JUMS

JUNIOR MANAGEMENT SCIENCE

Junior Management Science

journal homepage: www.jums.academy

On the Analysis of Moral Hazard Using Experimental Studies

Maria Huber

Ludwig-Maximilians-Universität München

Abstract

The term moral hazard generally implies individuals' tendency to exercise less effort into cost reduction if the negative consequences resulting from their actions are not borne by themselves. This paper analyzes using recent experimental studies under which circumstances moral hazard is likely to occur and how this problem could be mitigated or eliminated. A detailed overview and analysis of field and laboratory experiments from different areas are provided. At first, a description of the experimental process is presented. The paper then concentrates on findings and, additionally, on the discussion of the ethodology. Overall, the results suggest moral hazard to be an important problem in many markets. However, it is found out that experts without personal financial incentives do not respond to customers' insurance status. Besides, competition mitigates moral hazard on the supply side and evidence shows that moral hazard is less likely to occur in markets for natural disaster insurance where probabilities of damages are low. Additionally, peer pressure and pro-social preferences alleviate the problem of moral hazard in group schemes.

Keywords: First-degree moral hazard; second-degree moral hazard; experiments; analysis.

1. Introduction

Moral hazard is an issue that can occur in many different areas, but since information plays an important role, especially in insurance markets, numerous studies concerning moral hazard focus on those (Richter et al., 2014). For instance, evidence from empirical studies in the context of health insurance indicates a strong positive correlation between health insurance coverage and health expenditures while different possible explanations for this finding exist (Kerschbamer and Sutter, 2017): On the one hand, high-risk individuals are more likely to purchase insurance which refers to the problem of adverse selection. On the other hand, insured individuals demand more services or more expensive ones since their out-of-pocket costs are lower with insurance, a problem known as first-degree moral hazard. Another explanation is that physicians provide more services than necessary or more expensive ones to insured patients who are assumed to care less about the costs because costs are covered by insurance, a second-degree moral hazard problem. In order to analyze moral hazard, it is inevitable to differentiate between adverse selection, first-degree and second-degree moral hazard since the phenomena are equivalent in terms of final outcomes, but the underlying mechanisms are different (Balafoutas et al., 2017). Therefore, this paper an-

alyzes by means of recent experimental studies under which circumstances moral hazard emerges and which features mitigate or eliminate this issue completely.

The first section of the main part contains experiments on second-degree moral hazard i.e., supply side responses to first-degree moral hazard (Balafoutas et al., 2017). Particularly, it is investigated how reimbursement by a third party affects service providers' behavior and whether eliminating financial incentives for providers or allowing for competition influences the degree of moral hazard. The presented studies concentrate on markets for credence goods since moral hazard is assumed to be specifically relevant in such markets because of the high degree of informational asymmetry between expert sellers and consumers (e.g., markets for repair services or health care). Only experts know which quality of service is needed while customers can only observe ex post whether the problem is solved, but if so, they cannot be sure of having received adequate treatment (Balafoutas et al., 2017; Dulleck and Kerschbamer, 2006). It is assumed that this information asymmetry creates strong incentives for service providers to overtreat, undertreat and overcharge (Kerschbamer et al., 2016). Especially, if providers know that consumers do not bear the costs and are, consequently, less price sensitive. Overtreatment (or overprovision) means that sellers provide higher quality or quantity of the service than

needed to solve the customers' problem (e.g., taxi drivers taking passengers on detours) while undertreatment (or underprovision) relates to a situation where the service is insufficient (Kerschbamer et al., 2016). Overcharging is a case where experts charge for more than actually provided (e.g., computer repair experts charging the replacement of a module which has not been replaced) (Kerschbamer and Sutter, 2017). The results from the experiments show that experts without personal financial incentives did not respond to customers' insurance status (Lu, 2014). In addition, competition mitigated moral hazard on the supply side (Huck et al., 2016).

The second part of the paper concentrates on first-degree moral hazard i.e., individuals' tendency to exercise less effort if the negative consequences resulting from their actions are not borne by themselves (Balafoutas et al., 2017). It is investigated whether moral hazard exists in a market for natural distaster risk insurance. As in the case of second-degree moral hazard, first-degree moral hazard has not only been observed in insurance markets, but also in many different areas such as credit and labor markets. For instance, a person working in a team can free ride and trust on the other team members' performance when individuals are paid according to the team output (Holmstrom, 1982). Therefore, also experiments on joint liability group schemes are discussed. As a result, evidence suggests that moral hazard is less likely to occur in markets for natural disaster insurance where probabilities of damages are low (Mol and Botzen, 2018). In addition, experimenters found out that peer pressure and prosocial preferences alleviate the problem of moral hazard in group schemes (Corgnet et al., 2013; Biener et al., 2018).

The remainder of this paper is organized as follows: The next section briefly defines the term "moral hazard", explains the different types and distinguishes this problem from adverse selection. The aim of section 3 is to analyze moral hazard by using experimental studies from different areas. A detailed overview[1] of recent field and laboratory experiments is provided, due to structural reasons, first on second-degree and second on first-degree moral hazard. A description of the experimental process is for each experiment presented at first. The paper then concentrates on the results and, additionally, on the discussion of the experimental methodology. Finally, section 4 draws a conclusion and points out possible academic voids which can guide to future research topics.

2. Moral Hazard in Theory

2.1. Definition

The term "moral hazard" has its origin in the insurance literature. Arrow (1963, p. 961) defined it in the context of health insurance as the observation that "medical insurance increases the demand for medical care". Therefore, moral

hazard can be viewed as an insurance-induced behavior modification of individuals (Karten et al., 2018) – meaning that an individual with more insurance coverage has weaker incentives to prevent losses and therefore insured events will occur more often compared to an individual with less or no coverage (Balafoutas et al., 2017). However, since moral hazard is not only an issue in insurance markets, the term generally implies individuals' tendency to exercise less effort into cost reduction if the negative consequences resulting from their actions are not borne by themselves (Balafoutas et al., 2017). The phenomenon of moral hazard – as of adverse selection[2] – arises from an asymmetry of information between contracting parties (Holmström, 1979; Arnott and Stiglitz, 1991). Specifically, this asymmetry occurs ex post[3] (Finkelstein and McGarry, 2006) and is associated with hidden action i.e., the probability distribution of observable outcomes is dependent on agents' actions which are unobservable for the contracting party (Arnott and Stiglitz, 1991).

2.2. Types of Moral Hazard

According to the literature on insurance theory (e.g., Nell, 1998), moral hazard can be classified into several types which are represented in Figure 1. At first, it can be divided into the legal and the illegal moral hazard. The illegal type is insurance fraud which requires a material misrepresentation (e.g., lie or concealment), the intention to deceive and the aim to realize unauthorized benefits (Viaene and Dedene, 2004).[4]

The legal type is subdivided into the external and the internal moral hazard. External moral hazard (second-degree moral hazard) references to third parties who may change their behavior based on their customers' insurance coverage whereas the internal type corresponds to the insured individuals' behavior (first-degree moral hazard) (Karten et al., 2018). The former is defined as the supply side's tendency to increase the price or the extent of a service when moral hazard on the demand side is expected since the demand side is less price sensitive due to insurance (Balafoutas et al., 2017). The second legal form includes ex ante and ex post moral hazard: Ex ante moral hazard refers to an insured individual's behavior to spend less effort in reducing the likelihood of a loss (Einav and Finkelstein, 2018; Ehrlich and Becker, 1972). For instance, an individual with health insurance coverage may have fewer incentives to avoid an unhealthy lifestyle (e.g., smoking) since insurance covers the resulting financial costs. The degree to which a subject's demand for healthcare is influenced by the out-of-pocket price he has to pay for the care is described as ex post moral hazard (Pauly, 1968; Einav and Finkelstein, 2018) i.e., if an uninsured person would not have visited a doctor because of an innocuous disease, but he decided to do so because he was insured then his behavior is attributed to ex post moral hazard. Einav and

---

[1]Due to space constraints, it is not possible to present a broader overview in this paper since the research question requires a detailed discussion of experiments.

[2]The problem of adverse selection is briefly addressed in section 2.3.

[3]"Ex post" relates to the conclusion of the (insurance) contract.

[4]The exact specification may vary between different systems of justice (Viaene and Dedene, 2004).
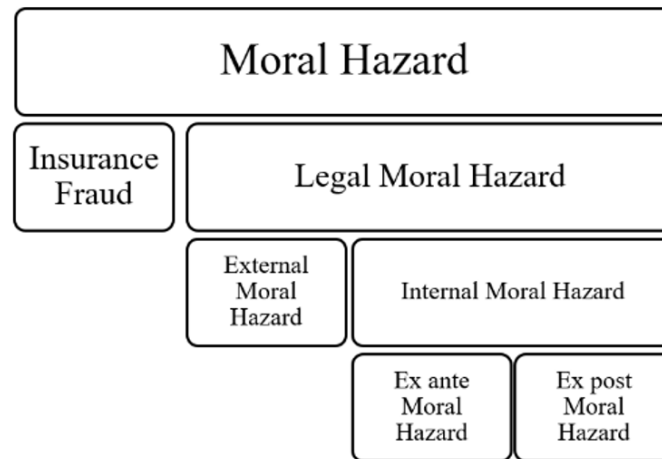
Figure 1: Types of Moral Hazard According to Nell (1998)[5]

Finkelstein (2018) argue that using "moral hazard" in this context is an abuse of the term since an individual's healthcare utilization (action) can be observed which means that there is more a problem of hidden information about the person's health status than a problem of hidden action.

2.3. Distinction from Adverse Selection

As already mentioned beforehand, the situation of asymmetrically distributed information can also lead to the problem of adverse selection which Arrow (1986) attributes to hidden information. Under adverse selection, a subject is assumed to have private information about his risk type prior to the insurance contract relative to the insurance company which creates an ex ante information asymmetry (Finkelstein and McGarry, 2006). A person with private information that he is a high risk is more likely to choose an insurance contract with a higher coverage level than a person who believes himself to be of a type of low risk (Finkelstein and McGarry, 2006). And consequently, the causality between coverage and riskiness is reversed compared to moral hazard (Finkelstein and McGarry, 2006). A positive correlation between the level of insurance coverage and the degree of riskiness can, therefore, result from both, adverse selection and moral hazard (Finkelstein and McGarry, 2006). This brings up difficulties in clearly disentangling these two problems empirically. However, this paper does neither concentrate on the analysis of adverse selection nor on approaches to clearly disentangle[6] moral hazard and adverse selection since the experiments presented in the following were designed in a way so that the problem of adverse selection was eliminated.

---

[5]Own representation based on Nell (1998).

[6]Cohen and Siegelman (2010) discuss three approaches to the disentanglement in their paper.

3. Experimental Evidence on Moral Hazard

3.1. Second-Degree Moral Hazard

Balafoutas et al. (2017) were the first to study moral hazard and its influence on market outcomes in a controlled field experiment concentrating on the effect of first-degree moral hazard on the behavior of the supply side. The authors provide evidence for second-degree moral hazard in a market for taxi rides where costs were reimbursed by a third party.

In the experiment, four research assistants, two men and two women, took undercover taxi rides in the capital city of Greece following a fixed script and secretly documented the drivers' driving and charging behavior. The rides were organized in quadruples meaning that all four assistants took a taxi from the same origin to the same destination in one or two-minute intervals and at random order. Overall, the experiment consisted of 400 rides while 200 were part of the control treatment (CONTROL) and 200 were assigned to the treatment with insurance (INS)[7]. The assistants explained to the taxi drivers in both treatments that they were not familiar with the city in order to create an information asymmetry. In CONTROL, the assistants asked the drivers shortly after the ride had begun for a receipt at the end of the ride (without mentioning the purpose of this question) while in INS, it was added that the receipt was needed since expenses were reimbursed by the passengers' employers.[8] At the end of the experiment, the actual fares paid by the assistants were compared to the correct prices. This was possible because charging fees for taxi rides are standardized in Greece: The tariff consists of a fixed fee per ride and a variable part. This variable part is either computed distance or duration-dependent contingent on what is more profitable for the driver and the

---

[7]Reimbursement from the employer and insurance have comparable financial consequences for the consumer (Kerschbamer and Sutter, 2017). Therefore, and for consistency reasons, the treatment will be declared as an insurance treatment.

[8]The authors state that, except for this additional information, both treatments were identical.

taximeter always applies the more profitable method automatically.

Balafoutas et al. (2017) measured overtreatment along two dimensions: The duration of the ride and the distance driven. Table 1 shows the values of duration and distance indices by gender and treatment. A comparison of the average duration index across treatments (1.14 in CONTROL and 1.13 in INS) and across genders (1.13 for male and 1.14 for female passengers) did not reveal any significant differences. In addition, the differences in values of the distance index were again insignificant across genders (1.06 for males and 1.07 for females), but marginally significant between CONTROL (1.06) and INS (1.08). Therefore, only a minor overtreatment effect along the distance dimension was found. The authors state that the reasons for the small differences in the overtreatment index between treatments could have been that overtreatment was associated with additional costs of service as for example fuel costs or opportunity costs of time. A passenger who does not have to bear the costs for the taxi ride would probably not mind a higher price but would complain about the duration of the ride longer than necessary.

In Table 2 one can observe overcharging frequencies and price indices across treatments and across genders. In CONTROL, 20% of taxi rides were overcharged while the overcharging frequency was 37% in INS. According to the authors, this indicates a statistically significant and causal effect of second-degree moral hazard. Additionally, the mean overcharging amount by which taxi drivers increased the fare was higher in INS (€ 1.43) compared to € 0.91 in CONTROL. Therefore, it is not surprising that the price index increased after the moral hazard manipulation as can be observed from Panel (b) in Table 2. This suggests that passengers' expenditures increased compared to the absence of second-degree moral hazard. The source of these results could be the drivers' assumption that, when employers reimburse the costs for the ride, passengers care less about higher costs and hence overcharging behavior will be undetected and not reported.

Another important finding was that in CONTROL, female passengers paid overcharged prices more frequently (26%) than male passengers (13%)[11] while the values were almost similar across genders (36% and 37%, respectively) in INS. Therefore, the difference in overcharging frequencies between both treatments was highly significant only for male passengers. Women could have been perceived as less likely to complain about being overcharged in general which explains the differences in overcharging between genders in CONTROL. And since the additional information for the driver about the employer paying for the ride did not change

this perception about women, the overcharging frequency increased only by an insignificant amount of 10 percentage points.

Passengers had to pay unjustified surcharges (e.g., transport to the airport) in 77% of all overcharging cases. In the remaining ones, drivers manipulated their taximeters, applied the night tariff during daytime or rounded up the price by more than 5% of the fare. Overall, the experiment stresses the importance for employers to reduce the extent of second-degree moral hazard. As one possible solution, vertical integration with service providers (e.g., firm's own chauffeur service) may be implemented.

In the following, the experimental methodology will be discussed. List (2006) argues that field experiments represent a bridge between laboratory and naturally-occurring data. In relation to an experiment in the laboratory, the experimenter potentially has less control over the environment in a field experiment since it is not possible to take all situational factors into account (Richter et al., 2014), but in exchange for more external validity i.e., realism increases (List and Reiley, 2008). In the presented experiment, taxi drivers were the population of interest being observed in their natural environment without knowing that they were being analyzed. This is important since different types of subjects may behave differently i.e., students in the laboratory may behave differently than real taxi drivers and subjects knowing that they are being observed may also change their behavior (List and Reiley, 2008). Another advantage of the methodology was that first-degree moral hazard and adverse selection can be excluded as sources for the found results since the assistants' behavior was exogenously controlled and held constant by exact instructions and passengers were randomly assigned to one of the two treatments (Balafoutas et al., 2017). An additional benefit was that all four assistants took taxis in short intervals from the same origin to the same destination in order to make the prices comparable. Thus, factors influencing the taxi driver's choice of route (and thereby the price) as, for instance, traffic or weather conditions were eliminated (Balafoutas et al., 2017). It is important to note that the results from this experiment may not represent results in other markets (or other countries) since the market for taxi rides is highly specific and the experiment was conducted only in the city of Athens.

Due to the fact that the market for taxi rides in Greece is highly regulated and over-treatment and overcharging may have different consequences for consumers in other markets, Kerschbamer et al. (2016) confirm the importance of second-degree moral hazard in a less specific market, the computer repair market. For that purpose, the impact of customers' insurance coverage on computer repair experts' provision and charging behavior was examined.

In the natural field experiment by Kerschbamer et al. (2016), equally manipulated computers were brought to 61 randomly selected repair shops in Austria for a reparation. One of the random-access memory modules was destructed in all computers which caused an unambiguous error

---

[9] Balafoutas et al. (2017, p. 9); The columns CTR and MOH represent results from CONTROL and INS, respectively.

[10] Balafoutas et al. (2017, p. 10); The columns CTR and MOH represent results from CONTROL and INS, respectively.

[11] According to the authors, women were, ceteris paribus, 18.1% more likely to face overcharging in CONTROL than men. This is shown in column 2 in Appendix 1.

Table 1: Overtreatment Indices[9]

Notes. Panel (a): The duration index is the ratio of time driven in each ride to time driven in the quickest ride in that particular quadruple. Panel (b): The distance index is the ratio of distance driven in each ride to distance driven in the shortest ride in that particular quadruple. CTR refers to the control treatment and MOH refers to the moral hazard treatment.

	CIR	MOH	Overall average
Panel (a): duration index			
Male passengers	1.124	1.135	1.130
Female passengers	1.152	1.126	1.139
Overall average	1.138	1.130	1.134
Panel (b): distance index			
Male passengers	1.056	1.071	1.064
Female passengers	1.053	1.084	1.068
Overall average	1.055	1.077	1.066

Table 2: Overcharging Frequency and Price Index[10]

Notes. Panel (a): Overcharging frequency refers to the share of rides that have been classified as cases of overcharging. In parentheses, we report the mean unconditional overcharging amount (which is zero if overcharging has not taken place). Panel (b): The price index is the ratio of total price paid in each ride to the lowest total price paid in that particular quadruple. CTR refers to the control treatment and MOH refers to the moral hazard treatment.

	CTR	MOH	Overall average
Panel (a): overcharging frequency (mean overcharging amount in parentheses, in €)			
Male passengers	0.13 (0.72)	0.37 (1.46)	0.25 (1.09)
Female passengers	0.26 (1.10)	0.36 (1.40)	0.31 (1.25)
Overall average	0.20 (0.91)	0.37 (1.43)	0.28 (1.17)
Panel (b): price index			
Male passengers	1.075	1.153	1.114
Female passengers	1.109	1.177	1.143
Overall average	1.092	1.165	1.129

message on the screen.[12] Therefore, every computer expert should have been able to diagnose and solve the problem. According to the IT department, the repair should have been completed within half an hour and for costs of € 60 to € 80. The customer, an undercover experimenter, entered the shop, asked for a repair and indicated that he was a non-computer expert by mentioning that he had no idea why the computer cannot be booted. Two different treatments were randomly assigned to the shops: In CONTROL, the customer explained before leaving the shop that he would need a bill after the repair while in INS, the customer added that the bill was needed for his insurance company because repair costs were covered.[13] After the reparation, the computers were checked in order to find out what had been done to solve the booting problem and whether the positions on the bill fit to the repair actually undertaken. Finally, to investigate the motives

for the differing behavior of sellers between treatments, the authors conducted a survey where they asked experts from 15 repair shops why insurance might lead to higher prices for customers.

The authors found out that the average price for the repair increased by 83% from € 70.17 in CONTROL to € 128.68 in INS indicating a highly significant effect of the insurance treatment. Figure 2 illustrates this large difference by means of the relative cumulative frequencies of repair prices. This finding is in line with what Balafoutas, Kerschbamer and Sutter found in the previous experiment.

Overtreatment yielded 29% of the price difference between the two treatments: In five cases, unnecessary repairs – additional to the replacement of the defective module – were carried out. The price for these five repairs was € 200.58 on average which was significantly higher than the average price for the other repairs in INS (€ 112.34). Interestingly, all these repairs were made in INS and since the computers were, except for the manipulation, in perfect con-

---

[12]According to the authors, the computers were in perfect condition aside from the manipulation.

[13]Both treatments were completely identical except for this difference (Kerschbamer et al., 2016).
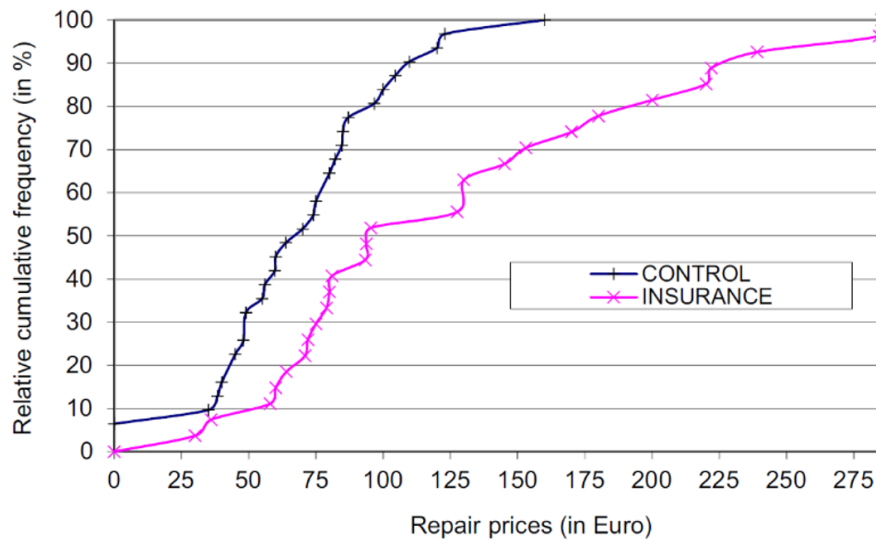
Figure 2: Relative Cumulative Frequency of Repair Prices[14]

dition, this can be interpreted as overtreatment.

Moreover, overcharging explained the remaining 71% of the price difference between CONTROL and INS: The authors found no difference due to charged repairs that had not been conducted,[15] but overcharging in the working-time dimension was found – probably since the customer was not present during diagnosis and repair. While no significant difference in hourly rates between treatments (€ 87.47 on average) occurred there was an increase in the charged working time of 85% from CONTROL (0.55h) to INS (1.02h).[16] This strong difference is also shown in Figure 3. The results from the survey on the motives for overcharging and overtreatment in the light of insurance – which are represented in Appendix 2 – showed that second-degree moral hazard was considered as the most likely explanation. Experts expected the customers to pay less attention to price minimization because of their insurance coverage.

With reference to the methodology it should be mentioned that just as in the experiment by Balafoutas, Kerschbamer and Sutter, first-degree moral hazard and adverse selection were ruled out in the experiment.[17] The computers were manipulated in a way that experts should have been able to easily find the problem and, therefore, incompetence was excluded as a reason for the differing behavior of experts. However, three shops out of 61 stated either that the computer was irreparable or that a repair would be more expensive than buying a new computer suggesting that finding

the error was probably not as simple as expected. This may be an issue because when experts spend more time on identifying the problem, the repair costs increase, consequently, due to incompetence and not because of intended misbehavior (Kerschbamer et al., 2016). Therefore, it is possible that parts of the overcharging effects in the working-time dimension were solely attributed to incompetence. Another problem may be that only 29 of all repair shops indicated the working time and the hourly rate as a position on the bill. It was not possible to observe whether the charged time was used for repair or not, but it may be arguable that the number of 27 observations (two excluded because of overtreatment) was too small for drawing a justifiable conclusion on the overcharging behavior. However, a positive course of action was that observations implying overtreatment were excluded from computing the effect of overcharging in the working-time dimension since the replacement of additional parts of the computer is positively correlated with the duration of the working time. In regard to the survey, one may criticize the number of interviewed experts.

In the previously discussed experimental studies, the sellers of credence goods had financial incentives for behaving fraudulently. In contrast, Lu (2014) investigated whether experts without personal monetary incentives also react to consumers' insurance status.

The author conducted a field experiment in which undercover testers visited doctors at hospitals in Beijing (China). These testers explained to the doctors that they were sent on the authority of a family member (patient) living in another region of the country who wanted a doctor in a high-rated hospital to have a look on his case.[18] Therefore, two hypothetical patients were designed and the testers brought their reference sheets with medical test results indicating

---

[14]Kerschbamer et al. (2016, p. 7456)

[15]In four cases, the experts billed replacements of parts which had actually not been replaced, but two of these cases occurred in CONTROL and two in INS (Kerschbamer et al., 2016).

[16]Since only 29 shops indicated the rate per hour and working time on the bill (and two observations were excluded due to overtreatment), these numbers were computed from 27 observations.

[17]See the discussion of the methodology of the previous experiment for an explanation.

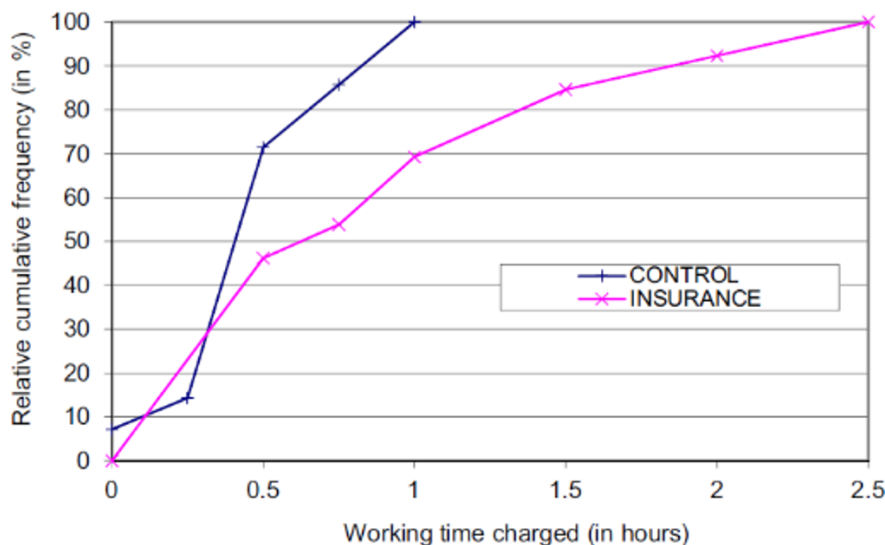[18]This procedure is very common in China (Lu, 2014).

Figure 3: Relative Cumulative Frequency of Working Time[19]

different health problems which either required medication or not. After the tester had described the patient's health problems according to a fixed script, the doctor had to decide whether to prescribe no drugs, generic drugs or more expensive brand-name drugs and the package size. The experiment was divided into four different treatments which were randomly assigned to the doctors: Insurance-incentive, no-insurance-incentive, insurance-no-incentive and no-insurance-no-incentive. In the treatments with insurance, doctors received the information that the patient was insured and, to the opposite, that he had no insurance coverage in the no-insurance treatments. Additionally, in the incentive treatments, doctors were informed that the tester would buy the prescribed drugs for the patient from the doctor's hospital. This case created personal financial incentives for physicians since their payments often depend on the revenue generated in their hospital.[20] The testers indicated that the drugs would be purchased elsewhere in the no-incentive treatments.

Evidence presented in Table 3 shows that when doctors had financial incentives to prescribe more drugs or more expensive ones, patients paid 522 Yuan on average in the insurance condition and 365 Yuan when they were not insured. Therefore, insured patients had to pay 43% more for drugs – which was highly statistically significant – since physicians prescribed more brand-name drugs (83% vs. 68%), a higher number (2.47 vs. 2.20) and more units of drugs (2.53 vs. 2.09) to insured.

These effects are displayed in the first column in Appendix 4. A possible reason for these results could have been

that doctors wanted to increase drug expenditures since their income was calculated in proportion. An important finding was that in the no-incentive treatments average outcomes were not statistically different between insurance statuses as can be seen from the second column in Appendix 4. Hence, physicians did not respond to the patients' insurance status when they did not expect any profits from prescriptions. By comparing the insurance-incentive to the insurance-no-incentive treatment, the author found out that doctors with incentives prescribed significantly more unnecessary drugs[21] to insured patients (second line in Table 3). The number and units of drugs were also significantly higher, but the share of branded drugs was almost equal in both treatments (83% and 81%). Overall, given insurance, financial incentives for doctors increased the average drug expenditures for patients significantly.

In this experimental study, adverse selection and first-degree moral hazard did not play a role either since testers, who were randomly assigned to the treatments, received exact instructions for their behavior. In addition, the testers indicated that they were not the patient who needed the doctor's advice and, therefore, the testers' characteristics should have had little impact on the doctors' behavior (Lu, 2014). However, Lu (2014) does not completely rule out the possibility that doctors' inferred information from the conversation with the tester may have influenced the results. For instance, although the author implemented two elements to make the doctors aware of the patients being neither poor nor rich, the doctors could have assumed that patients who did not want the drugs to be purchased in the hospital were more price sensitive since – according to Lu (2014) – prices

---

[19]Kerschbamer et al. (2016, p. 7456)

[20]In addition, hospitals in China often receive kickbacks from drug companies which also results in incentives for doctors to prescribe (Yip and Hsiao, 2008).

[21]One hypothetical patient had increased triglycerides, but, according to medical guidelines, the patient should have not received medication for this level of triglycerides (Lu, 2014). Therefore, it was possible to test for overtreatment.

Table 3: Average Treatment Outcomes[22]

Notes: "D&H" represents "for diabetes and hypertension only." Standard errors are in parentheses.

Dependent variables	Insurance incentive	No insurance incentive	Insurance no incentive	No insurance no incentive
For both patients				
Raw drug expenditure (Yuan)	522.11	365.14	-	
s.e.	(35.80)	(23.63)	-	
Prescription for triglycerides (0/1)	0.64	0.40	0.28	0.40
s.e.	(0.10)	(0.10)	(0.09)	(0.10)
Monthly drug expenditure D&H (Yuan)	424.78	298.71	324.50	307.03
s.e.	(23.54)	(15.84)	(18.95)	(15.44)
Number of drugs D&H	2.47	2.20	2.18	2.18
s.e.	(0.10)	(0.08)	(0.07)	(0.06)
Unit of drugs D&H	2.53	2.09	2.16	2.12
s.e.	(0.11)	(0.08)	(0.09)	(0.07)
Share of branded drugs D&H (0-1)	0.83	0.68	0.81	0.80
s.e.	(0.04)	(0.05)	(0.03)	(0.04)
Obs. for triglycerides	25	25	-	-
Obs. for other variables	49	49	49	49

at outside pharmacies can be below the ones in hospitals. Besides, it would have been optimal to visit each doctor for all four treatments to control for heterogeneity between the doctors, but Lu (2014) argument that presenting the same test results multiple times to each doctor would have caused suspiciousness among physicians seems plausible.

The first laboratory experiment discussed in this paper was conducted by Huck et al. (2016) who investigated the effects of medical insurance and competition on patients' and physicians' behavior with a focus on overtreatment. They did not only find second-degree moral hazard, but also evidence for first-degree moral hazard. The experiment consisted of four treatment variations – CONTROL, INS, treatment with competition (COMP) and treatment with insurance and competition (INS-COMP) – which are explained in the following.

In CONTROL, 336 students were randomly assigned to a fixed role as a physician or as a patient. The patients were confronted with a problem which required treatment. In each round, the patients were randomly matched to a physician and the severity of their problem (mild or severe) was determined. Then, patients had to choose – without knowing the severity of their problem – whether to consult their assigned physician or not. If a patient consulted a physician, he was able to observe the severity of the problem and chose the treatment (patients had to pay for treatments). In the case of a severe problem, the physician only had the option to provide a severe (and more costly) treatment to the patient as presented in Figure 4 whereas in case of a mild problem, the physician also had the opportunity and monetary incentives to overtreat. This means offering a severe treatment

although a mild treatment would have been sufficient for a cure and less expensive for the patient. The payoffs resulting from each condition can be observed from parentheses in Figure 4: The upper numbers are patients' payoffs while lower numbers are those of physicians. At the end of each round, patients who consulted their physician received information about their treatment, but still not about the severity of their problem. This type of information was only given to those who decided not to consult. All subjects also learned about their payoffs after each round.

INS was almost equal to the above-described process, but all patients – even the ones not consulting a physician – had to pay a fair premium to cover extra costs of overtreatment. An additional difference was that prices for severe and mild treatments were – in contrast to CONTROL – identical. Hence, a single patient did not have to bear higher costs resulting from unnecessary overtreatment alone. The premium was calculated dependent on the number of severe treatments provided i.e., the more severe treatments provided the higher the premium. Patients were aware of this calculation method and received information about the paid premium at the end of the rounds.

In COMP, patients were allowed to freely choose among all physicians which was defined as competition. In addition, patients and physicians observed the number of patients who had consulted the physician in previous rounds (i.e., market shares) from a history table. Finally, in INS-COMP, INS and COMP were combined to one treatment.

Huck et al. (2016) found out that insurance induced moral hazard on both sides of the market. Table 4 which summarizes average results from all rounds and for all treatments shows that 36% more patients consulted a physician in INS than in CONTROL because additional costs of treat-

[22]Lu (2014, p. 161); Due to space limitation, the complete table cannot be presented in this part of the paper, please see Appendix 3.
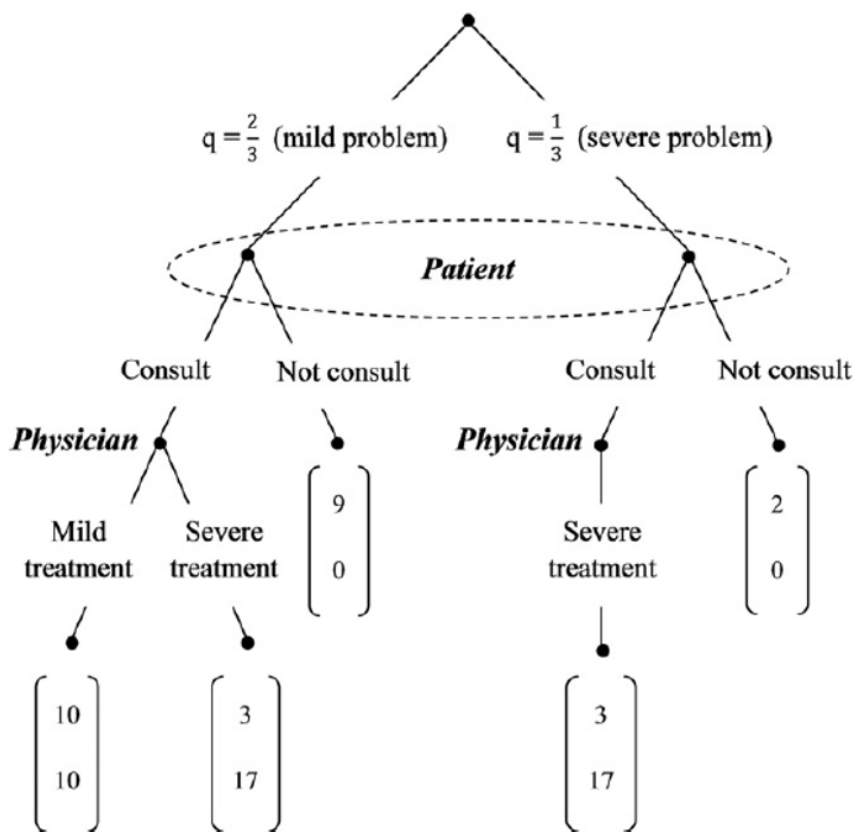
Figure 4: Game Tree of CONTROL with Actual Payoffs[23]

Table 4: Results from All Treatments[24]

Notes: The table shows averages over all 30 periods and 7 markets in the main treatments. The rates in the first four lines are indicated in percent: (1) is the share of consulting patients, (2) is the share of consulted physicians who give severe treatment when the problem is mild, where the average rate (2) is weighted by the number of consultations per session and period. (3) is the sum of actual earnings over the sum of potential earnings. (4) is the share of all interactions when appropriate treatment is provided. Average earnings in (5) and (6) are indicated in points.

	BASE	COMP	INS	INS-COMP
(1) consulting rate	40.7	54.7	55.3	83.1
(2) overtreatment rate	26.3	7.2	70.9	34.2
(3) efficiency rate	61.2	70.5	71.5	89.5
(4) correct treatment rate (CTR)	29.6	49.7	16.2	54.9
(5) average earnings physicians	9.1	11.5	14.4	19.1
(6) average earnings patients	6.8	7.2	5.7	6.4

ment were paid by the collective.[25] The overtreatment rate of 70.9% was about three times the level in CONTROL (26.3%). Physicians had additional incentives to overtreat because they assumed that patients were less concerned about the costs. Overall, only 16.2% of patients received a correct treatment in INS – compared to 29.6% in CONTROL.[26] However, the effect of insurance was stronger in the context of competition: Insurance increased the consulting rate by 52% from 54.7% to 83.1% and the overtreatment rate (34.2%) was about five times as high as in COMP (7.2%).

The lowest overtreatment rate (7.2%) was measured in

---

[23]Huck et al. (2016, p. 85)

[24]Huck et al. (2016, p. 87); The column BASE represents results from CONTROL.

[25]This effect was not significant according to a Wilcoxon-Mann-Whitney test (Huck et al., 2016). More information on the test results is provided in Appendix 5.

[26]To measure efficiency, the correct treatment rate instead of the efficiency rate is used since the latter does not take overtreatment as an inefficiency into account i.e., efficiency is high even when all patients consulted, but all were overtreated.

COMP compared to CONTROL (26.3%), INS (70.9%) and INS-COMP (34.2%) since competition provided incentives for physicians to avoid a severe treatment when a mild treatment would have been sufficient. Overtreating physicians were less likely to be consulted and therefore under pressure not to overtreat. Probably the most important finding was that competition[27] on the seller side outweighed some of the moral hazard effects: On the one hand, competition increased the consulting rate from 55.3% in INS to 83.1% in INS-COMP. But on the other hand, competition reduced physicians' overtreatment behavior by 48% from 70.9% to 34.2% yielding almost the level in CONTROL. As a result, the correct treatment rate raised from 16.2% in INS to 54.9% in INS-COMP.

In the following section, the methodology will be critically discussed. In general, an important advantage of laboratory experiments is the ability to control most aspects of the environment, but such experiments may have limited relevance for individuals' actual behavior in real-world situations (lack of external validity) since subjects typically know that they are part of an experiment and the environment might not be fully representative (e.g., students as subjects) (Richter et al., 2014; List and Reiley, 2008). Adverse selection was ruled out because patients' problems were drawn randomly for each round and there was no option to not insure or to choose different coverage levels. According to Huck et al. (2016), the findings should only be interpreted in a healthcare context with fee-for-service remuneration systems i.e., where physicians take advantage from offering high-level treatments. Hence, one drawback may be that the authors did not frame the experiment in a medical context (e.g., physicians were called "advisers"). The authors named several reasons for doing so, but this feature may have influenced patients' consulting decision since consumers are probably more sensitive about their decision when it comes to their health rather than in other contexts (e.g., problem with their car). Moreover, another disadvantage of the experiment is that patients did not suffer from physical consequences (e.g., pain) after not consulting a physician. Feeling such negative consequences may have had a stronger impact on the patients consulting decision for the following rounds than just learning the severity of the problem. Additionally, the difference in patients' payoffs between not consulting and consulting and receiving the right treatment was very small (see Figure 4). It should be mentioned that especially in the case of a severe problem this seems unrealistic although the severe treatment was very expensive. Contrariwise, the value for a person of being cured is difficult to measure and may differ from person to person. The experiment focused on over-treatment and, therefore, undertreatment and no treatment were excluded, but both cases may occur in real situations. Furthermore, it was assumed that physicians diagnose the problem correctly which is obviously an unrealistic assumption for the real world. Reputational incentives for physicians were weak since patients only knew whether their problem had been solved but had no idea about the necessity of the treatment. However, such incentives may be important to mitigate overtreatment since patients could have been more confident not to be overtreated, especially, in the context of competition. A doctor's reputation (e.g., internet portals like sanego) may have an influence on the number of consulting patients.

