 9: Überblick zur Bestätigung und Ablehnung der Hypothesen

	Inhalt der Hypothese	Ergebnis
H1	Die Bedeutung mindestens eines Treibers für eine Co-Creation-Teilnahme der Kunden ist für die Kundengruppen B2C,B2B-Mitarbeiter und B2B-Organisation unterschiedlich.	bestätigt (Ausnahme <i>Finanzieller Anreiz</i>)
H2a	Für B2B-Kunden (B2B-Organisation, B2B-Mitarbeiter) sind rationale Treiber relativ gesehen von höherer Bedeutung als emotionale Treiber. Das heißt, dass <i>Finanzieller Anreiz</i> , <i>Lerneffekt</i> und <i>Reputation</i> kundengruppenintern tendenziell als wichtiger wahrgenommen werden als <i>Altruismus</i> , <i>Spaß</i> und <i>Interesse an Innovationen</i> .	teilweise bestätigt (das heißt nur für B2B-Organisation)
H2b	Für B2B-Kunden (B2B-Organisation, B2B-Mitarbeiter) sind rationale Treiber relativ gesehen von höherer Bedeutung als für B2C-Kunden, das heißt <i>Finanzieller Anreiz</i> , <i>Lerneffekt</i> und <i>Reputation</i> betrachtet im kundengruppenübergreifenden Vergleich zwischen den Gruppen.	bestätigt (Ausnahme <i>Finanzieller Anreiz</i>)
H3a	Für B2C-Kunden sind emotionale Treiber relativ gesehen von höherer Bedeutung als rationale Treiber. Das heißt, dass <i>Altruismus</i> , <i>Spaß</i> und <i>Interesse an Innovationen</i> kundengruppenintern tendenziell als wichtiger wahrgenommen werden als <i>Finanzieller Anreiz</i> , <i>Lerneffekt</i> und <i>Reputation</i> .	bestätigt (Ausnahme <i>Finanzieller Anreiz</i>)
H3b	Für B2C-Kunden sind emotionale Treiber relativ gesehen von höherer Bedeutung als für B2B-Kunden (B2B-Organisation, B2B-Mitarbeiter), das heißt <i>Altruismus</i> , <i>Spaß</i> und <i>Interesse an Innovationen</i> betrachtet im kundengruppenübergreifenden Vergleich zwischen den Gruppen.	nicht bestätigt
H4	Aus privater Perspektive werden Onlineaktivitäten bevorzugt. Das heißt, dass die Gruppe B2C-Kunden CC-Aktivitäten über das Internet bevorzugt.	bestätigt
H5	Aus beruflicher Perspektive werden Offlineaktivitäten bevorzugt. Das heißt, dass die Gruppe B2B-Mitarbeiter und die Gruppe B2B-Organisation CC-Präsenzveranstaltungen vor Ort bevorzugen.	nicht bestätigt

B2C-Kunden bevorzugen die CC-Phase Testing. Privatkunden wollen also eher neue Produkte/Dienstleistungen verwenden und auf Fehler prüfen als neue Ideen zu generieren. Wahrscheinlich sehen viele darin einen Vorteil, neue Lösungen direkt für sich selbst anzuwenden. B2B-Mitarbeiter präferieren Development. Das Weiterentwickeln bereits existierender Lösungen passt zum Berufskontext der Mitarbeiter. Sie haben täglich mit implementierten Lösungen im Unternehmen zu tun und können daher am besten Vorschläge für B2B-Anbieter machen, welche Funktionalitäten sinnvoll zu ergänzen wären. Die Gruppe B2B-Organisation bevorzugt Ideation. Unternehmen als Geschäftskunden benötigen häufig individuelle und keine vorgefertigten Lösungen. Aus strategischer Kundensicht macht daher Ideation Sinn, um in partnerschaftlichen Anbieterbeziehungen zu grundlegend neuen Produkten oder Dienstleistungen anzuregen. Dies ist im Unternehmensumfeld nicht unüblich. Das Kundenunternehmen kann anschließend als Referenzkunde profitieren und gemeinsam erarbeitete Konzepte als Erstes („First-Mover“) anwenden.

6. Schlussbetrachtungen

6.1. Implikationen und Handlungsempfehlungen für Anbieter

Die Analyse der Umfrage liefert umfassende Ergebnisse. Hieraus lassen sich Schlüsse insbesondere für Anbieterunternehmen ziehen, die sowohl B2C- als auch B2B-Kunden als Zielgruppe haben und VCC zur gemeinsamen Wertschöpfung nutzen wollen. Grundsätzlich gilt für alle Kundengruppen, dass die Wichtigkeit der CC-Treiber individuell geprägt und kontextabhängig ist. Auch wenn die Bedeutung einzelner Treiberfaktoren für bestimmte Kundengruppen signifikant höher ist, heißt das nicht, dass die anderen weniger bedeutsamen Treiber völlig außer Acht gelassen werden dürfen. Diese Studie basiert nicht auf „entweder oder“-Abfragen zur Bedeutung der Treiber. Denn im Rahmen von VCC werden meist implizit verschiedene Kundentreiber zeitgleich angesprochen. Je nachdem welche CC-Aktivitäten wie genutzt

¹⁵Eigene Darstellung. MW auf Basis von 7er-Likert-Skalen.

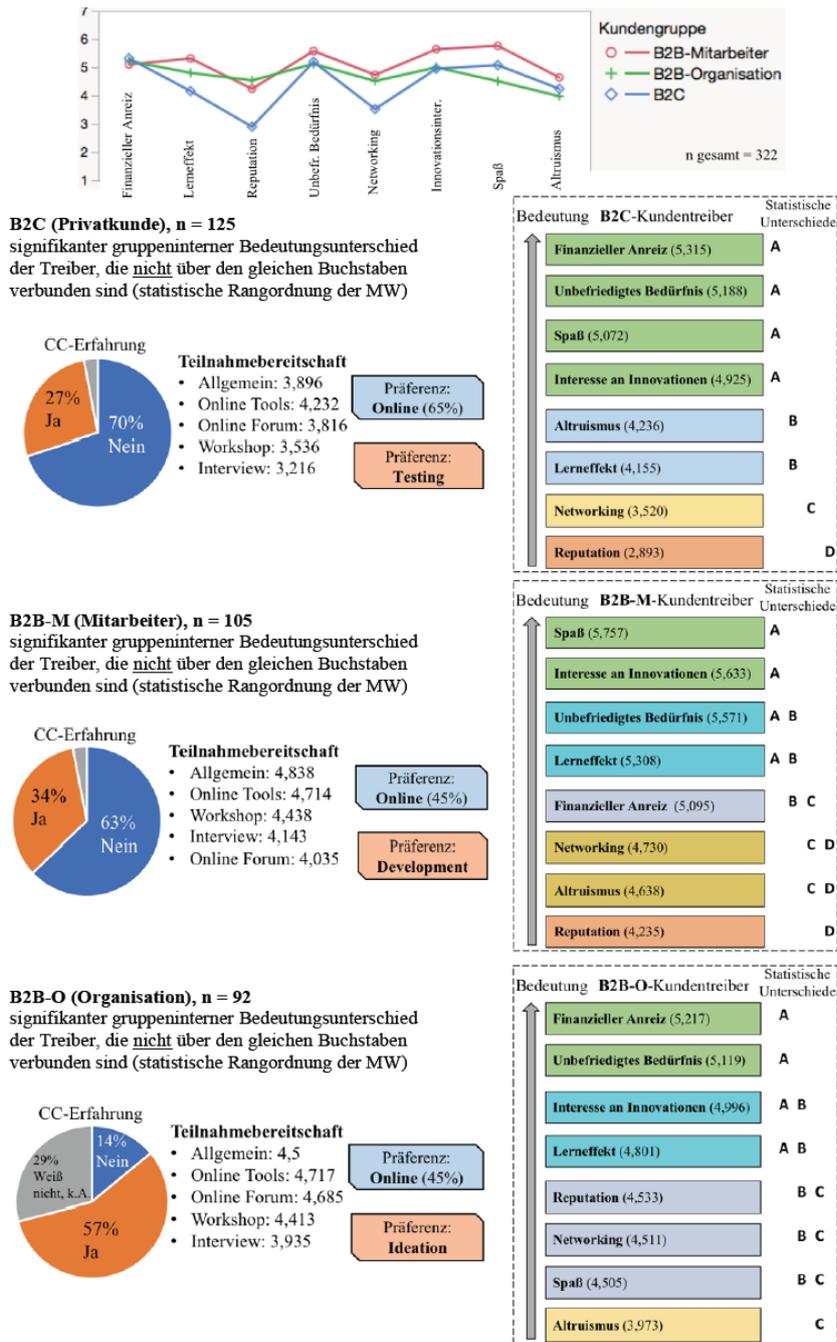


Abbildung 5: Visualisierung der Umfrageergebnisse ¹⁵

werden (z. B. Workshops, Onlinetools), gilt dies jedoch in unterschiedlicher Intensität. Insofern dienen die Erkenntnisse als Leitlinie in Hinblick darauf, mit welchem Fokus Anbieter CC-Maßnahmen für die Kundengruppen gestalten können.

CC-Programme sollten so initiiert werden, dass vor allem die bedeutendsten Treiber der Kundengruppen dadurch angesprochen werden. Für alle drei untersuchten Perspektiven gilt, dass ein Unbefriedigtes Bedürfnis und Interesse an Innovationen eine relativ hohe Bedeutung haben. Bei der Gestaltung aller CC-Aktivitäten ist folglich essentiell, Kunden anbieterseitig zu vermitteln, dass durch ihre CC-Teilnahme

ihre Bedürfnisse künftig besser vom Anbieter befriedigt werden können. Insofern sollten Anbieter alle geäußerten Probleme wertschätzen, darauf eingehen, gemeinsam priorisieren und anschließend erarbeitete Konzepte bestmöglich umsetzen. Dabei sind nicht immer alle Aspekte für jede Lösung realisierbar. Es kann auch zum Missbrauch von CC durch Akteure kommen, indem Konzepte eingebracht werden, die einen potentiell negativen Einfluss für Anbieterunternehmen haben können. Ein viel zitiertes Beispiel hierfür ist der Fall der Geschirrspülmittelmarke Pril. Hier wurde 2011 ein CC-Wettbewerb für das Labeldesign initiiert, bei

dem weniger ernst gemeinte „Spaßbeiträge“ viele Stimmen und Aufmerksamkeit bekommen haben (u. a. „tastes of chicken“, „PRIIIIIIIIIIIIIIL“). Aber auch andere Unternehmen wie General Motors und Ferrero haben negative Erfahrungen mit VCC-Beiträgen von Kunden gemacht (Baumgarth et al., 2016, S. 1 f.). Eine offene Kommunikation der Erwartungshaltung und möglicher Einschränkungen des Anbieters sind vorab wichtig, um sich aus Managementperspektive für solche Fälle abzusichern. Werden hingegen ernst gemeinte, priorisierte Bedürfnisse und Ideen von Kunden bei einem intensiven CC-Programm nicht umgesetzt, sollte dies auch vom Anbieter den einbringenden Kunden gegenüber verständlich begründet werden. Zumindest einfach zu bewerkstellende Anpassungen sollten aber definitiv stattfinden. Dem Innovationsinteresse kann entsprochen werden, wenn Anbieter offen kommunizieren und über CC detaillierte Informationen an vertraute Kunden weitergeben. Zusätzlich erscheint der Aufbau von Innovationscommunities mit verschiedenen Kunden sinnvoll. Dies kann über Onlineplattformen oder offline in Form von Veranstaltungen geschehen, die einen intensiven Austausch ermöglichen. So haben die Kunden den Eindruck, dass sie wirklich in den Innovationsprozess des Anbieters und darüber hinaus einbezogen werden, über Trends informiert sind und zeitgleich auch daraus Lernen können. Ein erfolgreiches Beispiel hierfür ist die Deutsche Notes User Group (DNUG), gegründet vor über 20 Jahren als Innovationsgemeinschaft für das IBM-Produktportfolio der digitalen Zusammenarbeit (E-Mails, Instant-Messaging, virtuelle Plattformen, etc.). Ursprünglich vom Anbieter IBM initiiert, hat sich die Nutzergemeinschaft inzwischen zu einem eigenständigen Verein (e.V.) entwickelt. Hier tauschen sich Mitarbeiter verschiedener Kundenunternehmen auf Events und in Onlineforen miteinander aus. Aber auch Privatpersonen sind willkommen. In enger Zusammenarbeit mit IBM werden Mitglieder mit neuesten Informationen versorgt und bringen regelmäßig Ideen ein, wie die Lösungen von IBM zukünftig noch besser zu gestalten sind (DNUG, 2018).

Mitarbeiter von Kundenunternehmen wollen sich beruflich vor allem an VCC beteiligen, wenn sie dabei Spaß haben. Folglich gilt es für Anbieter herauszufinden, welche Aktivitäten diesen Treiber bedienen können. Aufgrund der subjektiven Einschätzung macht es ggf. Sinn, verschiedene Formate anzubieten und Kunden (hier B2B-Mitarbeiter) selbst entscheiden zu lassen, woran sie sich beteiligen. Die Präferenz liegt in der CC-Phase Development und bei Onlineaktivitäten. Hierbei sollte in interaktive, ansprechende Formate bzw. Tools investiert werden, sodass am Ende Mehrwerte für beide Seiten, das heißt Anbieter- und Kundenseite, generiert werden. Dabei muss jedoch ein schwieriger „Spagat“ gelingen. Denn die organisationale, strategische Ebene von Kundenunternehmen (B2B-Organisation) ordnet Spaß der eigenen Mitarbeiter an hinterer Stelle als weniger bedeutenden CC-Treiber ein. Zunächst müssen Anbieter im B2B-Kontext das strategische Management von einer CC-Teilnahme überzeugen (B2B-Organisation). Hierfür ist der Finanzielle Anreiz von höchster Priorität. Andererseits gilt es, einzelne Mitarbeiter der Kundenunternehmen (B2B-Mitarbeiter) für die CC zu

gewinnen, die Spaß priorisieren. Folglich sollten vom Anbieter unterschiedliche Argumentationen zum Nutzen der CC, mit Blick auf die jeweilige Kundenperspektive, individuell hervorgehoben werden. Das Management von Geschäftskunden bevorzugt die CC-Phase Ideation. Kundenunternehmen können dabei als strategische Partner für Zukunftsprojekte gewonnen werden. So ist auch die Reputation von Anbieter- und Kundenunternehmen gemeinsam zu erhöhen.

Der hohen Bedeutung finanzieller Anreize für B2C-Kunden kann durch zukünftige Rabattcodes für eigene Lösungen oder Partner entsprochen werden. Eine direkte finanzielle Aufwandsentschädigung ist eher unüblich und kostenintensiv für die Anbieter. Diese kann in Einzelfällen Sinn machen, z. B. beim Start von CC-Projekten mit hohem Budget und strategischer Relevanz im B2C-Segment. Den Faktoren Spaß und Interesse an Innovationen kann wie bereits beschrieben begegnet werden. Offene Kommunikation mit dem Anbieter ist hierfür genauso entscheidend wie die Möglichkeit zum Austausch mit anderen Kunden im Innovationsprozess. Da die Gruppe B2C die CC-Phase Testing bevorzugt, sollten Anbieter insbesondere über Möglichkeiten der Kundenintegration bei Produkt- und Dienstleistungstests nachdenken. Das Unternehmen trnd hat sich bspw. als erfolgreicher Vermittler zwischen Anbietern und Privatkunden für solche Tests etabliert (trnd, 2018). Aber auch über eigene Kanäle können Anbieter Produkttests organisieren (z. B. Websites).

Letztendlich sind konkrete Anbietermaßnahmen situativ bestimmt. Die Bedeutung der Treiber für die Kundengruppen gibt eine erste Orientierung, worauf sich Anbieter bei der Gewinnung der Kunden zur CC-Teilnahme fokussieren können.

6.2. Restriktionen, Limitationen und zukünftige Forschung

Die Studie der Arbeit wurde so konzipiert, dass sie möglichst allgemeingültige Rückschlüsse auf Treiber für die unterschiedlichen Kundengruppen zulässt. Angewandte Instant Messaging-Szenarien sind sowohl für B2C- als auch für B2B-Kunden von hoher Relevanz; das heißt, Konstruktvalidität ist gegeben. Allerdings sind die gewonnenen Erkenntnisse entsprechend auf diese Szenarien im IT-Kontext beschränkt. Die Probanden wurden gebeten, sich in fiktive Szenarien hineinzuversetzen. Obwohl die Teilnehmer des Öfteren ein positives Feedback zum Fragebogen im Kommentarfeld abgegeben haben, gab es hierzu vereinzelt auch Kritik. So wurde bspw. bemängelt, dass die Szenarien und Fragen recht komplex seien oder für den konkreten Fall eines Probanden nicht zutreffend. Zwar ist davon auszugehen, dass sich die Teilnehmer grundsätzlich gut in die Szenarien hineinversetzen konnten und ehrlich geantwortet haben. Jedoch ist damit nicht auszuschließen, dass einzelne CC-Treiber bei einer Entscheidung unter realen Bedingungen noch einen anderen Stellenwert haben. Im Rahmen von Folgestudien erscheint daher ein Vergleich mittels weiterer CC-Szenarien im Produkt- und Dienstleistungsumfeld zwischen B2B- und B2C-Kundenperspektiven interessant. Hiermit kann überprüft werden, inwieweit die Erkenntnisse zur Bedeutung

der Treiber auch in anderen Situationen Gültigkeit behalten (externe Validität). Zudem kann untersucht werden, welche Interaktionseffekte und Abhängigkeiten der Treiber zueinander bestehen. Ferner ist zukünftig detaillierter zu betrachten, welche „negativen“ Aspekte Kundengruppen von einer CC-Teilnahme abhalten könnten. Hierauf wurde bisher in Studien kaum eingegangen. Ein weiterer Ansatz ist die Analyse, inwieweit Unterschiede bei VCC für Produkte und Dienstleistungen bestehen.

Obwohl die Items zur Messung der Treiberfaktoren zuverlässige Werte der internen Skalenzusammenhang ergeben, ist nicht auszuschließen, dass noch weitere Einflüsse eine entscheidende Rolle zur CC-Teilnahme für die Kundengruppen spielen. Um die Onlineumfrage kurz zu halten, wurde auf das Abfragen inhaltlich stark ähnlicher Teilaussagen verzichtet. Eine noch höhere Anzahl an Items kann jedoch die Validität und Reliabilität solcher Studien erhöhen. Mit Hilfe von „mixed-method“-Ansätzen aus qualitativer und quantitativer Forschung sollten die angewandten Fragen zur Erhebung der Bedeutung der Treiber erneut überprüft und zukünftig weiter ergänzt werden. So ist mittelfristig ein trennscharfes, allgemeingültiges Messmodell der CC-Treiber für Kunden zu etablieren. In Kombination mit der Überprüfung weiterer CC-Szenarien zur Bedeutung der Treiber lässt sich dabei ggf. auch ein Konstrukt bilden, mit dem noch höhere Effektstärken der Kundengruppen nachgewiesen werden (R^2 -Werte).

Natürlicherweise sind im Rahmen der quantitativen Studie keine detaillierten Rückschlüsse auf Einzelfälle oder zugrunde liegende Gedankenstrukturen der Probanden möglich. Um die Erkenntnisse dahingehend zu vertiefen, können zukünftig unterschiedliche qualitative Forschungsdesigns angewandt werden (Interviews, Beobachtungen, etc.) – auch in Hinblick auf verschiedene Szenarien und Gruppen.

Bei der Erhebung wurde aus Datenschutzgründen für die B2B-Perspektiven auf das Abfragen vom Arbeitgeberunternehmen, -größe und Branche verzichtet. Ferner stellen diese Informationen keinen Betrachtungsschwerpunkt der Arbeit dar. Insofern können die Ergebnisse für die Gruppe B2B-Organisation nur eine allgemeine Einschätzung bzw. Tendenz zu CC-Treibern für Kundenunternehmen wiedergeben. Ein Ansatz für Folgestudien wäre, CC-Treiber für B2B-Kundenperspektiven für sich gesehen noch differenzierter zu betrachten. Welche Unterschiede ergeben sich zwischen CC-Treibern größerer und kleinerer Kundenunternehmen? Gibt es Branchenunterschiede, das heißt, welche Treiber wirken in welcher Branche als besonders starker Anreiz für eine CC-Teilnahme? Hinsichtlich der Einschätzung der Unternehmensperspektive (B2B-Organisation) erscheint es sinnvoll, insbesondere höhere Manager zu befragen. Diese haben strategische Verantwortung und können unternehmerische Beweggründe einer CC-Teilnahme ihres Arbeitgebers somit noch besser einschätzen als „einfache“ Mitarbeiter. Letztere können aufgrund ihrer Berufserfahrung nur eine persönliche Einschätzung bzw. Vermutung für das Unternehmen abgeben – so wie in dieser Untersuchung geschehen. Dies eröffnet erste Einblicke. Dennoch wäre es anzustreben, künftige Untersuchungen der Kundenperspektive B2B-Organisation

mit Befragungen von Senior Managern (vorzugsweise auf C-Level-Ebene) durchzuführen, um zuverlässigere Ergebnisse zu strategischen Beweggründen einer CC-Teilnahme von Kundenunternehmen zu erhalten.

Auch die Perspektive der Mitarbeiter, die Anbieter kundenseitig für CC gewinnen müssen („individual level“, B2B-Mitarbeiter), kann in zukünftigen Untersuchungen noch detaillierter betrachtet werden; bspw. in Hinblick auf Verantwortungsbereiche, Tätigkeitsspektrum und Arbeitsumfeld. Zeigen sich bei den Mitarbeitern Unterschiede in der Bedeutung der CC-Treiber, wenn sie eine höhere oder niedrigere Stufe in der Unternehmenshierarchie innehaben? Wie werden Manager hierbei durch ihre Teams beeinflusst? Hierfür sind qualitative Ansätze (z. B. Ethnographie, Interviews) und quantitative Forschungsdesigns (z. B. Umfragen) geeignet. Interessant wäre auch zu erfahren, inwieweit sozio-kulturelle Prägungen die Einschätzung von Mitarbeitern zur eigenen CC-Teilnahme prägen.

Insgesamt entspricht das bereinigte Gesamtsample mit 322 Probanden einer akzeptablen Größe. Folglich ist davon auszugehen, dass die Wahrscheinlichkeit von zufällig entstandenen Ergebnissen und einem starken Einfluss von Ausreißern gering ist. So ergibt sich eine hohe Reliabilität der Ergebnisse. Deshalb lassen sich erste Rückschlüsse im Direktvergleich der Kundengruppen ziehen. Für die zukünftige Forschung zur Thematik wäre es aber wünschenswert, einen noch größeren und repräsentativen Stichprobenumfang in Umfragen zu erreichen.

6.3. Fazit

Die Konzeption dieser Arbeit zeigt, wie umfassend und aktuell das Themenfeld der VCC im B2C- und B2B-Kontext ist. Um im verstärkten Wettbewerb bestehen zu können, versuchen Anbieterunternehmen ihre Lösungen kundenorientiert zu gestalten und erarbeiten dabei zunehmend gemeinsam mit Kunden innovative Konzepte. Obwohl der Grundgedanke der CC ursprünglich aus dem B2B-Umfeld stammt, werden insbesondere über das Internet immer häufiger auch CC-Programme im B2C-Kontext initiiert. Entscheidend ist dabei die Frage, wie Anbieterunternehmen ihre Kunden dazu bewegen können, sich an CC-Aktivitäten zu beteiligen. Diese Kundentreiber, das heißt die Beweggründe von Kunden für eine Teilnahme, wurden durch die Onlineumfrage der Arbeit für B2C- und B2B-Kundenperspektiven detailliert untersucht.

Die Ergebnisse zeigen ein differenziertes Bild. Für die strategische Entscheidung, als Unternehmen an CC-Programmen teilzunehmen, sind tendenziell eher rationale Treiber von höherer Bedeutung, das heißt vor allem Finanzieller Anreiz. Aber auch ein bislang noch unbefriedigtes Bedürfnis mit dem aktuellen Angebot ist entscheidend, sich auf organisationaler Ebene als Geschäftskunde an VCC zu beteiligen. Die Mitarbeiter von Kundenunternehmen werten für sich jedoch völlig andere Treiber als bedeutsam für eine CC-Teilnahme im Rahmen ihres Jobs. Für die Mitarbeiter eines Kundenunternehmens haben eher emotionale Treiber eine höhere Bedeutung, allem voran der Faktor Spaß und das Interesse an Innovationen. Somit zeigt sich eine Diskrepanz zwischen

der Sicht des strategischen Managements von Kundenunternehmen (B2B-Organisation) und deren Mitarbeitern (B2B-Mitarbeiter), die im Rahmen von CC-Aktivitäten vom Anbieter zu adressieren sind. Für B2C-Kunden fallen vor allem jene Treiber auf, denen sie eine niedrige Bedeutung zuweisen: Reputation und Networking. Entgegen bisherigen Annahmen ist Finanzieller Anreiz mit der bedeutendste Treiberfaktor, aber auch ein Unbefriedigtes Bedürfnis, Spaß und Innovationsinteresse sind für Privatkunden wichtig. Grundsätzlich ist die Bedeutung der CC-Treiber für alle Kundengruppen jedoch immer vom spezifischen Kontext der Kundensituation abhängig.

Es zeigt sich, dass alle Kundengruppen Onlineaktivitäten der CC bevorzugen. Dies eröffnet zusätzliche Möglichkeiten für Anbieter, die sich im B2B-Kontext bisher auf Offlineaktivitäten fokussiert haben.

Im Kapitel 6.2 wurden einige Ansätze aufgezeigt, die für die weitere akademische Auseinandersetzung zu VCC mit Kunden von Bedeutung sind. Letztendlich bleibt abzuwarten, wie sich die Forschungsansätze in diesem Themenbereich entwickeln. Die zunehmende Praxisrelevanz lässt jedoch darauf schließen, dass sich Forscher auch zukünftig mit dem Thema intensiv auseinandersetzen. Gerade die Analyse von B2B- und B2C-Kudentreibern im direkten Vergleich ist ein breites Feld, welches bislang nur wenig erforscht wurde. Für zukünftige wissenschaftliche Arbeiten auf dem Gebiet hat diese Arbeit einen Ansatz und erste Erkenntnisse geliefert. Nun gilt es, diese noch weiter zu vertiefen.

Literatur

- Bagozzi, R. P. The role of social and self-conscious emotions in the regulation of business-to-business Abstract relationships in salesperson-customer interactions. *Journal of Business & Industrial Marketing*, 21(7):453–457, 2006.
- Baumgarth, C., Henseler, J., und Kristal, S. When co-creation goes wrong - Brand dilution through non-collaborative co-creation, April 2016. URL https://www.researchgate.net/publication/301677444_When_co-creation_goes_wrong-Brand_dilution_through_non-collaborative_co-creation.
- Breidbach, C. F. und Maglio, P. P. Technology-enabled value co-creation: an empirical analysis of actors, resources and practices. *Industrial Marketing Management*, 56:73–85, 2016.
- Brinkmann, J. und Voeth, M. An analysis of buying center decisions through the salesforce. *Industrial Marketing Management*, 36(7):998–1009, 2007.
- Broß, J. Director Co-Creation & Innovation (XU Exponential Game Changers GmbH), 2018. Persönliches Interview durch Steinhardt, Felix am 16. Mai 2018.
- Bruhn, M. und Strauss, B. *Kundenintegration im Dienstleistungsmanagement - Eine Einführung in die theoretischen und praktischen Problemstellungen*. Wiesbaden, Springer Gabler, 2009.
- Bundesamt, S. Dienstleistungen, 2018. URL <https://www.destatis.de/DE/ZahlenFakten/Wirtschaftsbereiche/Dienstleistungen/Dienstleistungen.html>.
- Cater, B. und Cater, T. Emotional and rational motivations for customer loyalty in business-to-business professional services. *The Service Industries Journal*, 29(8):1151–1169, 2009.
- Chen, L., Marsden, J. R., und Zhang, Z. Theory and Analysis of Company-Sponsored Value Co-Creation. *Journal of Management Information Systems*, 29(2):141–172, 2012.
- Corsaro, D. und Snehota, I. Searching for Relationship Value in Business Markets: Are we missing something? *Industrial Marketing Management*, 39(6):986–995, 2010.
- Cova, B. und Salle, R. The industrial/consumer marketing dichotomy revisited: a case of outdated justification? *Journal of Business & Industrial Marketing*, 23(1):3–11, 2007.
- Cunat, V. und Guadalupe, M. Globalization and the Provision of Incentives inside the Firm: The Effect of Foreign Competition. *Journal of Labor Economics*, 22(2):179–212, 2009.
- Dahl, D. W. und Moreau, P. Thinking Inside the Box: Why Consumers Enjoy Constrained Creative Experiences. *Journal of Marketing Research*, 44(3):357–369, 2007.
- Daiberl, C., Höckmayr, B., und Roth, A. How To Carry Out Co-Creation. *Online-Offline Co-Creation*, 4:12–13, 2016.
- de Winter, J. C. F. und Dodou, D. Five-Point Likert Items: t test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research & Evaluation*, 15(11):1–16, 2010.
- Dellarocas, C. The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms. *Management Science*, 49(10):1407–1424, 2003.
- Deutschland, U. Unilever - Unsere Marken, June 2017. URL <https://www.unilever.de/marken/>.
- DNUG. DNUG - Collaboration Professionals, August 2018. URL <https://dnug.de>.
- Dong, B. und Sivakumar, K. Customer participation in services: domain, scope and boundaries. *Journal of the Academy of Marketing Science*, 45(6):944–965, 2017.
- Döring, S. A. Seeing What to Do: Affective Perception and Rational Motivation. *dialectica*, pages 363–394, 2007.
- Eggert, A. und Ulaga, W. Customer perceived value: a substitute for satisfaction in business markets? *Journal of Business & Industrial Marketing*, 17(3/3):107–118, 2002.
- Egtar, M. A descriptive model of the consumer co-production process. *Journal of the Academy of Marketing Science*, 36(1):97–108, 2007.
- Füller, J. Refining Virtual Co-Creation from a Consumer Perspective. *California Management Review*, 52(2):98–122, 2010.
- Füller, J., Hutter, K., und Faullant, R. Why co-creation experience matters? Creative experience and its impact on the quantity and quality of creative contributions. *R&D Management*, 41(3):259–273, 2011.
- Flynn, N. Avoiding An Instant Messaging Nightmare. IM Policy Best Practices: A Business Guide to Clean and Compliant, Safe & Secure Instant Messaging. 2005.
- Foscht, T., Swoboda, B., und Schramm-Klein, H. *Käuferverhalten*. Wiesbaden, Springer Gabler, 2017.
- Franke, M.-K. *Hedonischer Konsum. Emotionen als Treiber im Konsumentenverhalten*. Wiesbaden, Springer Gabler, 2013.
- Franke, N. und Shah, S. How Communities Support Innovative Activities: Exploration of Assistance Among Users of Sporting Equipment. *Research Policy*, 32(1):157–178, 2003.
- Galvagno, M. und Dalli, D. Theory of value co-creation: a systematic literature review. *Managing Service Quality*, 24(6):643–683, 2014.
- Gemser, G. und Perks, H. Co-Creation with Customers: An Evolving Innovation Research Field. *The Journal of Product Innovation Management*, 32(5):660–665, 2015.
- Günter, B. und Kuhl, M. *Industrielles Beschaffungsmanagement*. Berlin, Springer, 2000.
- Grönroos, C. Value Co-creation in Service Logic: A Critical Analysis. *Marketing Theory*, 11(3):279–301, 2011.
- Groß, M. Wikis der Freien Universität Berlin, July 2018. URL <https://wikis.fu-berlin.de/pages/viewpage.action?pageId=689209609>.
- Gyrd-Jones, R. I. und Kornum, N. Managing the co-created brand: Value and cultural complementarity in online and offline multi-stakeholder ecosystems. *Journal of Business Research*, 66(9):1484–1493, 2013.
- Halfmann, M. *Der Konsument von morgen - Vom Homo oeconomicus zum Homo mysticus*. Wiesbaden, Springer Gabler, 2014.
- Handelsblatt. Bei McDonald's perlt es bald, July 2018. URL <https://www.handelsblatt.com/unternehmen/handel-konsumgueter/neuheiten-bei-der-fastfood-kette-mein-burger-aktion-erfolgreich/6674690-2.html?ticket=ST-1892063-ICddZeM6jNw6jkcXJivE-ap6>.
- Heckhausen, J. und Heckhausen, H. *Motivation und Handeln*. Berlin, Springer, 2006.
- Heinonen, K., Strandvik, T., Mickelsson, K.-J., Edvardsson, B., Sundström, E., und Andersson, P. A customer-dominant logic of service. *Journal of Service-Management*, 21(4):531–548, 2010.
- Hertel, G., Niedner, S., und Hermann, S. Motivation of Software Developers in Open Source Projects. *Research Policy*, 32:1159–1177, 2003.
- Hoyer, W. D., Chandy, R., Dorotic, M., Krafft, M., und Singh, S. Consumer Cocreation in New Product Development. *Journal of Service Research*, 13(3):283–296, 2010.
- Inc., W. WhatsApp Business, August 2018. URL <https://www.whatsapp.com/business/>.
- Jaakkola, E. und Alexander, M. The Role of Customer Engagement Behavior in Value Co-Creation: A System Perspective. *Journal of Service Research*, 17(3):274–261, 2014.
- Kaplan, A. M. und Haenlein, M. Toward a Parsimonious Definition of Traditional and Electronic Mass Customization. *The Journal of Product Innovation Management*, 23(2):168–182, 2006.
- Kleinaltenkamp, M. und Saab, S. *Einführung in das Business-to-Business-Marketing*. Berlin, Springer Verlag, 2009.
- Koch, M., Richter, A., und Schlosser, A. Produkte zum IT-gestützten Social Networking in Unternehmen. *Wirtschaftsinformatik*, 49(6):448–455, 2007.
- Kohtamäki, M. und Rajala, R. Theory and practice of value co-creation in B2B systems. *Industrial Marketing Management*, 56(7):4–13, 2016.
- Kollock, P. *The Economics of Online Cooperation*. Abingdon, Routledge, 1999.
- Kurzmann, H. und Reinecke, S. *Kundenintegration zwischen Kooperation und Delegation - Konzepte und Relevanz aus Managementsicht*. Wiesbaden, Springer Gabler, 2009.
- Lippold, D. *Marktorientierte Unternehmensführung und Digitalisierung. Management im digitalen Wandel*. Berlin, Walter de Gruyter GmbH, 2017.
- Lynch, J. und de Chernatony, L. The power of emotion: brand communication in business-to-business markets. *Journal of Brand Management*, 11(5):403–419, 2004.
- Marcos-Cuevas, J., Nätti, S., Palo, T., und Baumann, J. Value co-creation practices and capabilities: sustained purposeful engagement across B2B systems. *Industrial Marketing Management*, 56:97–107, 2016.
- Mencarelli, R. und Rivièrè, A. Perceived value in B2B and B2C: A comparative approach and cross-fertilization. *Marketing Theory*, 15(2):201–220, 2015.
- Mosely, G., Wright, N., und Wrigley, C. Facilitating design thinking: A com-