3.2. First-Degree Moral Hazard

Results from the prediscussed experiment also demonstrate evidence of first-degree moral hazard. Thus, the remaining part of this paper analyzes this phenomenon. Previous studies found, for instance, that moral hazard is less likely to occur under deterministic losses (Berger and Hershey, 1994) and with low probabilities of obtaining income loss compensation (Di Mauro, 2002).

Mol and Botzen (2018) were the first to experimentally study the existence of moral hazard in a market for natural disaster risk insurance. To be more specific, the causal effects of different financial incentives, probability levels, behavioral characteristics and deductibles on homeowners' damage reducing investments were examined.

In a laboratory experiment, participants played an investment game on computers for which they were randomly assigned to five different treatments: CONTROL, INS, treatment with premium discount (DISCOUNT), treatment with loan (LOAN) and treatment with loan and discount (LOAN-DISCOUNT). In CONTROL, subjects had no insurance coverage whereas in INS, participants were covered by insurance including a deductible. All treatments, except for CONTROL, implied insurance coverage and a deductible. In DISCOUNT, subjects were offered a premium discount proportional to their investment in damage reduction. In the fourth treatment – LOAN – participants were able to distribute their investment costs over multiple rounds at an interest rate of 1%. Subsequently, LOAN-DISCOUNT combined the previous two treatments.

At the beginning of the experiment, participants earned their starting capital through an effort task in order to purchase a virtual house which was prone to flood risk. The rest of the starting capital was subjects' savings which could have been used to pay for investments, insurance premiums and damages. Altogether, participants played 6 scenarios[28] consisting of 12 rounds with differing flood probabilities and deductible levels for each scenario (see Table 5). The scenarios started with the presentation of flood probability, estimated maximum flood damage and deductible level on subjects' screens. On the next page, which is displayed in Appendix 6, subjects were asked how much they wanted to invest in reducing the damage of a flood in the following rounds. For this purpose, five investment levels were proposed each specifying the costs for the investment, the amount by which the

---

[27]The authors state that the strong effect of competition was due to free choice of physician rather than to observability of market shares.

[28]After each scenario, the savings were automatically restored to the starting value (Mol and Botzen (2018).

Table 5: Overview of Scenarios[29]

Insurance treatments			
	Deductible		
	Extra Low (5%)	Low (5%)	High (20%)
Low probability (3%)	LxL	LL	LH
High probability (15%)	HxH	HL	HH

No Insurance treatments						
Probability	1%	3%	5%	10%	15%	20%

damage will be reduced and the resulting deductible if the house was flooded. After the subjects' decision, the premium and the investment costs were subtracted from the savings. Then, subjects observed a map indicating 100 houses under which their house was marked. The software randomly selected the flooded house(s) according to the flood probability and indicated the flooded house(s) in blue on the map (see Appendix 7). If a subject's house was flooded, the deductible – or the damage in CONTROL – was paid from his savings. In the following round, subjects could have either invested more or remain with their investment while a reduction was not allowed. Additionally, in LOAN an extra page to pay the loan costs was displayed to subjects. In the end, questions and decision tasks were presented to participants in order to obtain their behavioral characteristics.

The authors found out that average investments in damage reducing measures increased with the expected value of damage (i.e., higher probability of flood and/or higher deductible) for CONTROL and INS which can be seen in Figure 5. In the first round of INS, average investments were greater than zero for high and low-probability scenarios. Moreover, subjects invested significantly less in INS than in CONTROL in scenarios with high probabilities (15%) while such an effect was not found for low probabilities (3%). These results suggest the existence of moral hazard in scenarios with high probabilities, but not under low probabilities. Therefore, moral hazard may be less of a problem in natural disaster insurance markets with low probabilities of loss and high expected damages.

The results from Table 6 prove that premium discounts increased investments in damage mitigation significantly compared to INS. But in LOAN, participants were not more likely to invest more than in INS. Consequently, the combination of loan and discount did not generate the highest investments as hypothesized by the authors.[30]

Furthermore, participants' behavioral characteristics such as risk aversion, perceived effectiveness of protective methods and concern about flooding had a positive impact on the investment decision. However, females invested significantly more than males and individuals with high incomes in real life invested less compared to low-income participants. It was also found that subjects who had already experienced a flood invested extra in damage reduction afterwards, but this effect was not found when a neighbor's house had been flooded.

One drawback – as in many other laboratory experiments – is that study participants were largely students who may have not been representative subjects to study moral hazard in a flood insurance context because of less knowledge and experience with damage reducing investments compared to real homeowners. This could explain the result from LOAN since students may have generally disliked lending money or were put off by the interest rate (Mol and Botzen, 2018). In addition, investment costs were distributed across 12 rounds which lasted in the experiment at most several minutes while costs are spread over multiple years in the real world (Mol and Botzen, 2018). Smith (1982) stresses that salient payoffs – rewards for individuals' participation in the experiment that are related to participants' realized outcomes – are important in laboratory experiments.[31] Such payments need to be incentive compatible i.e., payments create incentives for subjects to behave according to their real preferences (Jaspersen, 2016). Therefore, an advantage of the experiment was the implementation of an incentive-compatible payment scheme: At the end of the experiment, all subjects were paid their final savings from one randomly selected scenario additionally to a participation fee (Mol and Botzen, 2018). Another point is that, for ethical and practical reasons, it is not possible to let subjects lose money for real in an experiment (Etchart-Vincent and l'Haridon, 2011; Jaspersen, 2016). In order to solve this problem, subjects had to perform an effort task in which they earned their initial endowment from which they lost without affecting their own money. It is important in experiments to make subjects believe that the earned (and lost) money is theirs in order to make them aware of losing instead of gaining money in the game. Otherwise, subjects may keep their endowment in mind when making decisions and consider their outcomes as gains causing biases in their

---

[29]Mol and Botzen (2018, p. 8)

[30]Indeed, premium discounts alone led to the highest investments in the game (Mol and Botzen, 2018).

[31]Camerer and Hogarth (1999) found out that salient rewards change participants' behavior in experiments.
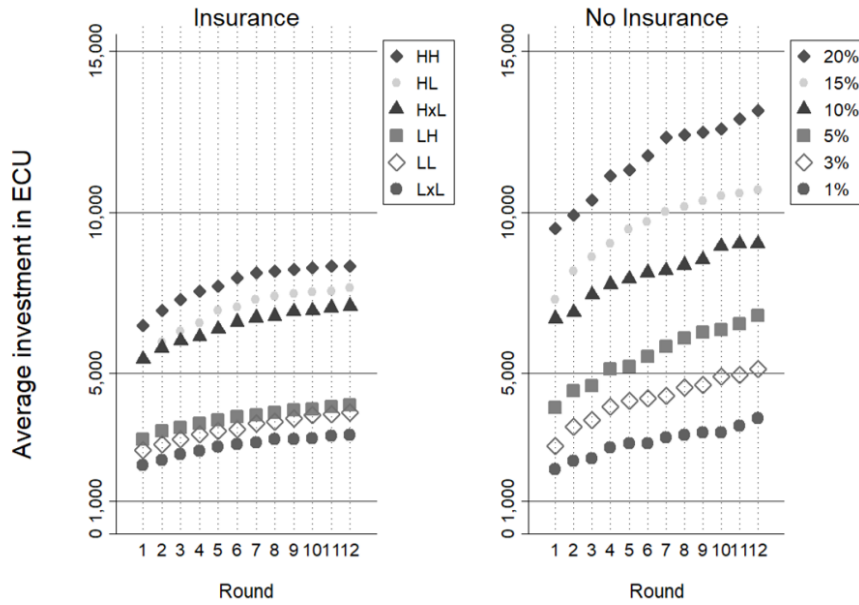
Figure 5: Average Investment in INS and CONTROL[32]

Table 6: Average Investment in the First Round for All Insurance Treatments[33]

	Baseline Insurance	Loan	Discount	Loan+Discount	Kruskall-Wallis test
scenario HH	5,421.49	3,711.86	9,233.33	8,614.04	$\chi^2 = 37.670$***
	(5,431.01)	(3,658.01)	(5,732.35)	(5,512.18)	
scenario HL	4,049.59	2,847.46	8,416.67	7,807.02	$\chi^2 = 43.713$***
	(4,843.98)	(3,916.43)	(5,681.64)	(5,717.89)	
scenario HxL	3,471.07	3,542.37	8,966.67	7,771.93	$\chi^2 = 46.829$***
	(5,010.11)	(5,032.04)	(5,971.59)	(5,840.19)	
scenario LH	2,727.27	1,661.02	3,850.00	3,719.30	$\chi^2 = 10.086$**
	(4,222.95)	(3,412.00)	(4,398.86)	(4,806.08)	
scenario LL	2,404.96	1,525.42	3,283.33	3,421.05	$\chi^2 = 10.842$**
	(4,253.58)	(3,650.02)	(4,584.76)	(5,119.81)	
scenario LxL	1,793.39	1,406.78	3,550.00	2,087.72	$\chi^2 = 19.308$***
	(3,976.84)	(3,312.04)	(4,560.05)	(3,434.49)	
Observations	121	59	60	58	

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$, st.dev in parentheses.

behavior (Harbaugh et al., 2010; Jaspersen, 2016).[34]

Moral hazard in teams[35] may arise when team members bear the costs of their effort for supplying inputs individually, but only the joint output is observable directly (Holmstrom, 1982). This can cause a free riding problem: Agents can cheat and rely on the performance of the remaining team members when they are paid according to the team output (Holmstrom, 1982; Corgnet et al., 2013). Corgnet et al.

(2013) investigated whether peer pressure[36] through monitoring is a solution to the problem of moral hazard in teams. For this purpose, four treatments were designed: CONTROL, treatment with team incentives (T), treatment with team incentives and visible peer monitoring (TVP) and treatment with team incentives and invisible peer monitoring (TIP).

In a laboratory experiment, participants had to do a long, repetitive and effortful work task which consisted of summing up tables. When a subject had completed a table, he received information about his accumulated individual production: The production increased by 40 Cents when the table

---

[32]Mol and Botzen (2018, p. 17)
[33]Mol and Botzen (2018, p. 19); The column Baseline Insurance represents results from INS.
[34]However, Etchart-Vincent and l'Haridon (2011) found the contrary.
[35]Holmstrom (1982) defines a team as a group of individuals whose productive inputs are related.

[36]Mas and Moretti, 2009 define social (or peer) pressure as the experience of disutility when being observed working less hard than others.

was completely correct and decreased by 20 Cents for each incorrect given answer in the table. Moreover, at the end of each period, participants were informed about the total profit generated by their team (10 group members). Anytime during the experiment, subjects could have surfed the Internet which was an alternative activity to the work task not generating any profits. Since both activities were undertaken on different screens completing tables while browsing simultaneously was not possible. Additional to the prementioned activities, subjects could have clicked on a yellow box on their screen which was the clicking task. The box appeared every 25 seconds on the screen and by clicking on it, 5 Cents were earned by the subject.[37] As a consequence, subjects could have earned money constantly without actually working on the working task and while being on the Internet. In CONTROL, subjects had individual incentives and received payments for the work task according to their individual production whereas in T, rewards were based on the total production of all group members (subjects obtained 10% of the total production). The third treatment variation – TVP – was similar to T except for the introduction of an option for peer monitoring in order to create peer pressure. Subjects were allowed to click on a watch option on their screen to observe other participants' activities in real time. During monitoring others, the working task and the leisure activity could not have been continued while proceeding with the clicking task was possible. After selecting the watch option, monitors were informed about each subject's activities (work task, browsing or watching), production and the individual input to the work task expressed as a percentage. Additionally, monitored subjects received a notification on their screen that they were currently being watched. In TIP, participants did not receive such a notification.

The results indicated that individual production[38] per period increased (except for period 3) under individual and team incentives showing evidence of a learning effect. Subjects evolved strategies to compute the tables more quickly. Figure 6 illustrates the interesting finding that average production per subject was significantly lower in T (2.83 tables) than in CONTROL (4.21 tables) yielding a difference of 49% between the two incentive schemes due to shirking behaviors.

The following results are important since a highly significant negative correlation between Internet usage and individual production for treatments with individual and team incentives was detected: A comparison of Internet usage revealed that the average time spent with browsing was significantly higher in T (28.5%) than in CONTROL (11.9%) which can also be seen in Figure 7.[39] Under team incentives, the

average proportion of time spent on the Internet of 28.5% decreased to 13.1% with the introduction of peer monitoring in TVP resulting in values almost similar to CONTROL (see Figure 7). This showed a clear impact of peer monitoring on subjects' choice of activity. Especially, visible monitoring was effective since Internet usage was significantly lower in TVP than in TIP.

Average production was 47% higher under peer pressure (in TVP) than in T which was interpreted as evidence of a strong peer pressure effect while no significant differences between TVP and CONTROL were found as shown in Figure 8. Therefore, visible peer monitoring combined with team incentives allowed production levels as high as under individual incentives supporting the authors' expectation that peer pressure eliminates the problem of moral hazard in teams. Social pressure was essential for the effectiveness of monitoring since production levels were significantly lower in TIP than in TVP and almost as high as under team incentives.

An advantage of the methodology was that subjects could have switched to the leisure activity since access to the Internet at the workplace is very common in organizations and according to a recent survey of Salary.com (see Appendix 8), 64% of employees visit websites which are not related to their working activity every day. The study also revealed that one of the most time-consuming activities employees waste their time with on the job is surfing the Internet. Corgnet et al. (2013) conducted the invisible monitoring treatment (TIP) with the objective of eliminating the role of social pressure in contrast to TVP. Yet, subjects knew about the possibility of monitoring others and may have felt watched even without a notification on their screen. Therefore, social pressure may not have been completely eliminated (Corgnet et al., 2013). This hypothesis was supported by the finding of a slight difference in Internet usage between TIP (19.8%) and T (28.5%). In the experiment, intrinsic motivation[40] was reduced through the introduction of a long and laborious work task because of the aim to investigate behavior under different incentive schemes. Corgnet et al. (2013) stated that intrinsic motivation would have been a confounding factor, but individual production may not always be driven only by extrinsic motivators such as the payment. For instance, the work itself or recognition should also be taken into consideration when conducting experiments on teamwork. Only large teams consisting of ten individuals were studied, but much work in organizations is performed by small teams. By keeping teams small it may be easier to increase the transparency of subjects' individual contribution to the output possibly even without monitoring.

In the absence of peer pressure, Biener et al. (2018) stud-

---

[37]This feature represented the payment an employee receives only for being present at his workplace (Corgnet et al., 2013).

[38]Production is defined as the monetary amount generated from working on the work task divided by 40 Cents. Thus, production is the number of correctly computed tables minus the number of false answers (Corgnet et al., 2013).

[39]According to the authors, 40.9% and 11.7% of subjects never surfed the Internet under individual and team incentives, respectively.

[40]Intrinsic motivation is defined as performing an activity because of the activity itself (perceived enjoyment) and not because of achieving valued outcomes (perceived usefulness) (Teo et al., 1999).

[41]Corgnet et al. (2013, p. 16)

[42]Corgnet et al. (2013, p. 26)
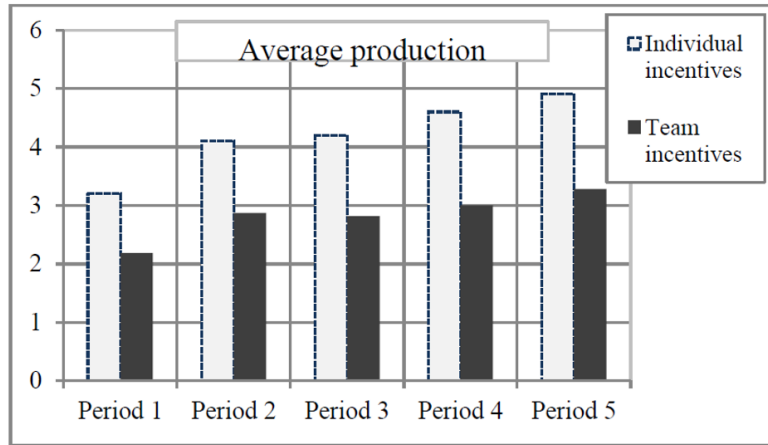
[43]Corgnet et al. (2013, p. 23)

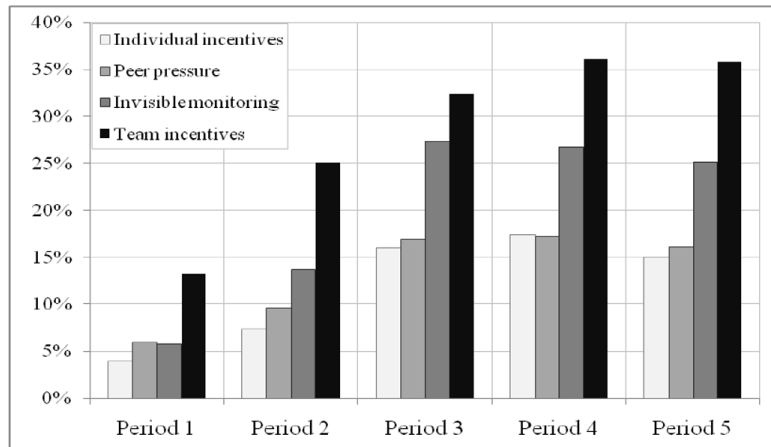**Figure 6:** Average Production in CONTROL and T[41]

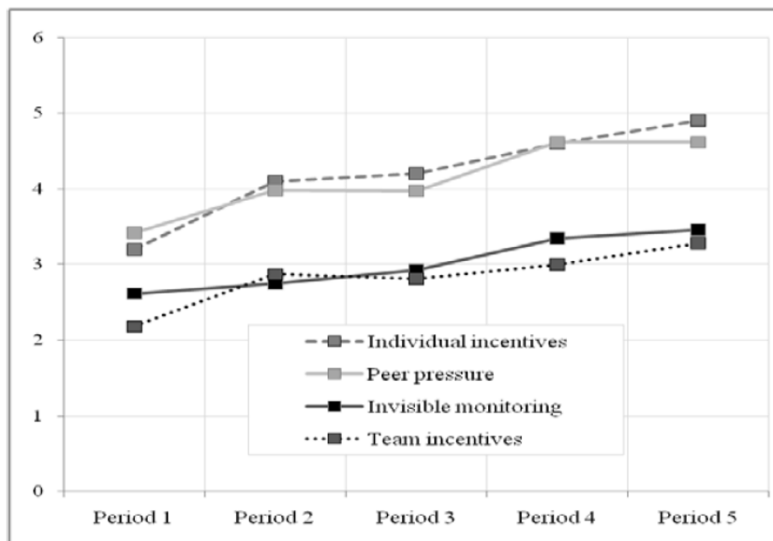**Figure 7:** Average Internet Usage for All Treatments[42]

Figure 8: Average Production for All Treatments[43]

ied whether pro-social preferences[44] between agents mitigate moral hazard in joint liability group contracts. At the beginning of the experiment, all participants played the same effort game for three rounds: Subjects obtained an initial endowment[45] W and were confronted with a lottery in which they had to draw one of ten balls from a bag containing four orange and six white balls. The orange balls indicated a loss of L while white balls represented no loss. Subjects were offered the opportunity to self-protect by replacing the bag with another bag with two orange and eight white balls in exchange for effort costs of e. The probabilities and payoffs for the basic and the self-protection game are presented in Table 7.

Subsequently, subjects were randomly assigned to one of five treatments for the remaining three rounds: CONTROL, treatment with individual insurance and low coverage (Ilow), treatment with individual insurance and high coverage (Ihigh), treatment with group insurance and private information ($G_{private}$) and treatment with group insurance and public information (G_{public}). In CONTROL, subjects played the above-described game for all rounds. The insurance treatments – in which subjects paid a premium in each round in order to only bear the deductible in case of a loss – were divided into individual insurance and group insurance. The individual insurance treatments Ilow and Ihigh included different coverage levels i.e., deductibles – a high deductible leading to low and a low deductible resulting in high coverage, respectively. The group insurance treatments $G_{private}$ and G_{public} were similar to Ilow concerning deductible and premium, but insurance covered groups of two individuals and losses which were not covered by insurance (i.e., losses below the deductible) were shared among them. In $G_{private}$, self-protection was private information which means that the peer was not able to observe the other's risk-taking behavior whereas in G_{public} this information was provided at the end of the round.

The experiment was conducted two times. First, with rural villagers from the Philippines who brought a friend or a relative of choice to the experiment. Therefore, it was possible for the experimenters to manipulate the network strength within the groups. Second, in a computer laboratory setting, the behavior of students from Germany was studied. In this variation, participants could not have identified their group members.

Biener et al. (2018) found strong evidence for moral hazard meaning that subjects' effort to self-protect decreased with increasing insurance coverage, i.e., when making the payoff less state-dependent as illustrated in Figure 9. From Figure 9 one can also observe that this effect was significantly stronger in the German sample. However, moral hazard was mitigated with joint liability in the group scheme. In $G_{private}$, self-protection increased compared to Ilow by 6.8 percentage points in the Philippine sample and relative to Ihigh by 15.6 and 27.3 percentage points in the Philippine and the German sample, respectively. According to the authors, this effect was driven by individuals with positive beliefs about their peer's self-protection behavior. Subjects with positive beliefs (bar (a)) were more likely to self-protect themselves compared to individuals with negative beliefs (bar (b)) as can be seen in Figure 10. This indicates that positive beliefs about the group member increased pro-social motives and trust in the peer influenced the subject's own behavior. Also, results suggested that strategic motives (such as fear of punishment) can further improve effort provision in non-anonymous groups: Slightly higher self-protection was detected in G_{public} compared to $G_{private}$, but only for the Philippine sample. For both samples, mean self-protection for subjects with positive beliefs was almost similar in both group insurance treatments while only in the Philippine experiment, a difference was found for subjects with negative beliefs. To be more specific, the mean proportion of self-protecting individuals was 18.3 percentage points higher in G_{public}. These results are presented in Appendix 9.

Network strength, i.e., group composition, was expected to influence the degree of pro-social and strategic motives to provide self-protection. Indeed, in the Philippine experiment, whether a group was formed out of friends and relatives (strong group) or of random persons (weak group) did not cause a significant difference in outcomes of $G_{private}$ and G_{public} (see Figure 11). It was only found that the outcome difference between G_{public} and $G_{private}$ was larger in strong groups. A possible reason may be that, if self-protection behavior is observable for the peer, image concerns play a larger role in non-anonymous groups. Anyway, it is important to mention that even random individuals from the same village may have known each other.

With regard to the methodology, it can be said that one major advantage was the execution of the experiment in two independent and different environments. Once, with low-income rural villagers from the Philippines and again with students from Germany. In consideration of the fact that the main results were found for both experiments and, therefore, hold independently of culture and experience with financial services, one can attribute them further credibility, according to Biener et al. (2018). In contrast to the experiment by Mol, Botzen and Blasch, subjects did not earn their initial endowment through an effort task. Instead, they received a so-called windfall payment at the beginning meaning that participants were endowed with money without consideration (Jaspersen, 2016). According to Jaspersen (2016), the experimental design was constructed incentive compatible through using a "random problem selection mechanism". This means that subjects had to take multiple self-protection decisions during the experiment, but earnings for subjects'

---

[44]Meaning that individuals are also concerned with other individuals' payoffs, not only with their own ones (Levitt and List, 2007).

[45]After each round, the initial endowment was restored to the starting value (Biener et al., 2018).

[46]Own representation based on Biener et al. (2018)

[47]Biener et al. (2018, p. 241); The bar C represents results from CONTROL.

[48]Biener et al. (2018, p. 241); The bars represent the proportion of individuals choosing self-protection who had positive (a) and negative (b) beliefs about their peer's self-protection effort.

Table 7: Payoffs and Probabilities[46]

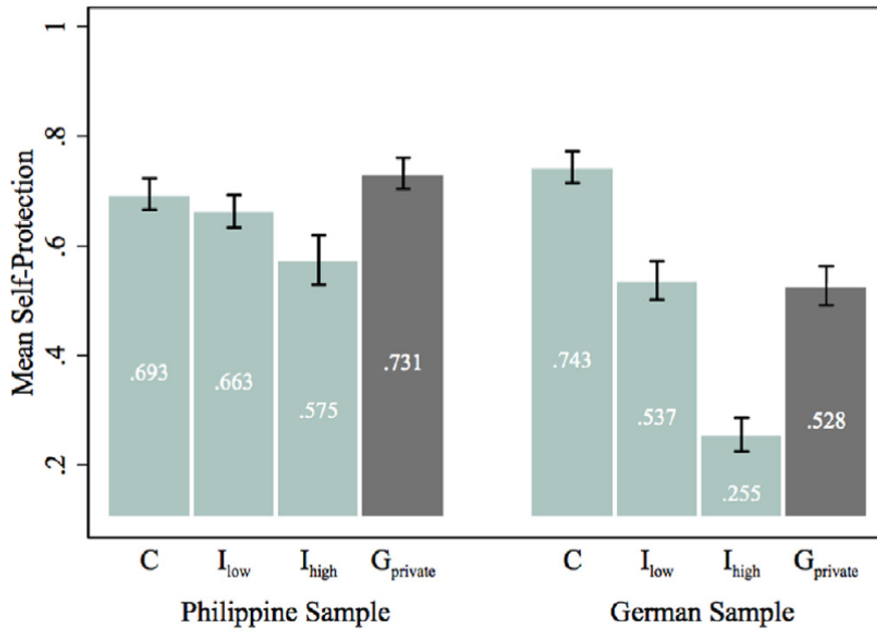|	Payoff (Probability)	
	Basic Game	Self-Protection Game
No Loss	W (60%)	W-e (80%)
Loss	W-L (40%)	W-e-L (20%)

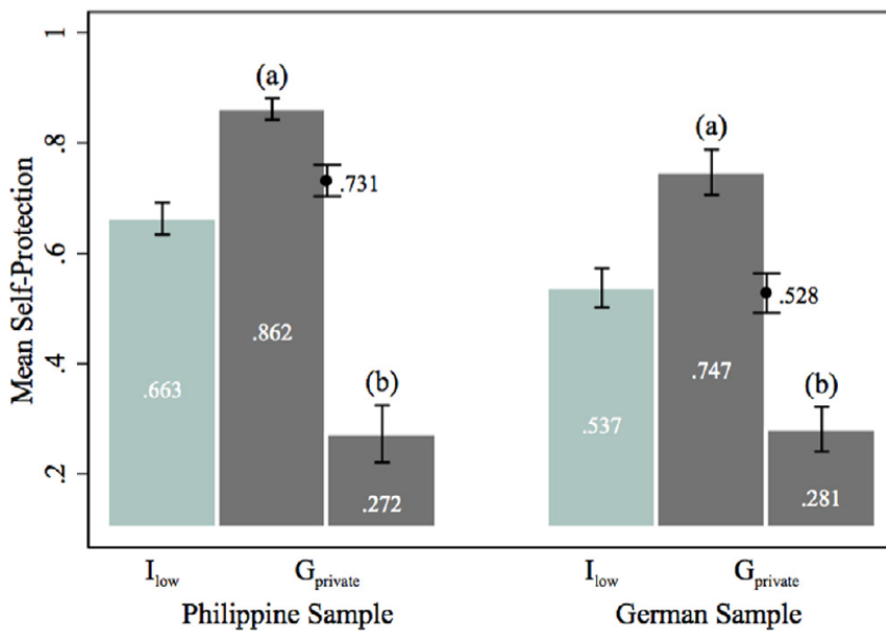**Figure 9:** Mean Proportion of Self-Protecting Participants[47]

Figure 10: Mean Proportion of Self-Protecting Participants by Beliefs[48]
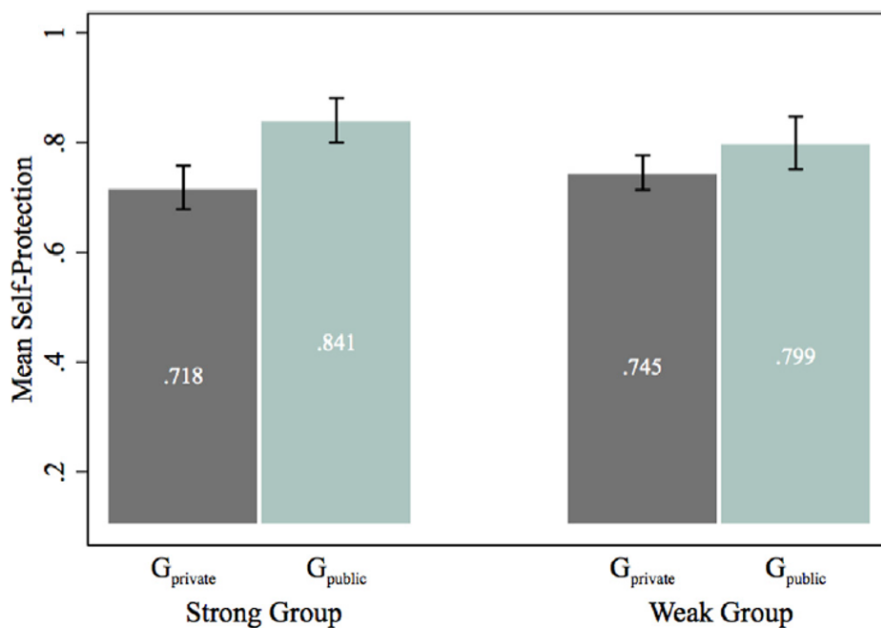
Figure 11: Mean Proportion of Self-Protecting Participants by Group Composition[49]

participation were determined by randomly selecting one round out of six (in addition to a show-up fee). Biener et al. (2018, p. 237) write in their paper that they "explained the [...] payout mechanism to all participants in advance". It remains unclear whether the term "in advance" means that participants received this information several days or minutes before the experiment. Actually, the point of time of receiving this information is crucial for participants' behavior in experiments since subjects who find out about receiving payment for their participation not before arriving at the experiment are less risk averse compared to individuals who are told one to five days before (Arkes Hal et al., 1994; Jaspersen, 2016). Another drawback may be that initial endowments varied between the two experiments. For the Philippine sample, the endowment was considerably above participants' average daily income while it was approximately 70% of the daily income of the German students (Biener et al., 2018). To stress the impact of such payments on behavior, consider that Ackert et al. (2006) found evidence in an experiment where they varied initial endowments that traders who obtained a higher initial endowment at the start were more risk-taking resulting in higher bids.

4. Conclusion

The purpose of this paper was to examine the circumstances under which moral hazard is likely to occur and how this problem could be mitigated or eliminated. To summarize, moral hazard had a significant impact on service provision in markets for credence goods. It was found that moral

hazard on the demand side, in addition to personal financial incentives for service providers to behave fraudulently, led to second-degree moral hazard. When taxi drivers were aware of the moral hazard problem between the passenger and his employer, they were more likely to overcharge (Balafoutas et al., 2017). This is in line with Kerschbamer et al. (2016) findings. However, Lu (2014) showed that the elimination of such personal incentives caused non-existence of second-degree moral hazard: Physicians who expected to receive a fraction of patients' drug expenditures wrote 43% more expensive prescriptions to insured while doctors without financial incentives did not respond to the insurance status. Moreover, Huck et al. (2016) findings revealed the powerful effects of competition between experts to mitigate second-degree moral hazard. Indeed, insurance induced moral hazard on both market sides, but competition outweighed some of these effects. On the one hand, the consulting rate was in fact significantly higher with competition, but on the other, overtreatment was moderated. As already explained, this experiment also showed evidence of first-degree moral hazard in a health insurance context. In contrast, moral hazard was found to be less of a problem in natural disaster insurance markets with low probabilities of damages (Mol and Botzen, 2018). Insurance premium discounts and higher expected damages (i.e., higher probabilities of loss and higher deductibles) increased homeowners' investments in damage reduction measures (Mol and Botzen, 2018). As expected, moral hazard was present when team members were paid according to their joint output. Average production levels were significantly lower and Internet usage was significantly higher under team incentives than under individual incentives. By introducing visible peer monitoring, subjects' performance under team incentives was as high as under individ-

---

[49]Biener et al. (2018, p. 244)

ual incentives supporting the authors' expectation that peer pressure eliminates the problem of moral hazard in teams (Corgnet et al., 2013). Also studying joint liability group contracts, but in an insurance context, Biener et al. (2018) stressed the role of pro-social preferences to alleviate moral hazard. The higher subjects' insurance coverage was, the more decreased mean effort to self-protect against losses, but when insuring groups of two, self-protection increased because subjects were motivated by pro-social concerns.

Overall, the results suggest moral hazard to be an important problem in many markets. From the considered experimental studies, circumstances for its occurrence and mitigation measures were derived (see Appendix 10 for an overview). However, as Cohen and Siegelman (2010) proposed for proceeding with the research on adverse selection, one should identify the circumstances under which moral hazard emerges – as it was done in this paper – instead of aiming at a once-and-for-all conclusion on its existence.

In the context of second-degree moral hazard, some general research questions are still unanswered. First, little attention has been paid on how different insurance schemes such as co-payments affect sellers' provision behavior (Kerschbamer and Sutter, 2017). Second, even though insurance companies often reimburse costs for services only if they were provided by their contract partners, Kerschbamer and Sutter (2017) are not aware of studies investigating whether insurance companies' partners actually behave less fraudulently. Third, although Lu (2014) found out that experts who did not profit from overtreatment or overcharging did not prescribe insured and uninsured individuals differently, the relationship between personal incentives and provision behavior and the impact of different incentive schemes for expert sellers have not been analyzed in detail (Kerschbamer and Sutter, 2017). Fourth, the experiment with competition only implemented free choice of physicians, but the effect of price-competition may be of importance as well. Fifth, in terms of the presence of the Digital Age, one topic for future research could also be the impact of digital technologies as for example platforms like "Uber" and "sanego" on mitigating moral hazard. Such platforms may allow for decreasing informational asymmetries between sellers and customers in credence goods markets because of customers' possibilities to rate their experiences with a specific physician or to real-time monitor their "Uber" driver. And, since the previous point is in a way related to reputational incentives it may also be important to study how those affect service provision behavior.

With regard to first-degree moral hazard, future research could address the behavior of actual homeowners, instead of students, in regions that are prone to flood risk in order to increase representativeness. An additional point could be the investigation of the effect of financial incentives in natural disaster insurance markets where insurance is not mandatory. Cohen and Siegelman (2010) write that the existence of moral hazard is an important reason for using deductibles in insurance contracts. However, in the experiment by Mol and Botzen (2018), deductibles had an influence on ex ante damage reduction, but this effect was rather small which

raises the question of the usefulness of high deductibles in such markets pointing out a further research topic. Furthermore, since behavior in group insurance schemes may depend on the group size it would be interesting to study the effect of different group sizes on moral hazard and, especially, whether moral hazard can also be mitigated with an increasing group size where free riding may be more likely to remain undetected and pro-sociality may decrease. Another academic void is the existence of ex post moral hazard in group insurance schemes. Finally, the finding that strategic motives can further improve subjects' effort provision in non-anonymous groups could have been driven by cultural factors since strong groups were only studied in the Philippine sample (Biener et al., 2018). Therefore, it would be interesting for follow-up research to investigate whether this effect is also robust across cultures and social classes.

References

Ackert, L. F., Charupat, N., Church, B. K., and Deaves, R. An experimental examination of the house money effect in a multi-period setting. *Experimental Economics*, 9(1):pp. 5–16, 2006.

Arkes Hal, R., Joyner Cynthia, A., Pezzo Mark, V, Gradwohl, N. J., Karen, S.-J., and Eric, S. The psychology of windfall gains. *Organizational Behavior and Human Decision Processes*, 59(3):pp. 331–347, 1994.

Arnott, R. and Stiglitz, J. E. Moral hazard and nonmarket institutions: Dysfunctional crowding out of peer monitoring? *The American Economic Review*, 81(1):pp. 179–190, 1991.

Arrow, K. Uncertainty and the welfare economics of medical care. *The American Economic Review*, 53(5):pp. 941–973, 1963.

Arrow, K. Agency and the market. In: K.J. Arrow and M.D. Intriligator, Publ. 1986. Handbook of Mathematical Economics. Vol. 3. [e-book] Amsterdam et al.: North-Holland Pub. Co. , 1986. URL http://www.ub.uni-muenchen.de/. accessed on 13.04.2019.

Balafoutas, L., Kerschbamer, R., and Sutter, M. Second-degree moral hazard in a real-world credence goods market. *The Economic Journal*, 127(599):pp. 1–18, 2017.

Berger, L. A. and Hershey, J. C. Moral hazard, risk seeking, and free riding. *Journal of Risk and Uncertainty*, 9(2):pp. 173–186, 1994.

Biener, C., Eling, M., Landmann, A., and Pradhan, S. Can group incentives alleviate moral hazard? The role of pro-social preferences. *European Economic Review*, 101:pp. 230–249, 2018.

Camerer, C. F. and Hogarth, R. M. The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19(1-3):pp. 7–42, 1999.

Cohen, A. and Siegelman, P. Testing for adverse selection in insurance markets. *Journal of Risk and Insurance*, 77(1):pp. 39–84, 2010.

Corgnet, B., Hérnan-Gonzalez, R., and Rassenti. Peer pressure and moral hazard in teams: Experimental evidence. Working Paper, No. 13-01, Orange, CA, Economic Science Institute at Chapman University, 2013.

Di Mauro, C. Ex ante and ex post moral hazard in compensation for income losses: results from an experiment. *The Journal of Socio-Economics*, 31(3):pp. 253–271, 2002.

Dulleck, U. and Kerschbamer, R. On doctors, mechanics, and computer specialists: The economics of credence goods. *Journal of Economic Literature*, 44(1):pp. 5–42, 2006.

Ehrlich, I. and Becker, G. S. Market insurance, self-insurance, and self-protection. *Journal of Political Economy*, 80(4):pp. 623–648, 1972.

Einav, L. and Finkelstein, A. Moral hazard in health insurance: What we know and how we know it. *Journal of the European Economic Association*, 16(4):pp. 957–982, 2018.

Etchart-Vincent, N. and l'Haridon, O. Monetary incentives in the loss domain and behavior toward risk: An experimental comparison of three reward schemes including real losses. *Journal of Risk and Uncertainty*, 42(1):pp. 61–83, 2011.

Finkelstein, A. and McGarry, K. Multiple dimensions of private information: evidence from the long-term care insurance market. *American Economic Review*, 96(4):pp. 938–958, 2006.

Harbaugh, W. T., Krause, K., and Vesterlund, L. The fourfold pattern of risk attitudes in choice and pricing tasks. *The Economic Journal*, 120(545):pp. 595–611, 2010.

Holmström, B. Moral hazard and observability. *The Bell Journal of Economics*, 10(1):pp. 74–91, 1979.

Holmstrom, B. Moral hazard in teams. *The Bell Journal of Economics*, 13(2):pp. 324–340, 1982.

Huck, S., Lünser, G., Spitzer, F., and Tyran, J.-R. Medical insurance and free choice of physician shape patient overtreatment: A laboratory experiment. *Journal of Economic Behavior & Organization*, 131(B):pp. 78–105, 2016.