- parison of design expertise. *Thinking Skills and Creativity*, 27:177–189, 2018.
- Mustak, M., Jaakkola, E., Halinen, A., und Kaartemo, V. Customer participation management. Developing a comprehensive framework and a research agenda. *Journal of Service Management*, 27(3):250–275, 2016.
- Nambisan, S. und Baron, R. A. Interactions in Virtual Customer Environments: Implications for Product Support and Customer Relationship Management. *Journal of Interactive Marketing*, 21(2):42–62, 2007.
- Nambisan, S. und Baron, R. A. Virtual Customer Environments: Testing a Model of Voluntary Participation in Value Co-creation Activities. *The Journal of Product Innovation Management*, 26(4):388–406, 2009.
- Nambisan, S. und Nambisan, S. How to Profit From a Better 'Virtual Customer Environment'. *MIT Sloan Management Review*, 49(3):53–61, 2008.
- Narver, J. C., Slater, S. F., und MacLachlan, D. L. Responsive and Proactive Market Orientation and New-Product Success. *The Journal of Product Innovation Management*, 21(5):334–347, 2004.
- O'Hern, M. S. und Rindfleisch, A. Customer Co-Creation: A Typology and Research Agenda, December 2008. URL <https://propellingbrands.files.wordpress.com/2009/01/ohern-rindfleisch-120108.pdf>.
- Payne, A. F., Storbacka, K., und Frow, P. Managing the co-creation of value. *Journal of the Academy of Marketing Science*, 36(1):83–96, 2008.
- Pee, L. G. Customer co-creation in B2C e-commerce: does it lead to better new products? *Electronic Commerce Research*, 16(2):217–243, 2016.
- Petri, J. und Jacob, F. The customer as enabler of value (co)-creation in the solution business. *Industrial Marketing Management*, 56:63–72, 2016.
- Peuster, J. Value Co-Creation, October 2016. URL <https://www.geno-kom.de/blog/blog/value-co-creation/>.
- Porter, M. E. *Wettbewerbsvorteile. Spitzenleistungen erreichen und behaupten*. Frankfurt am Main, Campus Verlag GmbH, 2010.
- Prahalad, C. K. und Ramaswamy, V. Co-opting Customer Competence. *Harvard Business Review*, 78(1):79–87, 2000.
- Prahalad, C. K. und Ramaswamy, V. Co-Creation Experiences: The Next Practice in Value Creation. *Journal of Interactive Marketing*, 18(3):5–14, 2004a.
- Prahalad, C. K. und Ramaswamy, V. *The Future Of Competition. Co-Creating Unique Value with Customers*. Brighton, Harvard Business School Press, 2004b.
- Ramaswamy, V. Co-Creation of Value – Towards an Expanded Paradigm of Value Creation. *Marketing Review St. Gallen*, 26(6):11–17, 2009.
- Ramaswamy, V. und Gouillart, F. Building the Co-Creative Enterprise. *Harvard Business Review*, 88(10):1–9, 2010.
- Ramaswamy, V. und Ozcan, K. What is co-creation? an interactional creation framework and its implications for value creation. *Journal of Business Research*, 84:196–205, 2018.
- Ranjan, K. R. und Read, S. Value co-creation: concept and measurement. *Journal of the Academy of Marketing Science*, 44(3):290–315, 2016.
- Rheinberg, F. *Intrinsische Motivation und Flow-Ebenen*. Berlin, Springer, 2006.
- Siepmann, R. Digital Reinvention Advisor (IBM Deutschland GmbH), 2018. Persönliches Interview durch Steinhardt, Felix am 16.05.2018.
- Sänn, A. *The Preference-Driven Lead User Method for New Product Development. A Comprehensive Way to Stimulate Innovations with Internal and External Sources*. Wiesbaden, Springer Gabler, 2017.
- Statista. Instant Messaging Nutzung in Deutschland, July 2018a. URL <https://de.statista.com/statistik/daten/studie/647656/umfrage/umfrage-zur-nutzungshaeufigkeit-von-instant-messaging-in-deutschland/>.
- Statista. Anteile der Wirtschaftssektoren am Bruttoinlandsprodukt (BIP) in den wichtigsten Industrie- und Schwellenländern im Jahr 2016, July 2018b. URL <https://de.statista.com/statistik/daten/studie/37088/umfrage/anteile-der-wirtschaftssektoren-am-bip-ausgewaehlter-laender/>.
- Storbacka, K., Brodie, R. J., Böhmman, T., Maglio, P. P., und Nenonen, S. Actor engagement as a microfoundation for value co-creation. *Journal of Business Research*, 69(8):3008–3017, 2016.
- Töllner, A., Blut, M., und Holzmüller, H. H. Customer solutions in the capital goods industry: examining the impact of the buying center. *Industrial Marketing Management*, 40(5):712–722, 2011.
- trnd. trnd - Lerne neue Produkte kennen, August 2018. URL <https://www.trnd.com/de/>.
- Vargo, S. L. und Lusch, R. F. It's all B2B... and beyond: towards a systems perspective of the market. *Industrial Marketing Management*, 40(2):181–187, 2011.
- Vargo, S. L. und Lusch, R. F. Evolving to a New Dominant Logic for Marketing. *Journal of Marketing*, 68(1):1–17, 2004.
- Vargo, S. L. und Lusch, R. F. Service-dominant logic: continuing the evolution. *Journal of the Academy of Marketing Science*, 36(1):1–10, 2008.
- Vargo, S. L. und Lusch, R. L. Institutions and axioms: an extension and update of service-dominant logic. *Journal of the Academy of Marketing Science*, 44:5–23, 2016.
- von Loewenfeld, F. *Involvement generieren, Kundenzufriedenheit schaffen. Neue Wege zur Kundenzufriedenheit*. Marburg, Tectum Verlag, 2003.
- Wasko, M. M. und Faraj, S. Why People Participate and Help Others in Electronic Communities of Practice. *Journal of Strategic Information Systems*, 9:155–173, 2000.
- Winship, C. und Mare, R. D. Regression Models With Ordinal Variables. *American Sociological Review*, 49(4):512–525, 1984.
- Witell, L., Kristensson, P., Gustafsson, A., und Löfgren, M. Idea generation: customer co-creation versus traditional market research techniques. *Journal of Service Management*, 22(2):149–159, 2011.
- Yin, R. K. *Case Study Research. Design and Methods*. Thousand Oaks, Sage Publications, 1994.



Die aktuelle Änderung der umsatzsteuerrechtlichen Behandlung wirtschaftlicher Tätigkeiten von juristischen Personen des öffentlichen Rechts

Felix Frauendorf

Universität Siegen

Abstract

Der § 2b UStG setzt für die juristischen Personen des öffentlichen Rechts die Mehrwertsteuer-Systemrichtlinie in deutsches Recht um und löste die alte systemwidrige Rechtslage des § 2 Abs. 3 UStG ab. Die juristischen Personen des öffentlichen Rechts müssen daher bei jeder Leistung prüfen, ob sich der Umfang ihrer Umsatzbesteuerung geändert hat bzw. ändern wird unter dem Gesichtspunkt der Wettbewerbsverzerrung. Die Arbeit untersucht daher, welche Änderung der § 2b UStG auf die umsatzsteuerliche Behandlung von wirtschaftlichen Tätigkeiten von juristischen Personen des öffentlichen Rechts hat und wie sich diese auswirken. Die Arbeit stellt die Grundzüge der Umsatzbesteuerung, die Umsatzbesteuerung nach alter und nach neuer Rechtslage dar. Abschließend wird der Umstellungsprozess zur neuen Rechtslage dargestellt. Als Ergebnis kann festgestellt werden, dass die unternehmerische Betätigung von juristischen Personen des öffentlichen Rechts sich vergrößern wird, da jede einzelne Leistung untersucht werden muss. Hierdurch kann es zu Fehleinschätzungen bei der Leistungszuordnung kommen, was zu einer falschen Umsatzsteuerdeklaration führt. Auf der anderen Seite kann das Volumen des Vorsteuerabzuges sich erhöhen. Es bleibt abzuwarten, ob der § 2b UStG europarechtskonform ist.

Keywords: Umsatzbesteuerung; juristische Person des öffentlichen Rechts; § 2b UStG; Mehrwertsteuer-Systemrichtlinie; Umstellungskonzept.

1. Einleitung

1.1. Problemstellung und Zielsetzung

Die juristischen Personen des öffentlichen Rechts (jPÖR) werden innerhalb des deutschen Steuerrechts bereits seit der Zeit der Weimarer Republik besonders behandelt.¹ Die Normen der Steuergesetze für jPÖR bilden ein Spezialgebiet und es zeigt sich insbesondere innerhalb des Umsatzsteuerrechts ein Problembereich durch den unionswidrigen § 2 Absatz (Abs.) 3 Umsatzsteuergesetz (UStG) alte Fassung (a.F.).² JPÖR üben nicht nur hoheitliche Tätigkeiten aus, sondern sind darüber hinaus auch noch wirtschaftlich tätig und stehen damit im direkten Wettbewerb zu privatrechtlichen Unternehmen. Im Bereich der Kommunen ergibt sich die wirtschaftliche Betätigung aus dem kommunalen Selbstverwaltungsrecht des Artikel 28 Abs. 2 Grundgesetz (GG) und wird durch die Regelungen in der jeweiligen Gemeinde-

ordnung³ konkretisiert. Durch die unionsrechtliche Vorgabe der Mehrwertsteuer-Systemrichtlinie (MwStSystRL)⁴ wird geregelt, dass jPÖR sich so behandeln lassen müssen wie ein privater Dritter. Dies bedeutet, dass es durch die Nichtbesteuerung von Leistungen nicht zu einer Wettbewerbsverzerrung gegenüber der Privatwirtschaft kommen darf. Die deutsche Rechtslage hat bisher das Umsatzsteuerrecht mit dem Körperschaftsteuerrecht verknüpft, sodass es bei der Einordnung von Leistungen zu Abweichungen zu den Unionsvorgaben kam, sodass der Gesetzgeber den Zustand durch die Umsetzung des § 2b UStG an Unionsvorgaben anpasste. Diese Norm ist deshalb von Bedeutung, da spätestens ab 2021 alle jPÖR verpflichtet sind, diese umzusetzen.

In der nachfolgenden Arbeit soll, obwohl die Regelungen des § 2b UStG überwiegend auf Kommunen angepasst

³In der Gemeindeordnung NRW sind dies für wirtschaftliche Tätigkeiten der § 107 Abs. 1 Gemeindeordnung (GO) Nordrhein-Westfalen (NRW) und für nicht-wirtschaftliche Tätigkeiten der § 107 Abs. 2 GO NRW.

⁴Wenn es nachfolgend um die Richtlinie 2006/112/Europäische Gemeinschaft (EG) des Rates vom 28. November 2006 über das gemeinsame Mehrwertsteuersystem geht, wird vereinfacht von der Mehrwertsteuer-Systemrichtlinie gesprochen.

¹Vgl. Müller, 2017, S. 9; Hüttermann, 2009, S. 308.

²Vgl. Hüttermann, 2017, S. 129; Hüttermann, 2009, S. 308.

worden sind⁵, von jPÖR gesprochen werden, soweit es keine Besonderheiten gibt.⁶ Kirchen werden nicht behandelt, da dies ein Spezialgebiet innerhalb der Umsatzbesteuerung der jPÖR ist. Daneben werden Beispiele und Anwendungsfälle in der Vielzahl der Fälle auf nordrhein-westfälisches Landesrecht beschränkt. Daher ergibt sich folgende Forschungsfrage: Inwiefern ändert sich durch § 2b UStG die umsatzsteuerliche Behandlung wirtschaftlicher Tätigkeiten von juristischen Personen des öffentlichen Rechts und welche Auswirkungen hat dies? Diese Forschungsfrage soll geklärt werden, indem die aktuelle steuerliche Literatur zu dem Thema ausgewertet und kritisch reflektiert wird.

1.2. Gang der Untersuchung

In einem ersten Schritt werden grundsätzliche Punkte geklärt, die sowohl für die alte Fassung des § 2 Abs. 3 UStG von Bedeutung sind, als auch für die neue Rechtslage des § 2b UStG. Dies sind neben dem System der Umsatzbesteuerung, die umsatzsteuerliche Unternehmereigenschaft des § 2 Abs. 1 UStG und die spezifische Definition der jPÖR sowie die unionsrechtliche Vorgabe durch die MwStSystRL. Danach werden innerhalb der alten Rechtslage die Abgrenzung der unternehmerischen Betätigung durch den Begriff des Betriebes gewerblicher Art sowie der land- und forstwirtschaftlichen Betriebe mit der jeweiligen Auslegung durch die Rechtsprechung vorgenommen und der nichtunternehmerischen Tätigkeit mit dem hoheitlichen Tätigkeiten, Vermögensverwaltung sowie der Kooperation von jPÖR.

Dadurch, dass die Rechtsprechung die Unternehmereigenschaft des § 2 Abs. 3 UStG a.F. ablehnte, wurde ein neuer Unternehmerbegriff für jPÖR durch den § 2b UStG eingeführt. Dazu sollen in einem ersten Schritt die rechtliche Grundlage des § 2b UStG geklärt werden. Darauf aufbauend sollen die Voraussetzungen des § 2b UStG kritisch beleuchtet werden, die eine enge Verzahnung mit der MwStSystRL haben. Darauf aufbauend kann ein Prüfschema helfen, die Tätigkeiten der jPÖR besser einzuordnen. In einem nächsten Schritt sollen die Möglichkeiten der Umstellung auf die neue Rechtslage durch die Optionserklärung und deren Folgen bzw. des Widerrufs dargestellt werden. Außerdem sollen die grundsätzlichen Unterschiede zur alten Rechtslage dargestellt werden. Da mit der neuen Rechtslage viele Veränderungen verbunden sind, sollen die Veränderungsprozesse helfen, den § 2b UStG bei jPÖR zu implementieren.

2. Grundzüge der Umsatzbesteuerung innerhalb des Untersuchungs-objektes der juristischen Personen des öffentlichen Rechts

2.1. Mehrwertsteuer-Systemrichtlinie

Die Grundlage für das Funktionieren des Europäischen Binnenmarktes und dem damit einhergehenden Prinzip der

Vermeidung der Wettbewerbsverzerrungen wird in dem Vertrag über die Arbeitsweise der europäischen Union (AEUV) in Artikel 26 Abs. 1 geregelt. Damit es zu einer Harmonisierung innerhalb des Rechtssystems der EU kommt, werden Richtlinien und Verordnungen erlassen, die entweder direkt von den Mitgliedsländern oder in modifizierter Form umgesetzt werden müssen. Durch Artikel 113 des AEUV wird geregelt, dass die indirekten Steuern⁷ innerhalb der EU angeglichen werden sollen. Diese Vorgabe wird durch die MwStSystRL umgesetzt. Die Unterscheidung zwischen Richtlinie und Verordnung ist insofern entscheidend, als dass EU-Richtlinien erst in staatliches Recht umgewandelt werden müssen und EU-Verordnungen unmittelbar in jedem Mitgliedsstaat gelten. Die Richtlinien müssen zwar umgesetzt werden, aber wie die Richtlinie umgesetzt wird, bleibt jedem Mitgliedsstaat gem. Artikel 288 AEUV selbst überlassen. Daher muss das deutsche Umsatzsteuerrecht an die MwStSystRL angepasst werden, um es europarechtskonform zu gestalten. Artikel 13 der MwStSystRL ist nicht nur eine Steuerbefreiungsvorschrift, sondern auch das Sonderrecht für die öffentliche Hand.⁸ Dabei sind Einrichtungen des öffentlichen Rechts nur dann steuerbefreit, wenn diese wirtschaftlich und im Rahmen der öffentlichen Gewalt (hoheitlich) tätig sind. Eine Ausnahme besteht, wenn die Nichtbesteuerung zu größeren Wettbewerbsverzerrungen führt.⁹

2.2. System der Umsatzbesteuerung

Die grundsätzliche Definition von Steuern ist in § 3 Abs. 1 Abgabenordnung (AO) zu finden. Dabei ist eine Steuer eine Geldleistung, die keine Gegenleistung darstellt. Dabei erhebt das öffentlich-rechtliche Gemeinwesen Steuern zur Erzielung von Einnahmen und jeder bei dem ein gesetzlich bestimmter Tatbestand erfüllt ist, muss die Steuer zahlen.¹⁰ Dabei knüpft die Umsatzsteuer an die Aufwendungen für den Verbrauch einer Leistung des Endverbrauchers an.¹¹ Die Umsatzsteuer ist gem. § 21 AO eine Verkehrssteuer, da laut § 21 Abs. 1 UStG nur die Einfuhrumsatzsteuer eine Verbrauchssteuer ist. Dagegen ist bei wirtschaftlicher Betrachtung die Umsatzsteuer eine Verbrauchssteuer¹², da der Aufwand des

⁷Umsatzsteuer, Verbrauchsabgaben und sonstige indirekte Steuern.

⁸Vgl. Wagner et al., 2017, S. 24.

⁹Dabei betrifft das erste Tatbestandsmerkmal die Rechtsform und das zweite Tatbestandsmerkmal soll Hoheitsbefugnisse vor einer Besteuerung schützen. Die Rückausnahme soll die Wettbewerbsneutralität schützen; Vgl. Wagner et al., 2017, S. 24.

¹⁰Dabei lassen sich die Abgaben (Sammelbegriff für alle Zahlungen, die zur Erzielung von Einnahmen genutzt werden) in drei Bereiche einteilen. Bei der Steuer spielt die Nutzung der öffentlichen Einrichtung keine Rolle. Die Gebühren sind Zahlungen für Leistungen, die für die Inanspruchnahme einer Körperschaft freiwillig oder erzwungenermaßen in Anspruch genommen werden, z.B. Verwaltungsgebühren (z.B. Beurkundungen, Passgebühren) oder Nutzungsgebühren (z.B. Rundfunk, Friedhof). Daneben gibt es noch Beiträge, die einen finanziellen Ersatz für die mögliche Inanspruchnahme einer Leistung der öffentlichen Einrichtung darstellen, wie z.B. Sozialversicherungsbeiträge; Vgl. Krömker, 2019, in: Lippross und Seibel, 2019, § 3 AO, Nr. 9-11.

¹¹Vgl. Stadie, in: Rau et al., 2018, Einführung Rn. 143.

¹²Der Begriff Aufwandsteuer wie in § 105 Abs. 2a GG kodifiziert passt besser, da die Gegenstände nicht unbedingt verbraucht werden (z.B. Kunst-

⁵Vgl. Wagner et al., 2017, S. 15; Widmann, 2016, S. 14.

⁶In der nachfolgenden Arbeit wird die maskuline Form verwendet, wobei dadurch auch Frauen berücksichtigt sind.

Verbrauchers für die vom Unternehmen erzeugte Dienstleistung oder Gegenstände, sowie für sonstige Leistungen erfasst wird.¹³ Dabei ist die Umsatzsteuer eine der wichtigsten Steuern und damit Einnahmequellen für den Staat, da 2017 die Umsatzsteuer mit einem Ergebnis von 170,5 Milliarden Euro damit 23,21% an den insgesamt 734 Milliarden Euro Steuereinnahmen von Bund und Ländern ausmacht.¹⁴ Dabei ist der Zweck der Besteuerung von jPöR nicht die Einnahmengenerierung, sondern die Gleichbehandlung mit privaten Wirtschaftsteilnehmern, da ansonsten die jPöR die Leistungen auf dem Markt wesentlich günstiger anbieten könnten als die privaten Wirtschaftsteilnehmer.¹⁵ Steuerbare Umsätze unterliegen der Umsatzsteuer gem. § 1 UStG. Damit ein Umsatz steuerbar ist, müssen die Haupttatbestände des § 1 Abs. 1 Nr. 1 UStG greifen. Dabei sind gem. § 1 Abs. 1 UStG Lieferungen und sonstige Leistungen, die ein Unternehmer gem. § 2 UStG im Inland gegen Entgelt im Rahmen seines Unternehmens ausführt, steuerbar. Daneben sind die Einfuhr von Gegenständen aus dem Drittland gem. § 1 Abs. 1 Nr. 4 UStG und der innergemeinschaftliche Erwerb gem. § 1 Abs. 1 Nr. 5 UStG steuerbare Vorgänge. Das grundsätzliche Prüfungsschema um die Umsatzsteuerschuld zu beurteilen, befindet sich im Anhang 1.

Die Unternehmen haben sowohl Rechte als auch Pflichten aus dem Umsatzsteuergesetz. Dabei muss der Unternehmer gem. § 22 UStG Aufzeichnungen führen und gem. § 18 Abs. 1-3 UStG eine Umsatzsteuererklärung abgeben. Dabei umfassen die Aufzeichnungen die Umsätze des Unternehmens, die abziehbaren Vorsteuerbeträge sowie weitere Aufzeichnungen über die Steuerberechnung.¹⁶ Dadurch, dass die Umsatzsteuer eine indirekte Steuer ist (Artikel 113 AEUV), belastet diese den Abnehmer der Leistung (Steuerträger), was im Regelfall der Endverbraucher ist.¹⁷ Der Umsatzsteuerbetrag ist grundsätzlich von dem Unternehmer als Steuerschuldner gem. § 13a Abs. 1 S. 1 UStG zu berechnen und danach auf elektronischem Wege der Finanzbehörde mitzuteilen. Ein Recht der Unternehmen ist bedingt durch das Wesen der Umsatzsteuer als Allphasen-Netto-Umsatzsteuer. Die Besteuerung der Umsätze findet auf jeder Wirtschaftsstufe statt, gekoppelt mit dem Recht der Unternehmen auf Vorsteuerabzug.¹⁸ Durch dieses System des Vorsteuerabzuges gem. § 15 UStG wird die Umsatzbesteuerung für die Unternehmen wettbewerbsneutral, da die Vorsteuer kein Kostenfaktor ist, sondern ein durchlaufender Posten.¹⁹ Dies funktioniert so, indem der leistende Unternehmer an einen anderen Unternehmer eine Leistung erbringt und die vom empfangenden

Unternehmer getragene Umsatzsteuer in Form der Vorsteuer gegenüber dem Finanzamt geltend gemacht wird. Damit belastet nur der letzte Umsatz den Endverbraucher, unabhängig davon, wie viele Vorumsätze bereits getätigt worden sind.²⁰ Dies ist an die Erteilung einer entsprechenden Rechnung gem. § 14 Abs. 2 UStG geknüpft.

2.3. Umsatzsteuerliche Unternehmereigenschaft gemäß § 2 Absatz 1 Umsatzsteuergesetz

Die Grundlage für die Unternehmereigenschaft findet sich in Artikel 9 der MwStSystRL und dieses ist im deutschen Umsatzsteuerrecht in § 2 UStG kodifiziert. Die Relevanz von § 2 Abs. 1 UStG ergibt sich daraus, dass die allgemeinen Unternehmereigenschaften erst einmal vorliegen müssen, bevor die Regelung des § 2 Abs. 3 UStG oder § 2b UStG greift. Dabei ist ein Unternehmer gem. § 2 Abs. 1 S. 1 UStG wer gewerblich oder beruflich²¹ auf nachhaltige Weise eine Tätigkeit zur Erzielung von Einnahmen selbständig ausübt.²²

Dabei ist ein Unternehmer gem. § 2 Abs. 1 S. 1 UStG wer gewerblich oder beruflich²³ auf nachhaltige Weise eine Tätigkeit zur Erzielung von Einnahmen selbständig ausübt.²⁴ Die Definition ergibt sich aus Artikel 9 der MwStSystRL, wonach bei der Betrachtung der selbständig ausgeführten wirtschaftlichen Tätigkeit weder auf den Ort, Zweck oder das Ergebnis abgestellt wird. Die Unternehmerfähigkeit liegt unabhängig von der Rechtsform oder der Rechtsfähigkeit vor.²⁵ Demnach müssen die drei Kriterien Selbständigkeit, Einnahmerezzielungsabsicht und Nachhaltigkeit erfüllt sein, damit die Unternehmereigenschaft gem. § 2 Abs. 1 UStG vorliegt. Das Kriterium der Selbständigkeit liegt vor, wenn diese eigenverantwortlich ausgeübt und die Selbständigkeit sowohl im eigenen Verantwortungsbereich als auch auf eigene Rechnung erfolgt. Dabei lassen sich einige Kriterien ableiten, bei denen die Selbständigkeit vorliegt.²⁶ Das ist zum einen das Unternehmerrisiko, dass jegliche Tätigkeiten Einfluss auf den Fortbestand des Unternehmens haben. Daneben muss das Unternehmen nicht lediglich für einen Auftraggeber tätig sein, sondern für mehrere und dies erfolgt unter Zuhilfenahme von Arbeitskräften, die in eigens dafür vorgesehenen Büros untergebracht sind.²⁷

Das zweite Kriterium der Einnahmerezzielungsabsicht gem. § 2 Abs. 1 S. 3 UStG ist erfüllt, solange der Zweck der

werke) und der Belastungsgrund der Umsatzsteuer die Verwendung von Einkommen oder Vermögen ist; Vgl. Stadie, in: Rau et al., 2018, Einführung, Rn. 143.

¹³Vgl. Stadie, in: Rau et al., 2018, Einführung Rn. 141.

¹⁴Vgl. BMF, 2017 Bundesministerium für Finanzen (BMF), Steuereinnahmen nach Steuerarten 2010-2017, Jahr 2017.

¹⁵Vgl. Baldauf, 2018a, S. 56.

¹⁶Vgl. Janzen, in: Seibold-Freund, 2015, 2015, § 22 UStG Rz. 2.

¹⁷Vgl. Stadie, in: Rau et al., 2018, Einführung Rn. 149 und 260.

¹⁸Vgl. Stadie, in: Rau et al., 2018, Einführung Rn. 250.

¹⁹Vgl. Stadie, in: Rau et al., 2018, Einführung Rn. 254.

²⁰Vgl. Stadie, in: Rau et al., 2018, Einführung Rn. 256.

²¹BFH, 05.09.1963, V 117/60 U, Sammlung der Entscheidungen des BFH (BFHE) 77, S. 550-558; BMF, 02.01.19, Bundessteuerblatt (BStBl.) I S. 17, Abschnitt 2.3. Abs. 1 S. 1; Der Begriff der gewerblichen oder beruflichen Tätigkeit ist zu unterscheiden vom dem Gewerbebetrieb nach dem Einkommensteuergesetz, da der Umsatzsteuerrechtliche Begriff weiter gefasst ist.

²²Vgl. Janzen, in: Lippross und Seibel, 2019, § 2 UStG Rz. 26.

²³BFH, 05.09.1963, V 117/60 U, Sammlung der Entscheidungen des BFH (BFHE) 77, S. 550-558; BMF, 02.01.19, Bundessteuerblatt (BStBl.) I S. 17, Abschnitt 2.3. Abs. 1 S. 1; Der Begriff der gewerblichen oder beruflichen Tätigkeit ist zu unterscheiden vom dem Gewerbebetrieb nach dem Einkommensteuergesetz, da der Umsatzsteuerrechtliche Begriff weiter gefasst ist.

²⁴Vgl. Janzen, in: Lippross und Seibel, 2019, § 2 UStG Rz. 26.

²⁵Vgl. BMF, 02.01.19, BStBl. I S. 17, Abschnitt 2.1 Abs. 1 S. 1 und 3.

²⁶Vgl. Janzen, in: Lippross und Seibel, 2019, § 2 UStG Rz. 68.

²⁷Vgl. Janzen, in: Lippross und Seibel, 2019, § 2 UStG Rz. 69.

Tätigkeit darin besteht, ein Entgelt zu erzielen.²⁸ Dazu ist es erforderlich, dass dies im Rahmen eines Leistungsaustausches passiert.²⁹ Dagegen ist es nicht erforderlich, dass die Einnahmerezzielungsabsicht der primäre Zweck der nachhaltigen Tätigkeit ist. Unter dem Begriff Einnahmen können Geld, Sachen oder sonstige Leistungen subsumiert werden.³⁰ Gerade bei Leistungen, die durch die öffentliche Hand durchgeführt werden, kann eine nicht vollständige Kostendeckung erfolgen, wenn die Tätigkeit gegen (teilweise) Kostenerstattung erfolgt.³¹

Diese dauerdefizitären Betätigungen³² können trotzdem eine Leistung gegen Entgelt darstellen, wenn ein Rechtsverhältnis zwischen Leistungserbringer und Leistungsnehmer vorliegt und es zu einem Leistungsaustausch kommt, wobei die Bemessungsgrundlage der Leistung die dafür empfangene Vergütung ist.³³ Wenn es sich um eine entgeltliche Tätigkeit handelt, dann ist in einem weiteren Schritt zu prüfen ob diese entgeltliche Leistung auch eine wirtschaftliche bzw. unternehmerische Tätigkeit ist.³⁴ Das dritte Kriterium der Nachhaltigkeit liegt vor, wenn die Tätigkeit auf Dauer angelegt ist.³⁵ Einzelne Merkmale sind nicht entscheidend, sondern das Gesamtbild der Verhältnisse.³⁶ Diese Merkmale sind ein planmäßiges Handeln, eine Wiederholungsabsicht in der vollziehenden Tätigkeit, Quantität der ausgeführten Umsätze, Zeitraum der Tätigkeit, Vornahme mehrerer gleichartiger Handlungen, Intensität der Betätigung (wirtschaftliche Größenordnung), Marktbeteiligung, Auftreten wie ein Händler gegenüber dem Auftraggeber oder gegenüber Behörden. Folgendes Beispiel soll das Kriterium verdeutlichen (vgl. Tabelle 1).

2.4. Juristische Personen des öffentlichen Rechts

JPöR sind nach ihrer Definition Einrichtungen, die aufgrund des öffentlichen Rechts eine Rechtspersönlichkeit besitzen³⁷ und im Zweifel dadurch definiert sind, dass diese durch eine Verfassung, durch Gesetz oder durch einen Verwaltungsakt anerkannt werden.³⁸ Grundsätzlich lassen sich JPöR gem. § 89 Bürgerliches Gesetzbuch (BGB) in Körperschaften des öffentlichen Rechts, Anstalten des öffentlichen Rechts und Stiftungen des öffentlichen Rechts einteilen. Durch das BMF-Schreiben findet eine nicht abschließende Aufzählung von JPöR statt, worunter die Gebietskörperschaften (Bund, Länder, Kommunen, Gemeindeverbände,

Zweckverbände), die Religionsgemeinschaften, die Innungen, Handwerkerkammern, Industrie- und Handelskammer, die staatlichen Hochschulen und sonstige Gebilde³⁹ fallen, die die Definition erfüllen.⁴⁰ Daneben gelten auch ausländische juristische Personen als JPöR soweit dies nach deutschem Standard so beurteilt wird.⁴¹ Die bisherige Rechtsprechung und Verwaltungsauffassung geht davon aus, dass eine JPöR insgesamt nur ein Unternehmen betreibt.⁴² Ein EuGH Urteil zweifelte diese Auffassung an, welche haushaltsgebundene Einheiten einer Gemeinde als nicht Steuerpflichtigen bzw. in dem deutschen Recht als Nicht-Unternehmer ansah.⁴³ Dies könnte im Umkehrschluss bedeuten, dass haushaltsungebundene Einrichtungen als Unternehmen anzusehen sind⁴⁴, wobei eine höchstrichterliche Rechtsprechung noch aussteht.⁴⁵ Für unter den Bundesländern stehende Gliederungsebenen wie z.B. Kommunen wird auch in Zukunft nur ein einheitliches Unternehmen existieren, sodass damit nur ein Umsatzsteuerbescheid einhergeht.⁴⁶ Die Bundesländer und der Bund werden aus Vereinfachungsgründen jede ihrer JPöR einzeln besteuern.⁴⁷

Die Bezeichnung als JPöR ist eine deutsche Bezeichnung und wird von Literaturmeinungen kritisiert⁴⁸, da nicht wie in Artikel 13 Abs. 1 MwStSystRl von Staaten, Ländern, Gemeinden und sonstigen Einrichtungen des öffentlichen Rechts gesprochen werden. Die deutsche Definition ist aus dem bisherigen Recht überführt worden⁴⁹ und es wäre besser gewesen, wenn auch die Definition an die umfangreichere Definition der Europäischen Union (EU) Vorgaben angepasst worden wäre. JPöR haben das Recht sich neben der Erfüllung ihrer hoheitlichen Aufgaben auch wirtschaftlich zu betätigen, sofern diese gewisse Voraussetzungen erfüllen.⁵⁰ Dabei hat die JPöR die Möglichkeit wirtschaftliche Tätigkeiten auf öffentlich-rechtlicher Grundlage oder auf privatrechtlicher Grundlage zu erbringen. Innerhalb der öffentlich-rechtlichen Betätigung ist dies mit und ohne eigene Rechtspersönlichkeit möglich.⁵¹ Wenn JPöR sich einer privatwirtschaftlich

²⁸BFH, Urteil vom 18.3.1988, V R 178/83, BFHE 153, S. 166-170.

²⁹Vgl. BMF, 02.01.19, BStBl. I S. 17, Abschnitt 2.3 Abs. 8 S. 3.

³⁰Vgl. Sterzinger, in: Küffner und Rust, 2016b, § 2 UStG Rn. 191.

³¹Vgl. Sterzinger, in: Küffner und Rust, 2016b, § 2 UStG Rn. 194.

³²Die umsatzsteuerlichen Folgen wurden vor dem Europäischen Gerichtshof (EuGH) behandelt EuGH, Urteil vom 12.5.2016, Rechtssache (Rs.) C-520/14, Gemeinde Borsele, UR 2016, S. 520-523; EuGH, Urteil vom 02.06.2016, Rs. C-263/15, Lajvér, UR 2016, S. 525-530.

³³Vgl. Sterzinger, in: Küffner und Rust, 2016b, § 2 UStG Rn. 195.

³⁴Vgl. Sterzinger, in: Küffner et al., 2018, § 2 UStG Rn. 196.

³⁵Vgl. BMF, 02.01.19, BStBl. I S. 17, Abschnitt 2.3 Abs. 5.

³⁶Vgl. Janzen, in: Lippross und Seibel, 2019, § 2 UStG, Rz. 61.

³⁷Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016; Tz. 3; Siehe Anhang 2.

³⁸Vgl. Arnold, in: Säcker et al., 2015, § 89 BGB, S. 1035.

³⁹Vgl. Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rn. 48.

⁴⁰Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016; Tz. 3.

⁴¹Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016; Tz. 3.

⁴²Vgl. BMF, 02.01.19, BStBl. I S. 17, Abschnitt 2.11 Abs. 2 S. 1; BFH, 18.08.1988, Urteil vom 18.8.1988, V R 194/83, Sammlung amtlich nicht veröffentlichter Entscheidungen des Bundesfinanzhofes (BFH/NV) 1989, S. 415-416.

⁴³EuGH, 29.09.2015, Rs. C-276/14, Gmina Wrocław, UR 2015, S. 829-834.

⁴⁴Vgl. Klenk, 2016, S. 182; Ismer/Baur-Rückert, MwStR 2016, S. 741.

⁴⁵Vgl. Weber, 2017, S. 77.

⁴⁶Vgl. Kirch, in: Fritsch et al., 2015, § 2 UStG Rn. 217; Kirch, in: Fritsch et al., UStG, 2017, § 2b UStG Rn. 28.

⁴⁷Vgl. Finanzministerium (FinMin) Bayern vom 4.4.1990, 36 - S 7106 - 32/25 - 22, UR 1991, S. 150.

⁴⁸Vgl. Heidner, 2018, S. 740.

⁴⁹Vgl. ? BMF, 02.01.19, BStBl. I S. 17, Abschnitt 2.11 Abs. 1. Deshalb ist die Vorschrift nur auf JPöR anwendbar, die auch unter die alte Rechtslage des § 2 Abs. 3 UStG a.F. fallen.

⁵⁰Dabei muss das Handeln des Staates am Gemeinwohl orientiert sein, wobei dadurch auch Einnahmen erzielt werden dürfen, soweit die nichterwerblich-wirtschaftliche Betätigung im Vordergrund steht; Vgl. Auerswald, 2016, S. 33-34.

⁵¹Mit eigener Rechtspersönlichkeit kann eine Landesbank oder eine Spar-

Tabelle 1: Blockheizkraftwerk

Quelle: BFH, 18.12.2008, V R 80/07, BStBl. II 2011, S. 292-296.

Betrieb eines Blockheizkraftwerks durch eine Kommune zur Erzeugung von Wärme und Strom, welcher ganz oder teilweise in regelmäßigen Abständen in das Stromnetz eingespeist wird.

Diese nachhaltige Betätigung dient der Einnahmeerzielung und begründet daher die Unternehmereigenschaft der Kommune, unabhängig davon, ob diese sonst unternehmerisch tätig geworden ist.

organisierten Gesellschaft (z.B. Gesellschaft mit beschränkter Haftung (GmbH), Aktiengesellschaft (AG)) bedienen, fällt diese Gesellschaft unabhängig von dem Tätigkeitfeld nicht in den Anwendungsbereich des § 2 Abs. 3 UStG a.F. oder § 2b UStG und gilt damit unabhängig davon, wer an dieser privatwirtschaftlich tätigen Gesellschaft beteiligt ist, als Unternehmen gem. § 2 Abs. 1 UStG.⁵² Das Abweichen der deutschen Definition von der EU Definition hat insofern Bedeutung, als dass der Anwendungsbereich des § 2b UStG auf jPÖR beschränkt ist, aber Artikel 13 MwStSystRl von Einrichtungen spricht und damit nach ständiger Rechtsprechung des EuGH und des BFH darunter auch Personengesellschaften und natürliche Personen fallen.⁵³ Damit die juristischen Personen des Privatrechts (Personengesellschaft, Kapitalgesellschaft⁵⁴) auch als Einrichtung des öffentlichen Rechts gelten, müssen diese neben dem Tätigwerden im Rahmen der öffentlichen Gewalt auch noch in die Organisation der öffentlichen Verwaltung eingegliedert sein.⁵⁵ Es bietet sich an, wenn juristische Personen des Privatrechts von jPÖR beherrscht und kontrolliert werden, diese auch zu den Einrichtungen des öffentlichen Rechts zu subsumieren⁵⁶, wobei diese Mischformen in dem deutschen Umsatzsteuerrecht eher keine Anwendung finden.⁵⁷

3. Umsatzbesteuerung der juristischen Personen des öffentlichen Rechts nach alter Rechtslage

3.1. Unternehmerische Tätigkeit von juristischen Personen des öffentlichen Rechts

3.1.1. Überblick über die unternehmerische Tätigkeit

JPÖR⁵⁸ sind als juristische Personen unternehmerfähig, sodass in einem ersten Schritt die Voraussetzungen des § 2 Abs. 1 UStG geprüft werden müssen.⁵⁹ Der § 2 Abs. 3 UStG a.F. ist in einem zweiten Schritt zu prüfen, da dieser

die Unternehmereigenschaft einer jPÖR einschränkt.⁶⁰ Daraus ergibt sich, dass jPÖR ihre wirtschaftliche Tätigkeiten nur im Rahmen ihrer Betriebe gewerblicher Art (BgA) gem. § 4 Körperschaftsteuergesetz (KStG), ihrer land- oder forstwirtschaftlichen Betriebe oder im Rahmen der in § 2 Abs. 3 S. 2 Nr. 2 bis 5 UStG a.F. aufgezählten Tätigkeiten ausüben.⁶¹ In Abgrenzung dazu dienen die hoheitlichen Tätigkeiten sowie die Vermögensverwaltung, die nicht umsatzsteuerbar sind.⁶²

Das Umsatzsteuerrecht greift über den Verweis in § 2 Abs. 3 S. 1 UStG a.F. auf den § 4 KStG zurück, sodass es nicht zu einer Doppelaufzählung kommt.⁶³ Dabei wird der Begriff BgA von der Verwaltung im Umsatzsteuerrecht und im Körperschaftsteuerrecht einheitlich ausgelegt.⁶⁴ Die Rechtsprechung hat dem schon früh widersprochen und legt den BgA Begriff richtlinienkonform aus in Übereinstimmung mit den Regelungen in Artikel 9 und 13 MwStSystRl⁶⁵, da die nationale Rechtsprechung verpflichtet ist, nationale Vorschriften im Sinne des Unionsrechts auszulegen.⁶⁶ Damit gab es durch die alte Rechtslage eine Beschränkung auf Anwendungsfälle, die durch Artikel 13 der MwStSystRl weiter gefasst sind.⁶⁷ Dadurch war eine europarechtskonforme Anwendung nicht mehr gewährleistet.⁶⁸ Dabei liegt ein BgA gem. § 4 Abs. 1 KStG vor, wenn es sich um eine Einrichtung, die einer nachhaltigen wirtschaftlichen Tätigkeit, zur Einnahmeerzielung, außerhalb der Land- und Forstwirtschaft dient und sich dabei innerhalb der Gesamtbetätigung der juristischen Person heraushebt. Dem BgA kommt im Umsatzsteuerrecht lediglich die Sammelbezeichnung für steuerbare Tätigkeit zu.⁶⁹

3.1.2. Definition Betriebe gewerblicher Art

Einrichtung

⁶⁰BFH, 28.10.2004, V R 19/04, BFH/NV 2005, S. 725-727; Hieraus ergibt sich, dass § 2 Abs. 3 UStG nicht *lex specialis* gegenüber der allgemeinen Unternehmereigenschaft des § 2 Abs. 1 UStG ist.

⁶¹Vgl. Korn, in: Bunjes et al., 2018, § 2 UStG Rn.191.

⁶²Vgl. Cloß, 2013, S. 257-258.

⁶³Vgl. Kirch, in: Fritsch et al., 2015, § 2 UStG Rn. 191; Stadie, in: Rau et al., 2018, § 2 UStG Rn. 1205.

⁶⁴Vgl. BMF, 02.01.19, BStBl. I S. 17, Abschnitt 2.11 Abs. 4.

⁶⁵BFH, 15.04.2010, V R 10/09, BStBl. II 2017, S. 863-869; BFH, 15.04.2010, XI R 17/08, BStBl. II 2017, S. 828-831.

⁶⁶Vgl. Baldauf, 2018b, S. 568.

⁶⁷Vgl. Baldauf, 2018b, S. 569; Belcke und Westermann, 2016, S. 87.

⁶⁸Vgl. Dziadkowski, 2017, S. 417.

⁶⁹BFH, 15.04.2010, V R 10/09, BStBl. II 2017, S. 863-869, Rn. 27.

kasse sein und ohne eigene Rechtspersönlichkeit ein Eigenbetrieb oder ein Regiebetrieb, die wiederum eigene jPÖR sind; Vgl. Baumgart, 2018, S. 6.

⁵²Vgl. Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rn. 72.

⁵³BFH, 08.07.1971, V R 1/68, BStBl. II 1972, S. 70; BFH, 13.03.1974, I R 7/71, BFHE 112, S. 61-70.

⁵⁴EuGH, 29.10.2015, Rs. C-174/14, S. 901-909.

⁵⁵Vgl. Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rn. 73.

⁵⁶Vgl. Ismer und Baur-Rückert, 2016, S. 745-746.

⁵⁷Vgl. Weber, 2017, S. 77.

⁵⁸Die Definition findet sich im Kapitel B III. sowohl für die alte als auch für die neue Rechtslage.

⁵⁹Vgl. Kirch, in: Fritsch et al., 2015, § 2 UStG Rn. 174.

Die Einrichtung ist weder in § 4 Abs. 1 KStG noch in Artikel 13 Abs. 1 MwStSystRl definiert. Die Einrichtung kann sich laut Finanzverwaltung entweder aus organisatorischen (z.B. gesonderte personelle oder sachliche Ausstattung, geschlossener Geschäftskreis, Buchführung oder ein ähnliches Merkmal) oder wirtschaftlichen Kriterien ergeben und sollte sich von dem nichtunternehmerischen hoheitlichen Bereich abgrenzen.⁷⁰ Das Merkmal der Einrichtung wird als überflüssig angesehen, da ein wirtschaftliches Handeln schon bei einer nachhaltigen und auf Einnahmeerzielung ausgerichteten Tätigkeit vorliegt.⁷¹

Nachhaltige Betätigung und Einnahmeerzielungsabsicht

Sowohl das zweite Kriterium der nachhaltigen Betätigung als auch das dritte Kriterium der Einnahmeerzielungsabsicht entsprechen der allgemeinen Definition der unternehmerischen Tätigkeit in § 2 Abs. 1 UStG. Zwischen der Rechtsprechung und der Ansicht der Finanzverwaltung sind keine gegenteiligen Auffassungen erkennbar.⁷²

Wirtschaftliche Bedeutsamkeit der unternehmerischen Betätigung

Ein weiteres Kriterium der wirtschaftlichen Bedeutsamkeit wird von der Finanzverwaltung anhand von gewissen Umsatzgrenzen bemessen. Dabei wird bei einem Jahresumsatz von 130.000 € immer und bei einem Jahresumsatz von mehr als 35.000 € regelmäßig von einer wirtschaftlich bedeutsamen Tätigkeit ausgegangen.⁷³ Diese Auffassung wird von der Rechtsprechung aus mehreren Gründen abgelehnt, da eine Umsatzschwelle nicht als Anhaltspunkt dienen kann, um die Richtlinienvorgaben umzusetzen.⁷⁴ Ein weiterer Punkt ist die Ablehnung des Vergleiches mit der Kleinunternehmerregelung, die private Unternehmen schützt, da sich diese Regelung auf das gesamte Unternehmen und nicht lediglich auf einen Teilbereich stützt.⁷⁵ Daneben werden die Umsatzschwellen als willkürlich festgelegt angesehen und stehen auch im Widerspruch zu dem Wettbewerbskriterium des Unionsrechts.⁷⁶ Dabei wird bereits von einem wettbewerbsrelevanten Verhalten ausgegangen, wenn eine jPöR auf privatrechtlicher Grundlage tätig wird, ohne auf die Besonderheiten des BgA Rücksicht zu nehmen.⁷⁷ Im Ergebnis

⁷⁰Vgl. Korn und Strahl, 2016, BStBl. I Sondernummer 1/2016, H 4.1 Einrichtung.

⁷¹BFH, 17.03.2010, XI R 17/08, BStBl. II 2017, S. 828-831.

⁷²Vgl. Kirch, in: Fritsch et al., 2015, § 2 UStG Rn. 194-195.

⁷³Vgl. BMF, 02.01.19, BStBl. I S. 17, Abschnitt 2.11 Abs. 4 S. 3; Korn und Strahl, 2016, BStBl. I Sondernummer 1/2016, R 4.1 Abs. 5.

⁷⁴Vgl. Schmitz und Möser, 2014, S. 609-610; BFH, Urteil vom 17.3.2010, XI R 17/08, BStBl. II 2017, S. 828-831.

⁷⁵BFH, BFH, 17.03.2010, XI R 17/08, BStBl. II 2017, S. 828-831, Rn. 32-33; Hüttermann, 2009, S. 313; Dieser führt an, dass bei Anwendung der Umsatzgrenze auf jeden BgA dies zu einer Ausweitung der Unternehmereigenschaft der öffentlichen Hand führt. Die Rechtsprechung hat nur Auswirkung auf den umsatzsteuerlichen BgA Begriff und nicht auf den körperschaftsteuerlichen; Vgl. OFD Nordrhein-Westfalen, 2014, S. 26.

⁷⁶Vgl. Hüllmann, 2009, S. 52.

⁷⁷BFH, 03.03.2011, V R 23/10, BStBl. II 2012, S. 74-77, Rn. 20-21.

ist die Umsatzregelung mit den unionsrechtlichen Vorgaben der gleichmäßigen Besteuerung und dem Wettbewerbsverzerrungsgedanken unvereinbar und deshalb bereits bei einer unter den Umsatzschwellen erzielten Tätigkeit steuerbar.⁷⁸

Versorgungsbetriebe und Verpachtungsbetriebe

BgA sind gem. § 4 Abs. 3 KStG auch Versorgungsbetriebe, welche die Bevölkerung mit Wasser, Gas, Elektrizität oder Wärme, dem öffentlichen Verkehr oder Hafendienst versorgen. Dabei sind Versorgungsbetriebe immer BgA, unabhängig, ob die Voraussetzungen des § 4 Abs. 5 KStG vorliegen.⁷⁹ Wenn hingegen die Versorgung in der Rechtsform eines privaten Unternehmens durchgeführt wird, dann greift nicht § 2 Abs. 3 UStG a.F., sondern die allgemeine Unternehmereigenschaft des § 2 Abs. 1 UStG.⁸⁰ Die Verpachtung eines BgA gilt nach § 4 Abs. 4 KStG als BgA. Durch die Verwaltung wurde daraus geschlossen, dass die Vermietung einzelner Immobilien zu einer Vermögensverwaltung führt und dies erst dann zu einem BgA führt, wenn dies über das bei einer langfristigen Vermietung übliche Maß hinausgeht.⁸¹ Der Auffassung der Finanzverwaltung wird widersprochen, da die Frage der Vermögensverwaltung nicht relevant sei und somit auch bei dem Verpachtungsbetrieb die umsatzsteuerliche Unternehmereigenschaft begründet wird.⁸²

3.1.3. Weitere unternehmerische Anwendungsfälle neben den Betrieben gewerblicher Art

Land- und forstwirtschaftliche Betätigung

Zwar handelt es sich bei einem land- und forstwirtschaftlichen Betrieb nicht um einen BgA, aber dieser ist dennoch steuerbar gem. § 2 Abs. 3 S. 1 UStG a.F. Die gesetzlichen Grundlagen befinden sich in § 24 Abs. 2 UStG und sind die Landwirtschaft, die Forstwirtschaft, der Wein-, Garten-, Obst-, und Gemüsebau, sowie Tierzucht- und Tierhaltungsbetriebe gem. § 24 Abs. 2 Nr. 2 UStG. Daneben ist nicht nur der Hauptbetrieb steuerbar, sondern auch der dem Hauptbetrieb dienliche Nebenbetrieb gem. § 24 Abs. 2 S. 2 UStG.⁸³ Zwar sollen laut Verwaltung die gleichen Merkmale wie im Einkommensteuer- und Gewerbesteuerrecht für den land- und forstwirtschaftlichen Betrieb gelten⁸⁴, aber diese Auffassung dürfte durch die BFH-Rechtsprechung keine Gültigkeit mehr haben.⁸⁵

⁷⁸Vgl. Stadie, in: Rau et al., 2018, § 2b UStG Rn. 1262, a.A. Korn, in: Bunjes et al., 2018, § 2 UStG Rn. 148.

⁷⁹BFH, 28.01.1988, V R 112/86, BFHE 152, S. 360-370.

⁸⁰Vgl. Kirch, in: Fritsch et al., 2015, § 2 UStG Rn. 199.

⁸¹Vgl. BayLfSt, 13.02.2008, S 7100 - 60 St 34 M, UR 2008, S. 397.

⁸²EuGH, 04.06.2009, Rs. C-102/08, SALIX Grundstücks-Vermietungsgesellschaft, Sammlung der Rechtsprechung des Gerichtshofes und des Gerichts Erster Instanz (Slg) 2009, S. I-4629.

⁸³Vgl. Kirch, in: Fritsch et al., 2015, § 2 UStG Rn. 204.

⁸⁴Vgl. BMF, 02.01.19, BStBl. I S. 17, Abschnitt 2.11 Abs. 5.

⁸⁵Vgl. Kirch, in: Fritsch et al., 2015, § 2 UStG Rn. 204.

Unternehmereigenschaft gemäß § 2 Abs. 3 S. 2 Umsatzsteuergesetzes alte Fassung

Neben den Voraussetzungen, um einen BgA zu begründen, gibt es in § 2 Abs. 3 S. 2 UStG a.F. aufgezählte Bestandteile, unter denen eine unternehmerische Betätigung immer gegeben ist. Die Tätigkeit in Ziffer 2, Tätigkeit der Notare im Landesdienst und Ratsschreiber im Land Baden-Württemberg, ist für den Bereich der Notare nicht mehr relevant, da diese bereits seit Anfang 2006 als Unternehmer gelten und diese historisch gewachsene Tätigkeit in Baden-Württemberg seit Anfang 2018 ausgelaufen ist.⁸⁶ Daneben hat Ziffer 3, die Selbstabgabestelle der Krankenkasse, keine Bedeutung mehr, da bereits 1981 ein Urteil gesprochen worden ist, dass Krankenkassen ab Mitte 1983 keine Brillen mehr in Selbstabgabestellen abzugeben haben.⁸⁷ Die Ziffer 5 für die Tätigkeit der Bundesanstalt für Landwirtschaft und Ernährung hat keinen eigenständigen Anwendungsbezug mehr, da diese Tätigkeit bereits in Artikel 13 Abs. 1 Unterabschnitt 3 MwStSystRl i.V.m. Anhang 1 Nr. 7 MwStSystRl auftaucht, worunter die Tätigkeiten aus den landwirtschaftlichen Interventionsstellen aus landwirtschaftlichen Erzeugnisse fallen. Lediglich Ziffer 4 für die Vermessungs- und Katasterbehörden hat noch einen praktischen Nutzen.⁸⁸

3.2. Nichtunternehmerische Tätigkeit von juristischen Personen des öffentlichen Rechts

3.2.1. Hoheitliche Tätigkeiten

Ein Hoheitsbetrieb ist kein BgA und liegt gem. § 4 Abs. 5 S. 1 KStG vor, wenn dieser überwiegend der Ausübung öffentlicher Gewalt dient. Innerhalb des Hoheitsbetriebes wird durch die jPöR keine gewerbliche oder berufliche Tätigkeit ausgeübt, sodass die hoheitlichen Tätigkeiten nicht steuerbar sind.⁸⁹ Es kommt bei dem Begriff überwiegend auf die Zweckbestimmung der Tätigkeit der jPöR an, wenn sich eine Zuordnung zum hoheitlichen oder unternehmerischen Bereich nicht eindeutig treffen lässt.⁹⁰ Der Begriff Ausübung der öffentlichen Gewalt ist gesetzlich nicht definiert, aber durch die BFH-Rechtsprechung und die Verwaltungsmeinung wird dieser Begriff ausgefüllt. Darunter sind Tätigkeiten zu verstehen, die jPöR eigentümlich und vorbehalten⁹¹ sind, sowie

aus der Staatsgewalt abgeleitet und staatlichen Zwecken dienen.⁹² Ein Indiz kann eine Leistung sein, zu deren Annahme der Leistungsempfänger verpflichtet ist.⁹³ Eine spezielle Form des Tätigwerdens im Hoheitsbereich sind Selbstversorgungsbetrieben. Dabei handelt es sich um Betriebe, die ihre Lieferungen und Leistungen nicht außerhalb des Hoheitsbereiches oder für eine andere BgA der eigenen Trägerkörperschaft erbringen, selbst wenn die Leistungserbringung auf wirtschaftlicher Tätigkeit erfolgt, begründet dies kein BgA.⁹⁴ Dies liegt darin begründet, dass ein BgA von der wirtschaftlichen Bedeutsamkeit abhängt und dies nach den Umsätzen bestimmt wird, aber Umsätze innerhalb einer jPöR erfüllen die Voraussetzungen des § 1 Abs. 1 Nr. 1 UStG nicht und daher ist auch kein BgA begründet.⁹⁵

Die neuere Rechtsprechung hat eine andere Auffassung und grenzt die Leistung nicht mehr anhand der Form der Aufgabenzuweisung oder dem Inhalt der Tätigkeit ab, sondern nur noch anhand der rechtlichen Grundlage.⁹⁶ Daher ist maßgeblich, ob eine Leistung innerhalb der öffentlich-rechtlichen Sonderregelungen ausgeübt wird oder ob die Tätigkeit unter den gleichen Rahmenbedingungen ausgeübt wird wie von privaten Wirtschaftsteilnehmern.⁹⁷ Im Ergebnis ist ein Handeln durch die jPöR auf privatrechtlicher Grundlage grundsätzlich immer unternehmerisch und bei einem Handeln auf öffentlich-rechtlicher Grundlage ist diese stets nichtunternehmerisch, außer eine Behandlung als Nichtunternehmer führt zu größeren Wettbewerbsverzerrungen.⁹⁸ Nur im Falle der Ausübung auf öffentlich-rechtlicher Tätigkeit kann ein hoheitliches Handeln vorliegen. Dies gilt nur, wenn eine Handlungsform ausgewählt worden ist, die öffentlich-rechtlich ist, da bei einer öffentlich-rechtlichen Aufgabe in der Form des Privatrechts niemals das Kriterium in Ausübung der öffentlichen Gewalt vorliegen kann, selbst wenn alle übrigen Voraussetzungen vorliegen.⁹⁹

3.2.2. Vermögensverwaltung

Kein BgA liegt vor, wenn es sich um vermögensverwaltende Tätigkeiten handelt, da der wirtschaftliche Geschäftsbetrieb in § 14 AO die Definition des BgA umfasst und Vermögensverwaltung ausschließt.¹⁰⁰ Unter die Vermögensverwaltung fallen die Vermietung und Verpachtung von

⁸⁶BVerfG, 24.02.2017 Bundesverfassungsgericht (BVerfG), Beschluss vom 24.2.2017, 2 Aktenzeichen einer Verfassungsbeschwerde zum Bundesverfassungsgericht (BvR) 2524/16, DNotZ 2017, S. 706-715.

⁸⁷BGH, 18.12.1981, I ZR 34/80, Entscheidungen des Bundesgerichtshofes in Zivilsachen (BGHZ) 82, S. 375-398.

⁸⁸Vgl. Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rz. 207.

⁸⁹Vgl. Kirch, in: Fritsch et al., 2015, § 2 UStG Rn. 201.

⁹⁰Vgl. Kirch, in: Fritsch et al., 2015, § 2 UStG Rn. 202; Sterzinger, in: Küffner et al., 2018, § 2 UStG, Rn. 799.

⁹¹BFH, 14.03.1990, I R 156/87, BStBl. II 1990, S. 866-868, a.A. Stadie, in: Rau et al., 2018, § 2 UStG Rn. 1283; Nach dem Gesetz reichen Zwangs- und Monopolrechte für die Annahme eines Hoheitsbetriebes nicht aus, wobei dies die wesentlichen Merkmale öffentlich-rechtlicher Pflichtaufgaben sind und dem Staat sind nur die klassischen Staatsaufgaben (ins. Polizei- und ordnungsrechtlichen Gefahrenabwehr, Strafverfolgung und Steuererhebung) vorbehalten. Dabei sind diese Tätigkeiten sowie nicht wirtschaftlich, sodass die Abgrenzungsformel keinen Sinn hat.

⁹²BFH, 21.11.1967, I 274/64, BStBl. II 1968 S. 218; Die Abgrenzung hoheitlicher Tätigkeiten von wirtschaftlichen Tätigkeiten befindet sich im Anhang 3.

⁹³Vgl. Bundesregierung, 2015, BStBl. I Sondernummer 1/2016, R 4.4 Abs. 1 S. 1.

⁹⁴Vgl. Stadie, in: Rau et al., 2018, § 2 UStG Rn. 1257; Oberfinanzdirektion NRW, Arbeitshilfe, 2014, S. 26.

⁹⁵Vgl. OFD Nordrhein-Westfalen, 2014, S. 26.

⁹⁶BFH, 15.04.2010, V R 10/09, BStBl. II 2010, S. 863-869; BFH, 20.08.2009, V R 30/06, BStBl. II 2010, S. 863-869; EuGH, 12.09.2000, Rs. C-276/97, Kommission/Frankreich, Slg 2000, S. I-6251; EuGH, Urteil vom 17. 10. 1989, Rs. C-231/87, Commune di Carpaneto Piacentino, Slg 1989, S. 3233.

⁹⁷Vgl. Stadie, in: Rau et al., 2018, § 2 UStG Rn. 794.

⁹⁸Vgl. Stadie, in: Rau et al., 2018, § 2 UStG Rn. 796-797.

⁹⁹Vgl. Sterzinger, 2009, S. 44.

¹⁰⁰Vgl. Becker, 2012, S. 2154; Für die Abgrenzung der Begriffe siehe Anhang 4.

Grundvermögen, die Kapitalanlage und die bloße nicht beeinflussende Beteiligung an einer Kapitalgesellschaft.¹⁰¹ Da die Rechtsprechung die Verknüpfung von Körperschaftsteuer - mit Umsatzsteuerrecht ablehnt¹⁰² müssen beide Rechtsgebiete getrennt behandelt werden mit der Rechtsfolge, dass vermögensverwaltende Tätigkeiten der Umsatzsteuer unterlegen können.¹⁰³ Die BFH-Rechtsprechung fußt auf der Rechtsprechung des EuGH, die klarstellt, dass der Gesetzgeber mithilfe des Artikel 13 Abs. 2 MwStSystRL vermögensverwaltende Tätigkeiten dem Hoheitsbereich zuweisen kann.¹⁰⁴ Dies ist durch den deutschen Gesetzgeber nicht umgesetzt worden, außerdem muss bei Vorliegen der Tatsache zum Hoheitsbereich, trotzdem noch eine Wettbewerbsverzerrung geprüft werden.¹⁰⁵ Davon zu unterscheiden ist die zwingende Umsetzung des Begriffes der wirtschaftlichen Tätigkeit, welcher durch § 2 Abs. 3 Satz 1 UStG a.F. i. V. m. § 4 Abs. 1 Satz 1 UStG kodifiziert worden ist.¹⁰⁶ Daher reicht die Zuordnung der vermögensverwaltenden Tätigkeiten nach dieser Systematik als hoheitlich einzustufen und bei fehlender Wettbewerbsverzerrung nicht zu besteuern, nicht aus.¹⁰⁷ Daneben müssen sogar nicht die Tatbestandsmerkmale der körperschaftsteuerlichen Voraussetzungen des § 4 KStG erfüllt sein, da dies im Umsatzsteuerrecht nur eine Sammelbezeichnung ist.¹⁰⁸ Im Ergebnis führt die richtlinienkonforme Auslegung des BFH dazu, dass die Leistungen der Vermögensverwaltung unabhängig von dem Vorliegen eines Wettbewerbs, nicht mehr von dem privatrechtlichen in den hoheitlichen Bereich überführt werden können und eine Zuordnung zu dem unternehmerischen Bereich stets erfolgt.¹⁰⁹ Damit erfolgt die Umsatzbesteuerung der Vermögensverwaltung rechtsformneutral.¹¹⁰ Dadurch wird auch in diesem Bereich das wirtschaftliche Tätigkeitsfeld der jPöR ausgeweitet.