Jaspersen, J. G. Hypothetical surveys and experimental studies of insurance demand: A review. *Journal of Risk and Insurance*, 83(1):pp. 217–255, 2016.

Karten, W., Nell, M., Richter, A., and Schiller, J. Risiko und Versicherungstechnik: Eine ökonomische Einführung. [e-book] Wiesbaden: Springer Gabler. Under: LMU Universitätsbibliothek, 2018. URL http://www.ub.uni-muenchen.de/. accessed on 13.04.2019.

Kerschbamer, R. and Sutter, M. The economics of credence goods - A survey of recent lab and field experiments. *CESifo Economic Studies*, 63(1):pp. 1–23, 2017.

Kerschbamer, R., Neururer, D., and Sutter, M. Insurance coverage of customers induces dishonesty of sellers in markets for credence goods. *Proceedings of the National Academy of Sciences*, 113(27):pp. 7454–7458, 2016.

Levitt, S. D. and List, J. A. What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, 21(2):pp. 153–174, 2007.

List, J. A. and Reiley, D. *Field Experiments. In: S.N. Durlauf and L.E. Blume, pp. 2117–2120*. The new Palgrave dictionary of economics. 2nd edition. London: Palgrave Macmillan, 2008.

List, J. A. Field experiments: A bridge between lab and naturally occurring data. *The B.E. Journal of Economic Analysis & Policy*, 6(2):pp. 1–45., 2006.

Lu, F. Insurance coverage and agency problems in doctor prescriptions: evidence from a field experiment in China. *Journal of Development Economics*, 106:pp. 156–167, 2014.

Mas, A. and Moretti, E. Peers at work. *American Economic Review*, 99(1):pp. 112–45, 2009.

Mol, J. and Botzen, J., W.J.W. and Blasch. Risk reduction in natural disaster insurance: Experimental evidence on moral hazard and financial incentives. Working Paper, No. 2018-10, University of Pennsylvania, Philadelphia, 2018.

Nell, M. Das moralische Risiko und seine Erscheinungsformen. In: Frankfurter Vorträge zum Versicherungswesen. Volume 29. Karlsruhe: Verlag Versicherungswirtschaft, pp. 7-33, 1998.

Pauly, M. V. The economics of moral hazard: Comment. *The American Economic Review*, 58(3):pp. 531–537, 1968.

Richter, A., Schiller, J., and Schlesinger, H. Behavioral insurance: Theory and experiments. *Journal of Risk and Uncertainty*, 48(2):pp. 85–96, 2014.

Salary.com. Articles. [online], 2018. URL https://www.salary.com/articles/why-how-your-employees-are-wasting-time-at-work/. accessed on 21.04.2019.

Smith, V. L. Microeconomic systems as an experimental science. *The American Economic Review*, 72(5):pp. 923–955, 1982.

Teo, T., Lim, V., and Lai, R. Intrinsic and extrinsic motivation in Internet usage. *The International Journal of Management Science*, 27:pp. 25–37, 1999.

Viaene, S. and Dedene, G. Insurance fraud: Issues and challenges. *The Geneva Papers on Risk and Insurance*, 29(2):pp. 313–333, 2004.

Yip, W. and Hsiao, W. C. The Chinese health system at a crossroads. *Health Affairs*, 27(2):pp. 460–468., 2008.

Junior Management Science

The Effect of Gratitude on Individuals' Effort – A Field Experiment

Oriana Wendenburg

Universität zu Köln

Abstract

This study uses a real-effort survey experiment to investigate whether expressions of gratitude induce reciprocal behaviour and hence significantly increase individuals' effort. I extend existent literature by exploring non-pecuniary gifts that signal different degrees of gratitude, all combined with an interpersonal element. Based on a formal model, I hypothesize that a greater amount of gratitude is accompanied by higher levels of provided effort. The results show that appreciation in form of a thank you note positively affects reciprocal effort choice, compared to receiving no gratitude. An even higher level of gratitude conveyed in form of a video clip, however, does not impel subjects to provide more effort. Moreover, while I detect women to behave more reciprocally than men, this effect is least present in the gratitude treatments. These insights provide valuable implications for experimental research as well as for organizations and modern labour markets, emphasizing that non-monetary gifts, such as expressions of appreciation, are a cost-effective tool for human resource management to determine workers' effort.

Keywords: Gratitude; non-pecuniary gifts; gift giving; reciprocity; personnel economics.

1. Introduction

Various research studies in personnel economics have highlighted the importance of deepening the understanding of what motivates people to devote time and effort to their work (Kube et al., 2012; Carpenter and Gong, 2016; Kovach, 1987; Lazear, 2000). This question is of great interest as highly motivated workers are the key for long-term success for many organizations. Data from numerous studies in management and organizational psychology have shown that individuals who feel valued and treated fairly by their employer are more encouraged to work well (Kovach, 1987; Baker et al., 1988; Bradler et al., 2016). This phenomenon can be explained by the consequential development of social-exchange relationships between supervisor and subordinate which triggers effective work behaviour and organizational commitment (Rhoades and Eisenberger, 2002). According to Akerlof (1982) famous gift-exchange theory, treating workers kindly (e.g. by paying above clearing market wages) leads to positive sentiments towards the employer and induces employees to reciprocate positively to the "gift" by working harder. Economists have so far primarily focused on purely financial gifts as a main mean for valuing individuals' work which is in line with early management theories that stress the role of financial incentives as a main drive to impel motivation and performance (Fehr et al., 1993; Fehr and

Falk, 1999). These theories lead back to the homo economicus model of humanity initiated by neoclassical economists in the later 20th century (Aspromourgos, 1986). The construct of the homo economicus is based on the idea that individuals act out of pure self-interest in order to achieve the highest possible well-being for him or herself. However, latest research streams argue that the homo economicus perspective cannot explain all observed behavioural regularities. More precisely, psychologists and sociologists challenged the view of neoclassical economists by demonstrating that there are other sources of employee motivation besides monetary rewards (Bandura, 1986; Stajkovic and Luthans, 2003). Indeed, a substantial body of recent empirical findings in behavioural economics strongly suggest that gift exchange can be triggered by non-monetary elements. In particular, mounting evidence anticipates that purely non-financial gifts such as individual expressions of appreciation, social recognition, and management attention significantly induce higher effort which in turn associates with increasing performance[1] (Bradler et al., 2016; Ellingsen and Johannesson,

---

[1]Remarkably, a small amount of literature contrarily reports detrimental effects of non-monetary incentives. Hammermann and Mohnen (2014), for example, show in a real-effort experiment that the performance of subjects who received a monetary prize exceed those of participants in pursuit of a non-monetary incentive.

2007; Kirchler and Palan, 2018; Kosfeld and Neckermann, 2011; Kube et al., 2012). A recent study by Kirchler and Palan (2018), for instance, states that compliments elicit reciprocity and might be preferable to financial gifts. Kube et al. (2012) provide strong evidence that individual's effort is lower for a financial, than for a non-financial gift of equal value, suggesting that non-pecuniary elements might be more effective when it comes to triggering reciprocity. Further evidence on the effectiveness of non-monetary gifts is provided by Vogelsang (2019), who claims that the gift of more leisure time has a positive effect on employee's performance. Similarly, evidence from a field experiment conducted in China reports that allowing employees to work from home may be seen as a gift and is as such reciprocated by working harder (Bloom et al., 2015). In a study on worker performance, Kosfeld and Neckermann (2011) highlight the motivating power of social recognition and public awards. This is in line with Bradler et al. (2016) who argue that recognition, especially when provided exclusively to a subset of high performers, can be a cost-effective tool for increasing average work effort. Nevertheless, as social awards provide selected recipients with public status, they contain an extrinsic component and differ therefore from purely intrinsic non-pecuniary rewards (Frey, 2007).

While the existing literature provides valuable insights on the effect of several non-financial gifts on the reinforcement of reciprocal behaviour, much less is known about the effect of differing levels of gratitude expressions in particular[2].

Subsequently, the present study attempts to fill this gap and addresses the following research question: Does the expression of gratitude induce reciprocal behaviour and hence increase individuals' effort, and more specifically, do efforts significantly differ for varying gratitude levels?

A field experiment was used to explore the extent to which different forms of gratitude expressions, in combination with a personal touch, elicit reciprocity. Gratitude is conveyed via a thank you video and a thank you message, both transmitted electronically to the recipient. A handmade element was included in both settings to signal the recipient that the experimenter invested time and effort to show appreciation[3].

The paper proceeds as follows: After providing a thorough understanding of the meaning of gratitude and its interrelation with the concept of motivation, a formal model outlining the behavioural mechanism behind gratitude is presented. Section 2 continues by demonstrating the relevance of gratitude, particularly in times of new work, and thereafter proceeds with a brief overview of present literature on

gratitude in the economic context and the corresponding development of the main hypotheses. Section 3 and 4 provide an overview of the experimental design and outline research results of the field experiment. In the subsequent discussion, the results are critically reviewed and assessed. Limitations and suggestions for future research follow. To emphasize the importance and added value of the study, the final part of the discussion poses implications for organizations modern labour markets.

2. Theoretical Foundation

The following section provides a definition and conceptualization of gratitude and further emphasizes its interrelationship with the concept of motivation. Subsequently, the study immerses into the underlying behavioural mechanism of reciprocity, conceptualized in a possible toy model. Thereupon, the relevance of gratitude in the context of new work is examined. The last part of the section provides a brief overview of empirical evidence and delineates the hypotheses of the underlying paper.

2.1. The meaning of gratitude

As a relatively new field of study, researchers still need to agree on a general definition for the construct of gratitude[4]. While it seems to be a well-known and common terminology, there are numerous concepts and behavioural descriptions that fall under the notion of gratitude. The Cambridge University Press (2019) defines gratitude as "a strong feeling of appreciation to someone or something for what the person has done to help you". As the word has its roots in the Latin term gratia meaning graciousness and gratefulness, all derivations must relate to kindness, generosity, the act of gift-giving, or receiving something as gratuity (Pruyser, 1976). From a psychological perspective, gratitude is defined as a positive and interpersonal emotion which strengthens people's social resources and preserves social relationships (Frederickson, 1998; Algoe et al., 2008). Being classified as an emotion, Emmons and McCullough (2004, 9) argue that gratitude is an attribution-dependent state that results from two stages of information processing: "(a) recognizing that one has obtained a positive outcome; and (b) recognizing that there is an external source for this positive outcome". According to this classification, positive benefits are determined by external parties, underlining gratitude's feature of being an other-oriented emotion which may imply "prosocial behaviour by focusing attention on the need and deservingness of the benefactor" (Tsang, 2006, 141).[5]

---

[2]In the underlying study, I strive to investigate the effect of gratitude expressions that are without any tangible or extrinsic value. To my knowledge, the only related study using an intervention of immaterial type is a field study by Grant and Gino (2010) who examined the effect of gratitude expressions on university fund raising.

[3]Investigating gratitude expressions in form of an electronic message in combination with an interpersonal element instead of face-to-face communication or e-messages without any personal touch is the main difference to the experimental setting of Grant and Gino (2010).

[4]Although the terms gratitude, recognition, thankfulness, and appreciation are often used interchangeably in the literature, I merely refer to the terms gratitude, (worker) appreciation, and recognition as substitutes throughout this paper.

[5]Prosocial behaviour can be understood as a social behaviour that occurs when individuals act to benefit other rather than themselves.

Apart from studying gratitude within the domain of emotions theory and thus referring to gratitude as a state of feeling grateful, psychologists further depict gratitude as a personality trait (Wood et al., 2008). Gratitude on the dispositional trait level is regarded as a strength of character which refers to a general orientation towards a positive perception of daily life events. Noteworthy, various scholars detected a positive association between perceiving positives in life and individuals' levels of subjective and psychological well-being (Hill and Allemand, 2011; Wood et al., 2008; McCullough et al., 2004).

Moreover, taking a closer look at the process of gratitude highlights the importance of differentiating between sender and receiver of gratitude (see figure 1). While the sender expresses gratitude and kindness, the latter experiences gratitude which ideally leads to the urge to reciprocate and behave in the benefactors' interest (Blau, 1964). According to Grant and Gino (2010, 947):

> "Expressions of gratitude signify that a beneficiary values, needs, appreciates, and accepts one's assistance rather than rejecting or devaluing it."

Last but not least, the management scholars Brun and Dugas (2008) propose a classification of the concept of recognition into four non-exclusive approaches, namely: the ethical dimension, the humanistic and existential view, the psychodynamic school, and the behavioural perspective. Differentiating between these dimensions does not only provide interesting insights about the respective elements of gratitude, but further elaborates on the sender's motive of expressing appreciation.

The ethical perspective for instance promotes the idea that gratitude stems from the thought of human dignity and social equity. Their concept of appreciation is founded on the theory that employees have to be noticed as persons and not as instrumental entities for the organization. While the humanistic and existential view is closely related to the ethical view, their discourse does not focus on organizational justice as a central theme. Rather, this approach emphasizes the need to create proper working conditions within the company to enable humanistic existential appreciation.

This is in contrast to the view of work psychodynamics who highlight the importance of recognising individual's contribution to a result – their work performance – and secondly appreciating people's actual effort – their job dedication – which is irrespective of the final results of their work. Lastly, the behavioural approach to appreciation considers expressing gratitude as an instrument for rewarding performance with the aim of highlighting the efficiency, the outcome, or the value of individuals' performed work for the organization. Thus, it takes an evaluative role which results in a form of conditioning human behaviour and can be as such assigned to the behavioural school of thought (Brun and Dugas, 2008).

In light of the research aim of the underlying study, the examination of gratitude in the present field study corresponds best to the psychodynamic outlook of gratitude and will be regarded accordingly in the present paper.

2.2. Gratitude and human motivation

In the past, motivation psychologists attempted to explain human behaviour and further strived to find reasons of why an individual takes a specific course of action (Rosenstiel and Nerdinger, 1980; Heckhausen and Heckhausen, 2018). In general, they differentiate between the two related concepts of motives and motivation. Motives are defined as characteristic value dispositions which are the concrete cause for taking actions and thus constitute an individual's enduring willingness to act. Some motives are congenital, while others evolve over the course of life. Although motives provide reasons for human actions and their direction, they have to be activated by a stimulus to initiate individual acts. The term motivation, which originates from the Latin word movere, refers to the current orientation towards an operational objective and builds the foundation for target oriented human behaviour (Nerdinger, 2014). A person's motivation emerges when he or she is in a situation that stimulates his or her personal motives and thus leads to action. Hence, the motivation of an individual does not only depend on the presence of situational influences and the respective individual itself but further is conditioned by the interaction of these two (Heckhausen and Heckhausen, 2018).

To gain an even better understanding of motivation, it is sensible to take a closer look at content theories which focus on the various categories of goals and needs which motivate people. In his two-factor theory about job factors, the American psychologist Herzberg proposed that there are two sets of factors affecting employee's attitude about work (Gawel (1997)). More precisely, he differentiates between hygiene factors (e.g. supervision, interpersonal relations, and salary) which can, if absent, cause dissatisfaction and motivators which enrich a person's job and provide positive satisfaction (e.g. recognition, achievement, and responsibility). Motivators hence arise from the intrinsic condition of the work while hygiene factors can be referred to as extrinsic elements to the job itself. With respect to this theory, it seems of high importance to eliminate job dissatisfiers while at the same time improve on motivating factors to increase subjects' motivation to provide higher effort and performance.

Another content theory of motivation includes Maslow's hierarchy of needs. The humanistic psychologist depicts that individuals have needs which can be brought into a hierarchical order (Maslow, 1987). The base of the pyramid form physiological needs such as nutrition and sleep, followed by security needs which include occupational certainty, health, and housing conditions. The next two levels comprise the need of love and belonging as well as the need for appreciation and respect. Self-actualization forms the top of the pyramid and refers to the desire of becoming "everything one is capable of becoming" (Maslow, 1987, 64). With respect to the present investigation, the need for interpersonal relationships as well as the desire for recognition and self-esteem
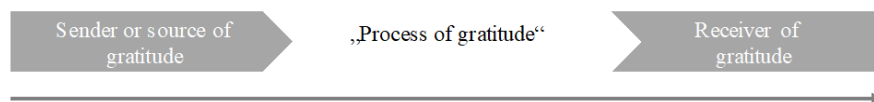
Figure 1: Process of gratitude

are particularly important and can, at least theoretically, explain human motivation for individuals' behaviour. Nevertheless, one question remains unanswered: Through which behavioural mechanism might the expression of gratitude spur individual's motivation to subsequently exert higher work effort?

This phenomenon can be explained by the renowned social norm of reciprocity. Reciprocity is a powerful determinant of human behaviour and social preference that refers to the social norm of responding towards (un)kind treatments likewise. The receipt of a benefit may therefore induce the norm of reciprocity and in this way lead to an act of kindness on the receiver's part (Falk and Fischbacher, 2006). In their extensive research on reciprocity experiments, Dufwenberg et al. (2001) outline the difference between direct reciprocity and indirect reciprocity. While direct reciprocity embodies the principle: "I am nice to you and you are nice to me", the latter involves a third party and thus involves benevolent acts towards a stranger. Scholars basically distinguish between two types of indirect reciprocity namely, downstream reciprocity and upstream reciprocity. Downstream reciprocity is built on reputation and refers to the assertion that an individual who was kind in the past has a higher chance of receiving kindness in the future. Upstream reciprocity, in contrast, is based on a recent positive experience. In particular, it captures the idea that the recipient of an altruistic act may feel motivated to reward the benevolent actor or a third party (Nowak and Sigmund, 2005). However, as I am interested in taking a closer look at gratitude and its effect on direct reciprocal behaviour towards the benefactor, I do not refer to indirect reciprocity in the underlying paper.

Following the vast amount of laboratory experiments and field studies emphasizing the omnipresence of reciprocal behaviour, reciprocity has become an established construct in the field of economics (Fehr et al., 1993). Notably, Emmons and McCullough (2004) outline reciprocity as the underlying principle behind the gift-exchange approach. As depicted earlier, the economic literature on manager-employee reciprocity has mainly focused on "wages as employer's means of exchange in reciprocal relationships with workers" (Bradler et al., 2016, 3088). However, as noted before, employees do not solely care about financial incentives but further value other non-monetary rewards such as recognition, respect, and private compliments communicated by the supervisor. Conveying these characteristics into a formal model, one may suggest the following utility function of a worker (Bradler et al., 2016; Dur et al., 2010; Sliwka and Werner, 2017):

$$U_i = w + \beta e + r\alpha e - C(e)$$

where w represents the base salary of the agent. Individual's provided effort is indicated with e, while β incorporates other intrinsic motives of the subject to exert effort. Worker's degree of reciprocal inclination is defined by r. With respect to the research subject of the underlying study, α indicates the amount of received gratitude. The final component of the equation refers to the associated effort costs (e.g. invested time or mental exertion) and is defined as $C(e) = \frac{1}{2}e^2$. Worker's utility increases when the parameter α is positive, which is the case when expressed gratitude is experienced. Contrarily, when no appreciation is expressed towards the individual, α will be zero. According to economic theories, employees will choose an effort level to maximize utility U_i. Differentiating the utility function with respect to e, results in the worker's optimal effort level $e^* = \beta + r\alpha$. Accordingly, an individual's optimal effort level is increasing with the amount of gratitude received, depending on the individuals' preference for reciprocity. Contrarily, when no appreciation is expressed towards the individual, α will be zero and individual's optimal effort will merely be determined by the parameter β. From a labour market perspective this would imply that individuals work harder and put forth additional effort when they are exposed to gratitude than when they receive no appreciation[6].

2.3. Relevance of gratitude in the context of new work

While there are numerous reasons for organizations to focus more on non-monetary bonus domains, appreciative expressions become particularly important in today's digital age and times of modern working societies. New communication technologies, changing workforce demographics, and the current rise of alternative work arrangements shape current and future workplace (Katz and Krueger, 2019). Increasing opportunities for flexible working arrangements (FWA), for example, enable employees to work where, when, and sometimes even how they choose (Lewis, 2003). Correspondingly, face-to-face communication in manager-employee relationships diminishes, while the use of online communication tools constantly accelerates. With respect to the conducted survey in the underlying field experiment, more than 67% of the participants ($n = 140$) have reported that their employer offers the opportunity of FWAs and 40% of them mentioned to make use of home office arrangements regularly ($n = 56$). However, the resulting physical distance may not only impede employee control and monitoring but further obstruct social relationship building between supervisor and subordinates (Lewis, 2003; Prutchno et al., 2000). In particular,

---

[6]Assuming that $r > 0$.

Gajendran and Harrison (2007) highlight the undermining effect of telecommuting on the depth of ties with colleagues and supervisors. This is particularly alarming, since good manager-worker relationships are claimed to be a vital incentive to motivate employees to work well (Agell, 2004).

To counteract this destructive effect, a sensible approach could be the use of gratitude. By appreciating workers effort verbally, individuals perceive that their actions matter and feel valued which in turn strengthens the social-exchange relationship between manager and employee (Grant and Gino, 2010). This is in line with research results from positive psychologists who state that:

> "Gratitude connects people ... gratitude not only creates and smoothens interpersonal relationships, it also fulfils important cohesive functions for society and culture as such" (Emmons and McCullough, 2004, 204).

While, to the best of my knowledge, there is no study investigating the relationship between gratitude and relationship building within the FWA context, I merely assume the positive association between the two constructs to emphasize its relevance in present work settings. With respect to the dynamic shift of today's working environment, it is hence crucial to understand the importance of non-pecuniary elements, especially gratitude, as not only a performance-enhancing substance to promote employee motivation but further its supportive function in social relationship building (Korzynski, 2013; Emmons and McCullough, 2004).

2.4. Empirical evidence and hypotheses development

Although it is undeniable that gratitude is omnipresent in social life, academic literature so far provides little evidence on responses to gratitude expressions within the gift-giving domain in organizations and its subsequent effect on employee performance. Moreover, the existing literature narrowly concentrates on the impact of financial gifts. This is not surprising, since rewarding an individual with a monetary payment is assumed to be the most explicit and unambiguous way of endowing. The act of giving a non-financial gift, on the contrary, is implicit and more context-specific which results in limited generalisability of empirical results (Bradler and Neckermann, 2019). Nevertheless, the rapid changes in technology and communication channels within modern labour markets as well as latest empirical findings underline the relevance of gratitude and demand for future research. Likewise, results of a variety of employee surveys propose that workers strive for appreciation and recognition from their employer and monetary compensation is only considered of secondary importance. This accords to Wiley (1997), who finds that "full appreciation for work done" was the only motivating factor which continuously ranked among the top two motivators for employees in the U.S. during the second half of the 20th century.

One of the early experimental studies investigating the effect of gratitude was conducted by the psychologist Clark (1975). In a field experiment, he demonstrates that female participants who were thanked for giving a "confederate" directions to the university library, were more likely to help on a subsequent occasion (e.g. to pick up books that the "confederate" dropped in the street) than were subjects who were cut off and told "nevermind, I'll ask another person" prior to the subject's completing the direction. This infers that subject's recent past experiences determine their prosocial behaviour in the future, as they are more likely to expect to be thanked (punished) again. Further field experiments emphasize that the reinforcement effect provoked by expressions of gratitude extends as well into the economic context. Rind and Bordia (1995), for instance, studied the effect of server's "thank you" on restaurant tipping. They conducted a field experiment in an upscale restaurant and find an 11% increase of tipping when the waiter showed gratitude by writing "thank you" on the back of the checks compared to when he didn't express gratitude at all. Another study experimentally tested the effectiveness of handwritten thank you notes on the response rates of physicians' mail surveys and reports an increase of response rates of 40.7% in the treatment group compared to the baseline group (Maheux et al., 1989).

While these studies demonstrate that individuals who are thanked for prior efforts are motivated to "work harder" in favour of others than are subjects who have not been thanked for their efforts of providing benefits, a few things remain uncertain. First, the results do not provide sufficient proof whether the severe effects arise through the mechanism of reciprocity; and secondly the influential impact interpersonal elements might have had on the outcomes remains unstudied. These issues are addressed by a study providing evidence on appreciation as a gift which is most closely related to the present investigation (Bradler and Neckermann, 2019). The scholars conducted two field experiments that explore effort levels of subjects in response to monetary gifts and worker appreciation. Their findings show that gratitude in form of a thank you note positively affects agent's reciprocal behaviour. More precisely, they detect that a combination of appreciation and monetary payment as a gift works less well than appreciation or money alone. Additionally, their experiments indicate that subjects' responses are highly sensitive to the presence of interpersonal elements since they report substantial performance increases when combining the gifts with a personal touch. Interestingly, they argue that it is not the interpersonal element alone which stimulates the sizable impact on performance but that the personal touch influences the way in which the presentation of the gift affects the behaviour of the individual. Thus, the personal touch rather functions as a signal and does as such trigger a framing effect[7]. This assumption accords with intention-based reciprocity models which assert that recipients evaluate the kindness of an action not only by its material consequences but also by what the action signals about the donors

---

[7]Based on Tversky and Kahnemann (1981), the concept of framing refers to a cognitive bias where individuals decide on options based on their labelling.

underlying intention (Falk and Fischbacher, 2006). The results of Bradler and Neckermann (2019) complement and underpin very well the findings of Kube et al. (2012) research on gift-exchange in the workplace with reciprocity as the underlying currency. Kube and his colleagues not only report a significant higher impact of non-financial gifts than of purely monetary gifts, but further claim that providing individuals with a cash gift in form of an origami (i.e. a banknote was nicely folded and wrapped) yields a 30% greater output compared to the baseline group who merely received the banknote without any interpersonal element attached to it. Given the outlined empirical findings and considering the earlier presented toy model, which assumes that subjects receive a positive utility from receiving gratitude, I firstly anticipate individuals to respond positively to receiving an expression of gratitude and secondly and more specifically expect a thank you note combined with a handwritten signature to increase their work effort[8]. Hence, I hypothesize:

> Hypothesis 1: The expression of gratitude, independent of its particular form, elicits reciprocity and leads to higher effort compared to no expression of gratitude.

> Hypothesis 2: The expression of gratitude in form of a thank you note combined with a handmade element elicits reciprocity and leads to higher effort compared to no expression of gratitude.

While the use of an appreciative note at the end of a survey is a common practice and hence may be considered as an expected gift, recent research streams emphasize the power of surprising gifts in gift-exchange relationships (Rogers and Frey, 2014; Macera and te Velde, 2018; Neckermann and Yang, 2017). Gilchrist et al. (2016), for instance, provide evidence that the way in which a monetary gift is structured is essential when eliciting reciprocity. Specifically, they delineate a significant productivity increase of 20% when providing workers with a higher wage that includes a salient gift ($3 + $1) compared to paying the same above market-wage ($4) per se. Likewise, using a laboratory experiment, Sliwka and Werner (2017) demonstrate that individuals provide higher effort under increasing wage profiles, that are not communicated in advance and hence are unexpected by the agent, than when wages stay constant over time. Similar results are reported by Gneezy and List (2006) who find that worker effort is substantially higher for the "gift" treatment, who received a surprising wage increase, than for the "non-gift" treatment. However, the positive effect is only present within the first period and vanishes after the initial few hours. With respect to the current study, implementing

a surprising gratitude expression in form of a personal video to induce reciprocal behaviour seems plausible. Latest insights on the personal nature of videos, reported by Beute and Pacinelli (2019), highlight the use of simple videos as the most innovative way of building trust and strengthening relationships within the context of customer and service experience. According to a survey they conducted, customers not only felt more valued but further stated that a video message is much more personal than a voice message or email. Furthermore, the authors claim that videos are a useful tool to exceed expectations on the receivers' part. By integrating a personal video message, one breaks the common pattern of being "...just another number, ..., or another phone call" (Beute and Pacinelli, 2019, 71). Although the authors study the impact of personal videos on accelerating sales and improving customer experience, I transfer their findings to the context of the underlying study and thus propose, that:

> Hypothesis 3: The expression of gratitude in form of a personal video elicits more reciprocity and hence leads to higher effort than a thank you note or no expression of gratitude.

Beyond evidence on the linkage between gratitude expressions and workers effort provision, latest work in behavioural economics have gathered substantial proof that make evident that gender is another relevant determinant of behaviour which should be considered in economic decision-making processes. In general, research results of gender differences in social differences are mixed. While some scholars assume women to behave more reciprocal than men (Croson and Buchan, 1999; Snijders and Keren, 2001; Eckel and Grossman, 1996; Heinz et al., 2012) other studies find no gender differences in the extent and form of social preferences (Bolton and Katok, 1995; Eckel and Wilson, 2004; Cox and Deck, 2006). A reasonable explanation for these competing findings relies on the observation that female participants are more sensitive to social cues in the experimental context than are male participants. Hence, already a little difference in the experimental design or in the subsequent implementation have a greater impact on the behaviour of female responders than respectively on male recipients (Croson and Gneezy, 2009).[9] In a dictator game experiment, scholars find that women are affected more strongly by the first-movers decision than men (Eckel and Grossman, 1996). More precisely, they demonstrate that females are more likely to accordingly punish or reward previous actions. Further evidence is provided by Ben-Ner et al. (2004a) who carried out a two-part dictator game to obtain evidence on the propensity towards reciprocity. They find that women primarily depend their decisions on the amount they received in the former round and moreover reciprocate significantly more than men. In support of this, a more recent work on the tendency of women to

---

[8]Kube et al. (2012) and Bradler and Neckermann (2019) used a thermos bottle and a thank you card, respectively, to express appreciation. Both gifts are considered as non-pecuniary but contain a tangible value. The set-up of the underlying experiment differs to the two above mentioned investigations as the present intervention is without any tangible or extrinsic value.

[9]Examples of differences include economic (e.g. size of payoff) and psychological variables (e.g. level of anonymity between participant and experimenter).

behave more reciprocally than men, provides sufficient proof of the existent gender difference (Heinz et al., 2012). The economists implemented a modified dictator game, in which recipients move before dictators by conducting a real-effort task. Recipients effort choice thus resembles the first move decision in the trust game (Berg et al., 1995). In a next step, dictators decide on how much of the generated money to take and likewise decide on the amount to return to the recipient. The experimental findings confirm their hypothesis that female dictators show more reciprocity by significantly decreasing their taking-rates than male dictators, who generally demonstrate to be more selfish. Notably, this treatment effect is considerably larger, when first movers decide for providing high effort in the initial step and thus generate more money for the dictator to decide on.

To complement the outlined laboratory studies, Dittrich (2015) conducted a large-scale online experiment with heterogeneous subjects. The behavioural economist questioned the general reliability of effects found in laboratory experiments due to their typically homogenous samples. He concludes that though these studies make a great contribution to the general understanding of behavioural differences between female and male subjects, their findings are not necessarily robust if other variables (e.g. income, age) are included. To combat the problem of robustness, he conducted an anonymous online experiment in a heterogenous population. Interestingly, his findings are in contrast to the previous outlined evidence from laboratory experiments. He does not only reject the assertion that women are more reciprocal than men, but rather elaborates that male subjects behave in fact more reciprocally than female subjects. Motivated by the varied outlined literature on gender differences above, it seems particularly interesting to investigate whether the underlying experiment may evince a significant gender difference in behaviour. Even though I do not conduct a laboratory experiment, I assume a rather homogenous sample in the underlying study and thus expect congruent results to the laboratory experiments outlined above (Heinz et al., 2012; Eckel and Grossman, 1996; Ben-Ner et al., 2004a). On the basis of this argumentation, I accordingly suggest that:

> Hypothesis 4: On average, women show more reciprocity than men and are thus more likely to respond to gratitude by putting forth more effort than men.

3. Research Methodology

The following chapter provides a detailed overview of the conducted empirical study. In a first step, the methodology and data collection as well as the experimental design of the online survey are outlined, followed by a detailed description of the sample pool.

3.1. Methodology and data collection

In order to validate or reject the above proposed hypotheses, a quantitative research design was chosen. The quantitative approach, in contrast to a qualitative research method,

is suitable when aiming to infer characteristics, behaviours, and attitudes of a large sample size (Kothari, 2004). Furthermore, as it yields quantitative data, it is usually associated with a deductive approach where the focus lies on using data to test general conclusions and theories (Saunders, 2011). Since the goal of the underlying paper is to observe and understand individual's behaviour with respect to receiving gratitude, this method of research seems appropriate. An online survey experiment was carried out, as field experiments in general allow for clearly identifying causal relationships, by manipulating the independent variable (here: level of expressed gratitude) and observing the ensuing effect on the dependent variable (here: level of induced effort) while at the same time not losing external validity (Harrison and List, 2004). Additionally, in contrast to gratitude scenarios in laboratory experiments, studying gratitude in a field experiment brings higher psychological realism, causes greater involvement of participants, and in total entails less artificialness than gratitude scenario studies (Tsang, 2006). Besides that, surveys present an ideal setting to study how non-pecuniary gifts elicit individual's willingness to exert additional effort, due to its underlying voluntary and time-consuming characteristic. To collect data for the analysis, participants were asked to follow a link to the online survey experiment on the platform Sosci Survey (www.soscisurvey.de). Since conducting behavioural experiments online requires technical reliability, I trusted in the software Sosci Survey as it is not only widely used among economic research but further provides its users with an excellent and time-efficient online support. The survey link combined with an introductory message about the research project was distributed via social platforms such as Facebook, LinkedIn, and Xing to ensure diversity within the sample and avoid any kind of selection bias. Approaching personal and business contacts to participate and further share the survey within their respective network complemented the data collection strategy. This sampling method is referred to as convenience sampling and belongs to the non-random sampling techniques, as it is often practically impossible to collect data from the entire population that is to be considered. Although convenience samples limit the extent to which findings can be representative for a population (Feild et al., 2006), it is often used in social and behavioural research as it is very feasible, prompt, and economical with respect to time and financial resources. Subjects were not told about the underlying experiment within the survey, but instead were simply provided with the information that the survey was part of a research project for a master thesis and that participation will take approximately 6-8 minutes. Employing an unobtrusive data collection method brings the major benefit of not having to be concerned about the Hawthorne effect, an effect that modifies participants behaviour when they are aware of being part in an experiment compared to their behaviour without this knowledge (Adair, 1984). Finally, to avoid priming effects, I did not indicate the main research theme of the project in the survey but instead framed the questionnaire under the topic of employee motivation in the context of new work. The online survey was

active for a little over 2 weeks from 14th November to 2nd December 2019.

3.2. Online survey experiment design

The underlying field study encompassed two successive questionnaires, whereby respondents had the choice to end or continue their participation after the first form. In between the two questionnaires, the gratitude intervention is implemented. The main interests of observation lie in the investigation whether subjects continued after the intervention, in the examination of disparities between the different treatment groups, and lastly in the comparison of behavioural differences within treatments, before and after implementing the gratitude manipulation. Thus, the underlying study employs both, a between-subject and a within-subject design.

The first questionnaire comprises a total of 15 questions[10]. It begins with a cover letter introducing the research institute, the aim of the research, and the resulting need of participation to support the experimenters' project. After the introduction, a few questions regarding respondents' employment status, flexible-work arrangements, and existing means of communication within their employers' organization follow.[11]

These questions are intended to arouse interest in and stimulate reflection on the topic of new forms of work and changing communication tools within organizations. Making respondents think about the tremendous shift of today's working environment, provides a good foundation for the subsequent real-effort task. To complete the task, respondents are asked to write down all keywords that come to their mind when thinking about work-life-balance. By this, subjects have to exert some degree of actual effort by thinking about the topic and correspondingly writing down notions. The task allows to measure the quantity of data entry by assessing the number of entries of each respondent, in which one entry corresponds to typing one word in the presented free-text field. Since I am interested in the level of effort provided, the quantity of data entry serves as an appropriate measure in the analysis. The first questionnaire ends with questions regarding individuals' intrinsic motivation, their inclination towards reciprocity, and items on socio demographics. The prior described phase is referred to as working period 1 in the following. At the end of working period 1, the treatment intervention takes place. Depending on the treatment, which is randomly distributed, the respondents are either thanked for participating in the survey, via a note or by

means of a personal video, or however receive no appreciation at all for their provided effort (see 3.2.1). Every participant is shown only one "end of survey" slide while all of them link to a second form that respondents were asked to fill out to support the experimenters project. The second questionnaire, which I refer to as working period 2 in the following, begins immediately with a second real-effort task. The task is very similar to the exercise in working period 1 with the only difference that subjects are urged to type notions that relate to the term employee motivation. It is assumed that the effort needed to come up with keywords that relate to the concepts of work-life-balance and employee motivation is similar high. Both terms are widespread and neither age-, gender- or subject-specific, which suggests same prerequisites for the entire subject pool. Similar to working period 1, the main effort measure is the quantity of data entries. The subsequent questions collected data on whether individuals' feel appreciated by their current supervisor and on how much salary they would be willing to sacrifice in exchange with more gratitude expressed by their supervisor. To ensure that the gratitude manipulation was effective, the final part of working period 2 asked respondents to indicate the extent to which the experimenters note or video, respectively, expressed gratitude and appreciation. This assertion estimates individuals' perception that the experimenters' communication expressed gratitude. Working period 2 ends with a final slide, which was the same for all treatment groups, thanking for participants' time and support.

3.2.1. Treatments

In total, the field study comprised three treatments, which have been randomly assigned to respondents[12]

Baseline ($n = 69$)

Participants in the control treatment receive a short "Your responses have been recorded" message at the end of working period 1. Below this message, a further note appears which links to an additional questionnaire that is run to validate results of the prior survey. Alternatively, respondents can end the participation by simply closing the web browser. No appreciation from the experimenter is communicated. However, to embed a personal touch throughout all treatments, including the baseline treatment, a handwritten signature of the experimenter is placed underneath the above-mentioned message[13].

GratitudeNote ($n = 65$)

Participants in the GratitudeNote treatment are shown the same message and handwritten signature as the control treatment, with the only difference that gratitude for the effort exerted to fill out the survey is expressed via the following note:

---

[10]Please refer to Appendix A1 for the complete version of the questionnaire.

[11]In case participants are not currently employed, they are asked to refer to their former employer. All subjects who stated to not have been employed before, are excluded from the study and are not further considered. I suggest individuals who have been part of a classical manager-employee relationship before to behave differently towards gratitude expressions than do subjects who have not had a supervisor before. Further, since I aim to deduce implications for management I decided to focus on this criterion when defining the subject pool.

[12]Please see Appendix A2 for a visual representation and the exact wording of the experimental interventions.

[13]According to Bradler and Neckermann (2019) a handwritten signature is enough to present a personal touch. A personal touch was embedded in all treatments to rule out the possibility that the treatment effects might be entirely driven by the interpersonal element.

"Thank you very much for participating in my survey. Your results will be of great use for my work project and are highly appreciated!". This appreciative remark is placed before referring to the next questionnaire, with the intent to induce reciprocity on the receivers' part and hence to increase the probability of them to participate in the follow-up survey.

GratitudeVideo ($n = 73$)

The GratitudeVideo treatment is the same as the GratitudeNote treatment. The only difference is that instead of receiving the above described thank you note, participants are shown a video in which the experimenter expresses her appreciation for participating in the study. It is worth noting that the wording used in the video is identical to the phrasing of the thank you note. Further, as the video itself presents already a high degree of personal touch, participants were not additionally shown a handwritten signature on the bottom of the page.