3.2.3. Beistandsleistungen

Die Beurteilung einer Beistandsleistung¹¹¹ durch eine jPöR im Bereich der hoheitlichen Aufgaben erfolgte stets nichtunternehmerisch, während eine Beistandsleistung im außerhalb des hoheitlichen Raums liegenden Betätigungsfeld als unternehmerisch zu behandeln war¹¹², wie dies

bspw. bei der Beistandsleistung von einem BgA an einen anderen BgA der Fall ist. In diesem Fall hatte der BFH auch eine andere Meinung, der die Beistandsleistung durch eine Kommune an eine andere Kommune als umsatzsteuerbar einstuft, obwohl es sich um eine hoheitliche Beistandsleistung handelte.¹¹³ In dem vorliegenden Fall hat eine Kommune einer anderen Kommune auf öffentlich-rechtlicher Grundlage entgeltlich eine Sporthalle zu schulischen Zwecken überlassen. Der BFH hat einen anteiligen Vorsteuerabzug toleriert, da die Hallenüberlassung gegenüber privaten Dritten in einer Wettbewerbssituation geschieht.¹¹⁴ Dieses richtlinienkonforme Urteil stellte die Anwenderpraxis vor große Herausforderungen, da es zu einer Ausweitung der Umsatzsteuerbarkeit kommt.¹¹⁵ Besonders die Gestaltungsmöglichkeit statt auf privatrechtlichen Verträgen auf öffentlich-rechtlicher Grundlage tätig zu werden, findet daher keine Anwendung mehr.¹¹⁶ Das Urteil eröffnet sowohl eine Chance als auch ein Risiko. Das Risiko kann darin bestehen, dass die Leistungen der Bestandsleistungen um den Betrag des Umsatzsteuersatzes, bei der gleichzeitigen Geltendmachung der Vorsteuer für die Eingangsleistung, erhöht werden, obwohl das ursprüngliche Ziel einer Kooperation damit zunichte gemacht wird und die Synergieeffekte nicht ausreichen, den erhöhten Betrag der Steuerlast auszugleichen.¹¹⁷ Dies zeigt sich besonders, wenn eine Leistung im nichtunternehmerischen hoheitlichen Bereich erbracht wird und kein Vorsteuerabzug durch den Empfänger der Leistung möglich ist.¹¹⁸ Ein Vorteil kann darin bestehen, dass bei hohen Investitionssummen der Beistandsleistungsempfänger die Vorsteuer geltend machen kann, wobei der Empfänger stets die Umsatzsteuer bezahlen muss.¹¹⁹ Daher lässt sich schlussfolgern, dass der Leistungsempfänger zwar immer die Umsatzsteuer bezahlen muss, aber anders als privatrechtliche Unternehmen nicht immer steuerpflichtige Ausgangsleistungen hat, sodass sich hieraus Liquiditätsnachteile ergeben, die durch höhere Abgaben aufgefangen werden müssen.¹²⁰

3.3. Zwischenfazit

Die Regelung des § 2 Abs. 3 UStG a.F. und des § 4 KStG kann wie folgt zusammengefasst werden: Die jPöR sind immer Nichtunternehmer. Unternehmer sind diese nur im Rahmen ihrer Einrichtungen, wenn sich diese durch eine nachhaltige wirtschaftliche Tätigkeit zur Einnahmenerzielung aus der Gesamtbetätigung der juristischen Person wirtschaftlich herausheben.¹²¹ Die Vorgaben aus dem Unionsrecht gehen von einem gegenteiligen Ansatz aus, da gem. Artikel

¹⁰¹Vgl. Becker, 2012, S. 2154.

¹⁰²BFH, 15.04.2010, V R 10/09, BStBl. II 2017, S. 863-869, Rn. 23-25.

¹⁰³Vgl. Bollweg und Küffner, 2010, S. 652.

¹⁰⁴EuGH, 04.06.2009, Rs. C-102/08, SALIX Grundstücks-Vermietungsgesellschaft, Slg 2009, S. I-4629, Rn. 37-40; Die Umsatzsteuerbefreiung für Vermietung und Verpachtung ergibt sich aus Artikel 135 MwStSystRL; Vgl. Küffner/Widmann, UR 2009, S. 491.

¹⁰⁵EuGH, 04.06.2009, Rs. C-102/08, SALIX Grundstücks-Vermietungsgesellschaft, Slg 2009, S. I-4629, Rn. 76.

¹⁰⁶BFH, 15.04.2010, V R 10/09, BStBl. II 2017, S. 863-869, Rn. 24.

¹⁰⁷BFH, 15.04.2010, V R 10/09, BStBl. II 2017, S. 863-869, Rn. 24.

¹⁰⁸BFH, 15.04.2010, V R 10/09, BStBl. II 2017, S. 863-869, Rn. 27.

¹⁰⁹Vgl. Iserer und Keyser, 2011, S. 83; Krausel, UR 2010, S. 488; BFH, 15.04.2010, V R 10/09, BStBl. II 2017, S. 863-869, Rn. 25.

¹¹⁰Vgl. Cloß, 2013, S. 331.

¹¹¹Vgl. Gröpl und Zukiwski, 2014, S. 79; Dies wird im Steuerrecht definiert als Unterstützungsleistungen einer jPöR an eine andere jPöR.

¹¹²Vgl. OFD Rostock, 21.11.2002, S 2706 – 04/01 – St 242, UR 2003, S. 303-304.

¹¹³Vgl. Heidner, 2016, S. 48; Baldauf, UR 2014, S. 554; BFH, 10.11.2011, V R 41/10, BStBl. II 2017, S. 869-873, Rn. 12-16.

¹¹⁴Vgl. Kußmaul et al., 2017b, S. 291.

¹¹⁵Vgl. Küffner, 2012, S. 277-278.

¹¹⁶Vgl. Carl, 2012, S. 1589-1590.

¹¹⁷Vgl. Suck, 2013, S. 207; Strahl, UR 2012, S. 387.

¹¹⁸Vgl. Schmitz und Möser, 2014, S. 610.

¹¹⁹BFH, 10.11.2011, V R 41/10, BStBl. II 2017, S. 869-873, Rn. 5; Strahl, 2012, S. 383.

¹²⁰Vgl. Suck, 2013, S. 207.

¹²¹Vgl. Heidner, 2018, S. 736; Ermittlung eines BgA siehe Anhang 5.

9 MwStSystRL jeder Steuerpflichtiger ist, der eine wirtschaftliche Tätigkeit unabhängig von den Rahmenbedingungen ausübt. Daher knüpft die Unternehmereigenschaft nicht an die Rechtsform an, sondern alleine an die Tätigkeit.¹²² Die Vorgaben aus der Rechtsprechung lassen sich wie folgt darstellen (vgl. Abbildung 1).

4. Umsatzbesteuerung der juristischen Personen des öffentlichen Rechts nach neuer Rechtslage

4.1. Änderung des nationalen Umsatzsteuerrechts der juristischen Personen des öffentlichen Rechts durch das Steueränderungsgesetz 2015

Durch die Kritik an der alten Regelung des § 2 Abs. 3 UStG a.F. sah sich der Gesetzgeber in der Verantwortung eine Neuregelung der Umsatzbesteuerung der öffentlichen Hand zu treffen.¹²³ Bereits vor der Einführung des § 2b UStG in seiner jetzigen Form, wurde die alte Rechtsgrundlage des § 2 Abs. 3 UStG a.F. durch den Bundesfinanzhof mithilfe der EuGH Rechtsprechung zu Artikel 13 der MwStSystRL ausgelegt.¹²⁴ Die gesetzgeberische Initiative ging von dem Koalitionsvertrag zwischen der Christlich Demokratischen Union (CDU) und der Christlich Sozialen Union (CSU) sowie der Sozialdemokratischen Partei Deutschlands (SPD) vom 16.12.2013 aus, mit dem Ziel interkommunale Kooperationen nicht zu belasten.¹²⁵ Daraufhin wurde eine Staatssekretär-Arbeitsgruppe gebildet, die im Mai 2014 einen Entwurf einer Neuregelung der Umsatzbesteuerung der öffentlichen Hand in Form des § 2b UStG vorlegt.¹²⁶ Der Entwurf wurde von kommunalen Spitzenverbänden und wirtschaftlichen Interessenverbänden entsprechend ihrer jeweiligen Interessen als zu weit gefasst oder zu begrenzt angesehen.¹²⁷ Bis auf kleinere Veränderungen wurde die Norm mit der Übergangsregelung des § 27 Abs. 22 UStG mit Artikel 12 des Steueränderungsgesetzes 2015 durch den Bundestag am 24.09.2015 beschlossen und vom Bundesrat ein paar Tage später am 16.10.2015.¹²⁸

Die alte Rechtsnorm des § 2 Abs. 3 UStG a.F. wurde durch den ab dem 01. Januar 2017 in Kraft getretenen § 2b UStG abgelöst. Durch den § 27 Abs. 22 UStG hat der Gesetzgeber eine „Eingewöhnungszeit“ von fünf Jahren geschaffen, sodass jPöR durch eine Optionserklärung gegenüber dem örtlich zuständigen Finanzamt erklären können, dass der § 2 Abs. 3 UStG a.F. für alle Umsätze bis zum 31.12.2020 weiterhin anzuwenden ist.¹²⁹ Der § 2b UStG wurde aufgrund eu-

roparechtlicher Vorgaben geschaffen, sodass die Rechtsprechung des Bundesfinanzhofes weiterhin anzuwenden ist.¹³⁰ Der § 2b Abs. 1 UStG adaptiert die Regelung des Artikel 13 Abs. 1 Unterabschnitt 1 und 2 MwStSystRL in das Umsatzsteuerrecht.¹³¹ Die bisherige Verknüpfung des Betriebes gewerblicher Art aus dem Körperschaftsteuerrecht, welche durch § 2 Abs. 3 S. 1 UStG a.F. mit dem Umsatzsteuerrecht verknüpft war, wurde durch das neue Gesetz aufgehoben.¹³² Durch die Aufhebung der Verknüpfung wird durch den neuen § 2b UStG eine deutliche Erweiterung der umsatzsteuerrelevanten Tätigkeitsfelder verursacht.¹³³ Nach der alten Fassung des § 2 Abs. 3 UStG a.F. waren die jPöR nur im Rahmen der BgA, ihrer landwirtschaftlichen bzw. forstwirtschaftlichen Betriebe und der aufgezählten Tätigkeiten gem. § 2 Abs. 3 S. 2 Nr. 2 bis 5 UStG a.F. unternehmerisch tätig.¹³⁴ Durch die Neuorientierung sind vermögensverwaltende Betätigungen gem. § 14 S. 3 AO und wirtschaftliche Tätigkeiten mit geringem Jahresumsatz¹³⁵ sowie ggf. Beistandsleistungen auch noch unter den Tatbestand des § 2b UStG zu subsumieren.¹³⁶

Der §2b UStG hat einen ähnlichen Aufbau wie Artikel 13 Abs. 1 Unterabschnitt 1 und 2 MwStSystRL.¹³⁷ Dazu stellt § 2 Abs. 1 S. 1 UStG sämtliche Tätigkeiten im Rahmen der öffentlichen Gewalt von einer jPöR von der Besteuerung frei, wobei direkt im Anschluss in dem nächsten Satz eine Eingrenzung gemacht wird, dass die Nichtbesteuerung keine größeren Wettbewerbsverzerrungen auslösen darf.¹³⁸ Abweichend von dem Unionsrecht enthält § 2b UStG durch den Absatz 2 und 3 eine beispielhafte Aufzählung, wann Wettbewerbsverzerrungen nicht gegeben sind.¹³⁹ § 2b Abs. 2 UStG ist eher allgemein gehalten, während der Absatz 3 besonders die Tätigkeiten der öffentlich-rechtlichen Zusammenarbeit behandelt.¹⁴⁰ Der letzte Absatz 4 enthält wie bereits die alte Fassung des § 2 Abs. 3 S. 2 UStG eine Tätigkeiten-Aufzählung, die automatisch eine Unternehmereigenschaft bewirkt.¹⁴¹ Grundsätzlich sind jPöR i.S.d. § 2 Abs. 1 UStG immer Unternehmer.¹⁴² Damit hat der § 2b UStG die Struktur, dass er aus einer Grundregel der Nichtbesteuerung im Rahmen der öffentlichen Gewalt mit den zwei Ausnahmeregelungen des Wettbewerbsvorbehaltes und der Sondertätigkeit besteht, wobei innerhalb des Wettbewerbsvorbehaltes zwei Rückausnahmen durch Absatz 2 und 3 existieren, die den Grundtatbestand erzeugen.¹⁴³

¹²²Vgl. Heidner, 2018, S. 737.

¹²³Der Gesetzgebungsprozess befindet sich in Anhang 6.

¹²⁴BFH, 10.11.2011, V R 41/10, BStBl. II 2017, S. 869-873; BFH, 03.03.2011, V R 23/10, BStBl. II 2012, S. 74-77; BFH, 15.04.2010, V R 10/09, BStBl. II 2017, S. 863-869; BFH, 20.08.2009, V R 30/06, BStBl. II 2010, S. 863-869.

¹²⁵Vgl. CDU, CSU, SPD, 2013, S. 64.

¹²⁶Vgl. Müller, 2017, S. 8.

¹²⁷Vgl. Weber, 2015, S. 365; Widmann, 2015b, S. 6.

¹²⁸Vgl. Baldauf, 2016; Widmann, 2015a, S. 883; Bundestag, 23.09.2015, Artikel 12.

¹²⁹Vgl. Hüttemann, 2017, S. 129-130.

¹³⁰Vgl. Jürgens und Menebröcker, 2018, S. 174.

¹³¹Vgl. Heidner, 2018, S. 739; Krogoll/Sloty-Harms, UVR 2016, S. 181.

¹³²Vgl. Baldauf, 2016, S. 358-359.

¹³³Vgl. Baldauf, 2016, S. 359.

¹³⁴Vgl. Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rn. 1; Wagner et al., 2017, S. 13.

¹³⁵Dies sind die BgA, die jährliche Umsätze weniger als 35.000 € erzielen.

¹³⁶Vgl. Baldauf, 2016, S. 358; Belcke und Westermann, 2016, S. 88.

¹³⁷Vgl. Kußmaul et al., 2017b, S. 293.

¹³⁸Vgl. Korn, in: Bunjes et al., 2018, § 2 UStG Rn.7.

¹³⁹Vgl. Weber, 2017, S. 263; Widmann, 2015a, S. 886.

¹⁴⁰Vgl. Korn, in: Bunjes et al., 2018, § 2 UStG Rn.7.

¹⁴¹Vgl. Fiand, 2016, S. 32.

¹⁴²Vgl. Heidner, 2018, S. 739.

¹⁴³Vgl. Korn, in: Bunjes et al., 2018, § 2 UStG, Rn.7; Ismer, 2016, S. 741.

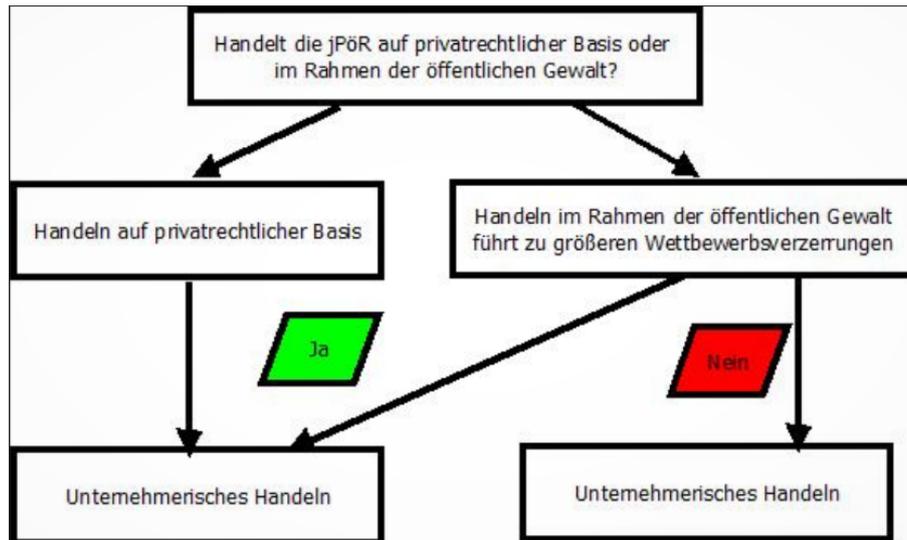


Abbildung 1: Prüfungsschema nach EuGH Rechtsprechung

Quelle: Eigene Darstellung in Anlehnung an Becker, 2012, S. 2155.

4.2. Erfassung von juristischen Personen des öffentlichen Rechts bei der Umsatzbesteuerung in der neuen Rechtslage

4.2.1. Überblick über die Voraussetzungen für die Anwendung der neuen Rechtslage

Der § 2b UStG schränkt den *lex generalis* des § 2 Abs. 1 UStG ein, was sich schon an der Platzierung hinter dem allgemeinen Paragraphen ergibt.¹⁴⁴ Damit der § 2b UStG angewendet werden kann, muss die jPöR unternehmerisch i.S.d. § 2 Abs. 1 UStG tätig sein, was sich auch schon durch den Wortlaut der Eingangsvorschrift des § 2b Abs. 1 UStG ergibt.¹⁴⁵ Trotz des Vorliegens der Voraussetzungen des § 2 Abs. 1 UStG sind die jPöR dann keine Unternehmer, wenn diese ihre Tätigkeit im Rahmen der Ausübung öffentlicher Gewalt vollziehen und wenn deren Nichtbesteuerung zu keiner größeren Wettbewerbsverzerrung führt.¹⁴⁶ Damit es für die jPöR zu keiner Einordnung als Unternehmer i.S.d. § 2 Abs. 1 UStG kommt, müssen nachfolgende vier Bedingungen gleichzeitig vorliegen. Als erste Bedingung muss die jPöR eine Tätigkeit ausführen und diese Tätigkeit muss zweitens im Rahmen der öffentlichen Gewalt ausgeführt werden. Daneben darf als dritte Voraussetzung eine Nichtbesteuerung zu keiner größeren Wettbewerbsverzerrungen führen.¹⁴⁷ Außerdem darf als letzte Voraussetzungen keine Katalogtätigkeit gem. § 2b Abs. 4 UStG vorliegen. Nachfolgend werden die Voraussetzungen kritisch erläutert.

4.2.2. Ausübung öffentlicher Gewalt

Die erste Voraussetzung, damit § 2b UStG greift, ist das Tätigwerden auf einer wirtschaftlichen Grundlage gem. § 2 Abs. 1 UStG. Die zweite Voraussetzung, um den § 2b UStG zu nutzen, ist das Tätigwerden im Rahmen der öffentlichen Gewalt. Eine jPöR hat verschiedene Möglichkeiten tätig zu werden, dies ist zum einen auf privatrechtlicher Grundlage oder zum anderen auf öffentlich-rechtlicher Grundlage.¹⁴⁸ Tätigkeiten (z.B. Miet- oder Pachtvertrag), die auf privatrechtlicher Grundlage erbracht werden sind stets unter Beachtung der weiteren Voraussetzungen zu einer unternehmerischen Tätigkeit des § 2 Abs. 1 UStG zu subsumieren.¹⁴⁹ Die Definition zur Ausübung auf öffentlich-rechtlicher Tätigkeitsgrundlage ist demnach schwierig, weil weder der Gesetzeswortlaut des § 2b Abs. 1 UStG diesen Begriff konkretisiert, noch eine Definition aus dem Unionsrecht entnommen werden kann.¹⁵⁰ Das BMF-Schreiben konkretisiert, dass weder das Ausführen einer Tätigkeit aufgrund gesetzlicher Anordnungen, noch die Erfüllung öffentlicher Aufgaben zu der Ausübung von öffentlicher Gewalt führt.¹⁵¹ Dieser Umstand kommt daher, dass öffentliche Aufgaben sowohl von den jPöR, als auch von privatrechtlichen Unternehmen durchgeführt werden können.¹⁵²

Das Tätigwerden im öffentlichen Bereich soll gem. BMF-Schreiben¹⁵³ dann vorliegen, wenn die jPöR im Rahmen einer öffentlich-rechtlichen Sonderregelung tätig wird, bei einer

¹⁴⁴Vgl. Müller, 2017, S. 10.

¹⁴⁵Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 3; EuGH, 29.09.2015, Rs. C-276/14, Gmina Wrocław, UR 2015, S. 829-834 Rz. 30; EuGH, 29.10.2015, Rs. C-174/14, Sadaçor, UR 2015, S. 901-909, Rz. 52.

¹⁴⁶Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 5; Baldauf, 2016, S. 359.

¹⁴⁷Vgl. Kirch, in: Fritsch et al., 2015, § 2b UStG Rz. 4; Huschens, 2015, S. 76.

¹⁴⁸Vgl. Küffner, DB 2016, S. 1470.

¹⁴⁹Vgl. Kirch, in: Fritsch et al., 2015, § 2b UStG Rz. 8; Sterzinger, 2017, S. 233; Baldauf, 2016, S. 359.

¹⁵⁰Vgl. Baldauf, 2016, S. 359.

¹⁵¹Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 18.

¹⁵²Vgl. Trost und Menebröcker, 2017, S. 39.

¹⁵³Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 6-18; siehe Kapitel C.II.1.

zulässigen öffentlich-rechtlichen Handlungsform. Diese Auffassung entspricht auch der ständigen Rechtsprechung des BFH.¹⁵⁴ Das Kriterium der Sonderregelung kann sich ergeben aus einem Gesetz oder einer Rechtsverordnung, einer Satzung. Die Satzungen werden von der jPöR erlassen, um die eigenen Angelegenheiten zu regeln, dabei kann eine wirtschaftliche Tätigkeit verneint werden, wenn die Handlung auf einer Satzung und einer öffentlich-rechtlichen Handlungsform beruht.¹⁵⁵ Daneben gelten auch Staatsverträge (z.B. Rundfunkstaatsvertrag), verfassungsrechtliche Verträge, Verwaltungsabkommen, Verwaltungsvereinbarungen und öffentlich-rechtliche Verträge zu der Sonderregelung.¹⁵⁶ Es gibt keine Staatsaufgabenlehre, sodass weiterhin Streitigkeiten bei der Zuordnung zu der öffentlich-rechtlichen Grundlage existieren.¹⁵⁷ Im Bereich des öffentlich-rechtlichen Vertrages kann es zu Abgrenzungsschwierigkeiten kommen, hinsichtlich des zugrunde liegenden Vertragswerkes, welches entweder ein Vertrag auf privatrechtlicher Grundlage ist oder ein Vertrag der dem öffentlichen Recht zugeordnet werden kann.¹⁵⁸ Dabei dient als Abgrenzungskriterium, ob der Vertrag dem öffentlichen Recht oder dem Zivilrecht zugeordnet werden kann, der Zweck und der Gegenstand der vertraglichen Regelung.¹⁵⁹ Ein Vertrag wird dann als öffentlich-rechtlich eingeordnet, wenn sowohl eine öffentlich-rechtliche Norm den Abschluss berechtigt und die Normen die Leistungspflichten regeln, sowie der Vertrag in Vollzug einer öffentlich-rechtlichen Norm geschlossen wird, der Vertrag als Substitution zu einem Verwaltungsakt dient und dieser sich auf eine öffentlich-rechtliche Berechtigung oder Verpflichtung des Bürgers bezieht.¹⁶⁰

Das zweite Kriterium der Handlungsform ist dann erfüllt, wenn die Handlungsform zulässig ist. Dabei gilt der Grundsatz, dass die Finanzbehörden von der Zulässigkeit der gewählten Handlungsform ausgehen.¹⁶¹ Das Handeln im Rahmen der öffentlichen Gewalt ist an eine rechtliche Grundlage geknüpft.¹⁶² Daneben gibt es auch Tätigkeiten, die nur in privatrechtlicher Handlungsform durchgeführt werden, z.B. die Überlassung von Werbeflächen¹⁶³, sodass in der Regel bei der Nutzung von der privatrechtlichen Form diese Tätigkeiten nicht von § 2b UStG erfasst werden.¹⁶⁴ Dies bietet für die jPöR die Möglichkeit, bei der Umsetzung einer öffentlich-rechtlichen Sonderregelung Tätigkeiten entweder in der unternehmerischen privatrechtlichen oder in der nicht-unternehmerischen öffentlich-rechtlichen Handlungsform zu erledigen unter Beachtung der Wettbewerbsverzerrungen.¹⁶⁵

Die Regelung, dass die jPöR im Rahmen sogenannter Hilfsgeschäfte, die verknüpft sind mit der eigentlichen Tätigkeit im nichtunternehmerischen Bereich der jPöR (z.B. Verkauf von zuvor genutzten hoheitlich tätigen Feuerwehrfahrzeugen) trotz einer nachhaltigen Betätigung noch in die nichtunternehmerische Sphäre gehören¹⁶⁶, wurde aus dem bisherigen Recht übernommen.¹⁶⁷ Diese Regelung hat zur Folge, dass eine Vielzahl von Leistungen weiterhin als nichtumsatzsteuerbar behandelt wird, obwohl die Befreiung auch bei nachhaltig ausgeübten Veräußerungsgeschäften gegen das Unionsrecht verstoßen kann.¹⁶⁸ Nach der Auffassung des EuGH¹⁶⁹ tritt bei der nachhaltigen Veräußerung von Gegenständen aus dem hoheitlichen Bereich durch den Veräußerer dieser jedoch wie ein Kaufmann auf, sodass dieser auch als solcher behandelt werden muss.¹⁷⁰ Voraussetzung dafür ist, dass die gelegentlich ausgeübten Tätigkeiten für sich eine wirtschaftliche Tätigkeit i.S.d. Artikel 9 der MwStSystRL sind und dass die Vermarktung der Gegenstände nicht rein passiv erfolgt und nach der Verkehrsanschauung die Verkäufe nicht händleruntypisch sind.¹⁷¹ Die folgenden Beispiele sollen die vorher geschilderten unterschiedlichen Fälle noch einmal verdeutlichen (vgl. Tabelle 2) .

4.2.3. Größere Wettbewerbsverzerrungen

Potentielle Wettbewerbsverzerrungen

Die dritte Voraussetzung für die Nutzung des § 2b UStG liegt darin, dass keine größere Wettbewerbsverzerrung vorliegt. Die Regelung des § 2b Abs. 1 S. 2 UStG zur Definition der größeren Wettbewerbsverzerrung ist aus Artikel 13 Abs. 1 Unterabschnitt 2 MwStSystRL abgeleitet. Dies wird auch durch den Gesetzgebungsprozess klar.¹⁷⁴ Das entsprechende BMF-Schreiben orientiert sich an der Rechtsprechung zu der Frage, ob die Nichtbesteuerung einer mithilfe der öffentlichen Gewalt ausgeübten Tätigkeit zu größeren Wettbewerbsverzerrungen führt.¹⁷⁵ Diese Ausführungen des BMF-Schreibens werden positiv gesehen, da damit keine Abweichungen der Rechtsprechung von der Verwaltungsmeinung bestehen.¹⁷⁶ Die jPöR kann sich nur als Nichtunternehmer

¹⁵⁴BFH, 10.02.2016, XI R 26/13, BStBl. II 2017, S. 857-863.

¹⁵⁵Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 8.

¹⁵⁶Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 10, 12.

¹⁵⁷Vgl. Widmann, 2015a, S. 886; Ringwald, 2015, S. 1.

¹⁵⁸Vgl. Baldauf, 2016, S. 359.

¹⁵⁹FG Münster, 16.04.2013, 15 K 227/10 U, EFG 2013, S. 1266-1270.

¹⁶⁰FG Münster, 16.04.2013, 15 K 227/10 U, EFG 2013, S. 1266-1270.

¹⁶¹Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 17.

¹⁶²Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 16.

¹⁶³Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 16.

¹⁶⁴Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 6.

¹⁶⁵Vgl. Weber, 2017, S. 78.

¹⁶⁶Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 19,20; a.A. Janzen, in: Lippross und Seibel, 2019, § 2b UStG Rn. 12; Hammerl/Fietz, MwStStR 2018, S. 58.

¹⁶⁷Vgl. Baldauf, 2018b, BStBl. I Sondernummer 1/2016, R 4.4 Abs. 2; BFH, 23.10.1996, I R 1/94, BStBl. II 1997, S. 139-143.

¹⁶⁸Vgl. Hammerl/Fietz, MwStStR 2018, S. 58.

¹⁶⁹EuGH, 09.07.2015, Rs. C-331/14, Petar Kezic, MwStStR 2015, S. 723-725, Rn. 24.

¹⁷⁰Vgl. Korn, in: Bunjes et al., 2018, § 2 UStG, Rn.13, 38; Heuermann, 2015, S. 725.

¹⁷¹Vgl. Baldauf, 2017, S. 875.

¹⁷²Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 25.

¹⁷³Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 19.

¹⁷⁴Vgl. Bundestag, 23.09.2015, S. 91.

¹⁷⁵Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 24,27,31; mit der Zitation von EuGH, 16.09.2008, Rs. C-288/07, Isle of Wight Council, Slg 2008, S. I-7203; BFH, 03.07.2008, V R 40/04, BStBl. II 2009, S. 208-213.

¹⁷⁶Vgl. Hammerl und Fietz, 2017, S. 58.

Tabelle 2: Beispiele Ausübung öffentlicher Gewalt

Quelle: Eigene Darstellung.

Beispiel 1: Eine Behörde stellte einen Reisepass gem. § 6 Passgesetz aus, aufgrund einer Aufgabenzuteilung des Landes NRW an diese Behörde. Der Antragsteller ist 25 Jahre alt und beantragt einen 10 Jahre gültigen 32 Seiten umfassenden Reisepass für eine Gebühr von 60,00 € gem. § 15 Abs. 1 Nr. 1a Passverordnung.

Beispiel 2: Die Polizei des Kreises Oberberg veräußert einen Einsatzwagen an einen Kfz-Gebrauchtwagenhändler, damit dieser die Teile noch verwerten kann.

Die kommunale Behörde wird hier im Rahmen der öffentlichen Gewalt tätig, da diese zur Handlung befugt ist und das Passgesetz dem öffentlichen Recht zugewiesen wird. Damit besteht ein Handeln auf originär hoheitlicher Basis, was keine Wettbewerbsrelevanz auslöst, da ein privater Dritter nicht tätig werden kann, sodass keine Umsatzsteuer anfallen kann.¹⁷²

Die Veräußerung unterliegt dem hoheitlichen Hilfsgeschäft, da eine Veräußerung eines Gegenstandes aus dem nichtunternehmerischen Tätigkeitsbereich diesem zugeordnet wird. Der Verkauf unterliegt zwar der allgemeinen Regelung des § 2 Abs. 1 UStG aber die Finanzverwaltung sieht auch bei regelmäßigen Hilfsgeschäften keinen steuerbaren Umsatz vorliegen.¹⁷³

behandeln lassen, wenn dies dafür sorgt, dass es zu keiner Wettbewerbsverzerrung gegenüber privaten Teilnehmern führt.¹⁷⁷

Der Wettbewerb kann eine Wettbewerbsverzerrung auslösen, denn ohne einen Wettbewerb kann Wettbewerbsverzerrung nicht existieren.¹⁷⁸ Dabei muss die Tätigkeit marktrelevant sein.¹⁷⁹ Zunächst ist es erforderlich, dass die jPöR eine Tätigkeit auf öffentlich-rechtlicher Grundlage ausübt, die ebenso von einem privaten Anbieter auf einem (potentiellen) Markt angeboten wird.¹⁸⁰ Damit ein Wettbewerb existiert, müssen demnach drei Bedingungen erfüllt sein: die Marktrelevanz der Leistung, mögliche marktrelevante rechtliche Rahmenbedingungen und die räumliche Marktrelevanz.