3.2.2. Measures

Different measures have been considered as appropriate to quantify reciprocal behaviour and effort provision in response to the intervention in between working period 1 and 2. In general, the click and response rate of the second questionnaire after the gratitude intervention serves as a main proxy for reciprocal behaviour and can be assessed with a dichotomous measure of whether respondents voluntarily proceeded to the second questionnaire. To additionally assess precise levels of reciprocal behaviour in the form of effort induced, the number of data entries of the real-effort task in working period 2 are compared between treatments and further within treatments by comparing the quantity of words between the two working periods[14]. According to the previous proposed hypotheses, one would not only assume a higher probability to continue for treatment GratitudeNote and GratitudeVideo, but further a higher quantity of words in working period 2 compared to the former period. Overall, I assume the effort of both gratitude treatments to outperform the effort of Baseline.

Furthermore, since previous studies claim that individuals' degree of reciprocal inclination influences subjects' actions, I elicited a measure to control for this behavioural mechanism. The measure used builds on the work of Perugini et al. (2003) and includes several statements such as "If someone does me a favour, I am ready to return it", "I go out of my way to help somebody who has been kind to me before", and "I am ready to undergo personal costs to help somebody who helped me before". Respondents were asked to indicate their level of agreement (or disagreement) using a 7-point Likert-type scale with 1 = strongly disagree and 7 = strongly agree.

To assess whether participant's general level of intrinsic motivation plays a role in the underlying investigation, a single-item measure was embedded asking participants about the number of surveys that have been completed voluntarily within the past 3-months. Respondents could choose between the response options "0", "$1 - 3$", "$4 - 6$", "$7 - 10$", "> 10", and "I haven't been asked to participate in a survey within the past 12 weeks". Although this measure is rather superficial and obviously does not poll the full set of individuals intrinsic motivation, it provides an idea of the direction of present intrinsic motivation within the subject pool and hence might offer great insights and opportunities for future research[15].

Furthermore, a manipulation check was incorporated in the questionnaire to test the effectiveness of gratitude expression in GratitudeNote and GratitudeVideo and thus to validate the online experiment. The manipulation check consisted of a single item, namely: "The thank you note (video) expressed gratitude and appreciation for my invested time and effort to fill out the survey." Once again, participants answered on the same 7-point Likert scale as described earlier.

In general, Likert scales are devised as "series of statements expressing either a favourable or unfavourable attitude toward the concept under study", where the respondent is "asked to indicate the level of her or his agreement or disagreement with each statement by assigning it a numerical score" (McDaniel and Gates, 2013, 315). In particular, 7-point scales were embedded since they provide a more accurate evaluation of respondents' true response compared to 5-point conditions in which participants may be more likely to interpolate (Finstad, 2010).

3.3. Sample

After removing participants who had aborted the survey before accomplishing the first survey, a total of 208 respondents remained in the data set[16]. 70% of the respondents were female participants ($n = 146$) and merely a percentage of 30% were male participants ($n = 61$). One subject indicated that he would prefer to not disclose his gender while no respondents stated the gender other. While the present study (inter alia) investigates gender differences in behaviour, participants who not clearly mentioned their gender ($n = 1$) were not further considered in the following analysis, leading to a total of 207 remaining respondents. All subjects were randomly allotted to one of the three treatment groups, presenting different degrees of gratitude (see 3.2.1). This resulted in three slightly different sample sizes. A total number of 65 participants were part of the thank you note treatment,

---

[14]It is important to stress that I assessed effort by counting words and not notions. I presume that taking down notions which consist of more than one word to be associated with higher effort for the individual. Thus, if a key term consisted of more than two words (e.g. home office), two words were measured.

[15]Asking participants about the number of surveys filled in voluntary in the past weeks, seems to be a suitable measure for general intrinsic motivation as I assume the majority of respondents to be students who usually often encounter survey requests.

[16]In total, $n = 60$ participants dropped out of the first part of the study and were thus not further considered in the analysis. Participants dropouts may result from the actual effort required to fill out the first effort task and the requirement to answer every single item in the questionnaire in order to proceed.

whereas 73 individuals were assigned to the thank you video group, leaving a sample size of 69 subjects in the baseline group. The higher number of women participating overall in the study is well mirrored in all three experimental groups, by keeping up an average ratio of 70:30. The average age of participants is between 25 and 34 years (n = 139) and the majority of subjects participated by using their smartphone device (n = 143). Within the subject pool, 37% (n = 77) of the respondents were employed full time, while only 15% (n = 31) remarked a part-time employment. Further, 36% (n = 76) of the sample declared a temporary position as an intern, working student, or student assistant. The remaining 12% (n = 24) indicated the employment status other. Only subjects who are currently or have previously been part of an employment relationship were considered in the underlying experiment (see 3.2 for a detailed declaration). Furthermore, more than 65% (n = 137) of the subjects come from the business and economics context, followed by legal students and jurists (n = 16) as well as subjects with a psychological background (n = 13). The main area of study of the residual participants is distributed over the fields of agriculture, health sciences, humanities, education, engineering, and maths. Finally, 24% of all participants reported to have voluntarily participated in more than 3 surveys within the past 12 weeks, a measure which is interpreted as individuals' general level of intrinsic motivation in the following[17]. 11% of the respondents (n = 22) stated to have not been asked to fill out any surveys in the past and thus were excluded from the measure, which has to be kept in mind for further analyses. Table 1 displays descriptive statistics of the considered sample in detail.

4. Results

Following the research methodology this section presents the empirical results gathered from the study. The results are used to verify (reject) the formulated hypotheses discussed in the first part of the paper. The collected data were exported from Sosci Survey to the statistical program Stata, which allows for a detailed examination of econometric model analyses. Before carrying out analyses, the data set was prepared and cleared up[18]. I start by examining whether participants proceeded to the second questionnaire after being exposed to the gratitude intervention (effort proxy 1) and afterwards investigate individual's level of provided effort, analysing between and within-subjects (effort proxy 2).

4.1. Effort proxy 1

Figure 2 provides an overview of the number of subjects who voluntarily clicked on "next" redirecting them to the second questionnaire and Appendix A3 displays descriptive statistics of the treatment effects.

At first sight, two outcomes become apparent. First of all, the treatment GratitudeNote seems to have a great influence on individuals' decision to proceed to the next questionnaire. More precisely, 56 out of 65 subjects in the GratitudeNote treatment decided to continue which significantly outperforms both, the continue click rate for the baseline group ($p = 0.038$, Fisher-exact test, two-sided[19]) as well as, opposed to my proposition, the click rate for GratitudeVideo ($p = 0.092$, Fisher-exact test). Secondly, contradictory to my hypothesis, no significant difference is observed when comparing the treatment for individuals' who were part of the GratitudeVideo treatment with the control group ($p = 0.711$, Fisher-exact test). Furthermore, pooling the two gratitude treatment groups[20] (this variable is considered as GratitudeAll in the following) and comparing it with Baseline yields no significant difference as well ($p = 0.167$, Fisher-exact test) which is liable to result from the considerable small number of individuals of the GratitudeVideo treatment who decided to continue. Moreover, the figure presented in Appendix A4 visualizes the mean continue click rate by gender. While 81% of all female participants proceeded to the second questionnaire, the continue click rate for male accounts only for 67% ($p = 0.046$, Fisher-exact test).

The ordinary least squares (OLS) linear probability model (LPM) reported in Table 2 complements these non-parametric findings. I estimate the results using a dummy for continue (1 if participants continue and 0 if not) as the dependent variable[21]. The independent variables are the specific treatments, a dummy for gender (1 for female and 0 for male), and interaction terms between the respective treatments and the gender dummy. I use Baseline, in which no gratitude was expressed, as a control group and compare the effects of each of the two other treatments to this setting.

The results indicate that subjects in the GratitudeNote treatment show a 15.9 percentage points higher likelihood of continuing the survey experiment than respondents in the control group ($p = 0.034$, column 4), while participants in the GratitudeVideo treatment only show a statistically insignificant higher probability of roughly 2 percentage points compared to Baseline ($p = 0.815$, column 4). Furthermore, GratitudeVideo is significantly different from GratitudeNote ($p = 0.056$, column 4, Wald test). Including the gender dummy results in minor changes of the treatment effects but reports that female respondents certainly are significantly more likely to proceed to the next questionnaire compared to male respondents ($p = 0.013$, column 5). This is in line with the earlier made proposition that women behave more

---

[17]The median response for this questionnaire item is "2" with the value "2" referring to "1-3" surveys completed within the past 12 weeks. With respect to a median split, I thus define any value below the median as "0" and every value above the median as "1".

[18]Please see section 3.3 for a detailed presentation of the adjusted sample pool.

[19]If not stated differently, all statistical tests in the present study are two-sided.

[20]By pooling gratitude, I aim to infer conclusions of the overall effect of appreciation on individual's effort.

[21]I applied LPM instead of logit or probit model for all dichotomous dependent variables. Although LPM assume linearity and thus may predict probabilities outside [0,1], they offer a more convenient interpretation of coefficients (Aldrich et al., 1984).

Table 1: Subject characteristics of all treatments

Note: The table displays means with standard deviations in parentheses of relevant subject characteristics of the treatments.

	Control	GratitudeNote	GratitudeVideo	All
Female	0.71	0.75	0.65	0.70
	(0.45)	(0.43)	(0.47)	(0.45)
Age (25-34 years)	0.62	0.72	0.67	0.67
	(0.48)	(0.45)	(0.47)	(0.47)
Smartphone	0.69	0.70	0.67	0.69
	(0.46)	(0.45)	(0.47)	(0.46)
Economics or business student	0.55	0.72	0.71	0.66
	(0.50)	(0.45)	(0.45)	(0.47)
Full-time employment	0.30	0.46	0.35	0.37
	(0.46)	(0.50)	(0.48)	(0.48)
Intrinsic motivation	0.24	0.27	0.20	0.24
	(0.43)	(0.45)	(0.40)	(0.42)
Total number of participants	69	65	73	207
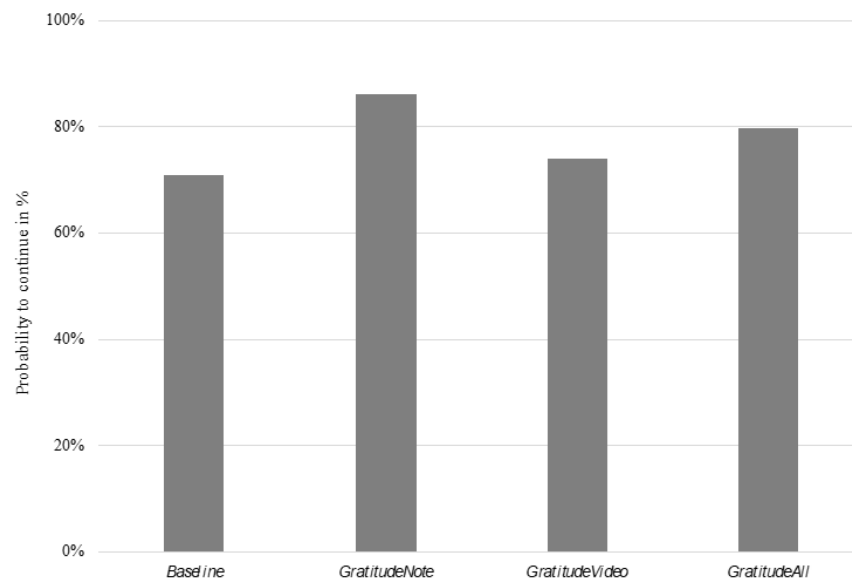
Figure 2: Overview of subjects click rate by treatments

reciprocally than men.

With respect to model 6, in which I included interaction terms, one can observe negative but insignificant coefficients for both interactions. Interestingly however, the probability estimates for all three main effects in this model, respectively GratitudeNote, GratitudeVideo, and gender have more than doubled in size. Thus, by implementing an interaction term one can deduct the partial effect which helps to understand the overall effects. These negative effects are included in the specific treatment probability coefficients in model 4 and 5, which explains the much smaller coefficients compared to model 6. Moreover, comparing coefficient estimates between model 3 and model 6 outlines that effect sizes are much larger when incorporating control variables. This may result from the inclusion of a dummy variable for intrinsic

motivation as a control variable which excludes 22 participants who stated to have not been asked to fill out any surveys in the past weeks, resulting in a reduced sample[22]. It is conceivable that the behaviour of these excluded participants varied widely from the average sample subjects in the experiments and have thereby distorted the results[23].

Dropping out intrinsic motivation as a control variable in model 7 supports this assumption as the given estimates

---

[22]Please refer to 3.3 for a detailed description of the dummy variable for intrinsic motivation.

[23]This effect is solely apparent for the regression model 6 including interaction terms and not already for model 3, in which I incorporate the exclusion of these participants already. Thus, the impact arises specifically by the interaction of the respective treatment and individual gender.

Table 2: Effort proxy 1 – Probability to continue

Note: Model 1-7 display coefficients from OLS regressions with robust standard errors in parentheses[24]. The dependent variable is the continue dummy. All results are compared to the Baseline treatment. Controls include several dummies, namely for being between 25-34 years old, for using a smartphone, for economics and business students, for intrinsic motivation, and for being fulltime employed[25]. Model 7 includes all controls besides intrinsic motivation. $p < 0.1^*$, $p < 0.05^{**}$, $p < 0.01^{***}$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GratitudeNote	0.151**	0.146**	0.263*	0.159*	0.149**	0.358**	0.246
	(0.079)	(0.069)	(0.150)	(0.074)	(0.072)	(0.162)	(0.151)
GratitudeVideo	0.029	0.036	0.130	0.019*	0.025	0.183	0.127
	(0.076)	(0.075)	(0.147)	(0.082)	(0.080)	(0.165)	(0.149)
Female		0.128*	0.226*		0.186**	0.357**	0.215*
		(0.068)	(0.128)		(0.074)	(0.135)	(0.128)
GratitudeNote x Female			-0.160			-0.283	-0.150
			(0.169)			(0.180)	(0.170)
GratitudeVideo x Female			-0.135			-0.221	-0.139
			(0.171)			(0.186)	(0.172)
Constant	0.710***	0.619***	0.550***	0.598***	0.464***	0.339**	0.484***
	(0.055)	(0.076)	(0.113)	(0.105)	(0.119)	(0.140)	(0.135)
Wald test: GratitudeNote = GratitudeVideo	$p = 0.072$	$p = 0.107$	$p = 0.335$	$p = 0.056$	$p = 0.093$	$p = 0.253$	$p = 0.395$
Controls	No	No	No	Yes	Yes	Yes	Yes
Observations	207	207	207	185	185	185	207
Adjusted R^2	0.014	0.028	0.024	0.016	0.048	0.053	0.017

hardly alter from model 3 in which I did not incorporate controls at all. However, taking a closer look at the adjusted R^2 delineates that the consideration of intrinsic motivation is essential for the explanatory power of the regression model (adjusted $R^2 = 0.053$ in column 6 and adjusted $R^2 = 0.017$ in column 7).

I replicated these OLS regressions with merely replacing the specific treatments by GratitudeAll to test whether the expression of gratitude, independent of its particular form, has an impact on individual's reciprocal behaviour. The respective regression results are displayed in Table 3.

It appears that GratitudeAll only has a significant effect on the continue dummy if the regression model includes the dummy for gender, the interaction term of GratitudeAll and gender, and additionally controls for all variables listed earlier ($p = 0.047$, column 6). Besides, one can observe a positive and significant effect for female subjects throughout all models, providing further support for the predicted gender difference.

With respect to the social norm of reciprocity and more specifically individual's preference for reciprocal behaviour, it is noteworthy that so far, I assumed reciprocal inclination to simply be greater than zero. However, relating to the for-

mal model delineated in section 2.2, it is worthwhile to take a closer look at whether subject's effort choices alter when including their explicit degree of positive reciprocity (Sliwka and Werner, 2017; Altmann et al., 2008). As depicted earlier, I implemented a three-item measure for reciprocal inclination at the end of the first survey. To assess individual's preference for reciprocity, I centered mean responses from each respondent and included this proxy for positive reciprocity in the following regression models[26]. Similar to the regression models in Table 2, I conducted OLS regressions with the continue dummy as the dependent and treatments as well as positive reciprocity as independent variables. Moreover, I included the two interaction terms: Treatment x positive reciprocity and Female x positive reciprocity. The outcomes are reported in Table 4.

The estimates prove that subject's reciprocal inclination positively and significantly influenced the probability to continue to the next questionnaire ($p = 0.071$, column 3). Furthermore, while the outputs provide evidence that positive reciprocity notably influences reciprocal behaviour of male subjects, inferring that men with a higher reciprocal inclination show more effort than less reciprocal men, this effect is severely weakened for female participants. More precisely, the effect size is almost quartered for women compared to men $(0.223 + (-0.167) = 0.056$, column 4). Interestingly,

---

[24]OLS makes the assumption that the variance of the error term is constant, meaning that they are homoscedastic. However, this condition is not always met, which is why I applied the Breusch-Pagan test for each regression model to test for heteroscedasticity and accordingly display robust or normal standard errors (Wilcox and Keselman, 2004).

[25]These control dummies were chosen based on the frequency distribution displayed in Table 1.

[26]Centering independent variables reduces multicollinearity in regressions that is caused by higher-order terms, such as interaction terms, and refers to the process of subtracting the overall mean. Though it is otherwise very common to standardize independent variables, I decided for centering the variable as resultant coefficients are more convenient to interpret.

Table 3: Effort proxy 1 – Probability to continue

Note: The table displays coefficients from OLS regressions with standard errors reported in parentheses. The dependent variable is the continue dummy. GratitudeAll is a dummy variable pooling both gratitude treatments. All results are compared to the Baseline treatment. Controls include several dummies, namely for being between 25-34 years old, for using a smartphone to participate, for economics and business students, for individuals' intrinsic motivation, and lastly for being fulltime employed. Model 7 includes all control variables besides intrinsic motivation. $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GratitudeAll	0.087	0.088	0.182	0.087	0.085	0.254**	0.171
	(0.062)	(0.062)	(0.114)	(0.069)	(0.068)	(0.127)	(0.115)
Female		0.137**	0.226*		0.195***	0.356***	0.216*
		(0.064)	(0.111)		(0.070)	(0.124)	(0.113)
GratitudeAll x Female			-0.132			-0.234	-0.130
			(0.136)			(0.149)	(0.136)
Constant	0.710***	0.613***	0.550***	0.588***	0.449***	0.330**	0.475***
	(0.051)	(0.067)	(0.094)	(0.098)	(0.108)	(0.132)	(0.119)
Controls	No	No	No	Yes	Yes	Yes	Yes
Observations	207	207	207	185	185	185	207
Adjusted R^2	0.005	0.022	0.022	0.003	0.039	0.047	0.015

Table 4: Effort proxy 1 – The impact of positive reciprocity

Note: The table displays coefficients from OLS regression with standard errors in parentheses. The treatments GratitudeNote and GratitudeVideo are both compared to the Baseline group. Positive reciprocity is the centered mean response of individuals to the reciprocity proxy. Controls include several dummies, namely for being between 25-34 years old, for using a smartphone to participate, for economics and business students, for individuals' intrinsic motivation and lastly for being fulltime employed. $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***.

	(1)	(2)	(3)	(4)
GratitudeNote	0.138*	0.135*	0.142*	0.135*
	(0.072)	(0.073)	(0.077)	(0.077)
GratitudeVideo	0.032	0.026	0.021	0.005
	(0.070)	(0.070)	(0.076)	(0.076)
Female	0.123*	0.117*	0.180***	0.174**
	(0.064)	(0.064)	(0.069)	(0.070)
Positive reciprocity	0.047	0.145*	0.080*	0.223**
	(0.038)	(0.081)	(0.044)	(0.090)
GratitudeNote x positive reciprocity		-0.059		-0.088
		(0.105)		(0.111)
GratitudeVideo x positive reciprocity		-0.065		-0.032
		(0.086)		(0.099)
Female x positive reciprocity		-0.098		-0.167*
		(0.080)		(0.093)
Constant	0.627***	0.638***	0.508***	0.517***
	(0.068)	(0.068)	(0.110)	(0.111)
Controls	No	No	Yes	Yes
Observations	207	207	185	185
Adjusted R^2	0.031	0.026	0.061	0.064

no discernible effect is found when interacting the treatment dummy and the reciprocity proxy in neither GratitudeNote nor GratitudeVideo treatment ($p = 0.430$ and $p = 0.749$, respectively, column 4), implying that the treatment effects are not driven by positively reciprocal respondents[27].

Concluding, examining effort proxy 1 provides empiri-

cal support for hypothesis 1 and hypothesis 2, meaning that gratitude per se increases subject's probability to behave in a reciprocal manner and more specifically, that gratitude in form of a note induces subjects to exert higher effort compared to no gratitude at all. However, contrary to my propo-

---

[27]Replacing the treatment variables by GratitudeAll reports comparable

results with respect to economical and statistical significance. Appendix A5 reports the respective results.

sition, the video intervention does not seem to have a statistically greater impact on individuals' effort than GratitudeNote or Baseline, resulting in no evidence for hypothesis 3. Lastly, throughout all tests and regressions I detect a positive relationship between female subjects and the continue dummy, indicating support for hypothesis 4, namely that women show more reciprocity than men and hence provide more effort than male participants.

4.2. Effort proxy 2

In the following, I examine whether individual's level of provided effort differs between treatment groups as well as within-subjects, hence before and after treatment, by comparing the results of the real-effort tasks. Starting off with the between subject comparison, I assess the quantity of data entries in working period 2 and compare these between Baseline, GratitudeNote, GratitudeVideo, and GratitudeAll.

Figure 3 shows the average quantity of words entered in the second real-effort task, by displaying means and 95% confidence bands for the considered groups.

Subjects of GratitudeNote enter, on average, 10.9 words in effort task 2 which is significantly more than 7.9 words, that are taken down by individuals in the Baseline treatment ($p = 0.022$, Mann-Whitney-U (MWU) test). Furthermore, the expression of gratitude in form of a video results in an average of 8.6 words in working period 2, which is not substantially different to the mean level of effort provided by GratitudeNote ($p = 0.333$, MWU) nor by Baseline ($p = 0.172$, MWU). Pooling the gratitude treatments results in an average quantity of 9.8 words, which is significantly more compared to the entries by the control group ($p = 0.036$, MWU), who received no gratitude at all. Furthermore, the figure in Appendix A6 depicts the average effort in working period 2 for female and male individuals. As expected, women provide, on average, more effort than men ($p = 0.065$, MWU).

To estimate the causal effect of gratitude on the level of effort provided, I additionally estimated OLS regressions with the quantity of words in working task 2 as the dependent variable. The key independent variables are the respective treatments, a dummy for gender, and two-way interaction terms between the specific treatments and gender. The underlying regression results are reported in Table 5[28].

The first key observation is that the expression of gratitude via a thank you note has a sizable effect on the quantity of words provided in the real-effort task. In fact, although statistically insignificant, subjects of the GratitudeNote treatment provide on average approximately 3.5 more words in working period 2 than individuals of the Baseline group ($p = 0.130$, column 4). By further considering the gender variable and the interaction terms, the effect of the GratitudeNote intervention on the quantity of words increases in magnitude and gains in statistical significance ($p = 0.020$, column 6). This phenomenon has already been observed for

effort proxy 1. One possible explanation might be that by including the dummy variable for intrinsic motivation, the total sample reduces by 19 participants[29]. Assuming those excluded participants to behave very differently from the average individual of the sample pool would explain these divergent estimates. Regression results predicted in column 7, in which intrinsic motivation was excluded from the controls, seem to support this thought.

Moreover, while coefficients for the GratitudeVideo treatment are also positive, indicating a positive influence of being exposed to the thank you video and effort provision, the effects are not significant and further not statistically distinguishable from the estimates of the thank you note intervention ($p > 0.1$, Wald test). Lastly, I find explicit evidence that female participants take down more words than male subjects, whereas I detect a negative but insignificant effect when studying the interplay of gender and treatments ($p = 0.061$, column 5 and $p > 0.1$, column 6).

Replicating the OLS regressions for GratitudeAll yields comparable estimates that are displayed in Table 6[30].

More precisely, pooling the gratitude treatments also yields an increase of subject's effort. However, this effect is only significant when including the gender variable, the interaction term between GratitudeAll and gender, and control variables (p = 0.052, column 6). The negative, though insignificant, estimate for the interaction term suggests that the magnitude of the treatment effect is at least partly driven by male respondents. Furthermore, as delineated earlier, the results offer substantial evidence for women behaving more reciprocally than men (p = 0.061, column 5). Nevertheless, the estimates in model 6 predict that the gender effect is mainly driven by female subjects who are part of the control group.

Similar to the analysis of the relationship between positive reciprocity and the provision of effort in section 4.1, I am curious about whether subject's inclination for reciprocity also played a role for the number of words taken down in task 2. Thus, I replicated the regression models displayed in Table 4 and replaced the dependent variable with the effort level in period 2. The key results are displayed in Table 7.

Surprisingly and in contrast to the findings of reciprocal inclination and effort proxy 1, I do not find that positive reciprocity significantly influences the level of effort exerted in working period 2[31]. With respect to the within-subjects observation, Figure 4 outlines the mean effort of task 2 compared to task 1 within-subjects by differentiating between Baseline, GratitudeNote, GratitudeVideo, and GratitudeAll[32].

---

[28]The table depicted in Appendix A7 replicates Table 5 using log values. No substantial differences are detected.

[29]The dummy for intrinsic motivation excludes 19 participants who continued to the second form and stated to have not been asked to fill out any surveys in the past weeks.

[30]The table depicted in Appendix A8 replicates Table 6 using log values.

[31]Estimating the regressions for GratitudeAll yields comparable insignificant outcomes. The regression outcomes are reported in Appendix A9.

[32]It is relevant to remark that the within-subjects analysis solely considers individuals who continued to the second survey and finished effort task 2. Therefore, the sample size reduces to n = 159. Descriptive characteristics of the main results are summarized in Appendix A3.
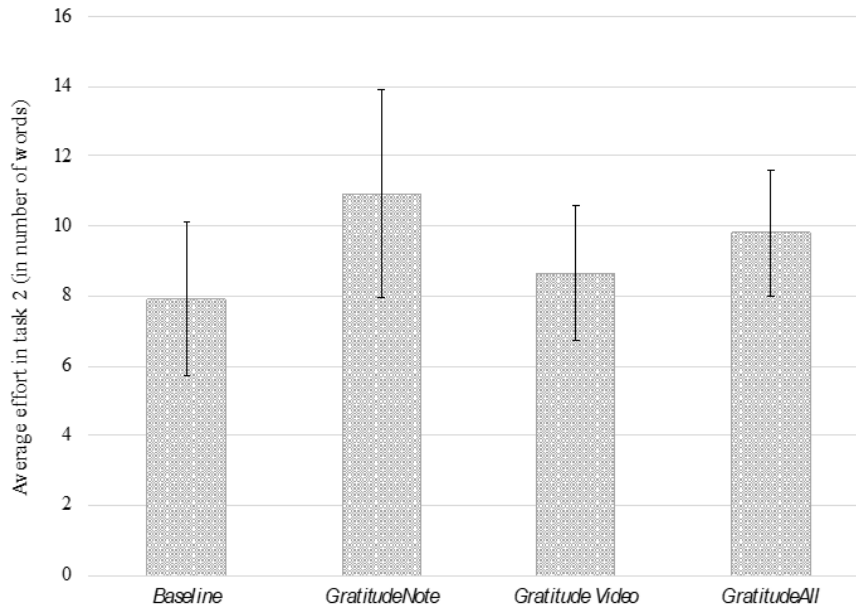
Figure 3: Average effort in working period 2 by treatment

Table 5: Effort proxy 2 – Quantity of words

Note: The table displays coefficients from OLS regressions with robust standard errors reported in parentheses. The dependent variable is the number of words provided in effort task 2. All results are compared to the Baseline treatment. Controls include several dummies, namely for being between 25-34 years old, for using a smartphone to participate, for economics and business students, for individuals' intrinsic motivation, and lastly for being fulltime employed. Model 7 includes all controls besides intrinsic motivation. $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GratitudeNote	3.013	3.033	2.874*	3.568	3.708	5.102**	4.102**
	(1.893)	(1.886)	(1.713)	(2.343)	(2.338)	(2.158)	(1.868)
GratitudeVideo	0.750	0.986	1.241	1.592	1.897	3.882	2.099
	(1.490)	(1.448)	(2.154)	(1.942)	(1.905)	(3.882)	(2.288)
Female		2.611*	2.682		2.723*	4.328**	3.767**
		(1.275)	(1.649)		(1.444)	(2.168)	(1.836)
GratitudeNote x Female			0.207			-1.662	-1.028
			(2.955)			(2.950)	(2.696)
GratitudeVideo x Female			-0.362			-2.502	-1.072
			(2.806)			(3.709)	(2.906)
Constant	7.898***	5.873***	5.818***	6.291***	3.787	2.297	3.979*
	(1.127)	(1.266)	(0.817)	(2.277)	(2.588)	(2.818)	(2.088)
Wald test: GratitudeNote = GratitudeVideo	p=0.212	p=0.246	p=0.514	p=0.305	p=0.337	p=0.664	p=0.393
Controls	No	No	No	Yes	Yes	Yes	Yes
Observations	159	159	159	140	140	140	159
Adjusted R^2	0.008	0.017	0.004	0.019	0.026	0.013	0.012

In addition, the figure displays 95% confidence bands, to represent the uncertainty of the estimates. First of all, one can observe that subjects who were part of the Gratitude-Note treatment provided approximately 2.5 more words in the second working period compared to the first working period ($p < 0.001$, Wilcoxon Signed-Rank (WSR) Test). Comparing the effort level of task 1 and 2 for GratitudeVideo appears to result in an increase as well, however the effort rises

by 1.4 words only ($p = 0.099$, WSR). Moreover, individuals of the Baseline treatment tend to provide less effort in the second task than in the first task ($p = 0.013$, WSR). Pooling gratitude treatments depicts an average effort increase of about 2 words in working period 2, suggesting that a gift in form of an expression of appreciation spurs individual's motivation to provide more effort, while not providing gratitude leads to a decrease of motivation to elicit effort ($p < 0.001$,

Table 6: Effort proxy 2 – Quantity of words (GratitudeAll)

Note: The table displays coefficients from OLS regressions with robust standard errors reported in parentheses. The dependent variable is the quantity of words in effort task 2. All results are compared to the Baseline treatment. Controls include several dummies, namely for being between 25-34 years old, for using a smartphone to participate, for economics and business students, for individuals' intrinsic motivation, and lastly for being fulltime employed. Model 7 includes all controls besides intrinsic motivation. $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GratitudeAll	1.902	2.035	1.948	2.635	2.862	4.425*	2.952*
	(1.448)	(1.439)	(1.536)	(1.936)	(1.982)	(2.252)	(1.768)
Female		2.764*	2.682		2.835*	4.309**	3.783**
		(1.330)	(1.638)		(1.498)	(2.141)	(1.827)
GratitudeAll x Female			0.114			-1.911	-0.836
			(2.392)			(2.797)	(2.343)
Constant	7.898***	5.755***	5.818***	6.283***	3.677	2.318	3.899*
	(1.123)	(1.300)	(0.811)	(2.260)	(2.613)	(2.758)	(2.068)
Controls	No	No	No	Yes	Yes	Yes	Yes
Observations	159	159	159	140	140	140	159
Adjusted R^2	0.003	0.015	0.008	0.018	0.027	0.021	0.016

Table 7: Effort proxy 1 – The impact of positive reciprocity

Note: The table displays coefficients from OLS regression with standard errors in parentheses. The treatments GratitudeNote and GratitudeVideo are both compared to the Baseline group. Positive reciprocity is the centered mean response of individuals to the reciprocity proxy. Controls include several dummies, namely for being between 25-34 years old, for using a smartphone to participate, for economics and business students, for individuals' intrinsic motivation and lastly for being fulltime employed. $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***.

	(1)	(2)	(3)	(4)
GratitudeNote	2.967	2.991	3.702	3.682
	(1.902)	(1.924)	(2.339)	(2.320)
GratitudeVideo	0.967	0.933	1.901	1.815
	(1.449)	(1.474)	(1.898)	(1.891)
Female	2.595**	2.586**	2.776*	2.896**
	(1.280)	(1.296)	(1.460)	(1.331)
Positive reciprocity	0.651	0.298	1.105	0.601
	(0.651)	(1.409)	(0.985)	(2.370)
GratitudeNote x positive reciprocity		-0.491		-0.419
		(1.615)		(2.124)
GratitudeVideo x positive reciprocity		0.433		1.790
		(1.488)		(2.499)
Female x positive reciprocity		0.469		0.319
		(1.335)		(2.377)
Constant	5.885***	5.906***	4.054	3.947
	(1.280)	(1.277)	(2.707)	(2.706)
Controls	No	No	Yes	Yes
Observations	159	159	140	140
Adjusted R^2	0.014	-0.005	0.025	0.006

WSR).

This picture is confirmed by regression analysis. To conclusively observe and compare subject's behaviour across time, it is reasonable to generate longitudinal data to run panel data regressions. Hence, I duplicated the dataset and specified panel and time variable. In a next step, I estimated fixed effects regressions to control for any time-invariant differences between the individuals to obtain the net effect of the independent variable on the dependent variable (Torres-Reyna, 2007)[33]. Since all variables besides the treatment variables (e.g. gender, studies, age) remain unchanged between working period 1 and working period 2, they are

---

[33]More precisely, fixed effects estimates are within estimates while random effects estimates are a linear combination of both, within and between estimates (Charness et al., 2012).
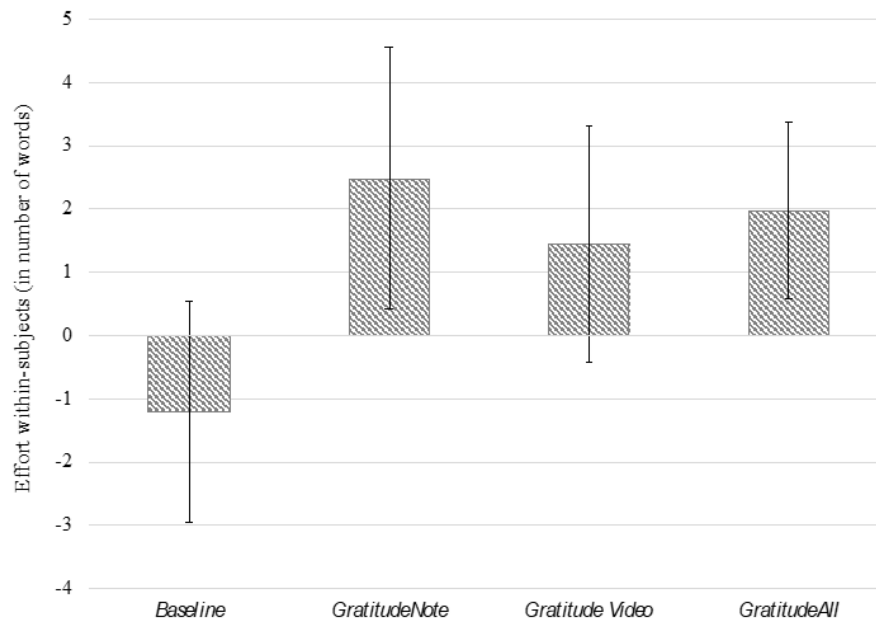
Figure 4: Effort within-subjects over time

omitted in the fixed effects model, which is why I only incorporate and report the results of the main treatment effects in Table 8.

Column 1 displays outcomes of the fixed effects regression with individual's effort in working period 2 as the dependent variable and the gratitude treatments as independent variables. Column 2 again accounts for fixed effects but replaces GratitudeNote and GratitudeVideo with GratitudeAll. Column 3 and 4 use the same specification however, instead of absolute figures, represent effort in logarithmized values. The results support the findings of the WSR test visualized in Figure 4 and show that individuals certainly behave reciprocally towards the thank you note by providing higher effort in the subsequent working task compared to their working effort in the period before being thanked by the experimenter ($p = 0.019$, column 1). However, expressing gratitude by means of a video leads to a smaller and statistically insignificant increase of effort in the following working period (p = 0.129, column 1). Overall it appears that individuals positively respond to expressions of gratitude, independent of note or video, in a reciprocal manner and put forth additional effort in return (p = 0.006, column 2). In sum, the key outcome of the analysis of effort proxy 2 are congruent to the results of effort proxy 1. Again, the statistical tests yield empirical support for hypothesis 1, hypothesis 2, and hypothesis 4 but, surprisingly provide no evidence for hypothesis 3. Accordingly, summarizing the results of the field experiment, I find the following:

1. Expressions of gratitude, combined with interpersonal elements, induce reciprocity and lead to the provision of higher effort compared to no expression of gratitude.
2. Expressing appreciation in form of a thank you note, signed by hand, results in a significantly higher recip-

rocal response than if the experimenter provided gratitude by means of a video or expressed no gratitude at all.
3. Women, on average, behave more reciprocally and accordingly exert higher effort, in response to an expression of gratitude, than men.

4.3. Robustness check

In order to validate the presented results, robustness checks are necessary to test whether the empirical findings are robust to different ways of measurement. Therefore, I investigated whether results for effort proxy 1 and 2 alter if additionally including the total time spent on the survey experiment. The respective results are displayed in Appendix A10 and show that results are, at least relating to their economical size, robust to adding an additional variable.

4.4. Manipulation check

In order to test the effectiveness of the gratitude intervention, participants' responses to the manipulation check were analysed. As delineated earlier, the manipulation check consisted of an item asking participants to indicate the extent to which the experimenter expressed appreciation on a 7-point Likert scale. The data show that the note worked as desired[34]. More precisely, subjects reported that the experimenter showed gratitude for the invested time and effort to fill out the survey as compared to the neutral response of 4 (mean = 5.7, SD = 2.3). Moreover, this mean value is significantly different from Baseline (mean = 4.4, SD = 1.5,

---

[34]Descriptive statistics for the manipulation check are subsumed in Appendix A11.

Table 8: Effort proxy 2 – Quantity of words (within-subjects)

Note: The table displays results from fixed effects regressions with robust standard errors reported in parentheses. The dependent variable is the quantity of words in working period 2 compared to Baseline and within-subjects in column 1-3. In column 4-6 the dependent variable is logarithmized. $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***.

	(1) Effort	(2) Effort	(3) ln (Effort)	(4) ln (Effort)
GratitudeNote	2.482**		0.238***	
	(1.052)		(0.092)	
GratitudeVideo	1.444		0.168	
	(0.946)		(0.104)	
GratitudeAll		1.973***		0.204***
		(0.709)		(0.069)
Constant	7.907***	7.907***	1.765***	1.765***
	(0.213)	(0.213)	(0.021)	(0.021)
Wald test: GratitudeNote = Gratitude-Video	p=0.464		p=0.609	
Observations	366	366	366	366
Within R^2	0.054	0.051	0.052	0.051
Overall R^2	0.017	0.012	0.015	0.013

$p < 0.001$, MWU). Participants of the GratitudeVideo treatment, who saw the thank you video and thus are expected to score higher on the mean of the manipulation check, exhibited the highest mean of all three treatments (mean = 5.8, SD = 1.1). However, contrary to my expectations, this mean response is only statistically different from the average response of the control treatment, but not from the mean value of GratitudeNote ($p < 0.001$ and $p = 0.947$, respectively, MWU). This suggests that the video did not succeed in expressing considerably more gratitude than the note, which might explain some of the surprising results stemming from the analysis.

5. Discussion

5.1. Interpretation of results

The empirical results presented in section 4 show that expressions of gratitude matter and positively affect individual's reciprocal behaviour. While at first glance, the statistical tests support the behavioural intuition behind the basic hypothesis that gratitude induces reciprocity, other results did not lead to explicit and persuasive conclusions. For instance, it seems surprising that the expression of gratitude by means of a personal video did not elicit substantially higher effort compared to not expressing gratitude at all. In addition, considering the psychological mechanism behind individuals' response to gratitude, the empirical results show that reciprocal inclination certainly plays a role and may, at least to some extent, explain subject's behaviour, which will be discussed in the following.

First of all, the findings of the underlying paper provide evidence for hypothesis 1, indicating that individuals assign value to gratitude and accordingly behave in a reciprocal manner. This is in line with the results of earlier studies about

gratitude, that have emphasized the role of gratitude expressions as a moral reinforcer in enhancing subject's prosocial behaviour (Grant and Gino, 2010; McCullough et al., 2001). More particular and in support of hypothesis 2, expressing gratitude by means of a thank you note, combined with a handwritten signature, elicits reciprocity and significantly increases effort compared to when no gratitude was expressed. Although it is not clearly apparent, whether this effect was stimulated by the interpersonal element attached to the note, I anticipate subjects to have perceived the handwritten signature as something exceptional and thoughtful which then in turn unconsciously affected the way how the gratitude expression was perceived by the individual (Bradler and Neckermann, 2019). In contrast to what literature on surprising gifts propose, the results provide no proof that expressing gratitude in form of a video clip induces significantly higher effort than when expressing appreciation via a thank you note nor when not expressing gratitude at all. One explanation for this effect could relate to the intensity of the video itself, in other words that the appreciative expression by means of the personal video was too weak to considerably affect behaviour. This is partly confirmed by the evaluation of the manipulation check which delineates that though subjects perceived the video as an expression of gratitude, this perception is not substantially different to the perception of the thank you note. However, I detect a significant difference between seized gratitude level of the video message and the control group, inferring that the manipulation only partly failed and that there must be another cause for subject's effort choice. Another reason for the lack of a significant effect may relate to the phenomenon of gender-pairing, implying that subjects vary their behaviour depending on the gender of the individual with whom they are interacting. This assertion relates to the work of Ben-Ner et al. (2004b) who conducted a dicta-

tor game experiment and found that women on average give less to female recipients than to male recipients or to individuals about whom no gender information was provided. A reason for this behaviour may be that women perceive other women as a potential threat which triggers the expression of covert and low-key aggressions towards individuals of the same gender (Campbell, 1999). With respect to the present study, the knowledge of the experimenters gender thus would affect participants effort choice. As the personal video indeed provided the participant with the information that the experimenter is a woman, it is conceivable that female participants who were part of GratitudeVideo intentionally acted in a non-reciprocal manner[35]. A brief look at the regression results in section 4 supports this possible explanation. Regarding effort proxy 1, being female and having seen the thank you video reduced the likelihood to continue to the next form by about 22 percentage points ($p = 0.237$, column 6, Table 2). Congruent findings provide the regression models for effort proxy 2. More precisely, the effect size of GratitudeVideo is reduced by 2.5 words for female subjects ($p = 0.501$, column 6, Table 5). Although these effects are both not statistically significant, they should not be neglected due to their considerably large magnitude. Thus, it is noteworthy that sharing information about the gender of the experimenter matters and potentially affects giving behaviour of female subjects.

Other reasons for the obtained results could relate to the issue that individuals might have felt that they are being manipulated or, even worse, might have perceived the personal video as a desperate and exorbitant exaggerated form of appreciating working effort of approximately 5 minutes and consequently felt kidded by the experimenter (Simonson et al., 1994). However, these are just assumptions that the present data cannot validate and thus remain potential causes.

As anticipated in hypothesis 4, women on average exerted more effort than men. Nevertheless, statistical tests show that this effect was most powerful and greatest in size for female subjects of the control treatment (p = 0.005, column 6, Table 3 and p = 0.046, column 6, Table 6, respectively)[36]. Though this behaviour can be partly explained by the above depicted determinant of gender-pairing, this explanation primarily holds for the video rather than for the note treatment. Hence, I assume there to be further determinants predicting individual's behaviour, especially for women. One possible cause could be a personal relationship with the experimenter that might be more widely distributed among Baseline compared to GratitudeNote. However, since I did not incorporate a survey item asking for whether individuals personally know the experimenter, I cannot further verify this notion.

Finally, a leading explanation for individual's reciprocal behaviour is their reciprocal inclination[37]. The key outcomes reported in section 4 provide evidence that subject's reciprocal inclination positively and significantly influenced the probability to continue to the next questionnaire. Furthermore, it became evident that the impact of positive reciprocity was stronger for male than for female subjects. This is contrary to the findings of Altmann et al. (2008) who observe no significant gender difference when investigating the relationship between person's reciprocal inclination and their trusting behaviour. In the present paper, no discernible interaction effect was found between the specific treatments and the measure for positive reciprocity, indicating that the impact of subject's reciprocal inclination on the probability to continue was more or less equivalent among treatments. This stands in contrast to the findings of Sliwka and Werner (2017), who demonstrate that certain treatment effects are driven by positively reciprocal workers. Surprisingly, the degree of reciprocal inclination is not significantly affecting effort choices in the second period, implying that more reciprocal subjects do not substantially exert higher effort in the second real-effort task than less reciprocal subjects[38]. This suggests that respondent's degree of reciprocal inclination might not have been their exclusive motivation to provide extra effort in working period 2. Alternative motivations such as individual's interest in the real-effort task and survey topic or, as already stated, a personal relationship with the experimenter are conceivable.

5.2. Limitations and future research

As with most empirical research, the current study is not without limitations that have to be well considered to guide directions for future research. A central weakness of this paper relates to the methodology of the study. While field experiments in general promote higher external validity than laboratory experiments, there are severe disadvantages entangled to this experimental method. First of all, it is difficult to keep external determinants of individual's behaviour constant among all respondents. These include, for instance, whether individuals are interacting or communicating with someone while participating in the survey experiment. As the study requires a great amount of attention to correctly perceive the gratitude intervention, not being able to control for this factor may pose a significant weakness. Second, it is unfeasible to obtain the precise identity of the participants. Thus, it is possible that subjects completed the experiment with the help of other individuals or alternatively participated more than once which raises the potential challenge of validity of results. Another powerful determinant

---

[35]Participants of GratitudeNote and Baseline were not explicitly informed about the gender of the experimenter. Moreover, they were merely presented the name of the experimenter at the beginning of the experiment, which I assume is not enough to substantially influence female behavior.

[36]Though the negative interaction effects between GratitudeAll and gender dummy are not significant, I perceive their impacts on behaviour relevant and hence take them into consideration in the present interpretation.

[37]Especially in a setting where effort is not enforceable, individual's reciprocal inclination gains in importance.

[38]This may have also resulted due to the limited observation of merely those subjects that have decided to continue when analyzing effort proxy 2. It is possible, that the sample considered in effort proxy 2 (n = 159) exhibits less differences with respect to their reciprocal inclination than when referring to the overall subject pool considered in effort proxy 1 (n = 207).

which I couldn't control for was the environmental situation that surrounded individuals while taking part in the survey experiment. According to Sukumaran et al. (2011, 3403):

> Everyday physical environments and objects possess re-established associations with normative behaviours: people tend to act differently when they are in a church, for instance, as opposed to a football stadium.

This infers that the design of the environment is a source of situational norms that is assumed to shape actions which especially affect subject's (pro)social behaviour (Aarts and Dijksterhuis, 2003). Another shortcoming of online compared to laboratory experiments relates to the differences in selection bias which assumes individuals to self-select considering their reservation wage and opportunity costs of working time. With respect to the voluntarily characteristics of online surveys, subjects who decided to take part in my study already demonstrate a certain degree of willingness to benefit others. This implies that subjects in the present sample might generally be more helpful and cooperative than individuals of sample pools of laboratory experiments, where respondents commonly receive at least a show-up fee (Arechar et al., 2018)[39]. This bias cannot be ruled out in the present study which limits external validity in terms of the representative nature of experimental participants. Given the fact that the majority of respondents are female students with a business or economics background and an age between 25 and 34 years further reduces subjects' representativeness. Since the above noted factors question the overall generalisability of depicted findings, it would be interesting to replicate the experiment in a controlled laboratory setting and compare results afterwards.

Another concrete limitation of this study refers to the narrow focus on subject's effort as the main outcome variable. While organizations usually not only attach importance to the quantity but further to the quality of work, I'd recommend researchers to additionally investigate the impact of gratitude on subject's performance to obtain a clearer picture of the overall effectiveness of appreciation as a non-monetary reward[40].

Further critical is the experimental feature of solely observing behaviour in a one-off interaction. Ample experimental studies demonstrate that performance effects, which have been found in one-shot settings, alter when considered in a repeated decision-making design (Carpenter and Gong, 2016; Gerhards, 2015). Transmitted to the context of the present study, one concern may be that recurring appreciation would no longer be perceived as personal and thoughtful and in worst case result in participants' doubt towards the sender's seriousness. Besides, decreasing effect sizes could

potentially arise due to the common condition of habituation, mental exhaustion, and increasing fatigue (McSweeney, 2004; Brachet et al., 2012). Contrary results are demonstrated by Kirchler and Palan (2018), who show that reciprocal behaviour conditional on repeated non-pecuniary gifts de facto grows and becomes stronger over time. To draw reliable conclusions and thus provide significant implications for management, it is therefore highly recommended to extend the current experiment to a repeated interaction setting to analyse the stability of gratitude as an immaterial gift over time.

Moreover, as this study only provides a brief insight into the causal relationship between appreciative expressions and individual's effort choice, further research is needed to analyse moderating and mediating variables which may significantly influence the way gratitude worked in the underlying experiment. For instance, Grant and Gino (2010) argue that the positive effect of gratitude is mediated by the psychological mechanism of social worth. Likewise, research on interactive effects of social responses has shown that differences in need for approval moderated individual's subsequent helping behaviour (Deutsch and Lamberti, 1986). Examining whether such personality variables mediate or moderate the investigation of the present study setting would depict an interesting avenue for future research.

Lastly, I advise future research to consider the distinction of gratitude according to their source. This differentiation has been largely ignored by existing research streams on gratitude but is receiving increasing attention on the labour market. According to Brun and Dugas (2008) there exist four different types of sources. While appreciation may be conveyed vertically (e.g. by the supervisor) or horizontally (e.g. by a colleague) within organizations, it can originate from immediate stakeholder groups outside the organization (e.g. by a client), or alternatively stem from any other stakeholder party that stands in contact with the organization. Exploring whether subjects respond differently to expressions of gratitude, depending on its sender, appears to be not only an interesting approach for future research work but additionally may bear revealing implications for management practices.

5.3. Implications

The findings of the above presented experiment provide essential insights for present research and depict valuable implications for organizations.

First, this study corresponds to existent research on non-financial gifts by providing additional evidence that gratitude in form of a short note positively induces reciprocal behaviour and thus affects subject's effort choice. Second, as one of the first studies in this field, I stress the importance of examining differing levels of gratitude expressions. While my empirical analysis rejects the theoretical driven assumption that a thank you video elicits more effort than a thank you note, I find partly significant differences, though to the other direction, between the respective treatment effects. This finding adds to present research by suggesting that

---

[39]The underlying survey experiment did not promise subjects to receive any kind of payment for their participation.

[40]For example, Kosfeld and Neckermann (2011) and Bradler et al. (2016) define performance instead of effort as their main outcome variable.

the type of presentation and communication of gratitude indeed may induce systematic changes in behaviour and should not be neglected in further research on gratitude[41]. Third, I emphasize the importance of taking a closer look at gender as an explaining variable for reciprocal behaviour. While there is an on-going debate on gender-specific outcomes in academic literature, the findings of the present field study corroborate to existing research from laboratory experiments indicating that women behave more reciprocally than men (see e.g. Heinz et al., 2012). Moreover, since behaviour of female respondents varied strongly among the treatment groups, this study manifests the assertion that female subjects are much more sensitive to the context of a situation than male respondents which can, at least to some extent, explain the mixed results of gender differences in social preferences (Croson and Gneezy, 2009).

Besides, this study poses clear implications for management. First and foremost, the findings indicate that an expression of gratitude can be a strong motivator for individuals to put forth additional effort and can as such constitute a particularly cost-effective and convenient tool for organizations to stimulate employees' average effort. Since the experimental study investigated the effectiveness of gratitude in form of an electronic message, this managerial tool becomes particularly powerful in times of digital age, where agile working and the use of FWAs are becoming more and more common and communication mainly takes place via digital media. Moreover, appreciative messages do not only represent a noteworthy alternative to monetary incentives but further are thought to have social effects. In their work on functions of gratitude, Algoe et al. (2008) suggest gratitude to be an important determinant for relationship building and maintenance. Well-established social relationships between supervisor and subordinate are vital for organizations, not only because they incentivize employees to work well but further because they are assumed to promote worker's happiness (Polak and McCullough, 2006; Watkins et al., 2003). Employing happy individuals is an essential cornerstone for organization's success, which has been underpinned by the latest movement in human resources of introducing "feel good managers", who are fully responsible for improving employees' happiness and general well-being in and outside the organization (Frenking, 2016).

6. Conclusion

Understanding the incentive effects of financial and non-monetary incentives on employee's motivation and performance is an omnipresent issue that expanded rapidly in the past decade. Although behavioural economists agree that individuals not merely hold selfish and materialist preferences, research on immaterial bonus domains is still scarce and incomplete. The underlying study is one of the first studies in this research field investigating whether different levels of gratitude expressions – conveyed via the Internet – lead to different levels of effort individuals are willing to provide. To analyse how appreciation incentivizes subjects, I depicted a formal model of gift-exchange, where individuals derive utility from reciprocating a non-financial gift of gratitude with higher effort. The results of the underlying field experiment support the proposed theoretical model. Particularly, receiving a gift of appreciation via an electronical transmitted note increased individual's effort by approximately 20% compared to a gift of gratitude in form of a video clip, and by more than 25% compared to not receiving a gift at all[42]. However, the underlying regression models demonstrate that the observed difference of descriptive figures is only statistically significant between the note and control treatment. Furthermore, this study delineates noteworthy gender differences in social behaviour, revealing that women per se behave in a more reciprocal manner than men, whereas individual's degree of reciprocal inclination has a greater impact on effort for male than for female respondents. In conclusion, the underlying paper makes essential contributions to present research and further emphasizes the urgency for organizations to understand the beneficial features of the manifestation of gratitude.

---

[41]Pre-testing subject's perception of different types and forms of appreciative expressions could be a reasonable approach to avoid false conclusions.

[42]These figures represent the average of the results of the two effort proxies reported in the descriptive statistics summarized in Appendix A3.

References

Aarts, H. and Dijksterhuis, A. The silence of the library: environment, situational norm, and social behavior. *Journal of Personality and Social Psychology*, 84(1):18, 2003.

Adair, J. G. The Hawthorne effect: a reconsideration of the methodological artifact. *Journal of Applied Psychology*, 69(2):334, 1984.

Agell, J. Why are small firms different? Managers' views. *Scandinavian Journal of Economics*, 106(3):437–452, 2004.

Akerlof, G. A. Labor contracts as partial gift exchange. *The Quarterly Journal of Economics*, 97(4):543–569, 1982.

Aldrich, J. H., Nelson, F. D., and Adler, E. S. Linear probability, logit, and probit models. *Sage*, 45, 1984.

Algoe, S. B., Haidt, J., and L, G. S. Beyond reciprocity: Gratitude and relationships in everyday life. *Emotion*, 8(1):425–429, 2008.

Altmann, S., Dohmen, T., and Wibral, M. Do the reciprocal trust less? *Economics Letters*, 99(3):454–457, 2008.

Arechar, A. A., Gächter, S., and Molleman, L. Conducting interactive experiments online. *Experimental Economics*, 21(1):99–131, 2018.

Aspromourgos, T. On the origins of the term 'neoclassical'. *Cambridge Journal of Economics*, 10(3):265–270, 1986.

Baker, G. P., Jensen, M. C., and Murphy, K. J. Compensation and incentives: Practice vs. theory. *The Journal of Finance*, 43(3):593–616, 1988.

Bandura, A. *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice-Hall, 1986.

Ben-Ner, A., Kong, F., Putterman, L., and Magan, D. Reciprocity in a two-part dictator game. *Journal of Economic Behavior & Organization*, 53(3): 333–352, 2004a.

Ben-Ner, A., Kong, F., and Putterman, L. Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving. *Journal of Economic Psychology*, 25(5):581–589, 2004b.

Berg, J., Dickhaut, J., and McCabe, K. Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1):122–42, 1995.

Beute, E. and Pacinelli, S. *Rehumanize Your Business: How Personal Videos Accelerate Sales and Improve Customer Experience*. N.J., Hoboken, John Wiley & Sons, 2019.

Blau, P. *Exchange and power in social life*. N.Y., New York, Wiley, 1964.

Bloom, N., Liang, J., Roberts, J., and Ying, Z. J. Does working from home work? Evidence from a Chinese experiment. *The Quarterly Journal of Economics*, 130(1):165–218, 2015.

Bolton, G. E. and Katok, E. An experimental test for gender differences in beneficent behavior. *Economics Letters*, 48(3-4):287–292, 1995.

Brachet, T., David, G., and Drechsler, A. M. The effect of shift structure on performance. *American Economic Journal: Applied Economics*, 4(2): 219–46, 2012.

Bradler, C. and Neckermann, S. The magic of the personal touch: Field experimental evidence on money and appreciation as gifts. *The Scandinavian Journal of Economics*, 121(3):1189–1221, 2019.

Bradler, C., Dur, R., Neckermann, S., and Non, A. Employee recognition and performance: A field experiment. *Management Science*, 62(11):3085–3099, 2016.

Brun, J. P. and Dugas, N. An analysis of employee recognition: perspectives on human resources and practices. *The International Journal of Human Resource Management*, 19(4):716–730, 2008.

Cambridge University Press. Cambridge Dictionary online, 2019. URL https://dictionary.cambridge.org/dictionary/english/gratitude. Retrieved on November 23, 2019.

Campbell, A. Staying alive: Evolution, culture, and women's intrasexual aggression. *Behavioral and Brain Sciences*, 22(2):203–214, 1999.

Carpenter, J. and Gong, E. Motivating agents: How much does the mission matter? *Journal of Labor Economics*, 34(1):211–236, 2016.

Charness, G., Gneezy, U., and Kuhn, M. A. Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1):1–8, 2012.

Clark, R. D. The effects of reinforcement, punishment and dependency on helping behavior. *Personality and Social Psychology Bulletin*, 1(4):596–599, 1975.

Cox, J. C. and Deck, C. A. When Are Women More Generous than Men? *Economic Inquiry*, 44(4):587–98, 2006.

Croson, R. and Buchan, N. Gender and culture: International experimental evidence from trust games. *American Economic Review*, 89(2):386–391, 1999.

Croson, R. and Gneezy, U. Gender Differences in Preferences. *Journal of Economic Literature*, 47(2):448–474, 2009.

Deutsch, F. M. and Lamberti, D. M. Does social approval increase helping? *Personality and Social Psychology Bulletin*, 12(2):149–157, 1986.

Dittrich, M. Gender differences in trust and reciprocity: evidence from a large-scale experiment with heterogeneous subjects. *Applied Economics*, 47(36):3825–3838, 2015.

Dufwenberg, M., Gneezy, U., Güth, W., and Van Damme, E. Direct vs indirect reciprocity: An experiment. *Homo Oeconomicus*, 18(1):19–30, 2001.

Dur, R., Non, A., and Roelfsema, H. Reciprocity and incentive pay in the workplace. *Journal of Economic Psychology*, 31(4):676–686, 2010.

Eckel, C. and Wilson, R. K. Whom to trust? Choice of partner in a trust game. *Department of Economics, Virginia Tech*, 2004.

Eckel, C. C. and Grossman, P. J. The relative price of fairness: Gender differences in a punishment game. *Journal of Economic Behavior & Organization*, 30(2):143–158, 1996.

Ellingsen, T. and Johannesson, M. Paying respect. *Journal of Economic Perspectives*, 21(4):135–150, 2007.

Emmons, R. A. and McCullough, M. E. The psychology of gratitude. *Oxford University Press*, 2004.

Falk, A. and Fischbacher, U. A theory of reciprocity. *Games and Economic Behavior*, 54(2):293–315, 2006.

Fehr, E. and Falk, A. Wage rigidity in a competitive incomplete contract market. *Journal of Political Economy*, 107(1):106–134, 1999.

Fehr, E., Kirchsteiger, G., and RIedl, A. Does fairness prevent market clearing? An experimental investigation. *The Quarterly Journal of Economics*, 108(2):437–459, 1993.

Feild, L., Pruchno, R. A., Bewley, J., Lemay, E. P., and Levinsky, N. G. Using probability vs. nonprobability sampling to identify hard-to-access participants for health-related research: costs and contrasts. *Journal of Aging and Health*, 18(4):565–583, 2006.

Finstad, K. Response interpolation and scale sensitivity: Evidence against 5-point scales. *Journal of Usability Studies*, 5(3):104–110, 2010.

Frederickson, B. L. What good are positive emotions? *Review of General Psychology*, 2(3):300–319, 1998.

Frenking, S. Feel good management as valuable tool to shape workplace culture and drive employee happiness. *Strategic HR Review*, 2016.

Frey, B. S. Awards as compensation. *European Management Review*, 4(1): 6–14, 2007.

Gajendran, R. S. and Harrison, D. A. The good, the bad and the unknown about telecommuting: Meta-analysis of psychological mediators and individual consequences. *Journal of Applied Psychology*, 92(1):1524–1541, 2007.

Gawel, J. E. Herzberg's theory of Motivation and Maslow's hierarchy of needs. *Practical Assessment, Research, and Evaluation*, 5(1):11, 1997.

Gerhards, L. The incentive effects of missions - Evidence from experimentswith NGO employees and students. *European Economic Review*, 79 (1):252–262, 2015.

Gilchrist, D. S., Luca, M., and Malhotra, D. When 3+ 1> 4: Gift structure and reciprocity in the field. *Management Science*, 62(9):2639–2650, 2016.

Gneezy, U. and List, J. Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, 74 (5):1365–1384, 2006.

Grant, A. M. and Gino, F. A little thanks goes a long way: Explaining why gratitude expressions motivate prosocial behavior. *Journal of Personality and Social Psychology*, 98(6):946, 2010.

Hammermann, A. and Mohnen, A. The pric(z)e of hard work: Different incentive effects of non-monetary and monetary prizes. *Journal of Economic Psychology*, 43(1):1–15, 2014.

Harrison, G. W. and List, J. A. Field experiments. *Journal of Economic Literature*, 42(4):1009–1055, 2004.

Heckhausen, J. and Heckhausen, H. *Motivation und Handeln: Einführung und Überblick. In Motivation und Handeln*. Berlin, Springer, 2018.

Heinz, M., Juranek, S., and Rau, H. A. Do women behave more reciprocally than men? Gender differences in real effort dictator games. *Journal of Economic Behavior & Organization*, 83(1):105–110, 2012.

Hill, P. L. and Allemand, M. Gratitude, forgivingness, and well-being in adulthood: Tests of moderation and incremental prediction. *Journal of Positive Psychology*, 16(1):397–407, 2011.

Katz, L. F. and Krueger, A. B. The rise and nature of alternative work arrangements in the United States, 1995-2015. *ILR Review*, 72(2):382–416,

2019.

Kirchler, M. and Palan, S. Immaterial and monetary gifts in economic transactions: evidence from the field. *Experimental Economics*, 21(1):205–230, 2018.

Korzynski, P. Employee motivation in new working environment. *International Journal of Academic Research*, 5(5):184–188, 2013.

Kosfeld, M. and Neckermann, S. Getting more work for nothing? Symbolic awards and worker performance. *American Economic Journal: Microeconomics*, 3(3):86–99, 2011.

Kothari, C. R. Research Methodology: Methods and Techniques. *New Age International*, 2004.

Kovach, K. A. What motivates employees? Workers and supervisors give different answers. *Business Horizons*, 30(5):58–65, 1987.

Kube, S., Maréchal, M. A., and Puppe, C. The currency of reciprocity: Gift exchange in the workplace. *American Economic Review*, 102(4):1644–1662, 2012.

Lazear, E. P. Performance pay and productivity. *American Economic Review*, 90(5):1346–1361, 2000.

Lewis, S. Flexible working arrangements: Implementation, outcomes, and management. *International Review of Industrial and Organizational Psychology*, 18(1):1–28, 2003.

Macera, R. and te Velde, V. On the Power of Surprising versus Anticipated Gifts in the Workplace. *Working Paper*, 2018.

Maheux, B., Legault, C., and Lambert, J. Increasing response rates in physicians' mail surveys: an experimental study. *American Journal of Public Health*, 79(5):638–639, 1989.

Maslow, A. H. *Motivation and personality*. Delhi, Pearson Education, 1987.

McCullough, M. E., Kilpatrick, S. D., Emmons, R. A., and Larson, D. B. Is gratitude a moral affect? *Psychological Bulletin*, 127(2):249–266, 2001.

McCullough, M. E., Tsang, J. A., and Emmons, R. A. Gratitude in Intermediate Affective Terrain: Links of Grateful Moods to Individual Differences and Daily Emotional Experience. *Journal of Personality and Social Psychology*, 86(2):295–309, 2004.

McDaniel, C. and Gates, R. *Marketing research*. Singapore, Wiley and Sons, 2013.

McSweeney, F. K. Dynamic changes in reinforcer effectiveness: Satiation and habituation have different implications for theory and practice. *The Behavior Analyst*, 27(2):171–188, 2004.

Neckermann, S. and Yang, X. Understanding the (unexpected) consequences of unexpected recognition. *Journal of Economic Behavior & Organization*, 135(1):131–142, 2017.

Nerdinger, F. *Arbeitsmotivation und Arbeitszufriedenheit. In Arbeits-und Organisationspsychologie*. Berlin, Springer, 2014.

Nowak, M. A. and Sigmund, K. Evolution of indirect reciprocity. *Nature*, 437 (7063):1291–1298, 2005.

Perugini, M., Gallucci, M., Presaghi, F., and Ercolani, A. P. The personal norm of reciprocity. *European Journal of Personality*, 17(4):251–283, 2003.

Polak, E. L. and McCullough, M. E. Is gratitude an alternative to materialism? *Journal of Happiness Studies*, 7(3):343, 2006.

Prutchno, R., Litchfield, L., and Fried, M. *Measuring the Impact of Workplace-Flexibility*. Boston: Boston College Center for Work and Family, 2000.

Pruyser, P. W. *The minister as diagnostician: Personal problems in pastoral perspective*. Louisville, KY, Westminster John Knox Press, 1976.

Rhoades, L. and Eisenberger, R. Perceived Organizational Support: A Review. *Journal of Applied Psychology*, 87(1):698–714, 2002.

Rind, B. and Bordia, P. Effect of Server's "Thank You" and Personalization on Restaurant Tipping 1. *Journal of Applied Social Psychology*, 25(9):745–751, 1995.

Rogers, T. and Frey, E. *Changing behavior beyond the here and now. The Wiley Blackwell Handbook of Judgment and Decision Making*. Chichester, Wiley and Sons, 2014.

Rosenstiel, L. and Nerdinger, F. *Grundlagen der Organisationspsychologie*. CE Poeschel Verlag, 1980.

Saunders, M. N. *Research methods for business students, 5/e*. Pearson Education India, 2011.

Simonson, I., Carmon, Z., and O'Curry, S. Experimental evidence on the negative effect of product features and sales promotions on brand choice. *Marketing Science*, 13(1):23–40, 1994.

Sliwka, D. and Werner, P. Wage Increases and the Dynamics of Reciprocity. *Journal of Labor Economics*, 35(2):299–344, 2017.

Snijders, C. and Keren, G. Do you trust? Whom do you trust? When do you trust? *Advances in Group Processes*, 18(1):129–60, 2001.

Stajkovic, A. and Luthans, F. Behavioral management and task performance in organizations: Conceptual background, meta-analysis, and test of alternative models. *Personnel Psychology*, 56(1):155–194, 2003.

Sukumaran, A., Vezich, S., McHugh, M., and Nass, C. Normative influences on thoughtful online participation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3401–3410, 2011.

Torres-Reyna, O. Panel data analysis fixed and random effects using Stata (v. 4.2). *Data & Statistical Services, Priceton University*, pages 1–40, 2007.

Tsang, J. A. Gratitude and prosocial behaviour: An experimental test of gratitude. *Cognition & Emotion*, 20(1):138–148, 2006.

Tversky, A. and Kahnemann, D. The framing of decisions and the psychology of choice. *Science*, 211(4481):453–458, 1981.

Vogelsang, T. Time is (Not) Money - Incentive Effects of Granting Leisure Time. Available at SSRN 3464318, 2019.

Watkins, P. C., Woodward, K., Stone, T., and Kolts, R. Gratitude and happiness: Development of a measure of gratitude, and relationships with subjective well-being. *Social Behavior and Personality: an international journal*, 31(5):431–451, 2003.

Wilcox, R. R. and Keselman, H. J. Robust regression methods: Achieving small standard errors when there is heteroscedasticity. *Understanding Statistics*, 3(4):349–364, 2004.

Wiley, C. What Motivates Employees According to over 40 Years of Motivation Surveys. *International Journal of Manpower*, 18(3):263–80, 1997.

Wood, A. M., Maltby, J., Stewart, N., Linley, P. A., and Joseph, S. A social-cognitive model of trait and state levels of gratitude. *Emotion*, 8(2):281, 2008.

To Be Is to Do: Exploring How Founder Social and Role Identities Shape Strategic Decisions in New Venture Creation Process

Magdalena Melonek

Technische Universität München

Abstract

Founder identity is essential in entrepreneurial decision making. A number of studies have sought to examine the relationship between founder identity and venture by taking a unilateral perspective of either social identity theory or identity theory. On the basis of an exploratory single-case study conducted with a German venture from the sustainable fashion industry, this bachelor thesis integrates both theories and explores how the synergy of a founder's social and role identities influences core strategic decisions in new venture-creation process. The results obtained in this analysis suggest that 1) founders' social and role identities have different yet complementary impact on the new venture-creation process, and 2) strategic decisions shaping a venture are a result of different interplays between multiple social and role identities, where either one of the identity types dominates a decision or both identities simultaneously reinforce it. My contributions broaden the understanding on the interrelation between founder and his or her venture by expanding the focus of founder identity theory to social as well as role identities.

Keywords: Sustainable entrepreneurship, entrepreneurial decision making, founder identity, social identity theory, identity theory, multi-founder ventures.

1. Introduction

> It is what I am, I am Mama Ocllo.
>
> – Martina Sturainer de Cueto, founder of Mama Ocllo, baby organic fashion venture.

2. Introduction

In recent years, research on entrepreneurship has increasingly dispelled any doubts that most of the entrepreneurs are only driven by the objective of profit-maximization when creating and shaping their ventures (Morton and Podolny, 2002; Cardon et al., 2009; Murnieks et al., 2014; York et al., 2016). There is a wide spectrum of underpinned factors that influence a founder's decisions about venture-creation and development – originating from his or her passion (Morton and Podolny, 2002; Cardon et al., 2009), pass through aspirations (Powell and Baker, 2011), and finishing by held identities (Fauchart and Gruber, 2011; Belz and Binder, 2017). Particularly, founders' identities derived from their values, beliefs, ambitions, professional background etc. are recognized as playing a very important role in the venture-creation and development process by impacting founders'

core strategic decisions in his or her day-to-day work. However, by investigating the impact of founders' identities on entrepreneurial decision making, most scholars have taken an unilateral perspective by drawing either from social identity theory (SIDT) – which focuses on "a person's knowledge that he or she belongs to a social category or group" (Stets and Burke, 2000) – or from identity theory (IDT) – which refers to "parts of a self composed of the meanings that persons attach to multiple roles they typically play in highly differentiated contemporary societies" (Stets and Burke, 2000). Consequently, taking the perspective of SIDT, researchers have demonstrated how founders' social identities imprint core strategic decisions in new venture creation process, regarding for instance: the venture's offerings, market segments, resources and capabilities, governance (Fauchart and Gruber, 2011; Belz and Binder, 2017) or venture goals and stakeholder approaches (York et al., 2016). On the other hand, from the IDT perspective scholars have looked into the effects of founders' role identities on selection of opportunities (Mathias and William, 2014), entrepreneurial self-efficacy (Murnieks et al., 2014) or successful transitions between work and founder role (Hoang and Gimeno, 2008).

Yet, remarkably few studies (Powell and Baker, 2011; Powell and Baker, 2017) have been designed to analyze the simultaneous impact of founders' social and role identities on entrepreneurial decision making in their ventures. With this in mind, in the present bachelor thesis I explore how the synergy of founders' social and role identities influences core strategic decisions in the new venture creation process. The main goal of this study is therefore to gain a better understanding of the relationship between the founder and his or her venture by extending the focus on founder identity, which is derived from both social and role identities.

This thesis is organized into four distinct sections. The first section provides background information on social identity theory and identity theory together with its separate application in the field of founder identity, which raises the issue of the interplay between social and role identities in organizational contexts. Then I introduce important contextual framework concerning entrepreneurial decision-making and sustainable entrepreneurship. In section 2, I outline the research methodology, which consists of an exploratory single-case study conducted with a German venture from the sustainable fashion industry. The empirical part details my findings on how founders' social and role identities imprint core strategic decisions in the new venture creation process. Finally, the discussion section is dedicated to providing arguments that first social and role identities have different yet complementary impact on the venture, and second that particular decisions shaping the venture may be interrelated to either social or role identities, or to a combination of both. By investigating the simultaneous impact of social and role identities on a venture created by more than one founder, I provide a more complete picture on founder identity but also raise questions on identity transitions over time and identity processes that take place between founders.

3. Literature Review

3.1. Social Identity Theory

Social identity theory (SIDT) was first introduced by Tajfel and Turner in the late 1970s when a growing interest in group processes and intergroup behavior arose among social psychologists (Tajfel and Turner, 1979; Tajfel, 1981; Tajfel and Turner (1986)). The basic idea behind SIDT is that all individuals classify themselves and others into social groups (e.g., German citizens, entrepreneurs, environmentalists) or categories (e.g., based on nationality, sex, political views, philosophical convictions) that influence their perceptions and actions.

Tajfel (1974) and Hogg and Abrams (1988) define social identity as a person's knowledge that he or she belongs to particular social groups and/or categories. In social psychology, a social group is understood as a set of individuals who recognize themselves as members of the same category and thus possess similar social identification (Tajfel and Turner (1986); Stets and Burke, 2000). Accordingly, all individuals possess a variety of social identities that describe them

in terms of the social groups or categories of which they are members, and they respond to the question of "Who am I?" within the social context (Hogg et al., 1995; Tajfel and Turner (1986)). For instance, a female German environmentalist may describe herself in terms of the categories of nationality, gender and political view as follows: "I am a German citizen; I am a woman; I am a nature-lover." However, belonging to a particular social group is not tantamount to interacting directly with members of the same group, because the basis for group membership lies in the uniformity of perception and behavior rather than in intraindividual relations within the group (Hogg et al., 1995; Stets and Burke, 2000). Consequently, an ecologist will engage in behavior to prevent climate change like other ecologists.

The process, through which people who belong to the same social group obtain perceptions, feelings and actions that are prototypical for that group is called depersonalization and is regarded as a central process to SIDT (Hogg, 2001). Depersonalized individuals act in accord with the group prototype rather than in accord with their own idiosyncrasies. Such group prototypes are considered cognitive representations that describe and prescribe attributes of the group by encompassing all of the beliefs, attitudes, and behaviors that are characteristic of average group members or of the ideal personification of that group (Hogg and Terry, 2000; Fiske and Taylor, 1991). Group prototypes maximize the uniformity of group members and, at the same time, maximize the distinctiveness between groups, which makes them strongly context relevant (Tajfel and Turner (1986); Hogg et al., 1995; Hogg and Terry, 2000). The context-dependency of prototypes means that the prototypical attributes of a social group may change with respect to the comparison group over time, as for example when environmentalists shift from defining themselves in opposition to capitalists and start to compare themselves with atheists. Therefore, Hogg and Abrams (1988) emphasize that a social group exists only in relation to another contrasting group: the so-called out-group. Through one of the most important processes in the formation of social identity – namely, the process of self-categorization – an individual categorizes himself or herself and others into in-groups and out-groups on the basis of attitudinal and behavioral matchings between a group and the individual. As a consequence of self-categorization, individuals accentuate their perceived similarity with the in-group and their separateness from the out-group (Hogg et al., 1995; Hogg and Terry, 2000). Individuals act so as to evaluate their in-group positively in comparison with the relevant out-group once their social identity is salient.

In SIDT, a salient identity is an identity that becomes activated in a particular situation (Oakes, 1987). Social identity theorists argue that the activation of a social identity is determined by two factors: namely, the accessibility and the fit of a given category. Accessibility refers to the connection between the individual's tasks and goals and the probability that a specific object or event occurs in the situation. A category fits the context when the individual perceives stronger intergroup differences than within-group differences (com-

parative fit) or when he or she perceives that a category explains a particular behavior. For instance, when an environmentalist speaks with a conservationist on climate change, it is likely that the category "political views" becomes salient, as it is accessible and best fits the context.

Originally, scholars claimed that the activation of a social identity is motivated to maintain and enhance one's self-esteem (Turner et al., 1987). Later research has introduced additional motivational factors such as self-knowledge (Abrams and Hogg (1990)) and uncertainty reduction (Hogg and Mullin, 1999).

3.2. Identity Theory

Identity theory (IDT) (McCall and Simmons, 1978; Stryker, 1968; Burke, 1980) is rooted in psychology and aims at understanding the role-related behavior of the self. Identity theorists base their principal assumptions on interactionism (Mead, 1934) and view the self as a composition of multiple role identities that emerge from the social interaction of the self and thus from the positions the self occupies within the society (Stryker, 1968, Stryker, 1980; Burke, 1980).

Consequently, each individual has as many group-based selves (role identities) as the social groups he or she interacts with (James, 1980). For instance, an individual may occupy various roles by being a father, a husband, and a caring boss at the same time. Conceptually, each individual categorizes himself or herself and others as role occupants, at the same time having knowledge of expectations about the role and its performance.

The set of meanings attached to oneself and others that results from a particular role performance is defined as a role identity (Stryker, 1968; Burke, 1980). McCall and Simmons (1978) conceptualize role identities as an imagination of the self in terms of how it perceives and how it would like to perceive itself in a particular position. Therefore, role identities provide the self with a frame of reference about concrete or potential behavior in a social situation (Burke and Reitzes, 1981; Hogg et al., 1995).

In IDT, the role is understood as a set of expectations associated with a particular position that one occupies within the society (Stryker and Burke, 2000). Roles can have a meaning only if they stand in relation to their counter-roles within a social network (Lindesmith and Strauss, 1956; Burke and Tully, 2001). For example, within the social network of marriage, the role identity "husband" exists only in relation to the role identity "wife." Within the university, the role identity "student" assumes the existence of the role identity "professor." As role identity prescribes behavior (Burke and Reitzes, 1981), the person occupying a role acts in accordance with his or her self-meanings and in accordance with the expectations attached to that role.

In contrast with SIDT, individuals do not perceive themselves as similar to others with whom they interact. Although the roles are interrelated, they are set apart, so members of social networks focus on their individual and complementary

roles to maintain the relationship between the self and the social structure (Hogg et al., 1995; Stets and Burke, 2000).