Ob eine Marktrelevanz der Leistung vorliegt, wird nach der Art der Leistung beurteilt, d.h. die Leistungen müssen ähnliche Eigenschaften besitzen.¹⁸¹ Dabei wird die Leistung aus der Sicht des Durchschnittsverbrauchers betrachtet, welche für diesen dieselben Bedürfnisse befriedigen soll.¹⁸² Der Wettbewerb wird dann festgestellt, wenn die Leistungsart bzw. die gleichartigen Leistungen auch von privaten Wirtschaftsteilnehmern im Angebotsportfolio auftauchen. Ein gegenwärtiger Wettbewerb ist immer wettbewerbsverzerrend. Damit der private Wirtschaftsteilnehmer beweisen kann, dass dieser in den relevanten Markt eintreten kann, sollte er dies

mit Tatsachen, objektiven Indizien oder einer Marktanalyse belegen.¹⁸³ Damit aber ein potentieller Wettbewerb zu Wettbewerbsverzerrungen führt, muss dieser Wettbewerb real denkbar sein und nicht bloß rein hypothetisch vorliegen.¹⁸⁴ Das Kriterium des potentiellen Marktes stellt darauf ab, dass der private Unternehmer in der Nutzung des Wettbewerbs auch nicht behindert werden darf.¹⁸⁵

Die rechtlichen Rahmenbedingungen als zweites Kriterium der Leistungen haben erst dann eine Bedeutung und sind damit marktrelevant, wenn diese den Leistungsempfänger durch ihren maßgeblichen Einfluss dazu bringen, die Leistung einer jPöR in Anspruch zu nehmen.¹⁸⁶ Wenn dies der Fall ist, dann ist auch das Kriterium für das Vorliegen eines Wettbewerbs erfüllt. Das dritte Kriterium der räumlichen Marktrelevanz besagt, dass die Beurteilung des Wettbewerbs nicht nur auf den lokalen Markt zu beschränken ist, sondern auf das Gebiet der Europäischen Union. Dabei kann dennoch von der Art der Leistung eine regionale Beschränkung sinnvoll sein.¹⁸⁷ Der relevante räumliche Markt kann nur außerhalb einer Abnahme - oder Annahmeverpflichtung (z.B. Annahme - und Benutzungszwang bei dem Anschluss mit Wasserleitung, Abwasserbeseitigung gem. § 9 GO NRW) gültig sein.¹⁸⁸ Damit liegt bei dem Annahme - und Benutzungszwang nie-

¹⁷⁷Vgl. Müller, 2017, S. 11; ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 22.

¹⁷⁸Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 24.

¹⁷⁹Vgl. Baldauf, 2018a, S. 56.

¹⁸⁰Vgl. Gohlke und Schmitz, 2016, S. 781.

¹⁸¹Vgl. Hammerl und Fietz, 2017, S. 58.

¹⁸²Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 24.

¹⁸³Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 24.

¹⁸⁴Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 23. Eine Leistung, die nicht von Privaten erbracht werden kann, sind z.B. die Erteilung einer verbindlichen Auskunft gem. § 89 Abs. 2 AO.

¹⁸⁵Vgl. Heidner, 2016, S. 48.

¹⁸⁶Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 26.

¹⁸⁷Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 29.

¹⁸⁸Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 27.

mals ein (potentieller) Wettbewerb vor.¹⁸⁹ Insgesamt ist das Kriterium des lokalen Marktes nur von Bedeutung, wenn einem privaten Unternehmer aufgrund der Besonderheiten des lokalen Marktes der Zugang für diesen verwehrt bleibt, aber dies für einen größeren Markt nicht zutrifft.¹⁹⁰

Wenn ein Wettbewerb vorliegt muss im zweiten Schritt geklärt werden, ob eine größere Verzerrung vorliegt. Das Kriterium „größer[er]“ Wettbewerbsverzerrung liegt bereits dann vor, wenn die Verzerrungen nicht unbedeutend sind.¹⁹¹ Dabei werden an den Begriff keine großen Anforderungen gestellt, damit dieser erfüllt ist.¹⁹² Die Definition des Ausschlusses größerer Wettbewerbsverzerrungen findet in § 2b Abs. 2 und Abs. 3 UStG Anwendung. Dort werden Fälle aufgezeigt, die zu keiner größeren Wettbewerbsverzerrung führen. Durch das im Gesetz stehende Wort „insbesondere“ wird deutlich, dass es sich hierbei nicht um eine abschließende Aufzählung handelt.¹⁹³ Die eigenmächtige Definition der größeren Wettbewerbsverzerrungen durch den deutschen Gesetzgeber wird kritisiert¹⁹⁴, da die Definition ein autonomer Begriff des Unionsrechts ist und in der Vergangenheit wurde festgestellt, dass autonome Begriffe nicht durch Mitgliedsstaaten definiert werden können.¹⁹⁵ Der BFH hat auch keine Möglichkeit die Vorschriften des § 2b Abs. 2 und 3 UStG zu „retten“, indem dieser diese Vorschriften richtlinienkonform auslegt, sodass der EuGH bezüglich dieser Norm schon bald ein Urteil fällen könnte.¹⁹⁶

Jährliche Umsatzgrenze

In § 2b Abs. 2 Nr. 1 UStG ist festgehalten, dass die Umsätze der jPöR aus gleichartigen Tätigkeiten voraussichtlich 17.500 € im Kalenderjahr nicht überschreiten dürfen, wobei die Regelung nur Tätigkeiten im Rahmen der öffentlichen Gewalt erfasst.¹⁹⁷ Diese Regelung orientiert sich vom Betrag an der Kleinunternehmerregelung, auch wenn dies nicht explizit im Gesetzesentwurf oder dem BMF-Schreiben steht.¹⁹⁸ Im Gegensatz zu der Kleinunternehmerregelung gibt es zum

einen keine Optionsmöglichkeit zu der Unternehmereigenschaft zu optieren bei dem Unterschreiten der Umsatzsumme.¹⁹⁹ Zum anderen bleibt offen, ob die Bagatellgrenze des § 2b Abs. 2 Nr. 1 UStG ein Netto - oder Bruttoumsatz²⁰⁰ ist.²⁰¹ Dabei muss auf den voraussichtlich zu vereinnahmenden Betrag abgestellt werden, der vor Beginn des Kalenderjahres eingeschätzt werden kann. Dies gilt auch, wenn der Betrag überstiegen worden ist, solange dies plausibel begründet werden kann.²⁰² Die einmalige Erstellung der Prognose reicht nicht, sondern ist für jedes Jahr neu zu erstellen, wobei der Vorjahresumsatz als Orientierungswert dienen kann.²⁰³ Hingegen bei einer Neuaufnahme einer Tätigkeit ist der voraussichtliche Umsatz des Jahres der Neuaufnahme maßgeblich, ohne das auf einen fiktiven Jahresumsatz hochgerechnet wird.²⁰⁴

Der dritte Unterschied zu der Kleinunternehmerregelung besteht darin, dass § 2b Abs. 2 Nr. 1 UStG nicht auf den Gesamtumsatz verweist, sondern auf gleichartige Tätigkeiten, welche die gesamte jPöR umfasst.²⁰⁵ Diese Regelung ist mehrfach anzuwenden, da gleichartige Tätigkeiten zusammenzufassen sind.²⁰⁶ Dieser relevante unbestimmte Rechtsbegriff wird aus der Sicht der Finanzverwaltung so definiert, dass Leistungen gleichartig sind, wenn diese aus der Sicht eines Durchschnittsverbrauchers dieselben Bedürfnisse befriedigen.²⁰⁷ Diese Auffassung basiert auf der EuGH und der BFH-Rechtsprechung²⁰⁸, wobei die Auffassung von der Literatur kritisch diskutiert wird.²⁰⁹ Laut BMF-Schreiben wird eine Gleichartigkeit bei der Parkraumüberlassung mittels eines Parkplatzes und eines Parkhauses gesehen.²¹⁰ Dieser Auffassung wird zugestimmt, dennoch wird das Abstellen auf potentielle Substitutionsleistungen vorgeschlagen.²¹¹ Dies liegt dann vor, wenn der Verbraucher sich nicht für eine Alternative einer weggefallenen potentiell gleichartigen Tätigkeit entscheidet. In der Konsequenz werden sowohl die weggefallene Leistung als auch die übrig verbliebenen anderen Leistung einzeln betrachtet.²¹² Es sind neben der Ansicht auf den

¹⁸⁹Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 27.

¹⁹⁰Vgl. Heidner, 2016, S. 48; Im Fall der Feuerbestattung kann diese in manchen Bundesländern nun von jPöR erbracht werden, aber in anderen Bundesländern durch die fehlende Abnahmeverpflichtung auch von privaten Dritten; Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, S. 1451, Tz. 28.

¹⁹¹Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 31; EuGH, 16.09.2008, Rs. C-288/07, Isle of Wight Council, Slg 2008, S. I-7203.

¹⁹²Vgl. Sauerland, 2017, S. 113.

¹⁹³Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 32; Küffner und Rust, 2016b, S. 1639; Pithan, 2015, S. 275; Pithan führt aus, wenn der Gesetzgeber eine abschließende negative Abgrenzung vorgenommen hätte, würde dies vermutlich gegen den Artikel 13 MwStSystRI verstoßen, der keine weiteren Definitionen zu größeren Wettbewerbsverzerrungen beinhaltet.

¹⁹⁴Vgl. Heidner, 2018, S. 740.

¹⁹⁵Bspw.: EuGH, 15.11.2012, Rs. C-532/11, Leichenich, UR 2013, S. 30-34, Rz. 17; EuGH, 26.01.2003, Rs. C-315/00, Maierhofer, Slg 2003, S. I-563, Rz. 25, 26.

¹⁹⁶Vgl. Heidner, 2018, S. 740.

¹⁹⁷Vgl. Hüttemann, 2017, S. 133; Aus der Regelung folgt, dass bereits bei einem Euro auf privatrechtlicher Grundlage dieser Umsatz unternehmerisch und damit steuerbar ist.

¹⁹⁸Vgl. Heidner, 2018, S. 741; Kirch, in: Fritsch et al., 2015, § 2b UStG Rz. 12; Hammerl und Fietz, 2017, S. 59; Baldauf, 2016, S. 361.

¹⁹⁹Vgl. Heidner, 2018, S. 741; Hüttemann, 2017, S. 133; Kronawitter, 2017, S. 18; ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 33.

²⁰⁰Bei einem Nettoumsatz handelt es sich um den Umsatz abzüglich der Umsatzsteuer.

²⁰¹Vgl. Heidner, 2018, S. 741; Müller, 2017, S. 11; Scharpenberg, 2017, S. 835; Müller plädiert für einen Bruttoumsatz; Scharpenberg hingegen plädiert für einen Nettoumsatz, da in § 19 Abs. 1 S. 1 UStG für die Prüfung der 17.500 € auf den Umsatz und die Steuer abzustellen ist. Die Regelung in § 2b UStG enthält eine solche Regelung nicht.

²⁰²Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 36; Küffner und Rust, 2016b, S. 1639.

²⁰³Vgl. Hammerl und Fietz, 2017, S. 59.

²⁰⁴Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 34.

²⁰⁵Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 36.

²⁰⁶Vgl. Müller, 2017, S. 11; Hammerl und Fietz, 2017, S. 59.

²⁰⁷Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 36.

²⁰⁸Einheitlichkeit der Leistung: BFH, 31.05.2001, V R 97/98, BStBl. II 2001, S. 658-660; Abgrenzung Lieferung und sonstige Leistung: EuGH, 02.05.1996, Rs. C-231/94, Faaborg-Gelting Linie A/S, Slg 1996, S. I-2395.

²⁰⁹Vgl. Heidner, 2018, S. 741; a.A. Hammerl und Fietz, 2017, S. 59; Sterzinger UR 2015, S. 662; Widmann UR 2015, S. 8.

²¹⁰Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 37.

²¹¹Vgl. Müller, 2017, S. 12.

²¹²Vgl. Müller, 2017, S. 12.

Durchschnittsverbraucher abzustellen, auch andere Ansätze möglich. Diese anderen Ansätze werden aufgrund der historischen Auslegung zu dem § 4 Abs. 6 S. 1 Nr. 1 KStG und der Verlautbarungen zu dem steuerlichen Querverbund gezogen.²¹³ Diese anderen Möglichkeiten können ein Betriebszweck sein, der sich entspricht (z.B. Nutzung von Hallen- und Freizeitbad)²¹⁴ oder auch die Nutzung von Bibliotheksbestand in Form von Büchern und Verleih von eBook Readern.²¹⁵ Daneben kann die Tätigkeit auch gleichartig sein, wenn diese in einem gleichen Gewerbegebiet ausgeübt wird oder wenn die Tätigkeiten sich ergänzen.²¹⁶ Die Regelung soll mit einem Beispiel in Anhang 7 verdeutlicht werden.

Im Ergebnis hat die Umsatzschwelle des § 2b Abs. 2 Nr. 1 UStG nicht nur positive Effekte. Es kann das Ergebnis eintreten, dass ein Vorsteuerabzug verwehrt bleibt, wenn der jährliche Umsatz während der Investitionsphase unter 17.500 € liegt.²¹⁷ Damit ist die Tätigkeit der jPÖR automatisch nichtunternehmerisch und der Vorsteuerabzug aus der Tätigkeit scheidet dadurch aus. Der Vorsteuerabzug ist dann nur möglich, soweit in dem Investitionsfeld ein Tätigwerden auf privatrechtlicher Grundlage erfolgt.²¹⁸ Insofern wäre eine Optionsmöglichkeit zur Regelbesteuerung für jPÖR angebracht, wobei es keine Unionsrechtliche Grundlage dafür gibt.²¹⁹ Nicht nur die jPÖR sind von dieser Regelung betroffen, sondern auch natürliche Personen oder eine privatrechtliche juristische Person, da letztgenannte keine Möglichkeit hat, gleichartige Tätigkeiten mit einem Umsatz unter 17.500 € zu kumulieren und somit mehrere gleichartige Gruppen jeweils von unter 17.500 € hätte.²²⁰ Deshalb ist zu erwarten, dass gegen die Umsatzschwelle geklagt wird, von jPÖR, die ein Vorsteuerabzugsinteresse haben, um eine Optionsmöglichkeit zu bekommen²²¹, aber auch auf privatrechtlicher Grundlage teilnehmende Wirtschaftsteilnehmer, damit die Regelung der Einzelbetrachtung von nicht gleichartigen hoheitlichen Tätigkeiten überarbeitet wird.²²²

Umsatzsteuerbefreiung

Auf der Grundlage des Artikel 13 Abs. 2 MwStSystRl haben die EU Mitgliedsstaaten das Recht bestimmte steuerbe-

freite Tätigkeiten von öffentlichen Einrichtungen dem öffentlichen Recht zuzuordnen, sodass diese damit nicht steuerbar sind. Der Hintergrund ist, dass das private Unternehmen von der Steuerfreiheit ebenso profitiert wie die Einrichtung des öffentlichen Rechts.²²³ Durch den § 2b Abs. 2 Nr. 2 UStG wurde diese Regelung in nationales Recht umgesetzt und die Aussage dieser Norm ist, dass von größeren Wettbewerbsverzerrungen nicht auszugehen ist, wenn vergleichbare privatrechtliche Leistungen ohne die Optionsmöglichkeit des § 9 UStG sowieso steuerbefreit sind. Der Gesetzgeber hat damit letztendlich alle von der Befreiung im Umsatzsteuerrecht umfassten Sachverhalte abgedeckt außer die in § 9 UStG genannten Fälle, bei denen ein Verzicht auf die Befreiung möglich ist. Diese Leistungen (§ 4 Nr. 8 a-g, Nr. 9a, Nr. 12 oder Nr. 13 UStG) werden von der Norm nicht erfasst, soweit die Leistung steuerbar ist und an einen anderen Unternehmer für dessen Unternehmen erbracht wird, unabhängig von der tatsächliche Inanspruchnahme der Optionsmöglichkeit des § 9 UStG.²²⁴ Diese Regelung ist insoweit durchdacht vom Gesetzgeber, als dass steuerfreie Leistungen und die damit einhergehende Behandlung als Nichtunternehmer nicht wettbewerbsverzerrend sein können, da es überhaupt keinen umsatzsteuerlich bedingten Wettbewerb gibt.²²⁵

Die Ausnahme innerhalb der Norm unter Verweis auf § 9 UStG soll verhindern, dass der jPÖR Wettbewerbsnachteile entstehen.²²⁶ Diese entstehen dann, wenn ein privatrechtlicher Teilnehmer von der Optionsmöglichkeit des § 9 UStG Gebrauch machen kann und damit auch einen Vorsteuerabzug hat und die jPÖR nicht. Diese Wettbewerbsnachteile werden durch die Regelung des § 2b Abs. 2 Nr. 2 UStG vermieden, da den jPÖR die Möglichkeit eines Vorsteuerabzuges offen gehalten wird.²²⁷ Die wichtigste Fallgruppe, der Vermietung und Verpachtung erfolgt in der Regel auf privatrechtlicher und nicht auf öffentlich-rechtlichen Verträgen, sodass die Anwendung von § 2b Abs. 2 Nr. 2 UStG damit nicht greift. Die anderen vermögensverwaltenden Tätigkeiten wie z.B. Grundstückveräußerungen oder Finanzgeschäfte laufen ebenso auf einem privatrechtlichen Vertrag und fallen damit nur unter den allgemeinen Unternehmerbegriff des § 2 Abs. 1 UStG.²²⁸ Die Funktionsweise der Norm lässt sich anhand von zwei Beispielen verdeutlichen (vgl. Tabelle 3).

Im Ergebnis wird die Norm nach ihrem Zweck für rätselhaft befunden und wird kaum für jPÖR Anwendung finden.²²⁹ Dennoch bleibt der tatsächliche praktische Nutzen abzuwarten bis die Norm für den Regelfall der jPÖR Anwendung findet.

Behandlung öffentlich-öffentlicher Zusammenarbeit

²¹³Vgl. BMF, 12.11.2009, IV C 7 - S 2706/08/10004, BStBl. I 2009, S. 1303.

²¹⁴Vgl. Märtens, in: Gosch, 2015, § 4 KStG Rz. 118b; Krämer, in: Dötsch und Achenbach, 2015, KStG, § 4 KStG Rz. 115.

²¹⁵Eigenes Beispiel abgeleitet aus dem Ursprungsbeispiel.

²¹⁶BFH, 04.09.2002, I R 42/01, BFH/NV 2003, S. 511-513.

²¹⁷Vgl. Hüttemann, 2017, S. 133; Sterzinger, 2016a, S. 2.

²¹⁸Vgl. Hammerl und Fietz, 2017, S. 59; Hüttemann, 2017, S. 134.

²¹⁹Vgl. Wiesch, 2016, S. 235.

²²⁰Vgl. Kirch, in: Fritsch et al., 2015, § 2b UStG Rz. 12; Baldauf, 2016, S. 362; Baldauf führt aus, dass dieser Umstand dazu führt, dass jPÖR z.B. 10 nicht gleichartige hoheitliche Tätigkeiten mit Wettbewerbsrelevanz haben können und damit insgesamt ein Umsatz von ca. 175.000 € umsatzsteuerfrei wäre.

²²¹Vgl. Hüttemann, 2017, S. 134; Müller, 2017, S. 12.

²²²Vgl. Meurer, 2016, S. 19; Belcke und Westermann, 2016, S. 88; a.A. Stadie, in: Rau et al., 2018, § 2b Rn. 164; Küffner und Rust, 2016b, S. 1639; Sterzinger, 2015, S. 662; Sterzinger führt an, dass Artikel 13 der MwStSystRl sich auf die Trägerkörperschaft bezieht und die tätigkeitsbezogene Betrachtung der gleichartigen Umsätze dem widerspricht.

²²³Vgl. Heidner, 2016, S. 48.

²²⁴Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 38.

²²⁵Vgl. Wiesch, 2016, S. 226.

²²⁶Vgl. Müller, 2017, S. 11; Bundestag, 23.09.2015, S. 95; EuGH, 04.06.2009, Rs. C-102/08, SALIX Grundstücks-Vermietungsgesellschaft, Slg 2009, S. I-4629.

²²⁷Vgl. Müller, 2017, S. 11; Küffner und Rust, 2016b, S. 1640.

²²⁸Vgl. Stadie, in: Rau et al., 2018, § 2b Rn. 173.

²²⁹Vgl. Stadie, in: Rau et al., 2018, § 2b, Rn 170-173.

Tabelle 3: Beispiele zur Umsatzsteuerbefreiung

Quelle: Eigene Darstellung in Anlehnung an [Kußmaul et al., 2017a](#), S. 327.

Beispiel	Lösungsvorschlag
Eine Kommune verfügt über ein Krankenhaus und behandelt dort Patienten und nimmt im Jahr 500.000 € ein.	Diese Leistungen sind als humanmedizinische Behandlung gem. § 4 Nr. 14b UStG ohnehin steuerfrei und es besteht keine Optionsmöglichkeit nach § 9 UStG.
Die Stadt A vermietet gewerbliche Flächen an lokale Unternehmer.	Diese Tätigkeit gehört zu den Vermietungsumsätzen, die gem. § 4 Nr. 12 UStG steuerbefreit sind, aber gem. § 9 Abs. 1 UStG besteht für diese Tätigkeit eine Optionsmöglichkeit. Damit sind diese Leistungen umsatzsteuerbar, was die Inanspruchnahme von Vorsteuerabzugsbeträgen ermöglicht.

Daneben entstehen keine größeren Wettbewerbsverzerrungen bei der vertikalen und horizontalen Zusammenarbeit²³⁰ von mehreren jPöR gem. § 2b Abs. 3 UStG. Die jPöR sind daneben nichtunternehmerisch tätig, wenn es zu einem Leistungsaustausch zwischen den jPöR kommt und dieser Austausch aufgrund von gesetzlichen Bestimmungen den jPöR vorbehalten ist gem. § 2b Abs. 3 Nr. 1 UStG oder das Ziel die Verfolgung gemeinsamer spezifischer öffentlicher Interessen an andere jPöR gem. § 2b Abs. 3 Nr. 2 UStG sind.²³¹ Die interkommunale Zusammenarbeit soll durch den § 2b UStG nicht tangiert werden, sodass keine Besteuerung der Zusammenarbeit erfolgt, was eine Verteuerung zur Folge hätte, die den Bürger belastet.²³²

§ 2b Abs. 3 Nr. 1 UStG umfasst Vorbehaltsaufgaben der jPöR bei Kooperationsleistungen, die entweder nur diesen vorbehalten und von diesen nachgefragt werden dürfen aufgrund gesetzlicher Bestimmungen oder bei denen der private Wirtschaftsteilnehmer ausgeschlossen ist.²³³ Die jPöR hat im Rahmen dieser Vorschrift eine Doppelrolle, da diese sowohl als Anbieter als auch als Empfänger auftritt, was zu einer fehlenden Marktrelevanz führt und damit zu keinem Wettbewerbsverhältnis gegenüber privaten Wirtschaftsteilnehmern.²³⁴ Diese Norm hat eher einen deklaratorischen Charakter, da ein Wettbewerb unter diesen Bedingungen nicht entstehen kann.²³⁵ Mithilfe dieser Norm soll die klassische hoheitliche Amtshilfe geschützt werden, die eine ergänzende meist kurzfristige Hilfestellung umfasst.²³⁶ Irrelevant ist es, ob die

empfangene Leistung im wirtschaftlichen oder im nichtwirtschaftlichen Bereich verwendet wird.²³⁷ Die Voraussetzung, damit die Vorschrift greift, ist ein Anknüpfen an gesetzliche Bestimmungen. Der Begriff gesetzliche Bestimmungen i.S.d. § 2b Abs. 3 Nr. 1 UStG ist enger definiert als der Begriff der Ausübung öffentlicher Gewalt des § 2b Abs. 1 S. 1 UStG.²³⁸ Unter den Begriff der gesetzlichen Bestimmungen fallen sowohl alle Gesetze und Rechtsverordnungen des Bundes und der Bundesländer sowie das besondere Kirchenrecht, soweit diese Regelungen sich im Geltungsbereich des UStG befinden.²³⁹ Ausgeschlossen von der Norm sind somit mit einer Satzungsautonomie ausgestattete jPöR (z.B. Kommunen, Sparkassen)²⁴⁰, damit diesen die Möglichkeit verwehrt bleibt, Vorbehaltsaufgaben zu schöpfen, indem diese die Satzungsbefugnis dafür missbrauchen.²⁴¹ Für das Vorliegen von § 2b Abs. 3 Nr. 1 UStG reicht es nicht, wenn ein Gesetz die Zusammenarbeit zwischen jPöR empfiehlt, sondern es muss die dabei erfüllende Aufgabe zuweisen.²⁴² Im Einzelfall wird es schwierig, wenn ein Gesetz zwar die Zusammenarbeit von zwei jPöR vorsieht, aber nicht regelt, ob die Leistung alternativ durch einen privaten Wirtschaftsteilnehmer erbracht werden darf.²⁴³ Durch die Subsumtion aus nicht eindeutigen Gesetzesformulierungen kann grundsätzlich ein faktischer Kooperationszwang abgeleitet werden durch die allgemeine Regelung des Wortes insbesondere, sodass auch wenn eine Vorschrift die Voraussetzungen des § 2b Abs. 3 Nr. 1 UStG nicht komplett erfüllt ist, dies zu einem Ausschluss von größeren Wettbewerbsverzerrungen führen.²⁴⁴ Diese Vorschrift

²³⁰Unter vertikaler Kooperation wird eine Kooperation über mehrere Gliederungsebenen hinweg definiert, z.B. Kommune kooperiert mit Zweckverband und unter horizontaler Kooperation wird die Kooperation auf einer Gliederungsebene, z.B. Kommune kooperiert mit Kommune, verstanden.

²³¹Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 40.

²³²Vgl. Bundestag, 23.09.2015, S. 91.

²³³Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 41.

²³⁴Vgl. Sauerland, 2017, S. 115; Müller, 2017, S. 12.

²³⁵Vgl. Müller, 2017, S. 12; Küffner und Rust, 2016b, S. 1641.

²³⁶Vgl. Kirch, in: Fritsch et al., 2015, § 2b UStG Rz. 20.

²³⁷Vgl. Sterzinger, 2017, S. 235.

²³⁸Vgl. Kußmaul et al., 2017a, S. 328.

²³⁹Vgl. Hammerl und Fietz, 2017, S. 60; ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 42.

²⁴⁰Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 42.

²⁴¹Vgl. Erdbrügger und Liegmann, 2018, S. 65.

²⁴²Vgl. Küffner und Rust, 2016b, S. 1641.

²⁴³Vgl. Küffner und Rust, 2016b, S. 1641.

²⁴⁴Vgl. Küffner und Rust, 2016b, S. 1637; a.A. Sterzinger, 2017, S. 236; Sterzinger wendet ein, dass bei einer Abweichung von dem Normwortlaut die Zuordnung nicht möglich sein soll.

ist praxiskonform, da diese weitgehend unproblematisch anwendbar ist und daneben kein Konfliktpotential mit dem Unionsrecht besteht.²⁴⁵ Das einzige Problem sind Fälle mit Leistungen, in denen unklar ist, ob ein privater Wirtschaftsteilnehmer die durch Gesetze bestimmte Zusammenarbeit von jPöR auf dem Gebiet der Vorbehaltsaufgaben alternativ durch sein Leistungsangebot erbringen kann. Dies gilt es noch zu klären.

Wenn die Tätigkeit zwar nicht alleine den jPöR vorbehalten sind, aber eine Zusammenarbeit zwischen jPöR durch ein spezifisches öffentliches Interesse bestimmt wird gem. § 2b Abs. 3 Nr. 2 UStG, führt dies ebenfalls nicht zu einer Verzerrung des Wettbewerbs. Dieser unbestimmte Rechtsbegriff hat seinen Ursprung in der EU-Vergaberichtlinie Artikel 12 Abs. 4 Buchst. b und ist dieser entnommen worden.²⁴⁶ Ob die Anlehnung an die Vergaberichtlinie überhaupt mit der Regelung des Artikels 13 MwStSystRL kompatibel ist, wird sowohl von dem EuGH als auch von der Literaturmeinung bezweifelt.²⁴⁷ Daneben kann die Begünstigung von Beistandsleistungen von jPöR außerdem ein Verstoß gegen das europaweit fixierte Beihilfeverbot sein.²⁴⁸ Damit es zum einem spezifisch öffentlichen Interesse kommt, müssen die nachfolgenden Bedingungen²⁴⁹ des § 2b Abs. 3 Nr. 2 UStG kumulativ vorliegen²⁵⁰:

- Langfristige öffentlich-rechtliche Vereinbarung (§ 2b Abs. 3 Nr. 2 S. 2 Buchst. a UStG)
- Erhalt der öffentlichen Infrastruktur (§ 2b Abs. 3 Nr. 2 S. 2 Buchst. b UStG)
- Wahrnehmung aller Beteiligten obliegende öffentliche Aufgabe (§ 2b Abs. 3 Nr. 2 S. 2 Buchst. b UStG)
- Gegenleistung besteht ausschließlich auf einer Kostenersatzung (§ 2b Abs. 3 Nr. 2 S. 2 Buchst. c UStG)
- Leistung wird im Wesentlichen an jPöR erbracht (§ 2b Abs. 3 Nr. 2 S. 2 Buchst. d UStG)

Die Norm des § 2b Abs. 3 Nr. 2 UStG führt auch bei grammatikalischer Auslegung²⁵¹ zu dem Ergebnis, dass die Voraussetzungen nur in kumulativer Weise vorliegen können.²⁵² Dies widerspricht der Vorgabe in der MwStSystRL, da die rechtliche Beurteilung der Tätigkeit maßgeblich von der Einrichtung der jPöR abhängt und nicht von deren Wettbewerbsrelevanz.²⁵³ Dieser Auffassung wird unterstützt, da nach §

2b Abs. 3 Nr. 2 S. 2 UStG durch das Wort regelmäßig, die Voraussetzungen der Buchstaben a bis d nicht immer komplett erfüllt sein müssen, damit eine Zusammenarbeit vorliegt.²⁵⁴ Die Anwenderpraxis sollte daher darauf achten, dass die Zusammenarbeit durch Gründe bestimmt wird, die nur den jPöR eigen sind.²⁵⁵

Die Bedingung der öffentlich-rechtlichen Vereinbarung besteht aus dem Bestandteil Langfristigkeit und öffentlich-rechtliche Vereinbarung. Die öffentlich-rechtlichen Vereinbarungen sind dabei insbesondere der öffentlich-rechtliche Vertrag, Verwaltungsabkommen und – Vereinbarungen sowie Staatsverträge.²⁵⁶ Hierbei wird deutlich, dass auch die Aufgabenübernahme aufgrund einer Satzung nicht unter die Norm fällt.²⁵⁷ Für das Kriterium der Langfristigkeit hat der Gesetzgeber keine Definition gefunden²⁵⁸, aber die Verwaltung geht davon aus, dass dies immer erfüllt sei, wenn eine unbefristete Vereinbarung getroffen worden ist. Diese Betrachtung wird kritisiert und es wird empfohlen auch auf den Einzelfall abzustellen.²⁵⁹ Bei befristeten Verträgen liegt eine Langfristigkeit bei einem mindestens fünfjährigen Zeitraum vor.²⁶⁰ Die Fünfjahresfrist ist eine Schöpfung im Rahmen des § 2b Abs. 3 Nr. 2 S. 2 Buchst. b UStG, da diese weder im Vergaberecht noch im Umsatzsteuerrecht (z.B. § 29 UStG mit einer Dauer von vier Monaten²⁶¹) im Rahmen der Wettbewerbs- und Steuerwürdigung einer Tätigkeit bzw. eines Umsatzes zu finden sind.²⁶² Es gibt innerhalb der Literatur Vorstellungen, dass die Langfristigkeit bereits bei einer nachhaltigen Tätigkeit vorliegt und da dies zu einer der Voraussetzungen von § 2 Abs. 1 UStG zählt, liegt diese immer vor.²⁶³ Als Kompromisslösung kann eine Anlehnung an die vergaberechtliche Rechtsprechung des EuGH vorgesehen werden, da dort eine vertragliche Höchstdauer von 16 Monaten bereits als langfristig angesehen wird.²⁶⁴ Dennoch können kürzere Zeiträume möglich sein, wenn dies bei der Art der ausgeübten Tätigkeit üblich ist.²⁶⁵ Die Üblichkeit kann als Kriterium nur schwer objektiv zu beurteilen sein und wird deshalb Streitpotential auslösen.²⁶⁶ Dabei muss die Entscheidung ex ante getroffen werden, damit bereits am Anfang des möglichen Besteuerungszeitraumes klar ist,

²⁴⁵Vgl. Küffner und Rust, 2016b, S. 1641.

²⁴⁶Vgl. Baldauf, 2016, S. 355; Müller, 2017, S. 13.

²⁴⁷Vgl. Hüttemann, 2017, S. 135; Gohlke und Schmitz, 2016, S. 782; Widmann, 2015a, S. 887; EuGH, 29.10.2015, Rs. C-174/14, Sudaçor, UR 2015, S. 901-909.

²⁴⁸Vgl. Hüttemann, 2017, S. 138-139.

²⁴⁹Vgl. Walkenhorst, UStB 2017, S. 20.

²⁵⁰Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 45.

²⁵¹Dies liegt daran, dass in § 2b Abs. 3 Nr. 2 S. 2 UStG von der Pluralform Leistungen und Vereinbarungen gesprochen wird, aber § 2b Abs. 3 S. 2 UStG nur im Singular von einer Leistung und in § 2b Abs. 3 S. 1 UStG von der Zusammenarbeit; Vgl. Müller, 2017, S. 15.

²⁵²Vgl. Müller, 2017, S. 15.

²⁵³Vgl. Müller, 2017, S. 15.

²⁵⁴Vgl. Küffner und Rust, 2016b, S. 1645.

²⁵⁵Vgl. Küffner und Rust, 2016b, S. 1645.

²⁵⁶Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 46; Das BMF-Schreiben verweist auf die Definitionen im Bereich des öffentlichen Rechts mit Ausnahme der öffentlich-rechtlichen Satzungen.

²⁵⁷Vgl. Küffner und Rust, 2016b, S. 1643; a.A. Hammerl und Fietz, 2017, Rn. 60; Diese gehen davon aus, dass die öffentlich-rechtliche Sonderregelung und die öffentlich-rechtliche Vereinbarung gleich zu interpretieren sind.

²⁵⁸Vgl. Scharpenberg, 2017, S. 837.

²⁵⁹Vgl. Scharpenberg, 2017, S. 837.

²⁶⁰Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 47.

²⁶¹Vgl. Küffner und Rust, 2016b, S. 1643.

²⁶²Vgl. Hidien und Schwarz, 2017, S. 349.

²⁶³Vgl. Maier und Übleiß, 2015, S. 209; Weber, 2015, S. 367; a.A. Meurer, 2016, S. 20; Meurer wendet ein, dass eine nachhaltige Tätigkeit auch aus mehreren kurzfristigen Tätigkeiten zusammengesetzt sein kann.

²⁶⁴EuGH, 19.12.2012, Rs. C-159/11, Ordine degli Ingegneri della Provincia di Lecce, NVwZ 2013, S. 710-712.

²⁶⁵Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 47.

²⁶⁶Vgl. Scharpenberg, 2017, S. 837; Sauerland, 2017, S. 115.

ob eine Besteuerung erfolgt oder eben nicht.²⁶⁷ Insgesamt bleibt bei dieser Norm offen, wie mit der vorzeitigen Kündigung langfristiger Vereinbarungen oder als Gegenstück dazu, die Verlängerung von kurzfristigen Vereinbarungen verfahren wird.²⁶⁸ Damit es nicht zu einem Streitpotential vor den Gerichten kommt, zwischen den jPöR und privatwirtschaftlichen Unternehmen oder der Finanzverwaltung über das Kriterium der Langfristigkeit, soll in Kooperationsverträgen ein Zeitraum von fünf Jahren festgehalten werden.²⁶⁹

Das zweite Kriterium des Erhalts der öffentlichen Infrastruktur gem. § 2b Abs. 3 Nr. 2 S. 2 Buchst. b UStG wird weder im Gesetz noch in der Gesetzesbegründung näher erläutert. Durch das BMF-Schreiben fallen alle Einrichtungen darunter, materieller/technischer/digitaler (z.B. Verkehrswegenetz, Entsorgung von Wasser, auf Kirchengrundstücken befindliche Pfarrgebäude und Gemeindehäuser), immaterieller/sozialer (z.B. Bildungswesen, innere Sicherheit, öffentlich-rechtlicher Rundfunk, Verkündigung, Seelsorge) und institutioneller Art (z.B. Rechtsordnung, Wirtschaftsordnung, Sozialordnung), die für die Ausübung der öffentlichen Gewalt i.S.d. § 2b Abs. 1 UStG notwendig sind.²⁷⁰ Neben dieser Definition fallen auch die Förderung, der Ausbau sowie die Errichtung unter den Begriff des Erhalts der öffentlichen Infrastruktur²⁷¹, sodass die Infrastruktur auch durch Dienstleistungen erhalten werden kann.²⁷² Die Bedeutung des Wortes Infrastruktur ist hingegen sehr unkonkret, dennoch bietet dies auch Vorteile. Dadurch ist es möglich, dass das Wort sehr weit ausgelegt werden kann und es bei Kooperationen der öffentlichen Hand kaum Einschränkungen bei der Erhaltung der öffentlichen Infrastruktur gibt.²⁷³ Eine weitere Voraussetzung gem. § 2b Abs. 3 Nr. 2 S. 2 Buchst. b ist die Wahrnehmung einer allen Beteiligten obliegende öffentliche Aufgabe. Darunter ist laut BMF eine Zusammenarbeit mehrerer jPöR zu verstehen, um das gemeinsame Ziel der einfachen oder mehrfachen Aufgabenerfüllung im Interesse der Allgemeinheit zu erfüllen. Dies liegt sogar dann vor, wenn die Aufgabe im Ganzen auf eine andere leistende jPöR übertragen wird oder wenn eine Zusammenarbeit auf horizontaler oder vertikaler Ebene erfolgt und sogar, wenn die rechtliche Zuständigkeit wechselt.²⁷⁴ Dabei muss die entsprechende Zuweisung aus einem Gesetz hervorgehen, das eine abstrakte Aufgabenzuweisung zuordnet, da eine konkrete Zuordnung zu einer jPöR die Zusammenarbeit und damit auch die Anwendung der Norm verhindert.²⁷⁵ Lediglich bei den verwaltungsunterstützenden

Hilfstätigkeiten (z.B. Gebäudereinigung) wird die Vorschrift nicht erfüllt.²⁷⁶ Die Geltung dieser Regelung soll anhand des Praxisbeispiels aus der Tabelle 4 verdeutlicht werden.

Die beiden Voraussetzungen des § 2b Abs. 3 Nr. 2 S. 2 Buchst. b sind nach aktuellem Rechtsstand noch sehr unklar im Hinblick auf den tatsächlichen Inhalt. Dadurch besteht ein großer Auslegungsspielraum, der für die Anwenderpraxis mit einer Rechtsunsicherheit behaftet ist.²⁷⁷

Die vierte Voraussetzung ist die Komponente der Kostenerstattung gem. § 2b Abs. 3 Nr. 2 S. 2 Buchst. c UStG. Die Gesetzesbegründung enthält weder eine Erläuterung zu dem Begriff der Kostenerstattung noch dazu, wie diese zu berechnen ist.²⁷⁸ Dabei dürfen zur Zusammenarbeit von jPöR nur kostendeckende Entgelte berechnet werden.²⁷⁹ Der Hintergrund dieser Regelung ist der, dass eine Gewinnerzielungsabsicht ein typisches Merkmal für privatwirtschaftliche Unternehmen ist und damit vermieden werden soll, dass die jPöR über das kostendeckende Entgelt hinaus Einnahmen erzielen und trotzdem nicht der Besteuerung unterliegen.²⁸⁰ Denn der Sinn bei einer Kooperation liegt darin, Synergien zu schaffen, um dadurch Kosten zu sparen.²⁸¹ Bezüglich der Kostenberechnung lässt die Finanzverwaltung jede sachgerechte Kostenermittlung zu, sodass damit sowohl die variablen, als auch die fixen Kosten abgedeckt sind, als auch Pauschalkostensätze. Ausgeschlossen sind hingegen die Berücksichtigung von Rücklagen und die Verzinsung von Eigenkapital.²⁸² Dies wirft Probleme auf, wenn eine Verzinsung gesetzlich oder in der Satzung fixiert ist.²⁸³ Damit es zu keinen Problemen bei der Kostenberechnung kommt, empfiehlt es sich, dass die Kostenbestandteile in der Zusammenarbeitsvereinbarung festgehalten werden.²⁸⁴ Dennoch kann es auch zu einer Kostenunterdeckung kommen, aber dies ist hinzunehmen. Das hat den Hintergrund, dass eine exakte Kostenerstattung praktisch kaum umsetzbar ist.²⁸⁵ Daher ist die Voraussetzung der Norm immer erfüllt, wenn keine Gewinne erzielt werden.²⁸⁶

Die letzte und fünfte Voraussetzung liegt darin, dass die leistende gleichartige Leistung im Wesentlichen an eine oder mehrere andere jPöR gem. § 2b Abs. 3 Nr. 2 S. 2 Buchst. d UStG erbringt. Dabei müssen zwei Bedingungen erfüllt sein, zum einen die Bedingung der gleichartigen Leistungen und zum anderen das Wesentlichkeitsmerkmal. Das Kriterium der gleichartigen Leistungen orientiert sich an der Sicht des

²⁶⁷Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 47.

²⁶⁸Vgl. Müller, 2017, S. 13.

²⁶⁹Vgl. Strahl, 2017, S. 20299; Hammerl und Fietz, 2017, S. 61; a.A. Müller, 2017, S. 13; Müller plädiert dafür, dass Verträge eine unbefristete Laufzeit haben sollen.

²⁷⁰Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 48; Einordnung der kirchlichen jPöR in das Schema gemäß der Dreiteilung; Burret, NWB 2017, S. 410; Kritische Sicht: Müller, 2017, S. 14; Küffner und Rust, 2016b, S. 1633; Die Autoren halten das Merkmal für sehr auslegungsbedürftig.

²⁷¹Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 48.

²⁷²Vgl. Hammerl und Fietz, 2017, S. 61.

²⁷³Vgl. Hammerl und Fietz, 2017, S. 61.

²⁷⁴Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 49.

²⁷⁵Vgl. Küffner und Rust, 2016b, S. 1645.

²⁷⁶Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 49.

²⁷⁷Vgl. Hammerl und Fietz, 2017, S. 61; Müller, 2017, S. 14.

²⁷⁸Vgl. Küffner und Rust, 2016b, S. 1645.

²⁷⁹Vgl. Weber, 2015, S. 367.

²⁸⁰Vgl. § 2 Abs. 1 Satz 3 UStG; Das Gewinnerzielungsmerkmal ist keine Voraussetzung für eine umsatzsteuerliche Unternehmereigenschaft; Gohlke und Schmitz, 2016, S. 785; Sterzinger, 2015, S. 660; Widmann, 2015b, S. 10-11.

²⁸¹Vgl. Gohlke und Schmitz, 2016, S. 785.

²⁸²Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 51.

²⁸³Vgl. Sterzinger, 2017; S. 237.

²⁸⁴Vgl. Hammerl und Fietz, 2017, S. 61; Küffner und Rust, 2016b, S. 1645.

²⁸⁵Vgl. Müller, 2017, S. 14.

²⁸⁶Vgl. Müller, 2017, S. 14.

Tabelle 4: Beispiele zur Abgrenzung Aufgabenübernahme

Quelle: Eigene Darstellung in Anlehnung an ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 50.

Aufgabe	Lösung
Die Kommune A betreibt ein Schwimmbad. Diese möchte zwar die Leistungen der Bevölkerung anbieten, aber möchte den Betrieb des Schwimmbades lieber der Kommune B überlassen.	Die Voraussetzung des § 2b Abs. 3 Nr. 2 S. 2 Buchst. b HS. 2 UStG ist erfüllt, da die Aufgabe im Ganzen von der anderen Kommune B übernommen wird.
Abwandlung: Die Kommune A bedient sich der Reinigungskräfte der Kommune B.	Die Voraussetzungen sind nicht erfüllt, da es sich nur um eine unterstützende Hilfstätigkeit handelt.

Durchschnittsverbrauchers gem. § 2b Abs. 2 Nr. 1 UStG.²⁸⁷ Der Begriff der Wesentlichkeit findet keine Erläuterung im Gesetz. Das BMF-Schreiben wendet die Regelung in Artikel 12 Abs. 4 der Vergaberichtlinie²⁸⁸ sinngemäß an und die Voraussetzung ist erfüllt, wenn mehr als 80 % der Leistungen an andere jPÖR erbracht werden und nicht bloß einen Innenumsatz darstellen.²⁸⁹ Dabei wird der durchschnittliche Gesamtumsatz der gleichartigen Tätigkeit aus den letzten drei Jahren ermittelt.²⁹⁰ Eine Klarstellung, dass bei einer Neuaufnahme einer Tätigkeit auf die geplanten Umsätze abzustellen ist wird gefordert.²⁹¹ Die Norm des § 2b Abs. 3 Nr. 2 S. 2 Buchst. d UStG erzeugt eine wettbewerbsstörende Isolierung der jPÖR gegenüber privatrechtlich agierenden Unternehmen, da jPÖR ein Anreiz gesetzt wird, nur untereinander zu agieren, sodass ein Preisvorteil ermöglicht wird und Markteintritte aus der Privatwirtschaft erschwert bzw. verhindert werden.²⁹²

Der § 2b Abs. 3 UStG gilt nur für jPÖR, wenn diese für die Kooperation öffentliche Gewalt ausüben, wobei in Absatz 3 dieses Merkmal nicht näher definiert ist.²⁹³ Damit fällt die Beistandsleistung des § 2b Abs. 3 Nr. 2 UStG nicht unter den Begriff der öffentlichen Gewalt und könnte keine Anwendung finden²⁹⁴, wobei der Gesetzgeber auch die Beistandsleistung der öffentlichen Gewalt hätte zuordnen können.²⁹⁵

²⁸⁷Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 53; siehe Kapitel D.II.3.b) Jährliche Umsatzgrenze.

²⁸⁸Vgl. Scharpenberg, 2017, S. 838.

²⁸⁹Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 54; Dies bedeutet, dass die monetäre Höhe der Umsätze entscheidend ist und nicht die Leistungseinheiten; a.A. Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rn. 195; FG Berlin-Brandenburg, Urteil vom 13.4.2016, 7 K 7160/13, EFG 2016, S. 1032-1036; Das Abstellen auf Leistungseinheiten scheint sinnvoll zu sein, da das FG Berlin-Brandenburg Leistungseinheiten ohne Leerstand für den Vorsteuerabzug zu Grunde legt. Dennoch gibt es auch Fälle, bei denen Leerstandszeiten in den Aufteilungsschlüssel einzubeziehen sind, wenn zu einer Zeit nur eine Nutzungsart möglich ist; Vgl. Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rn. 201.

²⁹⁰Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 54.

²⁹¹Vgl. Müller, 2017, S. 15.

²⁹²Vgl. Müller, 2017, S. 14; Widmann, 2015b, S. 11; Abschließendes Beispiel dieser Regelung in Anhang 8.

²⁹³Vgl. Müller, 2017, S. 15.

²⁹⁴Vgl. Sterzinger, 2015, S. 664.

²⁹⁵Vgl. Kuffner und Rust, 2014, S. 2535-2536; Sterzinger, 2015, S. 664.

Bezüglich der offenen Rechtsfrage wird es in Zukunft Gerichtsurteile geben.²⁹⁶ Insgesamt scheint die Norm des § 2b Abs. 3 Nr. 2 UStG noch viele offene Rechtsfragen zu besitzen, sodass eine Einordnung mithilfe eines Kooperationsvertrages unter § 2b Abs. 3 Nr. 1 UStG erfolgen sollte.²⁹⁷

4.2.4. Unternehmereigenschaft kraft Gesetzes

Die vierte Voraussetzung, um den § 2b UStG anwenden zu können, ist das Nichtvorliegen einer Katalogtätigkeit i.S.d. § 2b Abs. 4 UStG. Die Regelung des § 2b Abs. 4 Nr. 1 bis 4 UStG wurde aus der alten Vorschrift des § 2 Abs. 3 S. 2 UStG i.d.F. vom 31.12.2015 übernommen, sodass auch die Regelungen des Abschn. 2.11 Abs. 7 bis 11 UStAE laut BMF-Schreiben weiterhin angewendet werden.²⁹⁸ Die Katalogtätigkeiten i.S.d. § 2b Abs. 4 UStG sind stets unternehmerisch, sofern diese die Voraussetzungen des § 2 Abs. 1 UStG, also Selbständigkeit, Nachhaltigkeit und Einnahmeerzielungsabsicht erfüllen, obwohl eine der Ausnahmeregelungen des § 2b UStG vorliegt.²⁹⁹ Der Sinn dieser Vorschrift besteht darin, dass auch hoheitliche Tätigkeiten unabhängig davon, ob eine Wettbewerbssituation auftaucht, wegen der abstrakten Wettbewerbsrelevanz der Vorschrift immer das Wettbewerbskriterium erfüllen und damit als steuerpflichtig behandelt werden.³⁰⁰ Im Rahmen des Gesetzgebungsverfahrens kam es nicht zu einer Anpassung der Vorschrift, sodass auch viele veraltete Tätigkeiten in die Vorschrift mit aufgenommen worden sind.³⁰¹

Lediglich bei der neu eingeführten Nr. 5 dieser Vorschrift wird mithilfe eines dynamischen Verweises, also der jeweils aktuellen Regelung, auf den Anhang I der MwStSystRl verwiesen. Die beiden Bedingungen, damit § 2 Abs. 4 Nr. 5 UStG greift, sind gemäß BMF-Schreiben zum einen, dass der Umfang der Tätigkeit bedeutend ist, was bei einem Jahresumsatz

²⁹⁶Vgl. Müller, 2017, S. 15.

²⁹⁷Vgl. Strahl, 2017, S. 20300; Abschließendes Beispiel zum Kapitel siehe Anhang 9.

²⁹⁸Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 56.

²⁹⁹Vgl. Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rn. 207; Korn, in: Bunjes et al., 2018, § 2 UStG Rn. 51.

³⁰⁰Vgl. Ismer/Baur-Rückert, MwStR 2016, S. 745.

³⁰¹Vgl. Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rn. 207; Baldauf, 2016, S. 367; Es wird auf die Regelung im Kapitel C.I.3.b) verwiesen.

höher als 17.500 € der Fall ist und zum anderen, dass die übrigen Voraussetzungen des § 2 Abs. 1 UStG erfüllt sind.³⁰² Die Umsatzgrenze des § 2b Abs. 4 Nr. 5 orientiert sich dabei an § 2b Abs. 2 Nr. 1 UStG.³⁰³ Zu den in Anhang 1 genannten Tätigkeiten zählen insgesamt 13. Die 13 aufgezählten Tätigkeiten sind bereits nach der alten Rechtslage als unternehmerisch zu behandeln³⁰⁴, sodass damit keine Veränderung der Besteuerungspraxis einhergeht.³⁰⁵

4.2.5. Zwischenfazit

Es hat sich in dem Kapitel III gezeigt, dass viele Kriterien noch mit einer Unsicherheit behaftet sind³⁰⁶, sodass ein Prüfungsschema die umsatzsteuerliche Einordnung der Unternehmereigenschaft unterstützt. Die einzelnen Schritte werden im Folgenden näher betrachtet und erläutert. Prüfungsschemata zu dem § 2b UStG gibt es bereits³⁰⁷, aber das vorliegende Modell vereint die Eigenschaften der Modelle zu einem neuen Schema. Damit § 2b UStG überhaupt greift, muss die allgemeine Unternehmereigenschaft vorliegen. Dazu ist es erforderlich, dass sowohl die Fragen nach der Selbstständigkeit, Nachhaltigkeit, sowie der Einnahmeerzielungsabsicht positiv beantwortet werden. Wenn dies vorliegt, kann die Ausnahmenorm des § 2b UStG geprüft werden, ansonsten liegt bei einer Verneinung der Voraussetzung keine Unternehmereigenschaft vor. Anschließend wird die Katalogtätigkeit gem. § 2b Abs. 4 UStG geprüft, da bei Vorliegen dieser Vorschrift die weitere Prüfung des § 2b UStG entbehrlich ist.

Danach erfolgt die Prüfung, ob ein Handeln im Rahmen öffentlicher Gewalt festzustellen ist. Dazu müssen sowohl das Handeln auf öffentlich-rechtlicher Grundlage, als auch die öffentlich-rechtliche Handlungsform vorliegen. Wenn eine der beiden Voraussetzungen nicht erfüllt ist, dann liegt eine Tätigkeit auf privatrechtlicher Natur vor und ist damit steuerbar. Bei der Überprüfung der Wettbewerbsrelevanz der jPÖR im Rahmen der öffentlichen Gewalt sind zunächst die

beiden Soderbestände des § 2b Abs. 2 und 3 UStG zu prüfen, die nebeneinander geprüft werden müssen, da diese sich nicht ausschließen.³⁰⁸ Aus Anwendersicht sollte zuerst die eindeutigere Regelung des § 2b Abs. 2 Nr. 2 UStG geprüft werden³⁰⁹ und im nächsten Schritt, ob die Umsätze aus gleichartigen Tätigkeiten 17.500 € übersteigen. Wenn die Voraussetzungen des § 2b Abs. 2 UStG nicht vorliegen, dann ist ein Handeln in Zusammenarbeit mit anderen jPÖR zu prüfen. Bei Nichteinhaltung der Voraussetzung des § 2b Abs. 3 Nr. 1 UStG muss geprüft werden, ob die Zusammenarbeit durch gemeinsame spezifische öffentliche Interessen bestimmt sind. Dies liegt i.d.R. dann vor, wenn die fünf Voraussetzungen des § 2b Abs. 2 Nr. 2 a bis d UStG vorliegen. Dadurch, dass die Aufzählungen zur Vermeidung eines Wettbewerbs in § 2b Abs. 2 und 3 UStG nicht abschließend sind, kann eine Wettbewerbsverzerrung auch aus anderen Gründen verneint werden. Dazu ist als letzten Punkt das Vorliegen eines potentiellen Wettbewerbs gem. § 2b Abs. 1 S. 2 zu prüfen.³¹⁰ Dabei besteht die Unternehmereigenschaft dann, wenn die Tätigkeit der Art nach wettbewerbsrelevant ist und es keine lokalen Annahme - oder Abnahmeverpflichtungen gibt, die nicht umgangen werden können. Daneben dürfen auch keine rechtlichen Rahmenbedingungen bestehen, die leistungsbeeinflussend sind. Das Prüfungsschema befindet sich in Anhang 11.

4.3. Wechsel zur neuen Rechtslage

4.3.1. Anwendung der Optionserklärung

Durch die vom Gesetzgeber geschaffene Übergangsregelung des § 27 Abs. 22 UStG haben die jPÖR die Möglichkeit einen großen Übergangszeitraum in Anspruch zu nehmen. Dies ist auch notwendig, damit die mit der Neuregelung verbundenen Umstellungen rechtzeitig implementiert werden.³¹¹ Grundsätzlich wurde § 2b UStG zu Anfang 2016 in das Rechtssystem der Umsatzsteuer aufgenommen, wobei die neue Gesetzeslage gem. § 27 Abs. 22 S. 2 UStG erst ab dem 01.01.2017 angewendet werden muss. In der Übergangszeit zwischen Anfang 2016 und dem 31.12.2016 war die alte Rechtslage des § 2 Abs. 3 UStG a.F. anzuwenden.³¹²

Es änderte sich für jPÖR im Jahr 2016 zwar nichts an der Rechtslage³¹³, dennoch mussten diese bis zum 31.12.2016

³⁰²Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 57.

³⁰³Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 57.

³⁰⁴Dies sind bspw. Telekommunikation, Lieferung von Wasser, Gas, Elektrizität und thermischer Energie, den Hafen - und Flughafendienstleistungen, der Veranstaltung von Messen und Ausstellungen mit gewerblichen Charakter und die Umsätze von betriebseigenen Kantinen, Verkaufsstellen und Genossenschaften und ähnlichen Einrichtungen und die Leistungen der Rundfunkanstalten, soweit diese keiner Steuerbefreiung nach Artikel 132 Abs. 1 q MwStSyRl unterliegen. Die Voraussetzungen des § Artikel 132 Abs. 1 q MwStSyRl stellt öffentlich-rechtliche Rundfunkanstalten von der Steuer frei, sofern diese als nichtgewerblich eingestuft werden. In Deutschland haben diese keine Gewinnorientierung, sodass diese als nichtgewerblich angesehen werden und auch nach der Rechtslage des § 2b UStG von der Steuer freigestellt werden; Vgl. Stadie, in: Rau et al., 2018, § 2b, Rn. 319- 320.

³⁰⁵Vgl. Korn, in: Bunjes et al., 2018, § 2 UStG Rn. 53; Meurer, 2016, S. 21.

³⁰⁶Siehe Anhang 10.

³⁰⁷Vgl. Baldauf, 2016, S. 368; Dieses Prüfungsschema ist zum einen lückenhaft und zum anderen werden die Katalogtätigkeiten erst am Ende des Prüfungsschemas behandelt; Küffner und Rust, 2016b, S. 1634; Der Ablauf des Prüfungschemas ist zwar von der Reihenfolge sinnvoll und korrekt, aber die einzelnen Gliederungspunkte sind nicht ausführlich genug behandelt; Maier und Übleiß, 2016, S. 332; Diese Prüfungsschemaübersicht stellt weder eine Rangfolge der Prüfungskriterien dar, noch knüpft er als Grundvoraussetzung an die allgemeine Unternehmerschaft gem. § 2 Abs. 1 UStG an.

³⁰⁸Vgl. Hammerl und Fietz, 2017, S. 59; a.A. Müller, 2017, S.14; Widmann, 2015b, S. 9; Sterzinger, 2015, S. 664; Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rn. 154; Die Autoren plädieren dafür, dass § 2b Abs. 3 lex specialis zu § 2b Abs. 2 UStG ist, sodass zuerst die Vorschrift des Abs. 3 geprüft werden muss. Diese Ansicht verkennt, dass damit der Neutralitätsgrundsatz gebrochen wird; Vgl. Müller UR 2017, S.15; Widmann, 2015b, S. 9; Sterzinger, 2015, S. 664. Diese Frage wird auch nicht von der Finanzverwaltung geklärt, sodass diese Frage gerichtlich entschieden werden muss; Vgl. Hammerl und Fietz, 2017, S. 59; ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 30-55.

³⁰⁹Vgl. Müller, 2017, S. 11; Ringwald, 2015, S. 3.

³¹⁰Vgl. Korn, in: Bunjes et al., 2018, § 2 UStG Rn. 36, 43; Hammerl und Fietz, 2017, S. 59; Der Grundsatz „lex specialis derogat legi generali“ findet Anwendung, der besagt, dass eine speziellere Norm die allgemeinere Regelung verdrängt.

³¹¹Vgl. Meurer, 2017, S. 164; Maier und Übleiß, 2016, S. 331.

³¹²Vgl. BMF, III C 2 - S 7106/07/10012-06, UR 2016, S. 410.

³¹³Vgl. BMF, III C 2 - S 7106/07/10012-06, UR 2016, S. 410.

mit einer Optionserklärung gem. § 27 Abs. 22 S. 3 UStG gegenüber dem Finanzamt angeben, ob sie den Übergangszeitraum in Anspruch nehmen. Es handelt sich bei der Optionserklärung um eine nicht verlängerbare Ausschlussfrist.³¹⁴ Wenn eine jPöR ab dem Jahr 2017 neu entsteht, gilt für diese mit sofortiger Wirkung die Rechtslage des § 2b UStG.³¹⁵ Dabei muss die jPöR die Optionserklärung für sämtliche ausgeübte Tätigkeiten abgeben, sodass eine einzelne Organisationseinheit oder Einrichtung der jPöR dies für ihren Bereich nicht separat tun darf.³¹⁶ Diese Erklärung gibt der gesetzliche Vertreter oder ein Bevollmächtigter an das gem. § 21 AO zuständige Finanzamt ab, von dessen Bezirk der Unternehmer sein Unternehmen überwiegend betreibt. Da es sich bei der Optionserklärung um eine einseitig empfangsbedürftige Willenserklärung handelt, ist keine Antwort der Finanzverwaltung erforderlich.³¹⁷ Zwar sieht das Umsatzsteuergesetz für die Optionserklärung keine besondere Abgabeform vor, dennoch empfiehlt es sich dies in Schriftform zu erledigen.³¹⁸

Bei der Besteuerung von den Bundesländern sowie des Bundes ergibt sich eine andere Situation. Grundsätzlich werden die Umsätze abweichend von § 21 AO dezentral erfasst, sodass jede Organisationseinheit, die Umsätze bei dem für den Bezirk zuständigen Finanzamt deklariert.³¹⁹ Diese dezentrale Umsatzbesteuerung gilt auch noch bis Ende 2020 fort.³²⁰ Die Abgabe der Optionserklärung gem. § 27 Abs. 22 UStG kann bei den Bundesländern und beim Bund die Regierung oder ein anderes Verfassungsorgan durch einen Beschluss fassen³²¹, wobei dies für sämtliche von der Gebietskörperschaft Bundesland oder Bund ausgeübte entgeltliche Tätigkeiten durchzuführen ist.³²² Die einzelnen unselbständigen Organisationseinheiten (z.B. Behörde, Dienststelle) sind an das Votum des Bundes und Landes gebunden und der Bund oder die Bundesländer müssen sich entweder für die alte oder die neue Rechtslage entscheiden.³²³ Zwar ist die Optionserklärung auch durch den gesetzlichen Vertreter oder einen Bevollmächtigten an das durch § 21 AO benannte zuständige Finanzamt zu richten, fraglich bleibt aber, ob der Antrag durch den Bund oder ein Bundesland oder den gesetzlichen Vertreter der selbständigen Behörde zu stellen ist bzw. wie dies aussieht bei einer Organisationseinheit einer jPöR, die durch eine andere jPöR verwaltet wird.³²⁴ Da die Bestim-

mung des zuständigen Finanzamtes aufgrund der dezentralen Umsatzbesteuerung Schwierigkeiten bereitet, empfiehlt es sich, gem. § 27 AO auf den maßgebenden Verwaltungssitz (z.B. Sitz der Landesregierung) der Körperschaft abzustellen.³²⁵

4.3.2. Folgen und Widerruf der Optionserklärung

Wenn die jPöR die Optionserklärung gem. § 27 Abs. 22 UStG abgegeben hat, dann verzichtet diese auf die Anwendung der Neuregelung des § 2b UStG und nutzt stattdessen die alte Rechtslage des § 2 Abs. 3 UStG a.F. bis zum 31.12.2020 weiter. Durch das BMF-Schreiben hat die jPöR die Möglichkeit, die zwar bisher die Rechtsprechungsgrundsätze des BFH zur Unternehmereigenschaft angewendet hat, mit der Optionserklärung die Auslegung des § 2 Abs. 3 UStG i.S.d. Verwaltungsauffassung zu wählen.³²⁶ Wenn die jPöR sich nicht auf die Verwaltungsmeinung berufen wollte³²⁷, klärte das neue BMF-Schreiben diese Problemstellung, denn es besteht auch die Möglichkeit, dass die Rechtsprechungsgrundsätze angewendet werden können.³²⁸ Dies ist nur möglich, wenn dies einheitlich für das gesamte Unternehmen erfolgt.³²⁹ Die Berufung auf das Unionsrecht und die richtlinienkonforme Auslegung des § 2 Abs. 3 UStG a.F. zu der Vermögensverwaltung, Beistandsleistung und zur Anwendung der in KStR 2015 R 6 Abs. 5 geregelten Umsatzgrenze, war auch vor der Einführung des § 2b UStG möglich.³³⁰ Innerhalb des Optionszeitraumes besteht durch die Abschnittsbesteuerung die Möglichkeit zwischen der Auslegung des § 2 Abs. 3 UStG a.F. in der Verwaltungsfassung oder in der durch die BFH Rechtsprechung geprägte Auslegung zu wechseln.³³¹ Nachfolgend ein Beispiel hierzu (vgl. Tabelle 5).

In Anhang 12 ist eine Zusammenfassung der vier Optionen ersichtlich.

Der Widerruf der Optionserklärung kann mit Wirkung eines auf die Abgabe der Optionserklärung³³² folgenden Kalenderjahres gem. § 27 Abs. 22 S. 6 UStG widerrufen werden. Diese Handlung sollte überlegt werden, da nach einem Widerruf die Abgabe einer weiteren Optionserklärung ausgeschlossen ist, sodass es dann zur Anwendung von § 2b UStG kommt.³³³ Dennoch ist die jPöR auch während des Über-

³¹⁴Vgl. BMF, III C 2 - S 7106/07/10012-06, UR 2016, S. 410.

³¹⁵Vgl. Baldauf, 2018a, S. 60.

³¹⁶Vgl. BMF, III C 2 - S 7106/07/10012-06, UR 2016, S. 410.

³¹⁷Vgl. Sterzinger, 2016b, S. 2942.

³¹⁸Vgl. Sterzinger, 2016b, S. 2942.

³¹⁹BFH, 18.03.1988, V R 194/83, BFH/NV 1989, S. 415-416.; siehe zur Verwaltungspraxis FinMin. Hessen, 13.12.2006, S 7106 A - 096 - II 51, UR 2007, S. 661; FinMin. Bayern, 04.04.1990, 36 - S 7106 - 32/25 - 22, UR 1991, S. 150; a.A. EuGH, 29.09.2015, Rs. C-276/14, Gmina Wrocław, UR 2015, S. 829-834; Widmann, 2015a, S. 888; Diese sind der Auffassung, dass die Gebietskörperschaft der Steuerpflichtige ist.

³²⁰Vgl. Sterzinger, 2016b, S. 2942.

³²¹Alle Bundesländer und der Bund haben die Optionsregelung gewählt; Vgl. Halder und Lorenz, 2018, S. 784.

³²²Vgl. Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rn. 282; Sterzinger, 2016b, S. 2941.

³²³Vgl. Sterzinger, 2016b, S. 2943.