Initially, scholars agreed that role identities are context dependent. However, later research introduced the concept of a salience hierarchy, which is primarily relevant for predicting one's behavior in situations in which more than one role identity may be invoked (Stryker, 1968). Role identities are organized within the self in a salience hierarchy according to their probability of being activated in a particular situation (McCall and Simmons, 1978; Stryker, 1968). The likelihood that a person will activate one role identity rather than others and will simultaneously act in that role's congruent behavior is conceptualized as salience. Ergo, role identities at the top of the salience hierarchy have a higher probability of forming the basis for action in accordance with the expectations and meanings that are tied to that role. Two persons who hold the same role identities can make different behavioral choices in the same social situation because their role identities are organized differently in their salience hierarchies. For example, one may go to the university to the scheduled lecture while the other stays at home sleeping, although both have the "student" role identity.

Identity theorists argue that the salience of a role identity depends on the commitment of the self to that role. Stryker and Statham (1985) define the commitment as, "the degree to which the individual's relationships to particular others are dependent on being a given kind of person." Contemporary research on IDT (Stets and Burke, 2005) suggests that, apart from salient role identities that motivate role-congruent behavior, there are also internal control systems that explain the activation of a particular identity. According to the researchers, individuals look for situations in which they can activate their role identities to receive information from others about a well-performed role. The process in which the environment confirms a successful role performance is called identity verification, and it leads to positive emotions about the self. Thus, individuals are also motivated to behave according to a particular role identity by proving that identity among others to achieve higher levels of self-esteem. Other scholars introduce additional motivators for an identity activation in social situations: for example, self-consistency and self-regulation (Burke and Stets, 1999). Despite different motivators for identity activation, both the earlier and the latest research agree that role identities are drivers and influencers of one's action when they are invoked in social contexts.

3.3. Social Identity Theory and Identity Theory

As illustrated by Table 1, significant differences exist between SIDT and IDT. These differences originate from the disciplinary foundation, pass through heterogeneous bases of each identity, and finish by explaining the intergroup behavior. Despite the discrepancies in sources of meanings between both identity theories, researchers (Hogg et al., 1995; Stets and Burke, 2000) have slowly started to acknowledge the need of including role and social identities in the investigation of the self. In their work, Stets and Burke (2000)

Table 1: A Comparison of Identity Theory with Social Identity Theory

	Identity Theory	Social Identity Theory
Underpinned disciplinary context	Sociology	Social Psychology
Basis for identification	Role	Social group
Identity definition	A person's categorization as a role occupant together with a knowledge of expectations towards the role and the performance.	A person's knowledge that he or she belongs to particular social groups or categories.
The process of identity formation	Identification – in the process of identification an individual name himself or herself and others in terms of position/role occupants.	Self-categorization – the process emphasizes perceived similarities with an in-group and perceived differences with a relevant out-group.
Meaning of having an identity	Having a role identity means acting in accordance with expectations associated to that role and interacting with role partners (the occupants of counter-roles).	Having a social identity means belonging to a social group and behaving like other group members.
Group understanding	A social group is conceptualized as a collective of persons, where each of the group members occupies a different/unique role. The roles of each group member are interrelated with other group members' roles (counter-roles). Each individual possesses his or her own viewpoint.	A social group is defined as a set of individuals with homogenous perceptions, attitudes and behavior, who identify with each other and look at each other in a similar way in contrast to the out-group members.
Basis for group interrelation	Individuality and interrelatedness with other counter-roles, the so-called interconnected uniqueness, through which an individual by taking a role acts in accordance with his or her self-meanings and expectations that accompany the role. Individuals do not view themselves as similar to the others with whom they interact because each role is related but at the same time set apart.	In group-homogeneity or the so-called uniformity among social group members, which is expressed cognitively, attitudinally and behaviorally. An individual feels attraction to his or her in-group and behaves in accordance to it.
Involvement of third parties	Individual occupying a particular role as well as relevant others occupying counter-roles are directly involved in the individual's role performance.	Only the individual's perceptions and behavior are directly involved.
Intergroup relations	Focus on different but interrelated roles within a group.	Focus on similarities within the in-group and differences with relevant out-groups.
Salience definition	Salience is the probability that a person activates a particular role identity in a given situation and acts according to the expectations tied to that role.	Salience indicates the activation of an identity in a situation. A salient social identity is the one with the highest influence on a person's membership based on perception and behavior.
Activation definition	An activated role identity is the identity played out in a given situation.	Social psychologists merge the terms salience and activation.

(Continued)

Table 1—continued

Activation process	The activation of a role identity depends on the commitment (the degree to which the individual's relationships to particular others are dependent on being a given kind of person).
	
	   a. quantitative – the number of persons to whom one is tied through his or her role identity
	   b. qualitative – the depth and strength of ties with relevant others
		A social identity becomes activated (salient) as a result of the accessibility (the person's current goals and tasks as well as the probability that an event occurs in the situation) and fit (specifications and perceptions about the situation).

Salience vs. activation	Salience ≠ Activation	Salience = Activation
Hierarchy of identities	The salience hierarchy within the self-concept consists of organized role identities in relevance to their probability of being activated and thus forming a base for action.	The social identities are organized in a hierarchy of inclusiveness, which consists of three levels. For example, a woman may see herself in the first level as a human, in the second as a German and in the third as a member of the local community
Emphasis of the identity activation	Role performance.	Membership.
The core cognitive process	Self-verification – through the process of self-verification an individual perceives himself or herself as an embodiment of his or her role and acts in consistency with that role's expectations.	Depersonalization – through the process of depersonalization the person stops perceiving himself or herself as idiosyncratic but starts viewing himself or herself as a personification of the in-group prototype and acts according to the norms and values of that group.
The motivational foundation	Self-esteem, self-efficacy, self-consistency, self-regulation.	Self-esteem, self-evaluation, self-knowledge, self-consistency, self-efficacy, uncertainty-reduction, self-regulation.

conclude that, in general, the two theories have many similarities that can provide a basis for establishing a general theory of the self.

Particularly, according to both theories, each individual perceives himself or herself in terms of meanings conveyed by components – either roles or social groups – of a structured society. Consequently, the self exists within the society, is affected by the society in terms of norms tied to a social group or a role, and at the same time influences the society by behaving in accordance to its norms. Additionally, both kinds of identity theorists recognize a strong link between identity and behavior and mention a variety of internal motives that lead each individual to act in accordance with a social group or a role.

However, most importantly, each individual is always an occupant of multiple roles, and at the same time a member of diverse social groups, which makes role and social identities inseparable from the self (Stets and Burke, 2000). For instance, a person may be a vegan and identify with other vegans, while at the same time this person may occupy the role of being a founder of his or her venture. The social identity of being a vegan and the role identity of being a founder are not only inseparable from the person's self-perception but also may together be constantly relevant to the person's behavior. Taking into consideration that preliminary research has been limited to either social identities or role identities in the examination of identity-congruent behavior, this bachelor thesis aims to integrate both, roles and group memberships, to investigate their simultaneous influence on the behavioral choices of a founder in the new venture-creation process.

3.4. Founder Identity

The term founder identity has been used by Powell and Baker (2011) to refer to, "the set of identities that is chronically salient to a founder in her or his day-to-day work." Throughout this paper, I expand the term founder identity defined by Powell and Baker (2011) by emphasizing that the

identity of a founder is derived from both social and role identities.

In recent years, entrepreneurship and management researchers have focused their attention on the nexus between founder identity and entrepreneurship. A small but increasing amount of evidence exists to suggest that founder identity is essential for entrepreneurship by influencing core strategic decisions in new venture creation and running processes. However, prior studies that oscillate between founder identity and organizational contexts are based either on IDT, and therefore on role-congruent behavior, or on SIDT, which links an individual's action to his or her memberships in social groups.

3.4.1. Social Identity Theory Lens

By adopting a social identity lens, Fauchart and Gruber (2011) seek to understand both the central types of social identities held by the founders and the impact of founders' social identities on strategic decisions, such as market segments served, customer needs addressed, and resources and capabilities employed. Drawing on their investigation of Western European sports-equipment firms, the authors categorize founders into three primary social identity types, namely Darwinians, communitarians, and missionaries, each of which varies in social motivation, founder's self-evaluation and founder's frame of reference. Fauchart and Gruber (2011) reveal that different types of founders' social identities have different yet considerable impacts on the core strategic decisions while founding, but above all, that founders engage in behaviors that are congruent with their in-group memberships.

York et al. (2016) extend beyond the study of Fauchart and Gruber (2011) by exploring why individuals commit themselves to environmental entrepreneurship. Given that environmental entrepreneurship is a hybrid form of organization that combines two conflicting commercial and ecological logics, the authors attempt to identify how founders of such organizations define venture goals and approach stakeholders. York et al. (2016) claim that each entrepreneur is a holder of a hybrid identity that combines both commercial and ecological identities. Depending on the degree of coupling between the two mentioned identities, the authors distinguish three hybrid identity types: namely, the commercial dominant identity, ecological dominant identity or blended identity, each of which shapes the founder's decisions on venture goals and stakeholder incentives differently.

Drawing on SIDT, Belz and Binder (2017) study the relatively underexamined process of the founder's negational categorization, through which the identity is determined by the out-groups from which he or she differentiates. While numerous studies have focused on the affirmational categorization that defines one's identity in terms of in- group memberships, Belz and Binder (2017) contribute to the study of founder identity by demonstrating that out-groups – and thus the negational categorization – also influence new firm creation, especially with respect to the market segments served, the customer needs addressed, and the governance applied.

Furthermore, they find that affirmational and negational categorizations are only the two extreme forms of social categorizations, between which forms of social categorization exist that are mitigated and relevant for organizational contexts: namely, in-group differentiation, out-group mitigation, apathetic in-group confirmation, and apathetic out-group separation.

3.4.2. Identity Theory Lens

In contrast to the previously mentioned researchers, Hoang and Gimeno (2008) take the perspective of IDT and focus their study on the different structures of founder's role identities during venture founding. The authors view new venture creation as a role transition between a work role and a founder role. Taking into consideration that role transition is accompanied by various challenges, Hoang and Gimeno (2008) explore two dimensions of role identity: namely, identity centrality, expressed by the depth of attachment to the founder role; and identity complexity, revealing the diversity in a person's definition of the founder role, which significantly influences the behavioral persistence when receiving a negative feedback in a new venture-creation process. Additionally, the authors could link founder-role identities to long-term outcomes such as successful venture creation.

Mathias and William (2014) largely confirm other scholars' claims that founder identities have an impact on the venture-creation process by adopting the IDT lens. Their argument builds on research they conducted into how entrepreneurs' role identities influence the evaluation and selection of opportunities while founding a venture. Mathias and William (2014) extend past work that establishes a link between founder-role identities and venturing by taking into consideration the fact that each entrepreneur possesses different constellations of so-called within-work role identities, which may have different impacts on entrepreneurial decision making. Additionally, the scholars found that there is an association between entrepreneurs' role identities and their risk consideration. Their findings demonstrate that each entrepreneur in their sample holds three central, within-work role identities: 1) the entrepreneur role, 2) the manager role, and 3) the investor role. Each of these identities considers opportunities in line with, 1) an entrepreneurial focus, 2) a marketing focus, and 3) a long-term focus. Because most previous research has addressed interindividual comparisons (e.g., between entrepreneurs and managers), Mathias and William (2014) emphasize their investigation of interconnection between the plurality of role identities that each entrepreneur holds within his or her occupation and the entrepreneurial decision making.

3.4.3. Integration of SIDT and IDT in Organizational Contexts

Taking into consideration the latest call for the integration of IDT and SIDT, Powell and Baker (2011) made the first attempt to bridge both theories and to investigate why entrepreneurs perceive and act differently with respect to the same adversity. Their findings reveal that all founders hold

different structures of role and social identities that differently affect how they evaluate and act in response to the adversity involved in their respective ventures. Moreover, according to Powell and Baker (2011), identification with social identities is a source of aspiration for new role identities that lead founders to transform their firms. In other words, new venture creation and running processes allow founders to engage in desired roles (role identities) that express their social identities.

Table 2 summarizes the research on founder identity that has been conducted in the past ten years and thereby makes it evident that most scholars have applied either an identity theory or a social identity theory when exploring the relationship between a founder's identity and his or her venture. Yet, despite the call for the integration of the two theories, preliminary research, apart from that conducted by Powell and Baker (2011), has paid little attention to analyzing how a founder's role and social identities operate simultaneously in organizational contexts. To address this gap, in this bachelor thesis I seek to understand how the synergy of the founder's role and social identities imprint core strategic decisions in the new venture-creation process.

3.5. Entrepreneurial Decision Making

Shepherd et al. (2015) contribute to the broadly discussed research on decision-making process in psychology and marketing fields by extensively reviewing the process of entrepreneurial decision making, which according to the scholars has not received much attention in past reviews. Their review is congruent with previous work and summarizes entrepreneurial decisions regarding opportunity assessment, venture creation, opportunity exploitation, and exit, each of which is determined by multiple aspects of the economic and non-economic nature. However, while existing research has focused on the strictly economic perspective of entrepreneurship, Shepherd et al. (2015) draw attention to new types of entrepreneurs – especially to social and environmental entrepreneurs – who have to balance the economic, social and/or ecological elements when making strategic decisions. This study attempts to shed light on one of the core, non-economic drivers of entrepreneurial decision making in sustainable organizations: namely, the memberships with social groups (social identities) and the role occupancies (role identities) of the founder.

Moreover, scholars agree that the link between a founder and his or her venture is notably strong in the initial phase of founding, in which the organization is yet to be defined (Fauchart and Gruber, 2011). Therefore, the present research focuses on founders, on their social and role identities as motivational forces, and on the strategic decisions they make during venture founding.

According to Abell (1980), decisions that strategically define a venture and thus are permanent and provide the overall guidance for an enterprise tend to oscillate around which market to enter (who will be served by the venture), which products to offer (what customer needs are to be met), and which resources and capabilities to employ (how the venture

is going to meet these needs). Consequently, by exploring the impact of a founder's social and role identities on the new venture-creation process, this study will focus on core strategic decisions made in the initial phase of founding: namely, the market segments addressed, the venture's offerings, and the resources and capabilities employed by the venture.

3.5.1. Founder Identity and Entrepreneurial Decision Making

While the literature on founder identity generally acknowledges a strong association between a founder's self-concepts and entrepreneurial decision making, there are differences in the examination of decisions that are affected by the founder and particularly his or her identities.

For instance, Fauchart and Gruber (2011) demonstrate the impact of the founder's social identities on three core strategic decisions that are made initially in a new venture-creation process – namely, decisions about the market segments served, the customer needs addressed, and the resources and capabilities deployed.

Similarly, Belz and Binder (2017) focus their investigation on early stages of the entrepreneurial process and link a founder's social identity as expressed by his or her negational categorization to core strategic decisions about opportunity exploitation, such as which products to offer, which market segments to serve, and which governance to employ. However, by exploring founder identities in the context of entrepreneurial decision making, York et al. (2016) go even further and demonstrate that identities of a founder that are derived from commercial and ecological logics are directly associated with decisions about venture goals and approaches to stakeholder incentivizing.

Complementary to the previously mentioned studies, there is an extensive body of entrepreneurship research that identifies a linkage between founder identity and multiple decisions made in the entrepreneurial process, starting from the initial stages of opportunity selection and exploitation (Mathias and William, 2014), and finishing with strategic responses to adversity (Powell and Baker, 2011). In a similar vein, this bachelor thesis will deepen existing knowledge about the interface between founder identity and entrepreneurial decision making by investigating the influence of founder's identities derived from group memberships and roles on the core strategic decisions made in the initial phase of the entrepreneurial process.

3.5.2. Sustainable Entrepreneurship

A number of studies (Gibbs, 2009; Cohen and Winn, 2007; Cohen and Winn, 2007; Belz, 2013) have been designed to discuss how sustainable entrepreneurship can contribute to the positive transformation of societies and economies towards ecological, intergenerational and intragenerational consciousness expressed by more sustainable consumption and production practices. For instance, Belz (2013) seeks to problematize the role of sustainable innovation in overcoming – or rather, realistically mitigating – the

Table 2: A Review of Literature Focusing on Founder Identity in Entrepreneurship

Autors	Research Question / Research aim	Identity Theory	Identity Construct	Object of Affect	Main empirical findings
Fauchart and Gruber (2011)	1. What are the primary types of founder identities from a social identity perspective? 2. To what extent does a founder's social identity influence key dimensions of new firm creation?	Social Identity Theory	Social categorizations of a founder	Initial strategic decisions in new firm creation: market segments served, customer needs addressed, recourses and capabilities deployed	Founders behave in accordance to their social identity, which can be classified as Darwinian, communitarian or missionary identity. Depending on the social identity possessed by a founder, he or she has a different impact on new firm creation.
York et al. (2016)	Why and how do individuals engage in environmental entrepreneurship?	Social Identity Theory	Founder's hybrid identities derived from commercial and ecological logics	Venture goals and stakeholder incentive approach	Founder's hybrid identity coupling commercial and ecological logics shape the organization's goals, and the approach in which stakeholders are recruited.
Belz and Binder (2017)	How does negational categorization affect founder identity?	Social Identity Theory	Negational categorization of firm founder	Core strategic decisions in new venture creation: products, markets & governance	Negational categorization influences strategic decision making in the opportunity exploitation phase of the entrepreneurship process.
Hoang and Gimeno (2008)	How founder role identity affects entrepreneurial transitions and persistence in founding?	Identity Theory	Role identity centrality and complexity of a founder	Success of the transition from a work role to a founder role and successful venture creation	Identity centrality and complexity lower the impact of role novelty on the role transition, and provide different role identities of a founder, which positively influence his or her persistence while receiving negative feedback in new venture creation process.

Table 2—continued

Mathias and William (2014)	How do differences in role identities influence how entrepreneurs think about – and select – opportunities?	Identity Theory	Multiple within-work identities of a founder	Founder's consideration and selection of opportunities	Each founder possesses a variety of nested occupational role identities, e.g. founder, manager or investor. These multiple but interrelated role identities strongly influence the founder's thinking and behaving regarding the consideration and selection of opportunities.
Cardon et al. (2009)	What is the nature and experience of entrepreneurial passion?	Identity Theory	Founder's passion as a result of engaging in activities with identity meaning and salience	Founder's entrepreneurial behavior	Fonder's role identities form a foundation for the founder's passion for different entrepreneurial activities, which in turn has influence on creative problem solving, opportunity recognition, venture creation and venture growth.
Murnieks et al. (2014)	Exploration of the link between founder's identities and his or her passion.	Identity Theory	Founder's passion, which is positively related to his or her entrepreneurial identity centrality	Entrepreneurial self-efficacy and entrepreneurial behavior	Entrepreneurial identities drive entrepreneurial passion. Both role identities and passion have a strong impact on entrepreneurial behavior and self-efficacy.
Powell and Baker (2011)	How and why do resource-constrained founder-run firms respond to adversity in widely varied ways?	Social Identity Theory & Identity Theory	Structures of founder's salient role and social identities	Firm's strategic responses to adversity	Each founder holds different sets of social and role identities, which shape the founder's perception and action undertaken on adversity. Additionally, founder's within-work role identities are regarded as expression of his or her social identities.

challenges of the twenty-first century identified as overpopulation, ecosystem degradation, and climate change, each of which is regarded as a potential threat to humanity and Mother Earth. According to the author, sustainable innovations – which are defined by Belz and Peattie (2012) as "novel products and services that satisfy customer needs and, significantly, also continuously improve the socio-ecological performance along the whole life cycle in comparison to conventional or competing offerings" – have the potential to transform obsolete and harmful consumption and production patterns into environmentally and socially responsible practices. As noted by Gibbs (2009), organic fashion, wind and solar energy, electric cars and green buildings are only few examples of sustainable innovations that are revolutioning conventional capitalist economies. However, these are not necessarily established companies that come up with radical sustainable innovations due to deeply rooted organizational structures and strategies; instead, often sustainable entrepreneurs, and their start-ups are more likely to develop novel, sustainable customer solutions. Binder and Belz (2014) define sustainable entrepreneurship on the basis of a variety of existing definitions in the relatively nascent stream of research as, "the scholarly examination of how opportunities to bring into existance future goods and services are recognized, developed, and exploited by whom, and with what economic, social and ecological gains."

Current research on sustainable entrepreneurship seems to indicate that entrepreneurial actions pursuing the triple bottom line of ecological, social, and economic goals are often a consequence of the founder's identity, such as his or her love for the environment or for local community. For instance, Jakob Assmann founded Polarstern, a German venture providing 100% green power, because of his love for the mountains and nature, which is gradually suffering the consequences of human irresponsibility. Anna Yona, the founder of Wildling Shoes, decided to market organic and fair barefoot shoes for kids because of her great love for her children and their active and playful lifestyle. Choi and Gray (2008) investigate key decisions and actions taken trough the stages of the entrepreneurial processes in twenty-one sustainable ventures from diverse industries. By examining the second stage of the entrepreneurial process – namely, the concept-development stage – the authors have come to a realization that founders of sustainable enterprises are not driven by pure profit-maximization but by personal motivational factors that come to play when they define the business concept. York et al. (2016) supports this claim by suggesting that individuals who follow goals derived from ecological and commercial logics engage in entrepreneurship because of their compassion rather than for personal wealth accumulation.

Consequently, sustainable entrepreneurs differ from conventional entrepreneurs in terms of possessing a set of multiple motivations that are derived from their role and social identities and that are considered to be a major force in entrepreneurial decision making. Furthermore, entrepreneurs who pursue only economic goals while running a business tend to possess the singular role identity of a keeper of the bottom line and are driven by making money and sustaining the business financially (Powell and Baker, 2011). Because such entrepreneurs are not tied to any another role or social identity relevant for strategic decision making in the founding process, they will not provide interesting insights into the field of founder identity. Hence, to investigate the impact of the social and role identities of a founder on the new venture-creation process, I focus entirely on sustainable entrepreneurship, which, apart from providing the context for this investigation, promises to yield more interesting and more complete data.

4. Research Methodology

4.1. Research Design

In this thesis I seek to understand how social and role identities of a founder influence strategic decisions in new venture-creation process. Since entrepreneurial decision making at the interface between SIDT and IDT is still in a nascent phase I follow the recommendation of Yin (2014) and conduct an explorative, qualitative research to answer the research question. Particularly a case study research method has been applied as it is most suitable when a how question is being asked to gain a deep understanding about a contemporary phenomenon in real-world context, in which the researcher has no control over the relevant behaviors (Yin, 2014). Given the limited time available to conduct this study – approximately three months – and given that its purpose is to obtain a profound, real-world perspective on the founder's social and role identities in organizational contexts, where the boundaries between identities and organization are not clearly evident, I employ a single-case study design, as recommended by Yin (2014).

4.2. Study Setting and Sampling

Taking into consideration that the present thesis aims to explore founders' social and role identities during venture formation, my sampling approach bears a close resemblance to prior research on founder identity and the venture-creation process by using the following criteria to select the relevant case.

1. The venture is independently held and is a maximum of eight years old (Fauchart and Gruber, 2011; Belz and Binder, 2017) because this period of time enhances the probability of a stronger link between the founder and his or her firm (Fauchart and Gruber, 2011). Under other conditions, it might be challenging to capture how founders' identities have influenced strategic decisions as social and role identities are dynamic and may change over time (Hogg et al., 1995).

2. The venture meets the definition of a sustainable enterprise. Binder and Belz (2015) describe sustainable enterprises as for-profit organizations, which pursue the triple bottom line of objectives. Thus, apart from creating an economic value, they also aim to create a social

and an ecological value. For the purpose of this study, sustainable entrepreneurs are more promising than conventional entrepreneurs, as their entrepreneurial actions are driven by a wide spectrum of role and social identities. On the other hand, researchers (Fauchart and Gruber, 2011; Powell and Baker, 2011) agree that a conventional entrepreneur engage in the single-role identity of a profit-maximizer and views his or her venture in business terms rather than as a deeper reflection of the self.

3. The founder remains an active member of the management team and was decisive for choosing which products to offer, which market segments to serve, and which resources and capabilities to employ.

Accordingly, I gained access to the German venture Mama Ocllo from the sustainable fashion industry, which meets the above-mentioned criteria and is the single case of the present analysis.

Sustainable fashion enterprises emerge relatively quickly in western markets. Their founders are known for introducing green product innovations, including new organic fibres, extraordinary and multifunctional designs, and the highest quality properties, but also, for the upcycling and reusing techniques of existing textiles. Furthermore, while the mainstream textile industry proves to have fundamental flaws – especially its disrespectful approach to natural resources and human labor – sustainable fashion enterprises come up with innovative social and environmental projects to move society towards sustainability. The diversity of green innovations regarding products, resources and capabilities, but also addressing of outstanding market segments, is often a result of the founders' themselves, their values, beliefs, attitudes and behaviors, which are in turn derived from their identities.

My case, Mama Ocllo, dates its market entry back to October 2012. The company was created by three friends, who, despite their different origins (Germany, Peru, and France), share the same love for fashion, Peruvian culture, and Mother Earth. The main offering of Mama Ocllo is vegan, healthy, and fairly produced baby clothing with unique designs. Each piece is manufactured in Peru out of Peruvian's finest natural materials. The venture's philosophy is twofold: first, the cotton is cultivated 100% organically and processed with a complete renouncement of chemical additives and machinery improvements to ensure skin tolerability and well-being for babies. Second, Mama Ocllo aims to create a common value within the entire production chain and contributes at the same time to socio-economic development in marginal areas of Peru. Although Mama Ocllo represents the single case of this thesis, I relied on an embedded case study that involves the co-founders of the venture as the units of analysis (Yin, 2014). The embedded units were selected through a theoretical sampling technique, which allows for the selection of units that are appropriate for the elaboration of the emergent theory rather than for theory testing (Eisenhardt and Graebner, 2007). Mama Ocllo was founded by three founders – Martina, Gustavo and Telma – however, only two

of them, Martina and Gustavo, are taken into account as the main units of analysis because the core strategic decisions during founding were tied directly to them. Moreover, the venture has been continuously owned and managed by Martina and Gustavo, which allows me to gain access to first-hand stories about decisions made in the initial phase of venture creation. Telma, on the other hand, focuses entirely on product design and is therefore not involved in strategic decision making. Figure 1 presents an overview of the embedded single-case design for this thesis.

4.3. Data Collection

The data for this study comes from multiple primary and secondary sources in pursuance of data triangulation and construct validity (Yin, 2014). The following sources were employed in the present research:

Graphic elicitation. Graphic elicitation is an art-based research technique that includes visual, verbal or written stimuli to encourage respondents to share their ideas, especially when the topics of interviews are hard to express by words alone (Barton, 2015; Bagnoli, 2009). The employment of art-based techniques in the phase of data collection is particularly suitable for controversial topics such as politics, sexuality, and religion. but also for situations, in which the respondents are asked about higher level experiences, intrinsic motives or embedded identities that cannot be easily elaborated verbally despite the respondents' knowledge and awareness of their existence (Barton, 2015). Furthermore, opening an interview with a visual task may break down the ice and set up a comfortable conversation for the participants (Bagnoli, 2009). Following the recommendations of Bagnoli (2009) and Barton (2015), I employed a visual task as an introduction to further interviews with the founders of Mama Ocllo to identify each founder's social group memberships and to provide a basis for the conversation about individuals or groups of people who may be regarded as role models or inspirations for the interviewees. The illustrative task in this study follows the relational map developed by Josselson (1996) and consists of asking the interviewees to draw themselves in the middle of a paper and to include all the people or groups of people that come into their mind because they are important to them and their lives. The illustration of relationships between the interviewee and relevant others was based on the solar system, where the interviewee, represented by the sun, is surrounded by significant persons (or a group of persons) drawn as planets. The distance between the sun and a particular planet reflects how important this person or group is for the participant. Furthermore, the participants were asked to draw planets with dotted lines if there are some social groups or individuals who are not physically in their lives but remain important to their self- perception and thus occupy an abstract yet meaningful position. The relational maps drawn by the founders of Mama Ocllo provided a basis for another, 15-minute long conversation, which helped me to understand the drawings and allowed me to ask additional questions about the existence of role models included in the map and other individuals with whom the interviewees
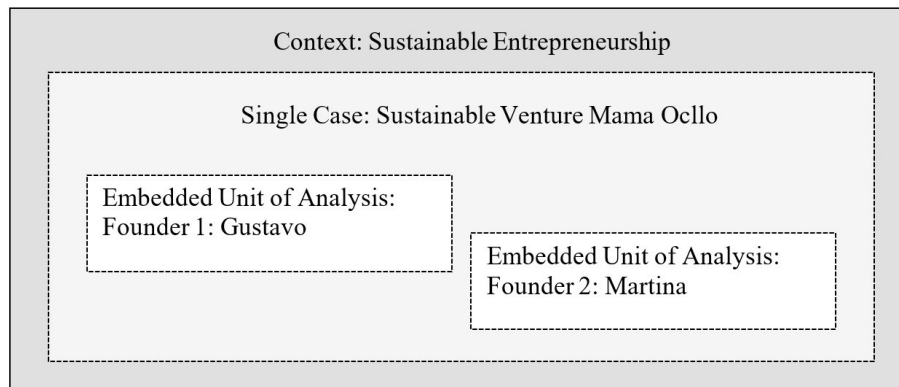
Figure 1: Embedded Single-Case Design – Overview

identify. At the end, I also asked participants to include vocational groups and passions in their maps, thereby to complement the overall picture of the founders' belonging to social groups together with the roles they play in their lives. Both participants drew their relational maps and described them in their mother languages: German and Spanish. Each relational map was translated into English and is included in the Appendix A.

Founder interviews. In-depth interviews provide the most important case-study evidence (Yin, 2014). Therefore, after initial identification of social and role identities of the founders through graphic elicitation, I conducted two face-to-face, semi- structured interviews, one with each founder. The interviews were conducted in the home offices of the participants. Each interview took approximately 60 minutes. Because one founder speaks German and another Spanish, the interviews were conducted in these languages to ensure liberty of expression. The interviews were transcribed verbatim. Full transcriptions in German are included in the Appendix B. The semi-structured interviews consisted of mostly open questions categorized into five thematic sections: (1) the founder, including his or her family and professional background together with the current roles, he or she occupies in his or her life and the roles he or she strives for; (2) the venture, particularly the link between the founder and his or her venture and the current roles he or she occupies within the company; (3) the venture's offerings and their interrelatedness with the founder's identities; (4) market segments addressed by the company and their interrelatedness with the founder's identities; and (5) resources and capabilites employed to produce the venture's offerings together with their interrelatedness with the founder's identities. Open questions from these five sections were designed to gain an understanding of the relationship between the founder's identities and each core strategic decision, but also to elicit how the synergy of both identity types influences the new venture-creation process. A guide underpinning the interview questions served as an indicator of themes that had to be addressed. It is enclosed in the Appendix C.

Secondary data sources (website, blog, Social Media presence, press releases). Apart from primary data sources, this thesis also uses data gathered from secondary, internal and external sources, including Mama Ocllo's website, one founder's blog, the social media presence of the company, and press releases to intensify the knowledge about how the founder's social and role identities are reflected in strategic decision making during firm creation. Thus, relying on multiple data sources allowed me to triangulate the collected interview data and to gain access to information about how founders talk about themselves, how they present their products, to whom they address the offerings, and how they introduce their suppliers and production processes.

4.4. Coding and Data Analysis

Data documentation and the coding process were conducted with the help of MAXQDA standard software, which allows for qualitative data analysis. The data was analyzed using the approach of Miles and Huberman (2014). First, the elicited raw data is processed and reduced by organizing field notes into formal write-ups and transcribing recorded interviews into comprehensible text that is free of errors. Secondly, individual codings are attached to the processed data through a careful line-by-line reading of interviews, which are later transformed into second-cycle codes: the so-called patterns. Following the approach of Miles and Huberman (2014), the coding and analysis process was divided into three main stages:

Stage one. First, using mainly descriptive and in vivo coding, I attached codes to all responses associated with social group memberships and with the actual or intended roles of each unit of analysis. When identifying founders' identities, I adopted first a SIDT and then an IDT lens and focused especially on each founder's self-descriptions, visions, motives, and beliefs. The first cycle of coding therefore resulted in a list of all the social and role identities held by each founder. In the second cycle of coding of the founder's identities, I developed pattern codes that summarized all the identities into meaningful and parsimonious categories and filtered out the identities that were not salient in the new venture-creation process. Table 3 provides a list of all the social and role identities that are relevant to new venture creation together with

Table 3: Representative Data for Founders' Social and Role Identities

Founder	Identity Type: Identity Description	Identity Refinement	Data Illustration
Martina	Social: Sustainable Living Devotee	-	"I don't want my target market to consist of old nobility who pays attention to only expensive goods. I don't want it because this is a society, I don't know, they are like the typical fur bearing people. I don't belong there. "
Martina	Social: Sustainable Living Devotee	LOHAS	"Of course I do [identify myself with LOHAS]. But I am still not fully consistent, as it [sustainable living] is too expensive for me. But at least from the point of view of my attitude, I do identify myself with LOHAS. If I had more money, I would introduce it [sustainable living] in more aspects of my life. Now I purchase products from small labels, I buy special products before buying something cheap, what doesn't have a good quality and what makes me feel bad. "
Martina	Social: Missionary	-	"But I would like to be perceived as a person, who makes a positive contribution to society. It is how I would like to be perceived. "
Martina	Social: Creative Philologist	-	"What fascinates me is (...) the creative work, it is lot's of fun. Everything what is creative, where I can visualize everything what we are doing. Writing, drawing, everything what stays in this area. "
Martina	Social: : Organic Industry Specialist	-	"Then I worked for a big exporter of organic grain products. I worked there in a purchase department, but I also got to know many other departments. (...) Then I was in Nicaragua on a company trip, where I also had contact with organic farmers."
Martina	Role: Voice of Mama Ocllo	Communicator, Networker, Marketer	"But I do everything related to the shop, also every product description. I am also responsible for pictures and I coordinate everything, what constitutes Mama Ocllo. (...) Also the whole marketing department and networking, quite a lot.
Martina	Role: Mother	-	"So I applied for jobs and I would have gotten jobs in purchase departments. I also received some job offers but I canceled all of them because the family was my biggest wish, not the profession itself. "
Martina	Role: Sustainable entrepreneur	Local Economy Supporter	"It would be contradictory, if I say, I want to do good for Peru but at the same time I work with conventional cotton. It wouldn't create any added value for Peru. "
Martina	Role: Sustainable entrepreneur	Peru Community Patron	"At the beginning I had too many ideas. Already initially I wanted to integrate a social project in Peru but the company has not been established yet. It was more a challenge of having to slowing down myself."
Martina	Role: Baby Organic Fashion Specialist	-	- "But I wish myself that people perceive us as specialists in high-quality baby fashion from pima cotton, everything sustainable, organic and fair. "

(Continued)

Table 3—continued

Gustavo	Social: Entrepreneur	Innovator	"I have mentioned innovators as the social group I partially identify with. They have the ability of taking good decisions, especially of taking fast decisions. They are also very responsible and are not afraid of new challenges. They do not fear risks and are very self-confident when they promote their products. These two aspects, risk appetite and introduction of novel goods on the market, play the most important role for me. "
Gustavo	Social: Entrepreneur	Creator	"Already some years before founding Mama Ocllo, I had many ideas in my head. I was searching for the right opportunity. I also had a phase that I was thinking a lot about my strengths and abilities. But in that moment, I didn't have anything clear. I wasn't any expert in biotechnology or the like. I have only known since always that I want to produce something. "
Gustavo	Social: Industrial Engineer	-	"I studied industrial engineering. I also behave similarly to other engineers because you should have particular qualities if you want to study a similar subject. These are especially a good planning ability, structured thinking and that you are able to sacrifice some of your life spheres. "
Gustavo	Social: Sustainable Living Devotee	LOHAS	"Yes, in some of aspects in my life, I am [a LOHAS]. But I think I am still in the transition process. We do not buy or consume 100% ecologically. "
Gustavo	Role: Peruvian Community Patron	-	"I can also say that I am a supporter of marginal regions in Peru. Generally, we support with Mama Ocllo the Peruvian economy, although the impact is very small yet. "
Gustavo	Role: Keeper of the Bottom Line	Specialist for Finance Accounting and Controlling Solutions	"Besides my involvement by Mama Ocllo, I also work in another company in the controlling department. We have small teams there and they are connected with every area in the company. Through the communication with different areas we develop projects for cost savings and better efficiency, and I am responsible for controlling these processes. "
Gustavo	Role: Keeper of the Bottom Line	Supply Chain Manager	"I am also responsible for the production, logistics and the import. "
Gustavo	Role: Sustainable Textile Manufacturer	-	"Also jetzt will ich auch einen Beitrag an der Entwicklung der nachhaltigen Textilbranche in Peru haben. Jetzt gibt es also diese Motivation, etwas gerade in diesem Sektor zu bewegen."

the representative data for each identity, which together provide the outcome of the first stage of coding.

Stage two. Stage two consisted of analyzing the core strategic decisions made during venture founding. I focused on the venture's offerings, the market segments addressed, and describe their venture's products, target market, resources and relationships with suppliers, including the intrinsic motives that drove or constrained each decision.

Stage three. In the last but most important stage of data coding, I looked for patterns regarding how social and role identities are mirrored in products, market segments, and applied resources and capabilities. During the search for the correspondence between both identity types and core strategic decisions in new venture creation, I made a systematic comparison of the social and role identities held by each unit of analysis to understand how their interplay impacted the entrepreneurial decisions the founder made during venture formation. Particularly, I focused on two aspects: the frequency of activation, and the relevance of each identity to every firm-related decision. The last step of the analysis pro-

cess consisted in including both units of analysis and comparing their overall influence on the venture's offerings, market segments, and resources and capabilities.

5. Results

The results section of this paper follows the structure of data analysis procedure and starts with a description of the founder's social and role identities that are relevant to organizational contexts. Furthermore, I illustrate the influence of the founder's identities derived from his or her membership to social groups and occupied roles on strategic decisions in new venture creation considering the venture's offerings, market segmentes served, and resources and capabilities employed.

5.1. Founder Identities and Meanings

Recalling the definition of founder identity provided by Powell and Baker (2011), which denominates "the set of identities that is chronically salient to founder in her or his day- to-day work," this paper extends the term founder identity to refer explicitly to the set of social and role identities that is chronically salient to the founder in her or his day-to-day work. Founder identity thus consists of both the social groups and roles to which he or she perceives himself or herself as tied. While investigating the social groups and roles to which the founders of Mama Ocllo feel tied through the attached expectations towards perception and behaviour, the data showed particularly interesting patterns that suggested how the social and role identities interplay in organizational contexts. This finding can be compared to the results obtained by Powell and Baker (2011), and it suggests that the founder's social identities constitute the foundation for engagement in desired roles within the venture. Moreover, it is especially the new venture-creation process that allows the founder to express his or her social group membership by his or her new role performance. In the following, I use the data derived from the interviews and from the visual tasks conducted with the founders of Mama Ocllo to illustrate the interplay between social and role identities in the new venture-creation process.