³²⁴Vgl. Sterzinger, 2016b, S. 2943.

³²⁵Vgl. Sterzinger, 2016b, S. 2943.

³²⁶Vgl. BMF, III C 2 - S 7106/07/10012-06, UR 2016, S. 410; Baldauf, 2016, S. 369; Sterzinger, 2016b, S. 2943.

³²⁷Bei der Berufung auf die Verwaltungsmeinung ist Vorsicht geboten, da der BFH und die Finanzgerichte nicht an die Verwaltungsmeinung gebunden sind; Vgl. Baldauf, 2018a, S. 61; Hammerl und Fietz, 2017, S. 63.

³²⁸Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 60.

³²⁹Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 60.

³³⁰Vgl. OFD Niedersachsen, S 7106-283-St 171, DStR 2012, S. 1923.

³³¹Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 60; Dies kann durch die jPöR auch rückwirkend ausgeübt werden, solange die materielle Bestandskraft noch nicht eingetreten ist; Bspw.: Die wirksame Optionserklärung wird im Jahr 2016 abgegeben. Ein Widerruf wird im Februar 2018 mit Wirkung zum 01.01.2017 zulässig, sofern die Umsatzsteuerfestsetzung für das Jahr 2017 noch nicht materiell bestandskräftig ist.

³³²Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 59; Küffner und Rust, 2016a, S. 1469; a.A. Weber, 2017, S. 83-84.

³³³Vgl. BMF, III C 2 - S 7106/07/10012-06, UR 2016, S. 410.

Tabelle 5: Unterschied Optionsmöglichkeit und § 2b UStG

Quelle: Eigene Darstellung in Anlehnung an Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rn. 287.

Eine jPöR erzielt mit Vermögensverwaltung Umsätze in Höhe von 15.000 € im Jahr 2016. Diese Tätigkeit wird in den Folgejahren unverändert fortgeführt mit der Optionserklärung, dann nur auf öffentlich-rechtlicher Sonderregelung.	
Anwendung § 2 Abs. 3 UStG nach BFH-Grundsätzen und Verzicht auf Kleinunternehmerregelung gem. § 19 UStG Diese Tätigkeit ist der Umsatzsteuer zu unterwerfen. Dies hat den Vorteil, dass die mit diesem Umsatz in Zusammenhang stehende Vorsteuer aus Vorleistungen abziehbar ist. Der Vorsteuerabzug bleibt längstens bis Ende 2020 erhalten.	Anwendung des § 2b UStG Der Umsatz wäre gem. § 2b UStG nicht steuerbar gewesen, da die Tätigkeit auf öffentlich-rechtlicher Sonderregelung ausgeführt und keine größeren Wettbewerbsverzerrungen durch den geringen Umsatz gem. § 2b Abs. 1 S. 2 i.V.m. Abs. 2 Nr. 1 UStG vorliegen. Dies führt aber auch dazu, dass kein Vorsteuerabzug möglich ist.

gangszeitraumes 2017 bis 2020 nicht im Korsett der Optionserklärung gefangen, sondern kann solange die Umsatzsteuerfestsetzung materiell nicht bestandskräftig ist, den Widerruf aussprechen.³³⁴

Deshalb ist es für die Anwenderpraxis von Vorteil, wenn diese eine Optionserklärung gem. § 27 Abs. 22 UStG abgeben, obwohl die Neuregelung des § 2b UStG, z.B. wegen vorsteuerintensiver Investitionen günstiger wäre.³³⁵ Denn die jPöR besitzt die Möglichkeit sich auf die Rechtsprechungsgrundsätze zu berufen und den Vorsteuerabzug zu erreichen. Vorteilhaft bei dieser Möglichkeit ist, dass bis zu dem Eintritt der materiellen Bestandskraft die Entscheidung geändert werden kann. Außerdem kann die jPöR für jeden Veranlagungszeitraum innerhalb der Optionsfrist entscheiden, ob diese die Anwendung des § 2 Abs. 3 UStG a.F. in der Fassung der Verwaltungsmeinung oder der Rechtsprechungsgrundlagen nutzt. Dies ist nicht möglich, wenn auf die Optionserklärung verzichtet wird oder eine Optionserklärung innerhalb der Übergangsfrist widerrufen wird.³³⁶

4.3.3. Änderung zur alten Rechtslage

Tätigkeiten, welche die jPöR im Rahmen der öffentlichen Gewalt ausübt, sind nach der alten Rechtslage gem. § 2 Abs. 3 S. 1 UStG nicht steuerbar und auch nach der neuen Rechtslage gem. § 2b Abs. 1 S. 1 UStG nicht, außer diese lösen gem. § 2b Abs. 1 S. 2 UStG größere Wettbewerbsverzerrungen aus. Durch die neue Rechtslage ergeben sich im Gegensatz zu der alten Rechtslage³³⁷ einige Veränderungen, die in der nachfolgenden Tabelle 6 dargestellt werden.

4.3.4. Umstellung auf die neue Rechtslage

Umstellungskonzept und organisatorische Auswirkungen

Der Übergangszeitraum sollte genutzt werden, damit spätestens zum Umstellungsdatum die Verwaltung die Möglichkeit hat, das neue Recht des § 2b UStG zu nutzen. Bezüglich der Umsetzung des § 2b UStG sind zum einen die organisatorischen Punkte zu klären, die einen kurzfristigen Charakter haben, damit der § 2b UStG überhaupt umgesetzt werden kann. Der § 2b UStG hat Auswirkungen auf den entsprechenden Haushalt und die Betrachtung der Vorsteuer der jPöR. Langfristig sind strukturelle Maßnahmen zu erfassen, damit ab dem Anwendungsjahr die korrekte Umsetzung der steuerlichen Deklarationspflicht gewährleistet wird. Als ersten Schritt ist eine Grundsatzentscheidung bezüglich des § 2b UStG zu treffen, ob diese strategisch umgesetzt wird oder es zu einer Einzelfallentscheidung kommt.³⁴² Zwar kann es zwischen der Regelung des Betriebes gewerblicher Art und der Neuregelung des § 2b UStG im Bereich der unternehmerischen Tätigkeit Überschneidungen geben, dennoch sollten beide Bereiche eigenständig gesehen werden.³⁴³ Anschließend ist eine Vielzahl an organisatorischen Punkten zu klären, wie das Projekt des § 2 b UStG umgesetzt werden kann, die nachfolgend in einer Übersicht dargestellt werden (vgl. Tabelle 7³⁴⁴).

Wenn diese Zuordnung des § 2b UStG erfolgt ist, steht sowohl die verantwortliche Dienststelle fest, als auch die Befugnisse, die diese hat. Der nächste organisatorische Punkt

³³⁴Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 59.

³³⁵Vgl. Sterzinger, 2016b, S. 2944; a.A. Maier und Übleiß, 2016, S. 331; Diese sind für eine Einzelfallbetrachtung.

³³⁶Vgl. Hammerl und Fietz, 2017, S. 63; Sterzinger, 2016b, S. 2944.

³³⁷Es wird auf die alte Rechtslage ohne die richtlinienkonforme Auslegung eingegangen.

³³⁸Vgl. Bundesregierung, 2015, BStBl. I Sondernummer 1/2016, R. 4.1 Abs. 5 S. 4.

³³⁹Vgl. BMF, 02.01.19, BStBl. I S. 17, Abschnitt 2.11 Abs. 4.

³⁴⁰Vgl. OFD Rostock, 21.11.2002, S 2706 – 04/01 – St 242, UR 2003, S. 303.

³⁴¹Vgl. BMF, 02.01.19, BStBl. I S. 17, Abschnitt 2.11 Abs. 4.

³⁴²Auf der Ebene der Bundesländer hat sich jedes davon für den strategi-

Tabelle 6: Auswirkungen der alten und der neuen Rechtslage

Quelle: Eigene Darstellung.

Alte Rechtslage (§ 2 Abs. 3 UStG)	Neue Rechtslage (§ 2b UStG)
Verknüpfung des Umsatzsteuerrechts mit dem Körperschaftssteuerrecht durch den BgA. Damit sind nur BgA steuerbar.	Es wurde eine eigene Vorschrift geschaffen, die sich vom den Voraussetzungen des Betriebes gewerblicher Art entkoppelt hat.
Wenn der Jahresumsatz 35.000 € nicht übersteigt, dann handelt es sich in der Regel nicht um einen Betrieb gewerblicher Art. ³³⁸	Es wird im Rahmen des § 2b UStG nur noch auf eine einzelne Tätigkeit durch § 2b Abs. 2 Nr. 1 UStG geschaut.
Die Vermögensverwaltung (z.B. Vermietung und Verpachtung) ist kein Betrieb gewerblicher Art und war bisher durch § 2 Abs. 3 UStG a.F. gedeckt und nicht steuerbar. ³³⁹	Die Leistungen aus Vermögensverwaltung sind steuerbar, außer diese liegen gem. § 2b Abs. 2 Nr. 1 UStG unter 17.500 € .
Öffentlich-rechtliche Beistandsleistungen auf hoheitlicher Basis sind dem nicht-unternehmerischen Bereich zuzuordnen und damit nicht steuerbar. ³⁴⁰	Die Beistandsleistung ist nicht steuerbar unter den Bedingungen des § 2 Abs. 3 UStG, außer diese erzeugen größere Wettbewerbsverzerrungen.
Privatrechtliche Leistungen sind nur im Rahmen des BgA steuerbar, aber ohne BgA nicht, außer es liegen größere Wettbewerbsverzerrungen vor. ³⁴¹	Sämtliche Tätigkeiten auf privatrechtlicher Basis werden aus dem Umkehrschluss zu § 2b Abs. 1 S. 1 UStG besteuert.

Tabelle 7: Umsetzung der organisatorischen Punkte

Quelle: Eigene Darstellung in Anlehnung an Halder und Lorenz, 2018, S. 784-785.

Organisatorischer Punkt	Anwendung in Hessen
§ 2b UStG soll entweder in Projektform oder in der Linie angegangen werden.	Es wird in Projektform angegangen.
Zentrale Ansiedlung in einer Behörde oder dezentrale Ansiedlung.	Gemischte Verantwortlichkeit, da es ein unterstützendes Projektteam gibt.
Geschäftsbereich Zuordnung (Organisations - und Steuerabteilung).	Projektleitung wird in der Zentralabteilung gestellt.
Mandatierung von Steuerberatern und wenn ja, welche Befugnisse hat dieser.	Die Landesverwaltung hat eine eigene landesinterne Steuerberatung, sodass keine externen Steuerberater mandatiert werden müssen.

betrifft die Frage, ob die Fallmeldungen entweder durch die Ressorts gemeldet oder aktiv von der Organisationsabteilung gesucht werden. Dazu müssen über das ERP-System sämt-

lichen Weg entschieden; Vgl. Halder und Lorenz, 2018, S. 784.

³⁴³Vgl. Kammerloher-Lis und Kirsch, 2018, S. 151.

³⁴⁴Dabei wird auf die hessische Landesverwaltung eingegangen, da diese in der Umsetzung des Projektes des § 2b UStG schon sehr weit fortgeschritten sind.

liche relevante Geschäftsvorfälle³⁴⁵ herausgefiltert werden, die einen Einfluss auf die Betrachtung von § 2b UStG haben. Dazu eignet sich folgendes Ablaufschema (vgl. Abbildung 2).

Das Clusterverfahren, d.h. die ersten drei Punkte, sind für jedes Ressort samt nachgeordneter Dienststelle bzw. jeden

³⁴⁵Unter dem Geschäftsvorfall wird eine wirtschaftliche Beziehung der jPöR verstanden, die mit sich selbst oder einem externen Unternehmen oder anderen jPöR eingegangen worden ist.

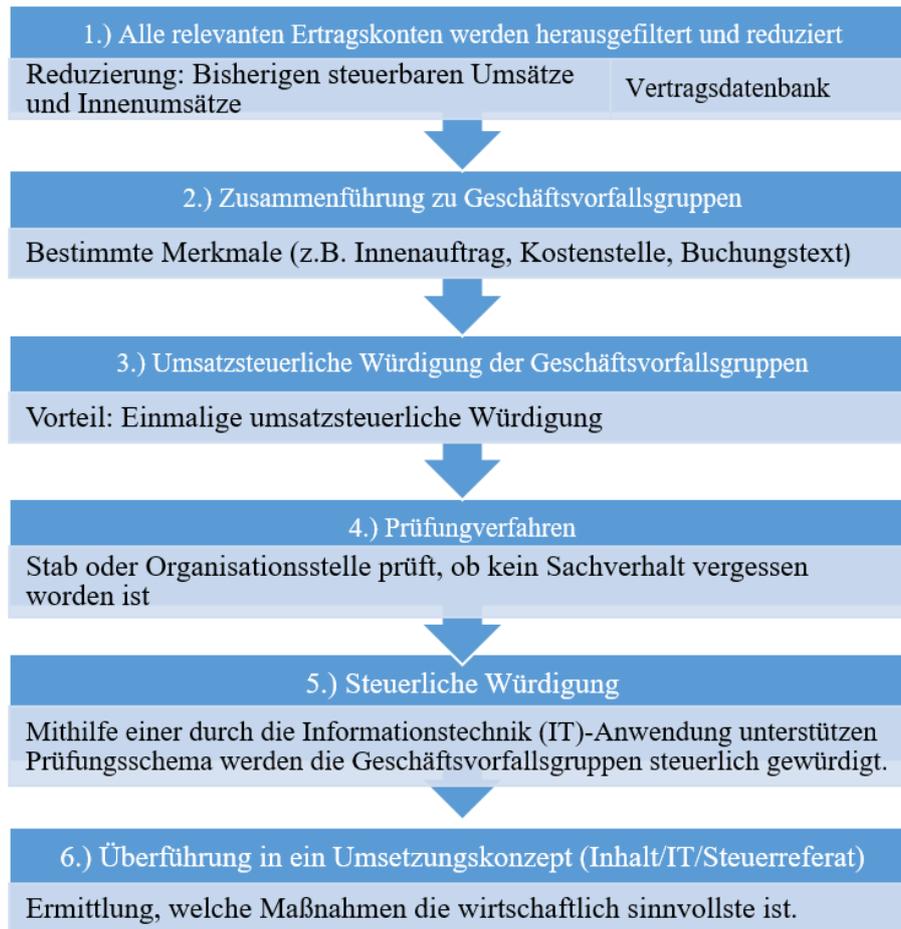


Abbildung 2: Umstellungskonzept

Quelle: Eigene Darstellung in Anlehnung an Halder und Lorenz, 2018, S. 785.

Fachbereich auf unter den Bundesländern liegende Gebietskörperschaften durchzuführen. Damit die Leistungen, Aufgaben und Vertragsbeziehungen erfasst werden, kann eine Vertragsdatenbank bei der späteren Kategorisierung helfen. Wenn die Vertragsdatenbank bisher nicht vorliegt, sollte dies nachgeholt werden, da die Datenbank neben der Bewertung von § 2b UStG Sachverhalten noch weitere Vorteile hat.³⁴⁶ Dabei kann eine Zusammenführung zu Geschäftsvorfallsgruppen gerade auf kommunaler Ebene bspw. nach nachfolgend beschriebenen drei Kategorien aufgebaut sein. Dies sind zum einen Leistungen zwischen den Organisationseinheiten bzw. des Sondervermögens einer Kommune unter Ausschluss von Zahlungsflüssen und zum anderen sind das Leistungen der Kommune gegenüber Sondervermögen, eigenen Beteiligungen, anderen jPÖR oder privaten Dritten.³⁴⁷ Daneben können

dies als dritte Position auch noch Leistungen sein, die von der Kommune von anderen jPÖR oder privaten Dritten eingekauft werden.³⁴⁸

Diese Zusammenstellung macht eine umsatzsteuerliche Analyse und Bewertung einfacher, sodass gleichartige Tätigkeiten mit denen ein Vorsteuerabzug einhergeht, identifiziert werden und kritische Vertragsbeziehungen gesammelt werden können.³⁴⁹ Darauf aufbauend kann ein Umsetzungskonzept helfen, dass ein geordneter und nachvollziehbarer Weg beschritten werden kann. Inhaltlich sind vor allem Maßnahmen (Baumaßnahmen mit dem Ziel der späteren umsatzsteuerpflichtigen Vermietung) relevant, damit geklärt werden kann, welche die wirtschaftlich sinnvollste Lösung für die jPÖR ist. Danach ist innerhalb des Umsetzungskonzeptes die IT Integration der Ergebnisse abzubilden, sowie zur Vermeidung von Complianceverstößen der Aufbau eines Steuerre-

³⁴⁶Die weiteren Vorteile sind die zentrale Transparenz über alle Vertragsbeziehungen, um damit auslaufende Verträge zu identifizieren und mögliche Vertragsdoppelungen zu erkennen und sämtliche Vertragsbestandteile; Vgl. Kubach et al., 2017, S. 35.

³⁴⁷Vgl. Kubach et al., 2017, S. 35.

³⁴⁸Vgl. Kubach et al., 2017, S. 35.

³⁴⁹Vgl. Kubach et al., 2017, S. 35.

ferates notwendig.³⁵⁰ Die Entscheidungen bezüglich des § 2b UStG sind noch mit Rechtsunsicherheit behaftet, sodass die ordentliche Dokumentation empfehlenswert ist.³⁵¹

Haushälterische Auswirkungen und Vorsteuerabzug

Der Vorsteuerabzug ist ein zentraler Bestandteil des deutschen Umsatzsteuersystems.³⁵² Die gesetzliche Grundlage für den Vorsteuerabzug ist § 15 UStG, wobei in den Genuss des Vorsteuerabzuges nur Unternehmen im Rahmen ihrer unternehmerischen Tätigkeit gem. Abschn. 15.1 UStAE kommen. Hierbei gilt ein Zuordnungsgebot zum unternehmerischen Bereich.³⁵³ Daher muss der Leistungsempfänger die Unternehmereigenschaft des § 2 Abs. 1 UStG erfüllen. Da der § 2b UStG die Anwendung der Unternehmereigenschaft vergrößert, ergibt sich für die jPöR ein Vorsteuerabzug, der besonders bei größeren Investitionen vorteilhaft sein kann und einen positiven Effekt auf den Haushalt bewirkt.³⁵⁴ Innerhalb der Optionen einen Vorsteuerabzug zu generieren, gibt es innerhalb des Optionszeitraumes wiederum drei Möglichkeiten.

Dies sind zum einen Leistungsbezüge, die teilweise von Anfang an nach dem Recht des § 2 Abs. 3 UStG a.F. unternehmerisch genutzt worden sind. Wenn im Moment des Leistungsbezuges der jPöR die beabsichtigte Verwendung schon auf der Grundlage der alten Rechtslage mindestens teilweise unternehmerisch zu qualifizieren war, dann kann daraus eine Zuordnung zum Unternehmen erfolgen.³⁵⁵ Durch die Billigkeitsregelung³⁵⁶ ist § 15a UStG anzuwenden, da der zum hoheitlichen und unternehmerischen Bereich zugeordnete Gegenstand innerhalb des Berichtigungszeitraumes gem. § 15a Abs. 1 sich erhöht hat.³⁵⁷ Diese Korrektur gilt nur, wenn sich durch die Einführung des § 2b UStG der unternehmerische Anteil erhöht hat und wenn der Gegenstand trotz gemischter Verwendung dem unternehmerischen Bereich zugeordnet wird.³⁵⁸

Als zweite Möglichkeit gibt es noch die erstmalige Verwendung von Leistungen die während des Optionszeitraumes bezogen werden, aber erst danach unternehmerisch verwendet werden. Dabei ist ein Vorsteuerabzug bereits zum Zeitpunkt des Leistungsbezuges möglich, soweit sowohl die beabsichtigte Verwendung, als auch die weiteren Voraussetzungen des § 15 UStG stimmen.³⁵⁹ Der Meinung der Finanzverwaltung wird zugestimmt.³⁶⁰ Bei der Zuordnungs-

entscheidung für die unternehmerische Verwendung handelt es sich um eine Prognoseentscheidung, da schon zum Zeitpunkt des Leistungsbezuges feststeht, dass die erstmalige Verwendung nach Anwendung des § 2b UStG eine unternehmerische Verwendung hat.³⁶¹ Dieser Ansicht wird auch durch die ständige Rechtsprechung des EuGH zugestimmt, da es dem Grundsatz der Neutralität der Mehrwertsteuer nicht entspricht, wenn die Vorsteuer erst mit der tatsächlichen Nutzung des Investitionsgutes einhergeht, da dann die Mehrwertsteuer anfällt, ohne dass diese durch die Vorsteuer neutralisiert wird.³⁶² Damit die geplante steuerliche Nutzung nachgewiesen werden kann (z.B. durch Zeitungsinserte nach gewerblichen Mietern) sollten zum einen die Verträge mit Steuerklauseln versehen werden, wonach nach Anwendung des § 2b UStG auf die vorsteuerabzugsschädliche Steuerfreiheit von Vermietungsumsätzen verzichtet wird.³⁶³ Außerdem ist es empfehlenswert, wenn eine Dokumentation über den konkreten Anwendungszeitpunkt erfolgt (z.B. Beschluss des Parlamentes oder eine formlose Mitteilung gegenüber dem Finanzamt).³⁶⁴

Die dritte Möglichkeit ist die erstmalige, nichtunternehmerische Verwendung von Leistungen während des Optionszeitraumes, die nach Ablauf des Optionszeitraumes unternehmerisch verwendet werden. Zwar ist ein Vorsteuerabzug zum Zeitpunkt des Leistungsbezuges ausgeschlossen, dennoch kann unter den Voraussetzungen des § 15a UStG eine spätere Berichtigung erfolgen, soweit der Gegenstand bei dem Leistungsbezug zu 10 % unternehmerisch genutzt wird.³⁶⁵ Die Billigkeitsregelung im UStAE kann nicht in Anspruch genommen werden, da der unternehmerische Nutzungsanteil überhaupt begründet worden ist, wobei die Finanzverwaltung die Vorsteuerberechtigung gem. § 15a UStG dennoch zulässt.³⁶⁶ Die Voraussetzung für die Nutzung von § 15a UStG ist, dass während des Leistungsbezuges im Optionszeitraum und der späteren Anwendung des § 2b UStG keine Nutzungsänderung eingetreten ist.³⁶⁷ Offen bleibt, ob die Regelung so auszulegen ist, dass geringfügige Nutzungsänderungen die Möglichkeit einer Vorsteuerkorrektur ausschließen.³⁶⁸

Die vorstehenden Grundsätze für Leistungsbezüge in dem Optionszeitraum sind auch für Leistungsbezüge vor dem 01.01.2017 entsprechend anzuwenden.³⁶⁹ Die Nutzung von § 15a UStG gilt unabhängig davon, ob während des Leistungsbezuges § 2b UStG bekannt war und ob die Absicht bestanden hat, eine spätere unternehmerische Tätigkeit i.S.d. § 2b UStG auszuüben.³⁷⁰ Daneben kann während der

³⁵⁰Vgl. Wagner et al., 2017, S. 19; Küffner und Rust, 2016b, S. 1647.

³⁵¹Vgl. Wagner et al., 2017, S. 19.

³⁵²EuGH, 15.07.2010, Rs. C-368/09, Pannon Gép Centrum, Slg 2010, S. I-7467.

³⁵³Vgl. BMF, 02.01.19, BStBl. I S. 17, Abschnitt 15.2c Abs. 1 S. 1.

³⁵⁴Vgl. Meurer, 2017, S. 165; Die jPöR können die Vorsteuer nach § 15 Abs. 1 UStG unter denselben Voraussetzungen wie private Wirtschaftsteilnehmer abziehen.

³⁵⁵Vgl. Sterzinger, 2018, S. 387.

³⁵⁶Vgl. BMF, 02.01.19, BStBl. I S. 17, Abschnitt 15a1 Abs. 7.

³⁵⁷Vgl. Sterzinger, 2018, S. 387.

³⁵⁸Vgl. Sterzinger, 2018, S. 387; Beispiel siehe Anhang 13.

³⁵⁹Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 65.

³⁶⁰Vgl. Meurer, 2017, S. 166.

³⁶¹Vgl. Meurer, 2017, S. 166.

³⁶²EuGH, 29.02.1996, Rs. C-110/94 INZO, Slg 1996, S. I-857; EuGH, 14.02.1985, Rs. C-268/83, Rempelmann, Slg 1985, S. 655.

³⁶³Vgl. Meurer, 2017, S. 167.

³⁶⁴Vgl. Meurer, 2017, S. 167; Beispiel siehe Anhang 13 Beispiel 1.

³⁶⁵Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 63.

³⁶⁶Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 63.

³⁶⁷Vgl. Sterzinger, 2018, S. 387.

³⁶⁸Vgl. Sterzinger, 2018, S. 387; Beispiel siehe Anhang 13 Beispiel 2.

³⁶⁹Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 67.

³⁷⁰Vgl. Sterzinger, 2018, S. 387.

Anwendung des § 2b UStG eine Berichtigung des Vorsteuerabzuges erfolgen, wenn die Umsatzgrenze des § 2b Abs. 2 Nr. 1 UStG während des Berichtigungszeitraumes innerhalb eines Tätigkeitsbereiches die Umsatzschwelle über - bzw. unterschreitet, sodass es zu einem Übergang zwischen der Besteuerung und der Nichtbesteuerung kommen kann.³⁷¹

Daneben fällt bei den Leistungsbezügen für eine Tätigkeit im nichtunternehmerischen Bereich als zweite Option, die durch § 2b UStG oder § 2 Abs. 3 UStG a.F.³⁷² ausgeschlossen wird, der Vorsteuerabzug mangels Unternehmereigenschaft komplett weg.³⁷³ Daneben können als dritte Option noch gemischte Leistungsbezüge auftreten³⁷⁴, die in zwei Bereiche eingeteilt werden. Der Bezug von vertretbaren Sachen oder sonstigen Leistungen ist entsprechend der unternehmerischen Nutzung aufzuteilen³⁷⁵ und die Vorsteuer kann auch uneingeschränkt nachgeholt werden.³⁷⁶ Ein einheitlicher Gegenstand, Nutzung im unternehmerischen Bereich und im nichtunternehmerischen Bereich, wird anteilig dem unternehmerischen Bereich zugeordnet, sofern die unternehmerische Nutzung gem. § 15 Abs. 1 S. 2 UStG größer als 10 % ist.³⁷⁷

Die Umsatzsteuer kann nicht immer an den Endverbraucher weitergegeben werden, doch wenn dies der Fall ist, dann wird das die Leistungen der jPöR erhöhen. Doch selbst, wenn die Weitergabe der Umsatzsteuer an den Endverbraucher möglich ist, aber die jPöR dies nicht möchte, wird das Entgelt zwar gleich bleiben, aber die Steuerlast würde auf die jPöR abgewälzt werden. Die Wirtschaftlichkeit einzelner Leistungen ist nicht mehr gegeben und die jPöR müssten die Aufgabenwahrnehmung einschränken.³⁷⁸ Daneben kann Fehlbeurteilung bei der Leistungseinordnung dazu führen, dass jPöR die falsch deklarierte Umsatzsteuer schulden oder eine Nachzahlung nebst Nachzahlungszinsen fällig wird.³⁷⁹ Die Risiken sind als Rückstellung im Haushalt zu berücksichtigen.³⁸⁰

Strukturelle Anpassungen mithilfe eines Tax Compliance Systems

Nachdem die jPöR durch organisatorische Maßnahmen den § 2b UStG erfasst hat, ist es auch wichtig, dass ein langfristiger Umstieg erfolgt bzw. Anpassung innerhalb der jPöR passiert. Dadurch, dass jPöR verpflichtet sind, inhaltlich richtige und vollständige Erklärungen für den unternehmerischen Bereich innerhalb des Umsatzsteuerfeldes abzuge-

ben³⁸¹, ergeben sich daraus diverse Anpassungsmaßnahmen. Denn die jPöR sollten alle ihre Deklarationspflichten erfüllen können, was die Einrichtung eines Tax Compliance Systems erfordert.³⁸² Daneben kann das Tax Compliance System³⁸³ auch helfen, die Haftungsrisiken des gesetzlichen Vertreters gem. § 69 AO zu mindern.³⁸⁴ Die Aufnahme aller Ist-Prozesse sowie der Soll-Prozesse (Anpassung an den geänderten Unternehmerbegriff) kann ideal durch ein Tax Compliance System abgebildet werden.³⁸⁵ Der Aufbau eines Tax Compliance Management Systems ist in § 153 AO kodifiziert, welcher durch den Institut der Wirtschaftsprüfer (IDW) Prüfungsstandard (PS) 980 ergänzt worden ist. Innerhalb des Systems gibt es sieben Bausteine. Dies sind die Tax Compliance- (Kultur, Ziele, Organisation, Risiken, Programm, Kommunikation, Überwachung und Verbesserung). Diese Bausteine werden in Tabelle 8 detaillierter beschrieben.

Die einzelnen Bausteine sind als Prozess zu verstehen, sodass die Compliance Kultur der Grundstein ist. In einem ersten Schritt wird eine Steuer-Prozessdokumentation durchgeführt. Aufbauend darauf werden die Risiken mithilfe einer Risiko-Kontroll-Matrix ermittelt, sodass die Risiken in den Zielen münden und in einem Maßnahmenpaket, welches von der Organisation umgesetzt wird.³⁸⁶ Die Tax Compliance Organisation bedient sich anderer Beteiligter im Unternehmen, indem diese das Tax Compliance System innerhalb des Unternehmens kommuniziert.³⁸⁷ Regelmäßig wird die Umsetzung der Maßnahmen geprüft³⁸⁸ und bei Bedarf verbessert.³⁸⁹ Damit das System geprüft werden kann, sind mindestens die folgenden Unterlagen zu erstellen (vgl. Tabelle 9).

Das Zusammenwirken wird in Abbildung 3 dargestellt.

Im Rahmen des Umsatzsteuerprozesses sind die Vorprozesse einzubeziehen, d.h. die Daten und Informationen, die in die Steuererklärung der jPöR übernommen werden. Es gibt einen direkten Zusammenhang zwischen dem Vorprozess und dem Umsatzsteuerprozess. Daher sollte geprüft werden, ob das interne Kontrollsystem als Datenbasis für den späteren steuerlichen Zweck ausreicht.³⁹⁰ Zu den Vorprozessen gehören die Rechnungswesen – und Buchhaltungsprozesse. Innerhalb dieses Prozesses kann die Kosten - und Leistungsrechnung an Bedeutung gewinnen, welche die Ermittlung von kostendeckenden Entgelten für die vertikale und ho-

³⁷¹Vgl. Liebgott, in: Birkenfeld und Wäger, 2018, § 2b UStG Rn. 264; Stadie, in: Rau et al., 2018, § 2b Rn. 451; Sterzinger, 2017, S. 238.

³⁷²Vgl. ?, III C 2 - S 7107/16/10001, BStBl. I 2016, Tz. 61.

³⁷³Vgl. Meurer, 2017, S. 165.

³⁷⁴Beispiele siehe Anhang 14.

³⁷⁵Vgl. BMF, 02.01.19, BStBl. I S. 17, Abschnitt 15.2c. Abs. 2 S. 1 Nr. 1.

³⁷⁶Vgl. Meyer, 2018, S. 268.

³⁷⁷Ein Durchführungsbeschluss (EU) 2015/2428 vom 10.12.2015, Amtsblatt der Europäischen Union (AbL) L 334/12, ermöglicht die weitere Nutzung.

³⁷⁸Vgl. Trost und Menebröcker, 2016, S. 2624.

³⁷⁹Vgl. Wagner et al., 2017, S. 11-12.

³⁸⁰Vgl. Erdbrügger und Liegmann, 2018, S. 63.

³⁸¹Vgl. Kammerloher-Lis und Kirsch, 2018, S. 151.

³⁸²Vgl. Kammerloher-Lis und Kirsch, 2018, S. 151.

³⁸³Alternativ kann dieses System auch Steuer-Internes Kontrollsystem (IKS) genannt werden; Vgl. Anwendungserlass zur Abgabenordnung (AEAO) vom 19.06.18, BStBl. I S. 706, Abschnitt 117a Nr. 2.6 S. 6.

³⁸⁴Vgl. Trost und Menebröcker, 2017, S. 277.

³⁸⁵Vgl. Kammerloher-Lis und Kirsch, 2018, S. 153.

³⁸⁶Vgl. Brill, 2017, S. 20407.

³⁸⁷Vgl. Behringer, 2017, S. 884.

³⁸⁸Die Angemessenheit und Wirksamkeit sollte durch einen Wirtschaftsprüfer bestätigt werden; Vgl. Ebbinghaus und Neu, 2016, S. 866.

³⁸⁹Vgl. Brill, 2017, S. 20408.