Educational and professional background are often a basis for social identity formation. Individuals who belong to the same academic or vocational group are characterized by similar qualities and behaviours typical for that group. Martina's career went in very different directions, starting from her job as banker trainee (which, as she says, "was not something that stole [her] heart") and ending with her master degree in Romance Philology. The philology career took her to Peru, where she specialized in Latin American history of literature. Today, Martina's belonging to the vocational group of creative philologists manifests itself in her creative writing, drawing and visualising qualities, which are significant for her everyday work at Mama Ocllo. Consider the following statements regarding her roles in the company:

> I am responsible for everything that is associated with our shop, including the whole product descriptions. I take care of the pictures, and I coordinate everything that shapes Mama Ocllo... including the whole marketing department and the networking quite a lot. I take care of everything apart from the Excel tables and calculations. (Martina)

> I assume responsibility for all communication, in Germany and in neighbouring countries as well. (Martina)

Martina expresses her social identity as a "creative philologist" by involving herself in the role I defined as "the voice of Mama Ocllo." The prototype qualities for philologists – such as the ability to communicate with others in a very comprehensive and effective way, both in written text and verbally – are fundamental to her and at the same time represent one of the most desired roles in the venture. As "the voice of Mama Ocllo" Martina draws customers and the venture closer together by communicating with them in every channel, especially including social media.

In a similar vein, Martina's categorizes herself as a "LOHAS devotee" (lifestyle of health and sustainability), as she expresses by following statements:

> I buy from smaller labels. I buy the special products before I buy the cheap ones, because the quality is questionable so the purchase makes me feel badly.(Martina)

> I belong to minimalistic thinking consumers. I say, 'I don't need the closet full of baby kimonos, what for?' (Martina)

The social identity of a "LOHAS devotee" is translated, as she says, into her lifetime role as a "sustainable entrepreneur," who, apart from aiming at creating economic value, also pursues social and ecologic objectives.

> Fairness, sustainability, health, well-being, vegan, social – everything that goes in the direction of a green lifestyle. Mama Ocllo should be a green-lifestyle label. It is what I would connect with my entrepreneurship. (Martina)

Throughout the interview, Martina articulated her wish to make a positive contribution to the society and emphasized how important it is for her to positively influence the well-being of others – especially including the babies who use Mama Ocllo's clothing, but also the Peruvian society. Take, for instance, the following statement:

> But I would like to be perceived as somebody who makes a valuable contribution to society. It is how I would like to be perceived by others. (Martina)

The more you enter there [the business world], the more you realize that there are a lot of social evils detected or that a sustainability department is being developed in companies that has always been irresponsible. And also, the BIO labels that promote something that they are not at all. There is so much illusion. And this is something I didn't want. I have always told myself to be transparent. (Martina)

Responsible and transparent behavior towards others is also consistent with the "missionary" social identity defined by Fauchart and Gruber (2011). Missionaries are driven by a social or environmental cause when they create a venture; they value their responsible behavior as founders and provide good examples for others in the society. Martina's missionary social identity is regarded as source of aspiration for the role she currently occupies as "Peruvian community patron." Venture creation is essential to translating her social identity into a congruent role identity, as the venture's offerings are connected with a social project that enables Martina to make a positive contribution to Peruvian society.

Similarly, Gustavo's membership in the social group of industrial engineers, which as he explains, is characterized by,

> ...good planning ability, structured thinking, and an ability to sacrifice some of your life spheres.

provides a basis for engaging in the role of the "keeper of the bottom line" in Mama Ocllo. Gustavo's function – to manage efficient financial planning and production in the venture – is congruent with the prototype qualities of his industrial-engineering social identity.

Moreover, Gustavo's strong connection to his mother country, Peru, which is embedded in the "patriot" social identity, has guided his aspiration to engage in the role of, as he puts it, "Peruvian marginal areas supporter." As in the case of Martina, so also for Gustavo, the venture creation was crucial to enabling him to express his social identity in a new role.

Social identities such as "creative philologist," "industrial engineer," or "LOHAS devotee" motivate founders to act consistently with those groups and to engage in desired role identities like the "voice of Mama Ocllo," "keeper of the bottom line" or "sustainable entrepreneur." Particularly, new venture creation process facilitates the founder to engage in new role identities in the venture, which are an expression of currently held social identities.

6. Core Strategic Decisions

6.1. Products

First of all, vocational groups that are represented by founder social identities are strongly motivational and may predetermine the engagement of the founder in the product development. The motivation to engage in product creation may be twofold. First, my data demonstrates that founders who have social identities associated with creative qualities such as drawing, designing, writing, or structuring, actively engage and are motivated in the creation of a product. Secondly, a founder's engagement in product development is also influenced by the degree to which the product reflects the founder and his or her necessities.

For instance, Martina actively participates in the product creation process because engaging herself in activities, which include her creative qualities is congruent with her social identity as a "creative philologist." On the other hand, baby clothing, the venture's offering, is strongly tied to Martina's salient role identity as a mother, which predetermines her engagement in the product development process.

> ...But I'm really fascinated with our product development ... I find it exciting to create a product totally on our own. I really enjoy the creative work—actually, I enjoy everything that is creative—and I have the opportunity to visualize what we are doing – writing, drawing, everything in this area. (Martina)

> ...It was my biggest wish to become a mother at one point in my life. It was just mine and that's why I immediately identified myself with the topic of baby clothing. (Martina)

I also observed Gustavo's motivation to participate in the product development process. His engagement was purely dictated by the fact that he, like other entrepreneurs with a strong creation focus, enjoyed the development process of a product. However, the engagement in product creation does not indicate that the final outcome with all of its characteristics is a reflection of himself.

> ...I have always known that I want to produce something. When I was in Peru, I was looking for diverse ideas. I even prepared provisional business plans in different areas. (Gustavo)

> I think so [I identify myself with the product] but not completely because I am a man. At the beginning, I had thought that I would develop my company in a more manly sector. (Gustavo)

Regarding the particular attributes of the products, the main offering of Mama Ocllo is organic baby clothing that ranges from kimono baby suits, rompers, shirts, trousers, dresses, and boots. Every product is manufactured exclusively from natural Peruvian fibres, including diverse, world renowned organic cotton types or ivory nuts from the Amazon rainforest, which contribute to the protection of Peruvian soil from exploitation by the conventional textile industry. Furthermore, each piece of clothing is entirely produced in Peru by family-led small businesses or non-profit organizations that support socio-economic development in marginal regions of Peru and the emancipation of women in political,

economic and family dimensions. Figure 2 illustrates the diversity of Mama Ocllo's offerings in terms of their main characteristics. The findings of this thesis reveal that none of these characteristics are implemented in the product incidentally but are instead an expression of the interplay between social and role identities.

First, the collected data constitute preliminary evidence to suggest that social identities inspire the initial decisions about which consumer needs to address, while role identities function as drivers to implement the products. Social identities may thus be regarded as more inspirational. However, when implementing particular product characteristics, founders are tied to the expectations of the roles that they occupy ("Who I am") or the roles that they currently strive for ("Who I want to be"). For instance, Martina was experienced and trained in the organic industry, and, like other organic-product specialists she is convinced of its advantages in comparison to conventional mass products, which provided an inspiration for developing particularly organic product. However, as she explains, the role of "mother," for which she was striving, motivated her to implement the product for her baby. From the moment that Martina and Gustavo decided to found Mama Ocllo, Martina has not doubted her decision to offer baby clothing because of her very strong "mother" role identity, to which she was tied to when taking the decision. As expected from this role, she has always wanted the best for her baby; that is why she decided to address the needs of the youngest by offering organic clothing consistently with the "mother" role identity.

> The first thought I had was, we have perfect baby clothing for our child ... Well, when we launched our first collection, Miguelito [Martina's and Gustavo's son] was not yet born, but it was my biggest wish to become a mother at one point in my life. It was just mine, and that's why I immediately identified myself with the topic of baby clothing. (Martina)

> The necessities of the youngest were most important for us. As I said, it fits also with our life phase. We can also be authentic through that [the product], if we try it simultaneously. (Martina)

In comparison to Martina, Gustavo was trained as an industrial engineer with experience in textiles. His vocational group gave him the direction of the venture's offering. However, what really drove him to realize the product consistent with his social identity, was the wish to occupy the role of a "textile industry manufacturer." Additionally, he underlines his strong wish to simply create products and to be independent, which is prototypical for the social group of entrepreneurs. Nevertheless, the entrepreneurial social identity did only determine the decision about developing a product rather than about which particular customer needs to address.

> To be honest, at the beginning I only wanted to produce something that I knew about. I just thought about venture creation, independence, and specialization in the textile industry. (Gustavo)

Taken together, the initial decision to produce baby clothing was inspired by social in- groups that the founders belong to but was driven by an interplay of role identities held or desired by both founders, Martina and Gustavo. While the decision to address the needs of the youngest was shaped by Martina's strong role identity as a future mother who wants the best for her baby, the selection of the textile industry was mainly driven by Gustavo's role identity as a "textile manufacturer," which he was striving for. Upon analyzing the interconnection between the venture's offerings and the founders' social and role identities still further, a significant association between Martina's identites and the product's characteristics can be observed. Expressed in quantitative terms, Martina's social and role identities shaped seven of seven main characteristics of Mama Ocllo's products. On the other hand, Gustavo influences only two of the main aspects. For instance, the decision to specialize in sustainable fashion arose from Martina's missionary social identity, as Martina highlights her wish to make a positive contribution to the society, which is now possible thanks to Mama Ocllo's offerings, which provide a positive social and environmental example for other fashion brands. However, the role of a "Peruvian community supporter," which is congruent with her missionary social identity, drove Martina to actively support Peruvian socio-economic development by deciding to produce strictly organic clothing without a negative impact on Peruvian agriculture and farmers.

> It would have been contradictory, having said that my wish is to make a positive contribution in Peru, if I at the same time manufactured conventional cotton. It wouldn't have any added value in Peru. (Martina)

> I have always been very enthusiastic about showing people here all the treasures that Peru has to offer. In Peru, there are not only ponchos or Andinos walking around ... That's why at the beginning we put the emphasis on Peru, which I now think was not really reasonable, because it's only me who has this kind of feeling. I know what meaning Peru has for me and what is so special there, but somebody else associates it with a developing country. (Martina)

Additionally, Martina's strong commitment to the expectations tied to her role as a mother supported her decision to produce clothing that is healthy for babies in terms of its breathability, silky softness, and thermal characteristics, which ensure maximal skin tolerability. The company's production of healthy clothing for the youngest consumers may be explained by Martina's desire for positive appraisals of her
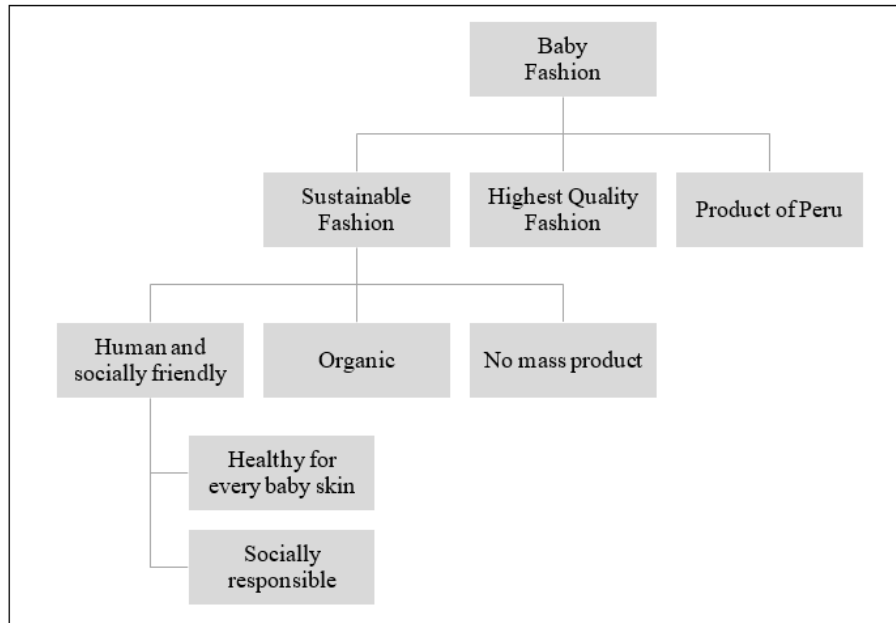
Figure 2: Main Characteristics of Mama Ocllo's Offerings

role performance, as is expected from a mother who wants to assure the best for her baby:

> ...We have to be consequent and we say, 'We want the best for the babies, the best for their skins ...' (Martina)

> We were sitting at the source and we could do what we wanted, what we really wished for our baby. This is a necessity of every mother. (Martina)

Moreover, the data points to an interesting pattern: a bidirectional relationship between founder identities and venture offerings. Not only did Martina's identities influence Mama Ocllo's products, but the product shaped the identities currently held by Martina. According to Martina, she wants to transmit through her product what she is currently living to be authentic to herself and her audience. Because in her private life she identifies herself with LOHAS consumers by purchasing mainly organic and fair products, she wanted her own product to be sustainable too. Engaging herself in the production of organic and socially responsible products has allowed Martina to translate her LOHAS social identity into the role of a "sustainable entrepreneur":

> I wish people to know that we are specialists in high-quality baby fashion from pima cotton. Everything is sustainable, organic and fair. (Martina)

> [Referring to products] The authenticity is for me most important. I am Mama Ocllo. It is not a company out of a hundred companies, because I am not a typical entrepreneur. It is now what you see, and this is a part of us. (Martina)

In the case of Gustavo, the data reveals that Martina's social and role identities have dominated those held or strived for by Gustavo. This finding is consistent with previous research on founder identity in sustainable entrepreneurship (Powell and Baker, 2011), which claims that sustainable entrepreneurs are motivated by a variety of social and role identities to pursue economic, social, and ecologic value. In contrast, conventional entrepreneurs are driven mainly by pure entrepreneurial identities such as investor, creator, profit-maximizer, etc. Their ventures are not necessarily reflections of themselves because their products are meant to be profitable and to address widely known customer needs. This argument applies well to Gustavo and his influence on the initial decision concerning Mama Ocllo's offerings. Apart from shaping the decision about producing clothing – which is consistent with the synergy of the congruent social and role identities he derived from experience in the textile area - Gustavo's wish to engage actively in the socio-economic development of Peru through the role of a "Peruvian community supporter" guided him to create a product representing the treasures of his country and allowing him to engage in the desired role.

> I can also say that I am a supporter of Peru's marginal areas. In general, we support the Peruvian economy with our venture... (Gustavo)

Interestingly, Gustavo's perception of the initial product differs from the actual offering and is strongly influenced by his self-categorization as an "entrepreneur," which is expressed by the congruent role of a "keeper of the bottom line" in the venture of Mama Ocllo. Although Martina has never doubted the final outcome of Mama Ocllo, Gustavo was neither sure about the baby aspect of the clothing, nor about the

organic aspect of the product. His entrepreneurial identity was about to shape his decision to create a mass product, for all age groups, from conventional cotton. Ultimately, it was Martina and her identities that shaped the final attributes of the product, which Gustavo himself confirms in the following statement:

> At the beginning I thought about offering both conventional and organic cotton. Later, Martina – actually both of us – decided to offer only organic clothing. Why baby clothing? I cannot really remember. I think it just arose through our long conversations. I actually thought about every age group, but we finally decided to specialize in baby clothing. And as we started developing the first collection, we wanted both cotton types for two different target markets. At one moment, Martina decided to sell exclusively organic baby clothing. (Gustavo)

All of the characteristics of the product that the consumers of Mama Ocllo know today are therefore a reflection of Martina's identities, which are derived from her profession but also from her sustainable lifestyle and family orientation. Gustavo's entrepreneurial identities remained in the initial business ideas. Today, the product reflects mainly his role identity as a "Peruvian community supporter." Take, for instance, following statements of the founders about their connection with the products:

> But the product line or what we want to do, it was actually all clear. We have never doubted the products. (Martina)

> Yes, they do [reflect] me. (Martina, speaking about the products)

> Well, I think yes, but not completely because I am a man. I thought at the beginning that I would develop a venture in a more male area. (Gustavo on whether the products are a reflection of himself)

> [I identify myself] especially with the fact that they are manufactured in Peru... (Gustavo)

When analyzing the impact of founder identities on the venture's products, I expected the results to be consistent with previous research on founder identity, which mainly describes the relationship as a straightforward path in which the founder with a strong confidence shapes his or her venture according to his or her self-definition. This does not seem to be the case with Gustavo's identities, which went through a transformation when the decision was made concerning Mama Ocllo's offerings. As previously mentioned, the product would have had different properties if the decision depended exclusively on Gustavo. However, it was Martina and her strong commitment to both social and role identities that

shaped the final product. Through perceived differences between Gustavo's and Martina's identities, and because of the knowledge about sustainability that Gustavo gathered from Martina, Gustavo started to adjust his identities towards a LOHAS social identity, which is also more congruent with the venture's offerings. Moreover, today Gustavo translates his LOHAS social identity into the role of a "sustainable textile manufacturer," which motivates him to implement Mama Ocllo's products:

> Well, if I have to be honest, the sustainability concept was new for me. I have never heard about it in Peru. I only had experience in the textile area, and I learned about organic agriculture here in Germany. Actually, thanks to Martina, I could learn about sustainability. (Gustavo)

> [Interviewer: After the decision about Mama Ocllo's offerings was made, what kept you motivated to realize the product?]

> It cost me a lot to understand the positive impact of ecological agriculture. All the time, I couldn't decide between the two product ideas ... And for me the sustainability concept was totally new ... However, after making the decision about the product, I understood that this is the right path. I learned about the positive impact of ecological agriculture and understood that it is the best for a Peruvian farmer. Now I want to contribute to the development of a sustainable textile industry in Peru. Now there is motivation to move something in this sector. (Gustavo)

6.1.1. Market Segments

Considering the association between market segments served and founder identities, my interview data first indicates that the target markets addressed consist of members of social groups that the founders belong to. Second, this indication applies to sustainable entrepreneurs rather than to for-profit entrepreneurs. Sustainable entrepreneurs dedicate their offerings to their in-groups; that is, they dedicate them to those people who share similar perceptions and behaviors. Such consumers give particular attention to the social and environmental performance of the product rather than to its price. On the other hand, conventional entrepreneurs identify the most profitable markets as the suitable targets for the company's offerings (Fauchart and Gruber, 2011); however, this does not mean that conventional entrepreneurs belong to those groups.

For instance, Martina has chosen the LOHAS consumer group, as she is also devoted to sustainable living. Therefore, she defines the adequate market segment as one that is congruent with her pro-social and environmental perceptions, behaviors, and vision.

Consider her following statements about the market segments addressed by Mama Ocllo:

> Yes, LOHAS [about the target market]. I want people to be interested in our concept. I want people to know what is so special about our products … I think of a target market, which apart from looking at the quality of the product, also gives attention to the actual idea, to the fact that it is not a mass product – a target market that wants something you cannot find in every department store. (Martina)

Interestingly, the decision to address LOHAS consumers was also driven by her negational categorization; thus, her separation from her respective out-group of affluent consumers, who do not put attention to the social and environmental performance of the product and are characterized by hyperconsumption:

> I don't want our market to consist of the old nobility who only pay attention to expensive products. I don't want it because it is a society of, I don't know, the typical fur-bearing people. I don't belong to them. (Martina)

This result supports the conclusion of Belz and Binder (2017), who indicate that sustainable entrepreneurs, apart from defining their venture's market segments, also underline the market segments that are not served by the offerings.

Consequently, Martina's social identity – particularly her affirmational categorization of the LOHAS in-group and her negational categorization of the out-group of the affluent consumer class – shaped the initial decision to address LOHAS consumers as the adequate market segment for Mama Ocllo's offerings. Interestingly, the decision to identify the LOHAS market segment was not that straightforward for Gustavo. Initially, his social identity as a conventional entrepreneur – defined by Fauchart and Gruber (2011) as a Darwinian identity – led him to the choice of the most profitable market segment: the consumer class with high purchasing power:

> Yes, at the beginning, I thought that this [affluent class] is the right market segment. (Gustavo)

> And when we started the first development, we wanted to offer both cotton types— organic and conventional – for different market segments. (Gustavo)

This finding is consistent with the previous results of Fauchart and Gruber (2011), who argue that traditional entrepreneurs do not necessarily produce for consumers like themselves, but instead identify the average or most profitable consumer.

Like the decision about Mama Ocllo's products, the decision about market segmets also implicated the incongruence of Gustavo's Darwinian identity with Martina's LOHAS social identity. Over time, Gustavo started to feel an attraction to the group of sustainable-living devotees by gaining awareness of the advantages for society and the environment that

are provided by this lifestyle. As a consequence, he attempted to gain a psychological entry to the social group of LOHAS by thinking and acting like the prototypical group members. Although the decision concerning which market to enter activated both of his social identities – namely, the Darwinian and later the LOHAS identities – Gustavo felt more tied to the social group of LOHAS and decided to serve more aware consumers rather than the most profitable ones. Like Martina's, Gustavo's negational categorization of the affluent consumer class is also reflected in the decision about which market segment not to address.

> Yes, in some aspects, yes [I identify myself with LOHAS]. But I think I am still in the process of transition. We do not buy or consume 100% organically yet. (Gustavo)

> But I went through different stages in my life, and now I see the things a little bit differently. I understand that it's good to identify yourself with the market segment that you offer your products to. If you identify yourself with your market segment, you can talk to the people from your market; you can interact with them. The market segment you mentioned [the affluent consumer class]—I don't belong to them. I don't know their wants, necessities, etc. And I even have the impression that this market segment is not interested in the organic or fair aspects of a product. I think branding is more important for them … I definitely do not identify with consumers with the highest purchasing power. That's why I don't think it is the right market segment for our products. (Gustavo)

Taken together, these findings suggest that Martina's strong pro-social and environmental orientation, represented by her LOHAS group membership, influenced the decision to address essentially socially and environmentally aware consumers. On the other hand, she also shaped her co-founder's social mobility into the group of LOHAS devotees, which makes him think and act consistent with the group's prototype, dominates his entrepreneurial identity, and finally convinced him to identify LOHAS consumers over the average or most profitable consumers as the best market segment for Mama Ocllo's offerings.

6.1.2. Resources and Capabilities

With regard to the resources and capabilities deployed to produce Mama Ocllo's offerings, the venture cooperates with small, family-run businesses in Peru and with a nonprofit organization that aims at women's emancipation in Peruvian marginal areas. The selection of suppliers was based on transparency, premium quality, and shared vision about socially and environmentally responsible production methods. As a sustainable venture, Mama Ocllo builds long-term relationships with its suppliers and aims at shared growth.

Considering the association between the resources and capabilities deployed by the venture and the founder identities, my data provides evidence to suggest that these aspects are a reflection of founder identities that are tied to the roles that the founders currently take on in life or strive for. Interestingly, the founders of Mama Ocllo felt commitment to divergent roles, which congruently reinforced the decision concerning which resources to deploy and what supplier cooperation to build. However, overall, the majority of aspects taken into account when deciding on resources and capabilities – such as shared vision, building sustainable relationships with suppliers, and working exclusively with organically cultivated and processed fibers – reflect Martina's role identities as "sustainable entrepreneur," "mother," and "Peruvian community patron." For instance, Martina, as a convinced "Peruvian community patron," decided to build valuable cooperation with Peruvian suppliers on the basis of shared growth and responsibility towards society and environment. Accordingly, Mama Ocllo cooperates exclusively with family-run businesses in Peru. Moreover, her life mission to do good for Peru encouraged Martina to cooperate with a non-governmental organization that supports socially excluded women in Peru:

> We have chosen our suppliers very carefully. It is again a family-run business... (Martina)

> And a feminist NGO that has existed for 30 years produces our toys. It is a big organization that stands for women rights. They want women to become more independent in society, especially women from the Andes who don't have much to say and do not have equal rights in politics. They developed a project called La Casa de la Mujer Artesana, which also aims to make women independent economically. We have worked with them since the beginning, and we will maintain this cooperation because there is a full transparency. This is a beautiful cooperation and the quality is perfect too. (Martina)

Additionally, the exclusive use of organically cultivated fibers and chemical-free processing are both a result of Martina's strong commitment to her role as a mother who wants the purest and healthiest materials for her baby. Martina's commitment to the "mother" role is not only influenced by the importance of her relationship with her son but also by the expectations of other mothers who buy Mama Ocllo's clothing:

> We have to be consistent. We say we have the best clothing for babies, the best clothing for baby skin and not a product that must be returned ... because the baby gets an allergy. (Martina)

The general picture that emerges from the data oscillating between Martina's identities and the resources and capabilities employed by Mama Ocllo is that Martina's role as a "sustainable entrepreneur" permeates basically every decision about which resources to employ and what criteria to use in selecting appropriate suppliers. Consistent with the expectations of that role, Martina drew special attention to her social responsibility in the production country but also to the minimal negative impact on the natural environment throughout the whole value chain. Take, for instance, her statement regarding the printing technique:

> Currently we want to start with textile printing. Plastic is the most important printing technique in the textile industry. This totally contradicts our concept ... Another printing technique is water-based. This printing technique is more expensive, but the water-based ink vaporizes. This means that it is neither harmful to the printer nor to the environment. And if there is some ink released, it is automatically sieved. So, nothing comes to the cycle. (Martina)

Comparing the influence of the founder identities on resources and the capabilities employed, the findings suggest again that Martina seemed to know the direction and destination in which she needed to go when choosing sustainable resources and focusing on long-term relationships. Gustavo however, experienced identity adjustment consistent with Martina's self-definition and vision for the venture.

With regard to the employment of recourses, the decision to manufacture exclusively from organic cotton was made by Martina. Gustavo wanted to act consistent with his previous experience in the role of a "textile manufacturer." As I explained previously, after Martina made the decision about natural fibers, Gustavo shifted his role as a "textile manufacturer" to that of a "sustainable textile manufacturer," which is inspired by the LOHAS social identity.

Gustavo's stronger commitment to sustainability recently became salient when the founders had to decide about the appropriate printing technique for the newest clothing collection. Taking two different printing techniques into consideration, the orientation towards social and environmental responsibility embedded in his role as a "sustainable textile manufacturer" dominated Gustavo's Darwinian identity:

> Well, the water-based printing technique has a smaller negative impact on the environment. There is always an impact on the environment, but water-based printing has a minimal impact in comparison to the plastic-based technique. (Gustavo)

Interestingly, although Gustavo mentions size as an important criterion for choosing the right supplier, which is consistent with his role identity as "keeper of the bottom line," he agreed to give a chance to a small family business in Peru. Consequently, despite the strong entrepreneurial orientation, Gustavo is more tied to his life mission to support Peruvian marginal areas and make a positive contribution to the textile industry by providing an example to follow.

7. Discussion

7.1. Discussion of the Results

I conducted this study for the purpose of my thesis by noticing the gap in previous research about founder identity (Fauchart and Gruber, 2011; York et al., 2016; Belz and Binder, 2017; Hoang and Gimeno, 2008), which treat social and role identities separately when investigating their impacts on a venture. Taking into consideration that each founder is at once a member of diverse social groups and a performer of multiple roles, I have aimed in this paper to explore how the social and role identities of a founder simultaneously influence core strategic decisions made in the venture-creation process, particularly concerning which customer needs to address, which market segments to serve, and which resources and capabilities to deploy. The general picture emerging from the analysis of the venture-creation process is that social and role identities have different yet complementary impacts on shaping the venture. Specifically, the identities complement each other in such a manner that founder's social identities stimulate his or her aspirations for engaging in roles, which in turn drive him or her to shape the venture through the process of decision making. Additionally, my findings suggest that each decision shaping the venture is a result of different interplays between social and role identities. While products are mainly a result of the complementary impact of social and role identities, the market segments served reflect the founder's social identities. Resources and capabilities are in turn driven by the expectations tied to the roles that are occupied or desired by founders, which in most cases are an expression of the social groups that they identify with.

Products. Previous studies (Fauchart and Gruber, 2011; Belz and Binder, 2017) demonstrate that the social identities of a founder are reflected in a venture's offerings, where affirmational (and negational) categorization provides a positive (and negative) frame of reference about a good (and a bad) products. Consistent with these results, my findings demonstrate that the initial decision about which products to offer is driven especially by the vocational groups to which the founders belong. Consider Gustavo and Martina, who decided to produce organic clothing as a result of Gustavo's identification with other textile specialists and Martina's vocation in the organic industry. Furthermore, my results extend the understanding of the interrelation between founder identities and a venture's offerings by suggesting that the roles that founders currently occupy or strive for also shape decisions about the final product. Particularly, I observed that founders feel highly committed to the roles they occupy ("Who I am") or desire ("Who I want to be") and behave in accordance with the expectations tied to those roles when they must implement the venture's final outcome. Martina, for instance, felt a strong attachment to her role identity as a mother when she designed organic clothing for babies, while Gustavo felt a commitment to his Peruvian community, which he wanted to support by developing a product that will be manufactured in his mother country. Given social and role identities together, the data presented in this study suggests that social identities fulfil an inspirational function in comparison with expectations towards the behavior tied to role identities, which have a motivating effect.

Market Segments. My key finding emphasizes that the decision about which market segments to serve is significantly related to the social identities held by the founder. First, founders tend to address market segments that are based on the uniformity of perceptions and behavior between the founder and customers. In other words, selected market segments consist of customers who belong to the same social groups as the founders. Take, for instance, Martina, who has a hybrid identity that consists of communitarian and missionary social identities. She addresses her offerings to LOHAS consumers who, like her, are also interested in the ecological and social performance of the product apart from its quality and price. This data gives additional support to the results obtained by Fauchart and Gruber (2011), who claim that founders with a communitarian social identity serve market segments that match the attitudinal and behavioral lines of the founders. Although Gustavo, first tied to his Darwinian identity, wanted to address the most profitable market segment, which is congruent with the findings of Fauchart and Gruber (2011), he later obtained a LOHAS social identity and realised the importance of addressing consumers who think and act like him, which in turn supports my claim about the uniformity of perceptions and behaviors as a criterion for choosing the right market segments in the context of sustainable entrepreneurship. Second, founders define their market segments based on their separation from social outgroups, which are characterized by unsustainable consumption patterns. For instance, consider Martina, who defines her customers in sharp contrast to the affluent consumer class that is characterized by mass consumption of luxury goods produced in a socially and environmentally irresponsible way. This statement is congruent with previous research of Belz and Binder (2017) on the negational categorization of sustainable entrepreneurs.

Resources and Capabilities. Interestingly, when making decisions about resources and capabilities that have to be deployed to produce a venture's offerings, founders feel especially commited to their roles in comparison with their social groups. My finding indicates that expectations tied to roles drive founders to behave in a manner that is consistent with those roles, as in the case of Martina, who, driven by her role of a Peruvian community supporter, emphasizes the importance of socially and environmentally responsible production methods in Peru. A possible interpretation of this finding may be the strong commitment to the roles defined by number and importance of relationships that are tied to the roles that founders engage in. Because the activation of a particular role depends on commitment to that role (Hogg et al., 1995; Stets and Burke, 2000), I indicated that founder's roles – especially roles as Peruvian community patrons or sustainable textile manufacturers – are associated with a very high number of persons – like the whole community or industry – to whom the founders are tied through these roles, which

makes these roles highly relevant to the decisions about re-
sources and capabilities. On the other hand, one must not
forget that these roles are often expressions of a founder's so-
cial identities. For example, Martina's missionary identity is
translated into the role of a "Peruvian community supporter,"
while Gustavo's increasing identification with LOHAS devo-
tees is expressed by his role as a "sustainable textile manu-
facturer." Moreover Martina's role identity as a mother is, as
she says, the most important role in her life, which also ex-
plains her commitment to that role when making decisions
about which resources and capabilities to deploy to produce
an offering that she will also use for her baby.

7.2. Theoretical Contributions

SIDT and IDT isolation versus SIDT and IDT integration.
My findings complement prior research on identity in en-
trepreneurship by integrating social and role identies as core
foundations of founder identity. While the vast majority of
previous studies (Hoang and Gimeno, 2008; Cardon et al.,
2009; Fauchart and Gruber, 2011; Mathias and William,
2014; Murnieks et al., 2014; Belz and Binder, 2017) adopt an
unilateral perspective on founder identity – such that either
social or role identities are investigated to have a potential
influence on founder's behavior in his or her day-to-day work
– my findings demonstrate that founders are simultaneously
tied to their social groups and to currently occupied or de-
sired roles when shaping their ventures. This result provided
me with a framework that is relevant to various observations.
First, I observed a more heterogeneous interrelation between
founder identity and the decision- making process where
core strategic decisions shaping the venture may reflect ei-
ther social identities or role identities or a synergy of the role
and social identities of a founder. This finding substantiates
previous results obtained by Powell and Baker (2011), who,
apart from highlighting the importance of bridging SIDT with
IDT, also demonstrate that entrepreneurial decision making,
particularly responses to adversity, is reflected by different
combinations of founder's social and role identities. Second,
I realized that a clear separation of founder's social and role
identities is not always possible because social categories
and roles complement each other such that social identities
provide an inspiration to engage in congruent roles in the
venture, which in turn drive the founder to act in a manner
that is consistent with those roles when implementing a de-
cision. However, this does not mean that my study opposes
previous studies; on the contrary, the social identities de-
rived by Fauchart and Gruber (2011), or the different shades
of founder's social identities delivered by Belz and Binder
(2017) should be regarded as amont the pillars of founder
identity theory. Finally, my case expands the understanding
of founder identity, which may also be composed of multi-
ple identities that are derived from the founder's vocation,
education, interests, relationships with others, world per-
ception, etc., and which guide or drive the founder to shape
his or her venture accordingly rather than being always con-
structed by a single identity (Fauchart and Gruber, 2011;

Mathias and William, 2014) that becomes salient in decision
making process.

Single versus multi-founder ventures. While previous re-
search (Fauchart and Gruber, 2011; Powell and Baker, 2011;
Belz and Binder, 2017; Mathias and William, 2014) has
relied primarily on companies that are created by a single
founder, my findings add to a growing body of literature
on founder identity by gaining insight into a multi-founder
venture in which the decision-making process is influenced
by diverse sets of social and role identities held by mul-
tiple founders. The investigation of a multi-founder ven-
ture provided me with a base for understanding the identity
processes taking place during the venture-creation process.
First, my findings suggest that a founder's path towards ven-
ture creation is not always a straightforward one in which
the founder, aware of his or her identities, simply translates
them into the venture through the decision-making process.
Rather, founders may experience various identity transitions
in the venture- creation process, particularly when a conflict
of identities arises between founders with regard to strategic
decisions that shape the venture. The conflict of identities
results in turn in each founder's individual efforts to shape
the venture in a manner that is consistent with held identi-
ties, as happened in the case of Martina and Gustavo, who
initially did not agree about producing conventional or or-
ganic clothing. Various identity conflicts that arose between
Martina and Gustavo in the initial phase of venture devel-
opment shaped the pattern through which Martina – with
her strong commitment to sustainability, Peruvian commu-
nity and her role as a mother – significantly dominated the
identities oriented towards conventional entrepreneurship
held by Gustavo. Finally, the incongruence in the strategic
decision-making process led Gustavo to adjust his identities
towards those held by Martina, which were then transmitted
into the venture. This finding is particularly congruent with
the newest research conducted by Powell and Baker (2017),
who demonstrate how individual social identities become
adjusted into a collective identity prototype that describes
"who we are" in a venture created by multiple founders
holding diverse identities. Secondly, the direct comparison
of the identity processes taking place between Martina and
Gustavo allowed me to notice that identities embedded in
entrepreneurs who pursue social, economic, and ecological
objectives are less flexible than those held by conventional
entrepreneurs.

*Unilateral versus bilateral relationship between founder
identity and venture.* While previous studies (Hoang and
Gimeno, 2008; Cardon et al., 2009; Fauchart and Gruber,
2011; Powell and Baker, 2011; Murnieks et al., 2014; Belz
and Binder, 2017) have shed light on the impact of founder
identities on ventures in their nascent stage of development,
my research provides additional insight into the link between
founder identity and the venture by providing preliminary
evidence for the existence of a bidirectional relationship such
that the venture has an impact on the founder's identities. In
essence, I have noticed that venture creation allows founders
to translate their social identities into new roles in the ven-

ture, as happened in case of Martina, who could express her belonging to the vocational group of philologists through her role identity as the "voice of Mama Ocllo." Powell and Baker (2011) emphasize that social identities usually provide a source of aspiration for new role construction if there is an incongruence between currently held social and role identities. My findings demonstrate that engagement in ventures provides a favorable context to create new roles that are congruent with currently held social identities, even if there is not an explicit identity conflict embedded in the founder. Moreover, involvement in the process of venture creation may also result in identity transitions, especially given a growing disparity between other founder's identities and the venture, which is what happened in case of Gustavo, who adjusted his role of a textile manufacturer into a role of a sustainable textile manufacturer congruently with Martina's identities and the venture.

7.3. Limitations and Opportunities for Future Research

The aim of this thesis is to explore how a founder's social and role identities influence core strategic decisions in the venture-creation process. As in every study, a number of potential weaknesses must be considered as well.

First, due to the exploratory nature of this study and the limited time available to conduct it, my dataset is limited to a single venture. This limitation makes my findings not generalizable beyond this case study. I acknowledge the lack of generalizability as a weakness of this thesis, though I focused on gaining an in-depth understanding of the link between founder identities and the venture, which justifies my selection of a single case (Yin, 2014). Although the results I have obtained demonstrate how idiosyncratic founder identities may be – a fact that makes the empirical study of the phenomenon more complex – my paper should be regarded as a stimulus for futures studies to continue exploring the dual impact of a founder's social and role identities on the venture by including more samples in the investigation. To gain insight into more ventures would not only guarantee the generalizability of the findings but may also identify different types of interplay between a founder's social and role identities than what are presented in this study. Second, when I decided to analyze the founder's social and role identities in the venture-creation process, I defined the context of this research as sustainable entrepreneurship because I assumed, as have previous researchers (Powell and Baker, 2011; Belz and Binder, 2017), that sustainable entrepreneurs hold a diversity of identities that they translate into their ventures such that they will provide more interesting insights into the phenomenon studied. However, the example of Gustavo – who at the beginning seemed to behave mostly in accordance with his Darwinian identity but later revealed a variety of social and role identities derived from his vocation, origin, and family that drive his decisions in day-to-day work – should encourage future research to extend the context of the investigation by taking into account not only diverse industries, but also non-sustainable entrepreneurs.

Furthermore, I believe that my thesis will serve as a basis for future studies on multi- founder ventures, particularly those that seek to determine how diverse founders that hold different sets of social and role identities influence each other mutually and influence the venture-creation process. Future research should also address identity transitions of founders and how identity dynamics are reflected in the venture. Finally, the results obtained in this thesis should also encourage researchers to meet the challenge of exploring a dual relationship between founder identities and the venture, where the venture may also influence the identities embeded in the founder's self- definition.

8. Conclusion

Sustainable ventures are a reflection of founder identities. Drawing on social identity theory and identity theory, my single-case analysis allowed me to explore how founder identities derived from his or her social group memberships and occupied roles simultaneously influence core strategic decisions about which customer needs to address, which market segments to serve, and which resources and capabilities to employ. Social and role identities have different yet complementary impact on new venture creation process. The difference in their impact refers to the fact that each strategic decision is interrelated to either social or role identity, or to an interplay of both. Social and role identities complement each other in their impact on new venture creation process in such a manner that founder's social identities stimulate his or her aspirations for engaging in roles, which in turn drive him or her to shape the venture through the process of decision making. With this bachelor thesis I would like to encourage researchers to further investigate the synergy of social and role identities in organizational contexts by putting special attention to broader study settings together with multi-founder ventures and identity transition processes.