³⁹⁰Vgl. Kammerloher-Lis und Kirsch, 2018, S. 151; Wenn es noch kein Steuer-IKS gibt, muss dies sowieso von Kommunen in Nordrhein-Westfalen durch den § 59 Abs. 3 S. 1 GO NRW umgesetzt werden.

Tabelle 8: Tax Compliance Systeme Bausteine

Quelle: Eigene Darstellung in Anlehnung an Behringer, 2017, S. 878.

Tax Compliance-Kultur	Hierbei wird das Wertesystem in Bezug auf steuerliche Fragen durch die Unternehmensleitung geklärt. Daneben müssen steuerliche Veränderungen erkannt und kommuniziert werden.
Tax Compliance-Ziele	Die Ziele sind in einer Steuerrichtlinie enthalten und geben den Handlungsrahmen vor. Diese Ziele sollen messbar sein, damit deren Darstellung transparent wird.
Tax Compliance-Organisation	Die Organisation stellt die Verantwortlichen für den Bereich Steuern mit einem Budget und Kompetenz auf und kommuniziert dies innerhalb der jPöR.
Tax Compliance-Risiken	Es werden Abweichungen von den steuerlichen Zielsetzungen der jPöR erfasst. Dazu dient ein systematisches Risikomanagement.
Tax Compliance-Programm	Das Programm baut auf der Risikoanalyse auf und soll verhindern, dass es wieder zu Verstößen kommt. Dazu dienen Maßnahmen der Risikobegrenzung. Diese sind einzuteilen in präventive Maßnahmen, wie z.B. Richtlinien, Anweisungen für die Mitarbeiter und detektive Maßnahmen wie z.B. Überwachungskonzepte.
Tax Compliance-Kommunikation	Das Programm muss sowohl anderen Abteilungen innerhalb der jPöR kommuniziert werden als auch außerhalb, z.B. Stadtrat, Landtag.
Tax Compliance-Überwachung und Verbesserung	Das System soll ständig verbessert werden, indem Schwachstellen erkannt und beseitigt werden.

Tabelle 9: Unterlagen des Tax Compliance Management Systems

Quelle: Eigene Darstellung in Anlehnung an Kammerloher-Lis und Kirsch, 2018, S. 156.

Steuer-Prozess-Dokumentation	Hierbei werden alle Ist-Arbeitsabläufe erfasst, welche die Grundlage für das Risikoinventar darstellen.
Risiko-Kontroll-Matrix	Die steuerlichen Risiken, ihre Beurteilung, sowie entsprechende Maßnahmen werden erfasst.
Steuerhandbuch	Es wird die Aufbau – und Ablauforganisation festgelegt, sowie alle wichtigen Aufgaben den Abteilungen zugewiesen und Aus – und Fortbildung werden abgebildet. Es umfasst das gesamte Tax Compliance Management System.

rizontale Zusammenarbeit gewährleistet. Daneben ist das Vertragsmanagement der Fachabteilungen wichtig, welches auf dem durch eine Vertragsdatenbank aufgebauten Bestand zurückgreift. Dabei sind die Verträge und die Kooperationsverträge sowie Gesetze, Verordnungen und Satzungen im Hinblick auf die Rechtsänderung durch § 2b UStG anzupassen.³⁹¹ Daneben hat das Beschaffungsmanagement bezüglich des Vorsteuerabzuges und aller Entscheidungsprozesse einen Einfluss auf den Vorprozess.³⁹² Der Vorsteuerschlüssel gem.

§ 15 Abs. 4 UStG wird auf einem mit der Finanzverwaltung abgestimmten Aufteilungsschlüssel beruhen.³⁹³ Im Rahmen dieses Punktes sind Prozesse zu entwickeln bzw. anzupassen, welche die Abläufe in den Geschäftsbetrieb der jPöR einordnen. Dazu gehören Prozesse, wie abteilungsübergreifend ermittelt werden kann, welche Daten an die Entscheidungsträger übermittelt werden können, damit die korrekte und pünktliche steuerliche Deklarationspflicht gewährleistet bleibt. Daneben müssen die technologischen Systeme, also Hard – und Software inklusive der Buchungskennzeichen

³⁹¹Vgl. Küffner und Rust, 2016b, S. 1647.³⁹²Vgl. Kammerloher-Lis und Kirsch, 2018, S. 153.³⁹³Vgl. Kammerloher-Lis und Kirsch, 2018, S. 153.

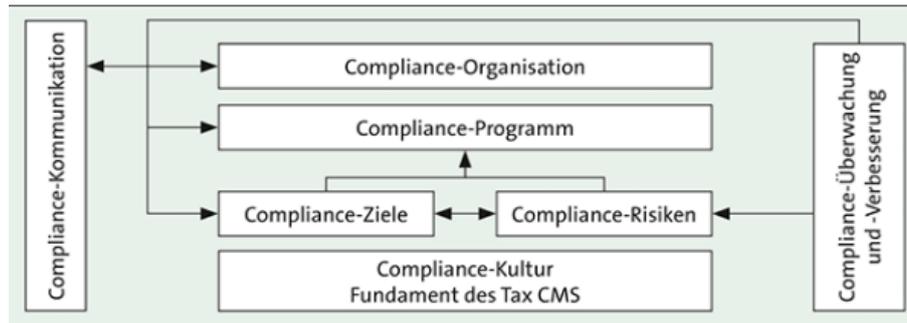


Abbildung 3: Zusammenwirken der Tax Compliance Bausteine

Quelle: Behringer, 2017, S. 884.

angepasst werden.³⁹⁴

5. Zusammenfassung

Die alte unionsrechtswidrige Rechtslage des § 2 Abs. 3 UStG a.F. wurde durch die Einführung des § 2b UStG an die Unionsvorgaben des Artikel 13 MwStSystRl angepasst und hat damit das Sonderrecht der jPöR innerhalb der Umsatzsteuer reformiert. Der größte Unterschied besteht darin, dass das Körperschaftsteuerrecht mit dem BgA keine Geltung mehr für das Umsatzsteuerrecht hat, da jPöR nach der neuen Rechtslage nur noch nach umsatzsteuerlichen Grundsätzen behandelt werden. Nach der neuen Rechtslage wird der Anwendungsbereich des § 2b UStG nur eröffnet, wenn die jPöR sowohl die Unternehmereigenschaft des § 2 Abs. 1 UStG hat als auch in Ausübung öffentlicher Gewalt tätig wird. Dadurch werden privatrechtliche Betätigungen von der Norm des § 2b UStG nicht eingeschlossen. Wenn die Norm des § 2b UStG genutzt werden kann, dann wird die Leistung der jPöR nicht-unternehmerisch behandelt, sodass damit alle umsatzsteuerrechtlichen Konsequenzen vermieden werden. Wenn die jPöR hingegen eine Tätigkeit mit Wettbewerbsverzerrung eingeht, dann wird die jPöR unternehmerisch tätig. Damit die jPöR nicht zu stark von der neuen Rechtslage betroffen sind, gibt es durch den § 2b Abs. 2 und 3 UStG wiederum Ausnahmetatbestände, wann Wettbewerbsverzerrung negiert werden kann. Die Ausnahmenvorschriften haben für die jPöR eine unterschiedliche Bedeutung. Die Vorschrift des § 2b Abs. 2 Nr. 1 UStG weist durch den nicht vorhandenen Bezug zu der MwStSystRl erhebliches Streitpotential aus. Die Regelungen zu den Optionsumsätzen stehen zwar im Einklang mit dem Unionsrecht, aber es scheint fraglich, ob überhaupt praktische Anwendungsfälle vorhanden sind. Unstrittig ist der § 2b Abs. 3 Nr. 1 UStG bei Vorbehaltsaufgaben der jPöR, wobei gleichwohl der § 2b Abs. 3 Nr. 2 UStG bei begünstigten Beistandsleistungen insgesamt gegen das Unionsrecht verstößt, als auch durch die vielen unklaren Rechtsbegriffe kaum praktische Anwendung finden wird. Dagegen klar ist wiederum

der § 2 Abs. 4 UStG, aber dies ist insofern nicht überraschend, da diese Norm aus der alten Rechtslage überführt worden ist.

Die Änderung der neuen Rechtslage sorgt dafür, dass die unternehmerische Betätigung von jPöR gegenüber der alten Rechtslage deutlich vergrößert wird. Dies liegt daran, dass jede einzelne Leistung untersucht werden muss mit dem Hintergrund der Wettbewerbsverzerrung und nicht wie in der alten Rechtslage erst ein BgA vorliegen muss, damit die Leistung unternehmerisch wird. Dadurch ergeben sich sowohl Chancen als auch Risiken. Fehlbeurteilungen bei der Leistungseinordnung durch viele vage Rechtsbegriffe des § 2b UStG können dazu führen, dass die jPöR die Tätigkeit zu der falschen Sphäre einordnet und diese entweder zu viel oder zu wenig Umsatzsteuer deklariert. Daneben scheint es fraglich, ob durch die Umsetzung des § 2b UStG das Unionsrecht umgesetzt worden ist oder ob einzelne Rechtsvorschriften diesem entgegenstehen. Auf der anderen Seite kann durch die Ausweitung des unternehmerischen Bereichs das Volumen des Vorsteuerabzugs erhöht werden, was eine positive haushälterische Auswirkung erzeugt. Die Auswirkungen durch die neue Rechtslage des § 2b UStG sind vielfältig. Empfehlenswert ist es, wenn die jPöR erst einmal eine Optionserklärung zur Verlängerung der Anwendung der alten Rechtslage tätigt, damit diese bis zum Jahr 2021 Zeit hat, sich auf die neue Rechtslage einzustellen. Daraufhin muss die jPöR sowohl kurzfristige organisatorische Maßnahmen (Zuordnung einer Umsetzungsstelle, Umstellungskonzept und Umsetzungskonzept) sowie langfristige strukturelle Maßnahmen (Aufbau eines Tax Compliance Systems und die damit einhergehenden Anpassungen) vollziehen, damit die steuerlichen Deklarationspflichten umgesetzt werden können. Diese Umsetzung erfordert sowohl eine gründliche Analyse als auch steuerliche Beratung für die jPöR.

Die vorliegende Masterarbeit ist insofern beschränkt, als dass das vorliegende Prüfungsschema helfen kann, eine Leistung einzuordnen, aber dies dennoch im Rahmen einer Einzelfallprüfung erfolgen muss. Außerdem bleibt durch viele Rechtsbegriffe gerade im Bereich der Zusammenarbeit von jPöR unklar, ob die gesetzlichen Vorgaben überhaupt umgesetzt werden können. Diese Vorgaben sind durch Gerichte zu klären und auch ob die gesetzliche Regelung des § 2b UStG

³⁹⁴Vgl. Küffner und Rust, 2016b, S. 1647.

wiederum Änderung erfährt. Aufbauend auf dieser Begrenzung schließen sich weitere Forschungsfelder an. Dadurch, dass der § 2b UStG erst flächendeckend ab 2021 angewendet werden muss, wird sich die Frage stellen, inwiefern der § 2b UStG die Vorgaben des Unionsrechts umgesetzt hat und wie mit den Diskrepanzen umgegangen werden wird. Daneben sind vielfältige Anpassungsmaßnahmen für jPöR erforderlich und es stellt sich die Frage, ob die Chancen durch die Nutzung des § 2b UStG den Aufwand rechtfertigen.

Literatur

- Auerswald, L. *Die Vermögensverwaltung im Ertragsteuerrecht*. Dissertation, Baden-Baden (Nomos Verlag), 2016.
- Baldauf, U. Neuregelung der Unternehmereigenschaft öffentlicher Trägerkörperschaften durch das Steueränderungsgesetz 2015 – Einführung eines neuen § 2b UStG. *DStZ*, pages 355–369, 2016.
- Baldauf, U. Besteuerung von Hilfsgeschäften bei Einrichtungen des öffentlichen Rechts. *DStZ*, pages 867–877, 2017.
- Baldauf, U. Umsatzbesteuerung von juristischen Personen des öffentlichen Rechts auf Grundlage der Neuregelung des § 2b UStG – Kurzanalysen und Beispielsammlung. *ZKF*, pages 55–61, 2018a.
- Baldauf, U. Unternehmereigenschaft öffentlich-rechtlicher Trägerkörperschaften: Administrierung des unionswidrigen § 2 Abs. 3 UStG a.F. bis zum 31.12.2020 – Bewertung des BMF-Schreibens vom 27.7.2017, BStBl I 2017, 1239. *DStZ*, pages 567–575, 2018b.
- Baumgart, M. Besteuerung der Betriebe gewerblicher Art, 2018.
- BayLfSt. Umsatzbesteuerung bei der Vermietung, 13.02.2008.
- Becker, S. Aktuelle Fragestellungen zur Umsatzbesteuerung der öffentlichen Hand. *BB*, pages 2154–2159, 2012.
- Behringer, S. IKS als Tax Compliance Management System gemäß IDW PS 980. *BBK*, pages 875–885, 2017.
- Belcke, E. und Westermann, E. C. Die Besteuerung öffentlicher Unternehmen – Die Umsatzsteuerreform und weitere aktuelle Praxishinweise zum Jahresbeginn 2016. *BB*, pages 87–93, 2016.
- BFH. Urteil vom 03.03.2011, 03.03.2011.
- BFH. Urteil vom 03.07.2008, 03.07.2008.
- BFH. Urteil vom 04.09.2002, 04.09.2002.
- BFH. Urteil vom 05.09.1963, 05.09.1963.
- BFH. Urteil vom 08.07.1971, 08.07.1971.
- BFH. Urteil vom 10.02.2016, 10.02.2016.
- BFH. Urteil vom 10.11.2011, 10.11.2011.
- BFH. Urteil vom 13.03.1974, 13.03.1974.
- BFH. Urteil vom 14.03.1990, 14.03.1990.
- BFH. Urteil vom 15.04.2010, 15.04.2010.
- BFH. Urteil vom 17.03.2010, 17.03.2010.
- BFH. Urteil vom 18.03.1988, 18.03.1988.
- BFH. Urteil vom 18.08.1988, 18.08.1988.
- BFH. Urteil vom 18.12.2008, 18.12.2008.
- BFH. Urteil vom 20.08.2009, 20.08.2009.
- BFH. Urteil vom 21.11.1967, 21.11.1967.
- BFH. Urteil vom 23.10.1996, 23.10.1996.
- BFH. Urteil vom 28.01.1988, 28.01.1988.
- BFH. Urteil vom 28.10.2004, 28.10.2004.
- BFH. Urteil vom 31.05.2001, 31.05.2001.
- BGH. Urteil vom 18.12.1981, 18.12.1981.
- Birkenfeld, W. und Wäger, C. *Das große Umsatzsteuer-Handbuch*. Köln, Otto Schmidt, Köln, 2018.
- BMF. Bmf-schreiben vom 19.04.2016.
- BMF. Umsatzsteuer-anwendungserlass: Ustae, 02.01.19.
- BMF. Bmf-schreiben vom 12.11.2009, 12.11.2009.
- BMF. Steuereinnahmen nach Steuerarten 2010-2017, 2017. URL https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Themen/Steuern/Steuerschaeetzungen_und_Steuer_einnahmen/2-kassenmaessige-steuereinnahmen-nach-steuerarten-1950-bis-2017.html.
- Bollweg, A. und Küffner, T. Privatrechtliche Vermögensverwaltung durch Gestattung einer Automatenaufstellung sowie öffentlich-rechtliche Überlassung von Personal und Sachmitteln als "Betrieb gewerblicher Art" einer Universität. *UR*, pages 646–655, 2010.
- Brill, M. W. Tax Compliance: Minimierung steuerstrafrechtlicher Risiken. *kösdj*, pages 1–38, 2017.
- Bundesregierung. Körperschaftsteuer-richtlinien 2015: Kstr, 2015.
- Bundestag. Bt-drs. 18/6094, 23.09.2015.
- Bunjes, J., Geist, R., Heidner, H.-H., Korn, C., Leonard, A., und Robisch, M. *Umsatzsteuergesetz: Kommentar*. C.H. Beck, München, 17., neubearbeitete Auflage edition, 2018.
- BVerfG. Beschluss vom 24.02.2017, 24.02.2017.
- Carl, T. Umsatzbesteuerung der öffentlichen Hand im Wandel: BFH erweitert unternehmerische Tätigkeit von juristischen Personen des öffentlichen Rechts. *NWB*, pages 1586–1590, 2012.
- CDU, CSU, SPD. Deutschlands Zukunft gestalten: Koalitionsvertrag zwischen CDU, CSU und SPD, 2013.
- Cloß, C. *Umsatzsteuerliche Behandlung von Finanz-, Versicherungs- und öffentlichen Dienstleistungen: Branchenspezifische Schwächen des Umsatzsteuersystems*. Berlin, 1. edition, 2013.
- Dötsch, E. und Achenbach, H. *Die Körperschaftsteuer: Kommentar zum Körperschaftsteuergesetz, Umwandlungssteuergesetz und zu den einkommensteuerrechtlichen Vorschriften der Anteilseignerbesteuerung*. Schäffer-Poeschel, Stuttgart, 2015.
- Dziadkowski, D. § 2b UStG und die "größeren Wettbewerbsverzerrungen". *UR*, pages 416–421, 2017.
- Ebbinghaus, M. und Neu, P. Aufbau von Tax Compliance Management Systemen: Warum sich die Mühe lohnt – auch für den Mittelstand! *StuB*, pages 862–866, 2016.
- Erdbrügger, A. und Liegmann, B. Umsatzbesteuerung der interkommunalen Zusammenarbeit (§ 2b UStG). *Kommp Spezial*, pages 63–66, 2018.
- EuGH. Urteil vom 2.05.1996, 02.05.1996.
- EuGH. Urteil vom 04.06.2009, 04.06.2009.
- EuGH. Urteil vom 9.07.2015, 09.07.2015.
- EuGH. Urteil vom 12.09.2000, 12.09.2000.
- EuGH. Urteil vom 14.02.1985, 14.02.1985.
- EuGH. Urteil vom 15.07.2010, 15.07.2010.
- EuGH. Urteil vom 15.11.2012, 15.11.2012.
- EuGH. Urteil vom 16.09.2008, 16.09.2008.
- EuGH. Urteil vom 19.12.2012, 19.12.2012.
- EuGH. Urteil vom 26.01.2003, 26.01.2003.
- EuGH. Urteil vom 29.02.1996, 29.02.1996.
- EuGH. Urteil vom 29.09.2015, 29.09.2015.
- EuGH. Urteil vom 29.10.2015, 29.10.2015.
- FG Münster. Urteil vom 16.04.2013, 16.04.2013.
- Fiand, A. Das Steueränderungsgesetz 2015 - Wichtige Regelungen für die juristischen Personen des öffentlichen Rechts im Bereich der Umsatzsteuer. *KStZ*, pages 29–33, 2016.
- FinMin. Bayern, 04.04.1990.
- FinMin. Hessen, 13.12.2006.
- Fritsch, F., Huschens, F., Koisiak, D., und Langer, M., editors. *UStG eKommentar*. 360 degree eKommentar. Bonn, Stollfuß Medien, Bonn, 2015.
- Gohlke, M. und Schmitz, J.-V. § 2b UStG unter dem Blickwinkel von Steuer- und Vergaberecht. *MwStR*, pages 780–786, 2016.
- Gosch, D. *Körperschaftsteuergesetz: Kommentar*. Beck-Online Bücher. Beck-Online and Beck, München, 3., völlig neubearb. Aufl. edition, 2015.
- Gröpl, C. und Zukiwski, M. Über das Ziel hinaus-Die neue BFH-Rechtsprechung zu Beistandsleistungen. *MwStR*, pages 77–86, 2014.
- Halder, G. und Lorenz, S. § 2b UStG und seine Umsetzung in den Landesverwaltungen – Die Vorgehensweise Hessens. *UR*, 67(20):783–786, 2018.
- Hammerl, R. und Fietz, A. Anwendungsfragen zu § 2b UStG: Darstellung und Anmerkungen zum BMF-Schreiben v. 16.12.2016 zur Umsatzbesteuerung der Leistungen der öffentlichen Hand. *MwStR*, pages 56–65, 2017.
- Heidner, H.-H. Umsatzbesteuerung der öffentlichen Hand: Auswirkung der Rechtsprechung auf die Neuregelung in § 2b UStG. *UR*, pages 45–49, 2016.
- Heidner, H.-H. Die Sonderstellung der juristischen Person des öffentlichen Rechts als umsatzsteuerlicher Unternehmer (§ 2b UStG). *UR*, pages 736–743, 2018.
- Heuermann, B. Verkauf von Grundstücken, die in das Privatvermögen einer als selbständiger Unternehmer tätigen natürlichen Person aufgenommen werden – Petar Kezić sp Trgovina Prizma/ Republika Slovenija. *MwStR*, pages 723–725, 2015.
- Hidien, J. W. und Schwarz, R. Ist die kommunale Straßenreinigung umsatzsteuerpflichtig? Zur Auslegung und Anwendung des § 2b UStG. *UR*, pages 338–353, 2017.
- Hüllmann, M. Die Unternehmereigenschaft der öffentlichen Hand im Umbruch? *UStB*, pages 46–52, 2009.
- Huschens, F. Anmerkung zu EuGH v. 29.10.2015 Rs. C-174/14. *eu-ustb*, pages 75–77, 2015.
- Hüttemann, R. Neuregelung der Umsatzbesteuerung der öffentlichen Hand – alles wird gut? *UR*, pages 129–139, 2017.
- Hüttermann, R. Die Besteuerung der öffentlichen Hand. *FR*, pages 308–314, 2009.
- Ismer, R. Die wirtschaftliche Tätigkeit der öffentlichen Hand im Lichte aktueller Rechtsprechung. *MwStR*, pages 654–660, 2016.

- Ismer, R. und Baur-Rückert, S. Die Besteuerung der öffentlichen Hand im Spannungsfeld zwischen unionaler Rationalität und mitgliedstaatlicher Freiheit. *MwStR*, pages 740-746, 2016.
- Ismer, R. und Keyser, A.-K. Umsatzbesteuerung der öffentlichen Hand: Aktuelle Entwicklungen und gesetzgeberische Handlungsoption. *UR*, pages 81-91, 2011.
- Jürgens, A. und Menebröcker, M. Ist der Vorsteuerabzug der öffentlichen Hand endlich geklärt? Eine systematische Betrachtung unter Berücksichtigung des BFH-Urteils v. 28.6.2017 (XI R 12/15) und der Schlussanträge der Generalanwältin am EuGH v. 19.4.2018 (C-140/17). *UStB*, pages 168-176, 2018.
- Kammerloher-Lis, S. und Kirsch, A. Tax Compliance Management Systeme bei Unternehmen der öffentlichen Hand am Beispiel der Kommunen. *DB*, pages 151-156, 2018.
- Klenk, F. Können einzelne Betriebe von Personen des öffentlichen Rechts Unternehmer sein? Folgerungen aus dem EuGH-Urteil Gmina Wrocław. *UR*, pages 180-183, 2016.
- Korn, K. und Strahl, M. Steuerliche Hinweise und Dispositionen zum Jahresende 2016- Teil 3: Handlungsbedarf aufgrund steuerlicher Änderungen zum Jahresende. *NWB*, pages 3667-3679, 2016.
- Krömker, U. § 3 ao. In Lippross, Otto-Gerd/Seibel, Wolfgang, editor, *Basis-Kommentar Steuerrecht*, Juris, pages Rn. 1-29. O. Schmidt, Köln, 2019.
- Kronawitter, M. Umsatzbesteuerung der Leistungen der öffentlichen Hand: Anwendungsfragen des § 2b UStG – mit Anmerkung Martin Kronawitter. *ZKF*, pages 12-18, 2017.
- Kubach, M., Leipnitz, T., und Porompka, B. Wie Verwaltungen das '2b-Monster' bezähmen. *Innovative Verwaltung*, 39(6):33-35, 2017.
- Küffner, T. Unternehmerische Tätigkeit einer Gemeinde beim Betrieb einer Sporthalle und Freizeithalle – Unmaßgeblichkeit rechtlicher Fehlvorstellungen beim Vorsteuerabzug. *UR*, pages 272-278, 2012.
- Küffner, T. und Rust, M. Kirchen im Fokus der Umsatzsteuer – muss das sein? Ausübung öffentlicher Gewalt durch kirchliche Rechtssträger nach dem geplanten § 2b UStG. *DStR*, pages 2533-2541, 2014.
- Küffner, T. und Rust, M. Wie sollen juristische Personen des öffentlichen Rechts jetzt handeln? *DB*, pages 1469-1470, 2016a.
- Küffner, T. und Rust, M. Reform der Umsatzbesteuerung der öffentlichen Hand: Probleme und Lösungsansätze bei Kooperationen zweier juristischer Personen des öffentlichen Rechts. *DStR*, pages 1633-1647, 2016b.
- Küffner, T., Stöcker, E.-E., und Zugmaier, O. *Umsatzsteuer-Kommentar*. Nummer 5 in NWB-Steuer-Kommentare. Herne, Verl. Neue Wirtschafts-Briefe, Herne, 5 edition, 2018.
- Kußmaul, H., Naumann, C., und Barke, F. Die juristische Person des öffentlichen Rechts in der Umsatzsteuer: Spezialregelungen der gesetzlichen Neukonzeption in § 2b UStG. *StB*, pages 326-334, 2017a.
- Kußmaul, H., Naumann, C., und Barke, F. Die juristische Person des öffentlichen Rechts in der Umsatzsteuer: Neukonzeption des § 2b UStG zur Angleichung der nationalen Regelungen an das Unionsrecht. *StB*, pages 288-295, 2017b.
- Lippross, O.-G. und Seibel, W. *Basis-Kommentar Steuerrecht: AO, AstG, BewG, EigZulG, ErbStG, EStG, FGO, GewStG, GrEStG, GrStG, InvZulG, KraftStG, KStG, SolZG, UmwStG, UStG*. Juris. Köln, O. Schmidt, Köln, 2019.
- Maier, S. und Übleiß, T. Beistandsleistungen und Umsatzsteuer - Zum Entwurf eines geplanten neuen § 2b UStG. *KommJur*, pages 206-210, 2015.
- Maier, S. und Übleiß, T. Neuregelung der Umsatzbesteuerung der öffentlichen Hand-die Finanzverwaltung rät zum Handeln. *KommJur*, pages 331-333, 2016.
- Meurer, T. Steueränderungsgesetz 2015. *UStB*, pages 14-22, 2016.
- Meurer, T. BMF-Schreiben zu § 2b UStG und die Auswirkungen auf den Vorsteuerabzug. *DStZ*, pages 163-171, 2017.
- Meyer, R. Nachholung unterlassener Vorsteuerabzüge im kommunalen Bereich. *ZKF*, pages 265-269, 2018.
- Müller, M. § 2b UStG – BMF-Schreiben klärt für juristische Personen des öffentlichen Rechts bestehende Unklarheiten (nicht). *UR*, 66(1):8-17, 2017.
- OFD Niedersachsen.
- OFD Nordrhein-Westfalen. *Besteuerung der juristischen Personen des öffentlichen Rechts: Arbeitshilfe*. OFD NRW, 2014.
- OFD Rostock, 21.11.2002.
- Pithan, C. P. Umfassender Systemwechsel bei der Umsatzbesteuerung der öffentlichen Hand: § 2b UStG – Was nun? Was tun? *ZKF*, pages 271-277, 2015.
- Rau, G., Dürrwächter, E., Flick, H., und Geist, R. *Umsatzsteuergesetz Kommentar*. Legios Kommentare. Köln, O. Schmidt, Köln, 2018.
- Ringwald, A. Umsatzbesteuerung von Leistungen der öffentlichen Hand –: Vorschlag einer Neuregelung des § 2b UStG - E. *UR*, pages 1-5, 2015.
- Säcker, J., Rixecker, R., Oetker, H., und Limperg, B. *Münchener Kommentar zum Bürgerlichen Gesetzbuch: BGB Band 1: Allgemeiner Teil §§ 1-240, ProStG, AGG*. München, 2015.
- Sauerland, T. Unternehmereigenschaft juristischer Personen des öffentlichen Rechts im neuen Umsatzsteuerrecht. *UStB*, pages 108-117, 2017.
- Scharpenberg, B. Größere Wettbewerbsverzerrungen durch Umsätze juristischer Personen des öffentlichen Rechts. *UR*, pages 834-838, 2017.
- Schmitz, M. und Möser, C. Umsatzbesteuerung der öffentlichen Hand: Zeit für eine einheitliche Neukonzeption. *UR*, pages 607-618, 2014.
- Seibold-Freund, S. *Besteuerung von Kommunen: Regiebetrieb, Eigenbetrieb und kommunale GmbH im Ertragsteuer-, Grundsteuer-, Umsatzsteuer- und Gemeinnützigkeitsrecht*, volume Bd. 12 of *Finanzwesen der Gemeinden*. Berlin, E. Schmidt, Berlin, 2., völlig neu bearbeitete und wesentlich erweiterte auflage edition, 2015.
- Sterzinger, C. Unternehmereigenschaft einer juristischen Person des öffentlichen Rechts. *UR*, pages 37-44, 2009.
- Sterzinger, C. Neuregelung der Umsatzbesteuerung der öffentlichen Hand. *UR*, pages 655-671, 2015.
- Sterzinger, C. Neuregelung der Umsatzbesteuerung der öffentlichen Hand. *SteuK*, pages 1-4, 2016a.
- Sterzinger, C. Einzelfragen zur Optionserklärung nach § 27 Abs. 22 UStG. *DStR*, pages 2941-2944, 2016b.
- Sterzinger, C. Anwendungsschreiben der Finanzverwaltung v. 16.12.2016 zur Neuregelung des § 2b UStG. *DStR*, pages 233-238, 2017.
- Sterzinger, C. Vorsteuerabzug und Vorsteuerberichtigung der öffentlichen Hand unter Berücksichtigung der Schlussanträge v. 19.4.2018 in der Rs. C-140/17. *UR*, pages 377-388, 2018.
- Strahl, M. Kooperationen von Körperschaften des öffentlichen Rechts: – Vor dem Hintergrund der jüngeren Rechtsprechungsentwicklung zur Umsatzsteuer –. *UR*, pages 381-387, 2012.
- Strahl, M. Neues zur Unternehmereigenschaft der öffentlichen Hand – Umstellungsnötigkeit bei Kooperationsverhältnissen. *kösdI*, pages 1-25, 2017.
- Suck, J. Zukunft von Beistandsleistungen unter umsatzsteuerrechtlichen Aspekten: – Rechtsprechung des BFH und ihre Folgen für die öffentliche Hand: Unionsrechtliche Grenzen für ein von den Kommunen gefordertes Nichtanwendungsgesetz –. *UR*, pages 205-211, 2013.
- Trost, C. und Menebröcker, M. Neuregelung der Besteuerung der öffentlichen Hand: Optionserklärung juristischer Personen des öffentlichen Rechts. *DB*, pages 2622-2626, 2016.
- Trost, C. und Menebröcker, M. *Umsatzsteuer in der öffentlichen Verwaltung - inkl. Arbeitshilfen online*, volume v.13202 of *Haufe Fachbuch*. Freiburg, Haufe Lexware Verlag, München, 2017.
- Wagner, Wilfried, Dziadkowski, D., Birkenfeld, W., Widmann, W., Küffner, T., Tillmanns, J., Schäfer, R., Hidién, J. W., und Desens, M. Wandel bei der Umsatzbesteuerung der öffentlichen Hand: Mehr Wettbewerbsgerechtigkeit oder neue Belastungen für die Kommunen? *ifo Schnelldienst*, 70(2): 3-32, 2017.
- Weber, C. Steueränderungsgesetz 2015: Änderungen des Umsatzsteuerrechts. *UVR*, pages 364-370, 2015.
- Weber, C. Umsatzbesteuerung der öffentlichen Hand – Das Anwendungsschreiben des BMF zu § 2b UStG. *UVR*, pages 75-85, 2017.
- Widmann, W. Die durch das Steueränderungsgesetz 2015 angeordneten umsatzsteuerlichen Änderungen. *MwStR*, pages 883-889, 2015a.
- Widmann, W. Umsatzbesteuerung der juristischen Person des öffentlichen Rechts: Entwurf der Finanzministerkonferenz eines neuen § 2b UStG. *UR*, pages 5-11, 2015b.
- Widmann, W. Anwendung des neuen § 2b UStG gem. § 27 Abs. 22 UStG. *UR*, pages 13-16, 2016.
- Wiesch, T. *Die umsatzsteuerliche Behandlung der öffentlichen Hand*, volume Band 49 of *Rechtsordnung und Steuerwesen*. Köln, Otto Schmidt, Köln, 2016. URL <http://gbv.ebib.com/patron/FullRecord.asp?x?p=4733951>.