References

Abell, D. F. *Defining the business: The starting point of strategic planning*. London, Prentice Hall, 1980.

Abrams, D. M. and Hogg, M. A. *Social Identity Theory: Constructive and Critical Advances*. London, Harvester/Wheatsheaf, 1990.

Bagnoli, A. Beyond the standard interview: the use of graphic elicitation and arts-based methods. *Qualitative Research*, 9(5):547–570, 2009.

Barton, K. C. Elicitation techniques: Getting people to talk about ideas they don't usually talk about. *Theory & Research in Social Education*, 43(2): 179–205, 2015.

Belz, F.-M. Shaping the future: Sustainable innovation and entrepreneurship. *Social Business*, 3(4):311–324, 2013.

Belz, F.-M. and Peattie, K. *Sustainability marketing: A Global Perspective*. Chichester, Wiley, 2012.

Belz, F. M. and Binder, K. Sustainable Entrepreneurship: A Convergent Process Model. *Business Strategy and the Environment*, 26(1):1–17, 2017. doi: 10.1002/bse.1887.

Binder, J. and Belz, F.-M. *Sustainable entrepreneurship: What it is. Handbook of Entrepreneurship and Sustainable Development Research. Kyrö P. (ed.)*. Cheltenham, Edward Elgar Publishing, 2015. doi: https://doi.org/10.4337/9781849808248.00010.

Binder, J. K. and Belz, F. M. Mission possible: Recognizing entrepreneurial opportunities in social and ecological problems. *Academy of Management Proceedings*, 2014(1):16256, 2014.

Burke, P. J. The self: Measurement requirements from an interactionist perspective. *Social Psychology Quarterly*, 43:18–29, 1980.

Burke, P. J. and Reitzes, D. C. The link between identity and role performance. *Social Psychology Quarterly*, 44:83–92, 1981.

Burke, P. J. and Stets, J. E. Trust and commitment through self-verification. *Social Psychology Quarterly*, 62:347–366, 1999.

Burke, P. J. and Tully, J. C. The measurement of role identity. *Social Forces*, 55(4):881–897, 2001.

Cardon, M., Wincent, J., and Drnovsek, M. The Nature and Experience of Entrepreneurial Passion. The Academy of Management Review. *The Academy of Management Review*, 34:511–532, 2009. doi: 10.5465/AMR.2009.40633190.

Choi, D. Y. and Gray, E. R. The venture development process of "sustainable" entrepreneurs. *Management Research News*, 31(8):558–570, 2008.

Cohen, B. and Winn, M. I. Market imperfections, opportunity and sustainable entrepreneurship. *Journal of Business Venturing*, 22:29–49, 2007.

Eisenhardt, K. M. and Graebner, M. E. Theory building from cases: Opportunities and challenges. *Academy of Management Journal*, 50(1):25–32, 2007.

Fauchart, E. and Gruber, M. Darwinians, communitarians and missionaries: The role of founder identity in entrepreneurship. *Academy of Management Journal*, 54(5):935–957, 2011.

Fiske, S. T. and Taylor, S. E. *Social Cognition*. New York, McGraw-Hill, 1991.

Gibbs, D. Sustainability entrepreneurs, ecopreneurs and the development of a sustainable economy. *Greener Management International*, 55:63–78, 2009.

Hoang, H. and Gimeno, J. Becoming a Founder: How Founder Role-Identity Affects Entrepreneurial Transitions and Persistence in Founding. *Journal of Business Venturing*, 25(1):41–53, 2008. doi: 10.1016/j.jbusvent.2008.07.002.

Hogg, M. A. *Social categorization, depersonalization and group behavior. InM.A. Hogg & R.S. Tindale (ed.). Blackwell Handbook of Social Psychology: Group Processes*. Malden, Blackwell Publishers, 2001.

Hogg, M. A. and Abrams, D. *Social identifications: A social psychology of intergroup relations and group processes*. London, Routledge, 1988.

Hogg, M. A. and Mullin, B. A. *Joining groups to reduce uncertainty: Subjective uncertainty reduction and group identification. In D. Abrams & M.A. Hogg (ed.). Social Identity and Social Cognitions*. Oxford, Blackwell, 1999.

Hogg, M. A. and Terry, D. J. Social identity and self-categorization processes in organizational contexts. *Academy of Management Review*, 25:121–140, 2000.

Hogg, M. A., Terry, D. J., and White, K. M. A tale of two theories: A critical comparison of identity theory with social identity theory. *Social Psychology Quarterly*, 58:255–269, 1995.

James, W. *The Principles of Psychology*. New York, Dover, 1980.

Josselson, R. *The space between us. Exploring the dimensions of human relationships*. London, Sage, 1996.

Lindesmith, A. R. and Strauss, A. L. *Social Psychology*. New York, Holt Rinehart and Winston, 1956.

Mathias, B. D. and William, D. W. The impact of role identities on entrepreneurs'evaluation and selection of opportunities. *Journal of Management*, 20(10):1–27, 2014.

McCall, G. and Simmons, J. *Identities and interactions*. New York, Free Press, 1978.

Mead, G. H. *Mind, Self and Society*. Chicago, University of Chicago Press, 1934.

Miles, M. B. and Huberman, A. M. *Qualitative data analysis: an expanded sourcebook*. London, Sage, 2014.

Morton, S. and Podolny, J. Love or Money? The Effects of Owner Motivation in the California Wine Industry. *Journal of Industrial Economics*, 50:431–56, 2002. doi: 10.2139/ssrn.133631.

Murnieks, C. Y., Mosakowski, E., and Cardon, M. S. Pathways of Passion: Identity Centrality, Passion, and Behavior Among Entrepreneurs. *Journal of Management*, 40(6):1583–1606, 2014. URL https://doi.org/10.1177/0149206311433855.

Oakes, P. *The salience of social categories. In J.C. Turner (ed.). Rediscovering Social Group*. New York, Basil Blackwell, 1987.

Powell, E. E. and Baker, T. It's what you make of it: Founder identity and enacting strategic responses to adversity. *Academy of Management Journal*, 57(5):1406–1433, 2011.

Powell, E. E. and Baker, T. In The Beginning: Identity Processes and Organizing in Multi-Founder Nascent Ventures. Academy of Management Journal. Advance Online Publication. amj.2015.0175, 2017.

Shepherd, D. A., Williams, T. A., and Patzelt, H. Thinking about entrepreneurial decision making: Review and research agenda. *Journal of Management*, 41(1):11–46, 2015.

Stets, J. E. and Burke, P. J. Identity Theory and Social Identity Theory. *Social Psychology Quarterly*, 63(3):224–237, 2000.

Stets, J. E. and Burke, P. J. New directions in identity control theory. *Advances in Group Processes*, 22(1):43–64, 2005.

Stryker, S. Identity salience and role performance: The importance of symbolic interaction theory for family research. *Journal of Marriage and the Family*, 30:558–564, 1968.

Stryker, S. *Symbolic Interactionism: A Social Structural Version*. Palo Alto, Benjamin/Cummings, 1980.

Stryker, S. and Burke, P. J. The past, present, and future of an identity theory. *Social Psychology Quarterly*, 63(4):284–297, 2000.

Stryker, S. and Statham, A. *Symbolic interaction and role theory. In G. Lindzey & E. Aronson (ed.). Handbook of Social Psychology*. New York, Random House, 1985.

Tajfel, H. Social identity and intergroup behavior. *Social Science Information*, 13:65–93, 1974.

Tajfel, H. Human groups and social categories: Studies in social psychology.CUP Archive, 1981.

Tajfel, H. and Turner, J. C. *An integrative theory of intergroup conflict. In W.G. Austin & S. Worchel (ed.). The Social Psychology of Intergroup Relations: 33-47*. 1979.

Tajfel, H. and Turner, J. C. *The social identity theory of intergroup behavior. In: S. Worchel en WG Austin (red.), The Psychology of Intergroup Relations, 7-24*. 1986.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., and Whetherell, M. S. *Rediscovering the social group: A self-categorization theory*. New York, Basil Blackwell, 1987.

Yin, R. K. *Case Study Research. Design and Methods*. Thousand Oaks, CA, Sage, 2014.

York, J. G., O'Neil, I., and Sarasvathy, S. D. Exploring environmental entrepreneurship: Identity coupling, venture goals, and stakeholder incentives. *Journal of Management Studies*, 53(5):695–737, 2016.

Junior Management Science

journal homepage: www.jums.academy

Withholding Tax on Digital Transactions – Status Quo and Potential Alternative Courses of Action

Die Quellenbesteuerung bei digitalen Transaktionen – Status quo und mögliche Handlungsalternativen

Markus Sebastian Gebhart

Ludwig-Maximilians-Universität München

Abstract

The providers of digital advertising services are criticised for being omnipresent worldwide, but to pay no significant corporate tax in many countries. In addition to supranational efforts at the OECD and EU level to solve the taxation problems of digital business models, some literature argues that providers of foreign digital advertising services are already subject to limited tax liability in Germany on the basis of existing tax laws. This paper therefore deals with the question of how German taxation rights for digital advertising services which are provided domestically by foreign corporations with the help of the internet can be ensured. To this end, the paper first examines whether German tax laws are capable of taxing cross-border digital advertising services. Subsequently, it will examine whether the draft directives discussed at EU level on the taxation of digital business models would be suitable for ensuring taxation in Germany. The study shows that the existing German tax laws cannot ensure the taxation of digital advertising services in most cases and that the introduction of the EU draft directives in their current form cannot be recommended.

Zusammenfassung

Die Erbringer digitaler Werbeleistungen stehen in der Kritik, weltweit omnipräsent zu sein, jedoch in vielen Staaten kaum nennenswerte Unternehmenssteuerabgaben zu leisten. Neben supranationalen Bemühungen auf Ebene der OECD und EU, die Besteuerungsprobleme digitaler Geschäftsmodelle zu lösen, wird von Teilen der Literatur vertreten, im Ausland ansässige Erbringer digitaler Werbeleistungen unterlägen in Deutschland bereits auf Grundlage bestehender Steuergesetze der beschränkten Steuerpflicht. Die Arbeit beschäftigt sich daher mit der Frage, wie deutsche Besteuerungsrechte an digitalen Werbeleistungen sichergestellt werden können, welche von im Ausland ansässigen Körperschaften mithilfe des Internets im Inland erbracht werden. Hierzu wird zunächst untersucht, ob die deutschen Steuergesetze in der Lage sind, grenzüberschreitende digitale Werbeleistungen steuerlich zu erfassen. Anschließend wird geprüft, ob die auf Ebene der EU diskutierten Richtlinienentwürfe zur Besteuerung digitaler Geschäftsmodelle geeignet wären, eine Besteuerung in Deutschland sicherzustellen. Die Arbeit zeigt, dass die bestehenden deutschen Steuergesetze in den meisten Fällen eine Besteuerung digitaler Werbeleistungen nicht sicherstellen können und die Einführung der EU-Richtlinienentwürfe in derzeitiger Ausgestaltung nicht empfohlen werden kann. *Rechtsstand 01.04.2019*

Keywords: Digitale Werbeleistungen; Quellenbesteuerung; Rechteüberlassung; Digitalsteuer; signifikante digitale Präsenz.

1. Einleitung

Die Digitalisierung schreitet immer weiter voran und nimmt großen Einfluss auf das gesamte Wirtschaftsleben. Im Zuge der Digitalisierung wird daher auch von der vierten industriellen Revolution oder Industrie 4.0 gesprochen.[1] Nach der Mechanisierung durch Wasser- und Dampfkraft im späten 18. Jahrhundert (erste industrielle Revolution), der

---

[1] So etwa Kagermann, Lukas & Wahlster, 2011; Forschungsunion & acatech, 2013.

Elektrifizierung im frühen 20. Jahrhundert (zweite industrielle Revolution) und der Informatisierung und Automatisierung der 1970er-Jahre (dritte industrielle Revolution) stellt die Digitalisierung die nächste zentrale Entwicklungsstufe der industriellen Produktion dar.[2] Diese „digitale Wirtschaft" hat eine Reihe neuer Geschäftsmodelle ermöglicht. Eines dieser neuen digitalen Geschäftsmodelle ist die Erbringung digitaler Werbeleistungen.[3]

Viele bekannte und hoch profitable digitale Geschäftsmodelle basieren auf der Erbringung digitaler Werbeleistungen. Als prominente Beispiele dieser Geschäftsmodelle sind z. B. die Konzerne Alphabet (vormals Google) und Facebook zu nennen. Die bekannteste Alphabet Tochter Google hat im Jahr 2018 einen Umsatz von ca. 116 Mrd. USD durch die Erbringung digitaler Werbeleistungen erwirtschaftet, was ca. 85 % der Gesamtumsätze entspricht.[4] Facebook erzielte in 2018 nahezu seine gesamten Umsätze i. H. v. ca. 56 Mrd. USD aus der Erbringung digitaler Werbeleistungen.[5] Die in Deutschland erwirtschafteten Umsätze werden bei Google im Jahr 2018 auf ca. 3,9 Mrd. EUR und bei Facebook auf ca. 850 Mio. EUR geschätzt.[6] Gleichzeitig stehen diese Konzerne regelmäßig in der Kritik, zwar in vielen Staaten der Welt wirtschaftlich omnipräsent zu sein, aber in den wenigsten dieser Staaten nennenswerte Unternehmenssteuerabgaben zu leisten.[7] So wies Google im Jahr 2017 in Deutschland „nur" ca. 102 Mio. EUR Gewinn aus und entrichtete darauf Steuerzahlungen i. H. v. ca. 33 Mio. EUR.[8] Im selben Zeitraum belief sich der Gewinn bei Facebook in Deutschland auf ca. 2,5 Mio. EUR, worauf ca. 1 Mio. EUR Steuern entrichtet wurden.[9] Daher wird mittlerweile angezweifelt, dass das derzeitige internationale Steuersystem mit seinen Wurzeln in der ersten Hälfte des 20. Jahrhunderts noch in der Lage ist diese (und weitere) neuen digitalen Geschäftsmodelle des 21. Jahrhunderts steuerlich sachgerecht zu erfassen.[10]

Die Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) hat es sich daher im Jahr 2013 zur Aufgabe gemacht, die mit der Erbringung digitaler Werbeleistungen und weiterer digitaler Geschäftsmodelle verbundenen Besteuerungsprobleme auf internationaler Ebene zu lösen.[11] Die Bewältigung dieser Aufgabe beschäftigt die OECD derzeit im Zuge des Aktionspunktes 1 des Projektes zur Bekämpfung von Gewinnverkürzung und Gewinnverlagerung (engl. Base Erosion and Profit Shifting, kurz BEPS).[12] Allerdings konnten auf Ebene der OECD bislang keine Hand-

lungsempfehlungen ausgesprochen werden, um den Herausforderungen zu begegnen, welche digitale Geschäftsmodelle für das internationale Steuersystem darstellen.[13]

In Ermangelung einer international konsensfähigen Lösung auf Ebene der OECD sind mittlerweile viele Staaten dazu übergegangen unilaterale Maßnahmen zu implementieren, um die Besteuerung bestimmter grenzüberschreitender digitaler Transaktionen aus digitalen Geschäftsmodellen sicherzustellen. In Europa planen derzeit z. B. Frankreich, Großbritannien, Italien, Österreich, Spanien und die Türkei die Einführung solcher unilateraler Maßnahmen.[14] Ein Ziel dieser Maßnahmen ist dabei auch, die Erbringer digitaler Werbeleistungen wie z. B. Google und Facebook der nationalen Besteuerung zu unterwerfen. Auf Grundlage eines gemeinsamen Vorstoßes der Finanzminister Deutschlands, Frankreichs, Spaniens und Italiens und daraus resultierend des Rates der Europäischen Union werden seit September 2017 nunmehr auch auf europäischer Ebene konkrete Maßnahmen geprüft, welche die Besteuerung bestimmter digitaler Transaktionen wie die Erbringung digitaler Werbeleistungen sicherstellen sollen.[15] Seit dem 21.03.2018 stehen dazu auf Ebene der Europäischen Union (EU) zwei Richtlinienentwürfe zur Diskussion. Die Besteuerung soll kurzfristig mit einer sogenannten Digitalsteuer sichergestellt werden. Langfristig soll eine sachgerechte Besteuerung durch die Einführung eines neuen steuerlichen Tatbestands anhand einer signifikanten digitalen Präsenz sichergestellt werden.[16]

Deutschland hat sich bei der Planung/Einführung unilateraler Maßnahmen bisher zugunsten einer Lösung auf Ebene der OECD bzw. der EU zurückgehalten. Seit Kurzem wird jedoch von Vorstößen aus Teilen der Finanzverwaltung berichtet, welche eine Steuerpflicht von im Ausland ansässigen Unternehmen, die im Inland digitale Werbeleistungen erbringen, bereits auf Grundlage bestehender Steuergesetze als gegeben sehen.[17] Dies würde jedoch nicht nur die Erbringer digitaler Werbeleistungen treffen. Nach der Argumentation dieses Vorstoßes wären im Ergebnis alle Unternehmen, welche solche Werbeleistungen in Deutschland in Anspruch nehmen, zum Einbehalt und zur Abführung dieser Steuer im Namen der Werbeleistungserbringer verpflichtet und würden (auch rückwirkend) für den Steuereinbehalt haftbar gemacht werden. Dies würde auf der Seite deutscher Werbetreibender zu großen finanziellen Belastungen führen, die im Falle einer solchen Verpflichtung teilweise existenzbedrohend hohe

---

[2]Vgl. Ittermann & Niehaus, 2015, S. 33-35.

[3]Vgl. OECD, 2015, Tz. 116.

[4]Vgl. Alphabet Inc., 2019, S. 27.

[5]Vgl. Facebook Inc., 2019, S. 33, 42.

[6]Vgl. Paperlein, 2018.

[7]Vgl. anstatt vieler z. B. Berschens, 2017; Dams & Heuzeroth, 2017.

[8]Vgl. Google Germany GmbH, 2019.

[9]Vgl. Facebook Germany GmbH, 2018; die Werte für das Jahr 2018 liegen derzeit noch nicht vor.

[10]Vgl. z. B. Devereux & Vella, 2014; Olbert & Spengel, 2017.

[11]Vgl. OECD, 2014, S. 17.

[12]Vgl. OECD, 2014, S. 17. Insgesamt besteht das BEPS Projekt auf 15 verschiedenen Aktionspunkten.

[13]Vgl. OECD, 2015, Tz. 383; auch in einem weiteren Zwischenbericht finden sich keine solchen Handlungsempfehlungen OECD, 2018.

[14]Vgl. für Frankreich Bunn, 2019; Französisches Ministerium für Wirtschaft und Finanzen, 2019; für Großbritannien British Broadcasting Corporation (BBC), 2018; Frankfurter Allgemeine Zeitung, 2018; für Italien Ernst and Young (EY), 2019b; für Österreich Ernst and Young (EY), 2019a; für Spanien Ernst and Young (EY), 2019c; für die Türkei Ernst and Young (EY), 2018.

[15]Vgl. Le Maire, Schäuble, Padoan & De Guindos, 2017; Rat der Europäischen Union, 2017.

[16]Vgl. Europäische Kommission, 2018b.

[17]Von diesen Vorstößen berichten u. a. Schlotter & Hruschka, 2018, S. 711-712; Diffring, 2019; Diffring & Saft, 2019, S. 387; Linn, 2019, S. 419; Hoke, 2019; KPMG, 2019; Heil & Pupeter, 2019, S. 604-605.

Steuerbeträge an die Finanzverwaltung abzuführen hätten.[18] Wäre dieser Vorstoß jedoch gesetzeskonform, so bestünde aus deutscher Sicht zumindest im Hinblick auf die Besteuerung digitaler Werbeleistungen überhaupt keine Notwendigkeit zur Implementierung weiterer steuerlicher Maßnahmen. Allerdings besteht derzeit große Unsicherheit, inwiefern diesem Vorstoß der Finanzverwaltung zu folgen ist.

Vor dem Hintergrund der auf nationaler und internationaler Ebene kontrovers geführten Diskussion stellt sich daher die Frage, wie deutsche Besteuerungsrechte an digitalen Werbeleistungen sichergestellt werden können, welche von im Ausland ansässigen Körperschaften mithilfe des Internets im Inland erbracht werden. Die Beantwortung dieser Frage ist das Ziel der vorliegenden Arbeit. Der Fokus liegt dabei alleinig auf der direkten Unternehmensbesteuerung in Form der Körperschaftsteuer.[19] Hierzu soll zunächst geklärt werden, ob überhaupt Handlungsbedarf zur Sicherstellung dieser Besteuerungsrechte besteht oder ob die bestehenden deutschen Steuergesetze bereits in der Lage sind, grenzüberschreitende digitale Werbeleistungen zu erfassen. Im Anschluss daran soll in Ermangelung internationaler Handlungsempfehlungen auf Ebene der OECD geprüft werden, ob die von der EU geplanten Maßnahmen geeignet wären, die Besteuerung digitaler Werbeleistungen in Deutschland sicherzustellen.

Die Arbeit trägt dazu bei, die bestehende Unsicherheit bezüglich der Besteuerung digitaler Werbeleistungen nach derzeit geltendem deutschen Steuerrecht zu reduzieren. Sie soll zur Klärung steuerlicher Pflichten sowohl aufseiten der Erbringer digitaler Werbeleistungen als auch aufseiten von Unternehmen beitragen, die diese Werbeleistungen in Anspruch nehmen. Damit sind die Ergebnisse dieser Arbeit für Anbieter und Nachfrager digitaler Werbeleistungen gleichermaßen relevant. Weiterhin wird zur Diskussion möglicher Handlungsalternativen zur Sicherstellung von Besteuerungsrechten an digitalen Geschäftsmodellen beigetragen. Damit können die Ergebnisse der Arbeit Entscheidungsträger bei der Findung einer (langfristigen) Lösung der Besteuerungsproblematik unterstützen.

Der Aufbau der Arbeit gestaltet sich folgendermaßen: Kapitel 2 befasst sich zunächst mit digitalen Werbeleistungen im Allgemeinen sowie den Grundzügen des derzeitigen internationalen Steuerrechts und stellt die grundlegenden Probleme dar, welche digitale Werbeleistungen im internationalen Steuerrecht verursachen. Daran anschließend werden in Kapitel 3 die in Deutschland aktuell bestehenden nationalen Anknüpfungspunkte im Rahmen der beschränkten Steuerpflicht vorgestellt, welche für die steuerliche Erfassung digitaler Werbeleistungen grundsätzlich in Frage kommen könnten. Weiterhin wird auf die Arten der Steuererhebung bei beschränkter Steuerpflicht eingegangen. Kapitel 4 analysiert, inwiefern eventuell bestehende deutsche Steuerrechte an digitalen Werbeleistungen durch bestehende Doppelbesteuerungsabkommen eingeschränkt werden. Im Anschluss wird

in Kapitel 5 am Beispiel von Suchmaschinenwerbung des Unternehmens Google geprüft, ob die bestehenden steuerlichen Rahmenbedingungen tatsächlich in der Lage sind, deutsche Besteuerungsrechte an digitalen Werbeleistungen sicherzustellen, welche von im Ausland ansässigen Körperschaften im Inland erbracht werden. Kapitel 6 zieht ein Zwischenfazit über den Status quo der Besteuerung digitaler Werbeleistungen in Deutschland und schließt damit, ob vor diesem Hintergrund auch in Deutschland Handlungsbedarf besteht. Abschließend wird in Kapitel 7 geprüft, ob die derzeit diskutierten Handlungsalternativen auf Ebene der EU dazu geeignet wären, die Besteuerung digitaler Werbeleistungen in Zukunft sicherzustellen. Kapitel 8 zieht ein abschließendes Fazit.

Die Arbeit wird zeigen, dass die derzeit bestehenden nationalen Steuergesetze in Deutschland in den meisten Fällen nicht dazu in der Lage sind Besteuerungsrechte an digitalen Werbeleistungen sicherzustellen, welche von im Ausland ansässigen Körperschaften mithilfe des Internets im Inland erbracht werden. Insbesondere bestehen in der Regel keine Verpflichtungen zur Steuererhebung in Form eines Steuerabzugs durch den Vergütungsschuldner und damit auch keine Haftungsrisiken für Unternehmen, die solche digitalen Werbeleistungen in Anspruch nehmen. Auch in Deutschland scheint damit Handlungsbedarf bezüglich der Besteuerung digitaler Werbeleistungen gegeben. Da die auf Ebene der EU diskutierten Maßnahmen mit vielen Problemen verbunden sind, kann deren Implementierung nach hier vertretener Ansicht nicht empfohlen werden, auch wenn sie generell dazu in der Lage wären digitale Werbeleistungen der Besteuerung zu unterwerfen.

2. Digitale Werbeleistungen als Beispiel digitaler Transaktionen im internationalen Steuerrecht

Im Folgenden werden zunächst digitale Werbeleistungen als Beispiel digitaler Transaktionen (Kapitel 2.1) sowie die Grundzüge des internationalen Steuerrechts (Kapitel 2.2) vorgestellt. Im Anschluss befasst sich Kapitel 2.3 mit den grundlegenden Problemen, die sich aufgrund der Eigenheiten digitaler Werbeleistungen bei deren Besteuerung ergeben können.

2.1. Digitale Werbeleistungen als Beispiel digitaler Transaktionen

Mit der voranschreitenden Digitalisierung der Wirtschaft hat sich die Erbringung digitaler Werbeleistungen als Teil vieler neuer Geschäftsmodelle etabliert.[20] Digitale Werbeleistungen unterscheiden sich von traditioneller Werbung[21] darin, dass für die Erbringung der Leistung auf das Internet zurückgegriffen und die Leistung vollständig digital erbracht

---

[18]Vgl. z. B. Haselrieder & Pontzen, 2019.

[19]Der Terminus Steuerpflicht bezieht sich im Folgenden daher, soweit nicht anders kenntlich gemacht, auf die Körperschaftsteuerpflicht.

[20]So zählt z. B. OECD (2015, Rn. 116, 136-139) die Erbringung digitaler Werbeleistungen mit zu den am meisten verbreiteten neuen digitalen Geschäftsmodellen.

[21]Unter Werbung versteht man „die Beeinflussung von verhaltensrelevanten Einstellungen mittels spezifischer Kommunikationsmittel, die über Kommunikationsmedien verbreitet werden"; Bruhn et al., 2013, S. 150.

wird. Sowohl der Vertragsabschluss als auch die Erbringung der Leistung erfolgen digital. Vorteile gegenüber traditioneller Werbeformen ergeben sich insbesondere aus der Möglichkeit, Werbung durch gezieltes Sammeln von (persönlichen) Daten in einem viel höheren Maße personalisiert auf Werbekunden zuzuschneiden.[22] Die Werbeleistung wird dabei i. d. R. nicht von den Werbetreibenden selbst, sondern von Dritten erbracht, die als Intermediäre zwischen Werbetreibenden und Werbekunden fungieren und digitale Werbeflächen bereitstellen.[23] Diese Intermediäre, im Folgenden auch als Erbringer digitaler Werbeleistungen bezeichnet, bedienen dabei oft mehrseitige Märkte.[24] Sie stellen auf der einen Seite den potenziellen Werbekunden regelmäßig selbst kostenlose Dienstleistungen zur Verfügung, z. B. in Form von Suchmaschinen oder Social-Media-Plattformen (z. B. Google oder Facebook). Auf der anderen Seite können sie Werbetreibenden damit eine möglichst große Reichweite ihrer Werbeleistungen kostenpflichtig anbieten.[25]

Es gibt verschiedenste Formen digitaler Werbeleistungen. Die umsatzstärksten Formen in Deutschland waren im Jahr 2018 Suchmaschinen- (ca. 3 Mrd. EUR Umsatz, 42 % Marktanteil), Banner- (ca. 1,6 Mrd. EUR Umsatz, 23 % Marktanteil) und Social-Media-Werbung (ca. 1,4 Mrd. EUR Umsatz, 20 % Marktanteil).[26] Bei Suchmaschinenwerbung bezahlen Werbetreibende dafür, bei bestimmten Suchanfragen in Suchmaschinen priorisiert in der Ergebnisliste aufgeführt zu werden. Im Rahmen von Bannerwerbung wird für digitale Werbeflächen auf Webseiten bezahlt. Durch Social-Media-Werbung werden gezielt Nutzer von Social-Media-Plattformen beworben.

Digitale Werbeleistungen werden sowohl im Rahmen des BEPS-Projekts der OECD als auch in Diskussionen auf Ebene der EU als prominentes Beispiel für digitale Transaktionen genannt, ohne wiederum letzteren Begriff abschließend zu definieren.[27] Digitale Transaktionen umfassen nach Ansicht der OECD den Kauf von (digitalen) Gütern und Dienstleistungen über das Internet,[28] also den Kauf „immaterieller Mittel zur Bedürfnisbefriedigung, die sich mit Hilfe von Informationssystemen entwickeln, vertreiben und anwenden lassen"[29].

Digitale Transaktionen im Allgemeinen und digitale Werbeleistungen im Speziellen können der obigen Definition folgend damit als Teilbereich des elektronischen Geschäftsverkehrs (E-Commerce) angesehen werden. Der E-Commerce umfasst den Kauf und Verkauf von Gütern und Dienstleistungen, (1) bei denen die Auftragserteilung online (z. B. über Computernetzwerke wie das Internet) erfolgt, die Lieferung und Bezahlung der Güter und Dienstleistungen aber (2a)

sowohl online als auch (2b) offline erfolgen kann.[30] Werden sowohl Auftragserteilung als auch Lieferung komplett digitalisiert erbracht, spricht man auch von direktem oder online E-Commerce, während bei indirektem, auch offline E-Commerce genannt, die Lieferung in physischer Form erfolgt.[31] Digitale Transaktionen und somit auch digitale Werbeleistungen können damit dem Teilbereich des online E-Commerce zugeordnet werden.

2.2. Grundzüge des internationalen Steuerrechts

Ein weltweit einheitliches internationales Steuerrecht als solches existiert nicht.[32] Der Begriff umfasst im weiteren Sinne vielmehr alle Normen, welche die Besteuerung von grenzüberschreitenden Sachverhalten regeln. Dazu gehören insbesondere das nationale Steuerrecht von Einzelstaaten sowie Teile des Gewohnheitsrechts der Völker und internationale Steuerabkommen.[33] Das internationale Steuerrecht gestaltet sich somit in verschiedenen Ländern unterschiedlich.[34]

Alle Einzelstaaten sind grundsätzlich bei der Gestaltung ihres Steuersystems nach dem Souveränitätsprinzip selbstbestimmt.[35] Allerdings wird diese Autonomie durch das Gewohnheitsrecht der Völker dahingehend beschränkt, dass Staaten ausländische Sachverhalte nur bei Vorliegen eines hinreichenden steuerlichen Anknüpfungspunktes besteuern dürfen (steuerlicher Nexus oder auch genuine link).[36] Völkerrechtlich akzeptierte steuerliche Anknüpfungspunkte stellen dabei insbesondere das Wohnsitzstaat- und Quellenstaatsprinzip dar,[37] welche ihren Ursprung in der ersten Hälfte des 20. Jahrhunderts haben.[38]

Das Wohnsitzstaatsprinzip[39] knüpft das Besteuerungsrecht an die Ansässigkeit von juristischen Personen.[40] Diese Ansässigkeit bestimmt sich dabei üblicherweise durch juristisch geprägte Anknüpfungsmerkmale wie den Sitz oder Ort der Geschäftsleitung der juristischen Person.[41] Nach dem Wohnsitzstaatsprinzip verfügt also der Staat (im Folgenden Ansässigkeitsstaat), in dem eine juristische Person ansässig ist, über einen hinreichenden steuerlichen Anknüpfungspunkt und damit ein Besteuerungsrecht dem Grunde nach.

---

[22]Vgl. OECD, 2015, Tz. 136.

[23]Vgl. OECD, 2015, Tz. 137.

[24]Für eine ausführliche Beschreibung digitaler mehrseitiger Märkte siehe OECD (2018, Rn. 48-54).

[25]Vgl. OECD, 2015, Tz. 138.

[26]Vgl. Statista, 2018.

[27]So etwa DST-RL-E, Begründung, S. 9-10; OECD, 2018, Rn. 359.

[28]Vgl. OECD, 2015, Rn. 278, 292, 294.

[29]Stelzer, 2004, S. 235.

[30]Vgl. OECD, 2011, S. 72.

[31]Vgl. OECD, 2015, Tz. 117.

[32]Vgl. Miller & Oats, 2016, S. 25; Korts, 2016, S. 11.

[33]Vgl. Miller & Oats, 2016, S. 26-27; Brähler, 2014, S. 1-2; Korts, 2016, S. 11; für eine Übersicht der relevanten Normen aus deutscher Sicht siehe Schmidt, Sigloch & Henselmann, 2005, 55-63.

[34]Vgl. Korts, 2016, S. 11.

[35]Vgl. Brähler, 2014, S. 3.

[36]Vgl. Brähler, 2014, 3; Schmidt et al., 2005, 61; Loschelder, 2018a, § 49 EStG, Rz. 12.

[37]Vgl. Miller & Oats, 2016, S. 26-27; Prussak, 2013, S. 55-56.

[38]Bruins, Einaudi, Seligman & Stamp, 1923, S. 22-26; Prussak, 2013, S. 55-56.

[39]Auch als Wohnsitzprinzip, Ansässigkeitsprinzip oder residence principle bekannt.

[40]Vgl. Brähler, 2014, S. 3-4. Das internationale Steuerrecht ist grundsätzlich auch auf natürliche Personen anwendbar. Der Fokus soll im Folgenden jedoch der Forschungsfrage folgend auf Körperschaften und damit juristischen Personen liegen.

[41]Vgl. Lehner & Reimer, 2005, S. 542; Brähler, 2014, S. 4; OECD, 2015, Tz. 22.

Dem Umfang nach folgt die Besteuerung dem Welteinkommensprinzip[42], d. h., das gesamte in- und ausländische Einkommen des Steuerpflichtigen unterliegt der Besteuerung im Ansässigkeitsstaat.[43]

Das Quellenstaatsprinzip[44] knüpft hingegen an inländischen Einkunftsquellen an. Ein Staat (im folgenden Quellenstaat) verfügt nach dem Quellenstaatsprinzip über einen hinreichenden steuerlichen Anknüpfungspunkt und somit ein Besteuerungsrecht dem Grunde nach, wenn Einkünfte einer juristischen Person innerhalb seines Territoriums anfallen; auf eine persönliche Verbindung der Person zum Staat kommt es nicht an.[45] Obwohl die Definition, wann Einkünfte als im Inland belegen gelten, zwischen einzelnen Staaten variiert, ist das Prinzip der Betriebsstätte im Rahmen von Unternehmensgewinnen weithin akzeptiert.[46] Das Vorhandensein einer Betriebsstätte setzt dabei für gewöhnlich ein Mindestmaß an physischer Präsenz des nicht ansässigen Unternehmens im Quellenstaat voraus.[47] Der Umfang der Besteuerung richtet sich nach dem Territorialitätsprinzip[48].[49] Der Quellenstaat kann nur Sachverhalte des Steuerpflichtigen der Besteuerung unterwerfen, die innerhalb seines Staatsgebietes verwirklicht werden.[50]

Regelmäßig wenden Staaten sowohl das Wohnsitzstaats- als auch das Quellenstaatsprinzip parallel zur Begründung steuerlicher Anknüpfungspunkte an. Dies kann zu juristischer Doppelbesteuerung führen, wenn mehr als ein Staat Steueransprüche auf dasselbe Einkommen eines Steuerpflichtigen erhebt.[51] Um negative Effekte aus einer möglichen Doppelbesteuerung zu vermeiden bzw. zu verringern, verfügen viele Staaten über ein Netz von bilateralen Steuerabkommen (Doppelbesteuerungsabkommen, kurz DBA). Diese Abkommen haben zum Ziel, die Besteuerungsrechte zwischen den Abkommensstaaten zu verteilen und Doppelbesteuerung zu vermeiden.[52] Die meisten Doppelbesteuerungsabkommen basieren dabei auf dem Musterabkommen der OECD oder in besonderen Fällen auf dem Musterabkommen der UN.[53] Diese Musterabkommen folgen dabei dem Grundsatz der Gewinnbesteuerung am Ort der Wertschöpfung, d. h., Gewinne sollen dort der Besteuerung unterliegen, wo die Wertschöpfung dieser Gewinne stattfindet.[54] Weiterhin existieren auf Ebene von Einzelstaaten unilaterale, sowie vereinzelt multilaterale und supranationale Maßnahmen zur Vermeidung von Doppelbesteuerung.[55]

2.3. Problemfelder bei der Besteuerung digitaler Werbeleistungen

Bei der Erbringung digitaler Werbeleistungen handelt es sich regelmäßig um stark digitalisierte Geschäftsmodelle.[56] Als solche teilen sie oftmals folgende Eigenschaften, welche eine angemessene steuerliche Erfassung durch das derzeit bestehende internationale Steuersystem erschweren:[57] (1) grenzüberschreitende Ausdehnung ohne Masse, (2) hohe Bedeutung von Daten und Nutzern und (3) Abhängigkeit von immateriellen Werten wie z. B. geistigem Eigentum.

Zunächst ist festzuhalten, dass es den Erbringern digitaler Werbeleistungen durch den Einsatz moderner Informations- und Kommunikationstechnologien (IKT)[58] wie z. B. dem Internet stark erleichtert wird grenzüberschreitend in verschiedenen Absatzmärkten wirtschaftlich aktiv zu sein, ohne dafür auf eine physische Präsenz in diesen Absatzmärkten angewiesen zu sein (Ausdehnung ohne Masse).[59] Jedoch stellt insbesondere das Quellenstaatsprinzip zur Begründung eines steuerlichen Anknüpfungspunktes auf eine physische Präsenz nicht ansässiger Unternehmen ab. Dadurch besteht die Gefahr, dass eine Besteuerung nicht ansässiger Erbringer digitaler Werbeleistungen in verschiedenen Quellenstaaten häufig ins Leere läuft, da es schon an einem steuerlichen Anknüpfungspunkt mangelt.

Ein weiteres Problem für eine angemessene steuerliche Erfassung digitaler Werbeleistungen stellt die hohe Bedeutung von Daten und Nutzern dar. Zum einen sammeln und nutzen Erbringer digitaler Werbeleistungen oftmals umfassende Datensätze potenzieller Werbekunden, um Werbeleistungen personalisiert erbringen zu können.[60] Zum anderen spielen Nutzer und deren Aktivitäten gerade für Erbringer digitaler Werbeleistungen, die in mehrseitigen Märkten aktiv sind, aufgrund von Netzwerkeffekten eine sehr große Rolle.[61] Je größer z. B. die (aktive) Nutzerbasis auf der Marktseite der potenziellen Werbekunden ist, desto größer ist der

---

[42]Auch Universalitätsprinzip genannt.

[43]Vgl. Brähler, 2014, S. 5-6; Miller & Oats, 2016, S. 27-28; Lehner & Reimer, 2005, S. 542; Lübbehüsen & Kahle, 2016, Rz. 1.9; Schreiber, 2017, S. 445.

[44]Auch als Belegenheitsprinzip oder als territorial oder source principle bekannt.

[45]Vgl. Brähler, 2014, S. 4; Lehner & Reimer, 2005, S. 542.

[46]Vgl. OECD, 2015, Tz. 26; Lübbehüsen & Kahle, 2016, Rz. 1.14-1.26.

[47]Vgl. OECD, 2018, Tz. 378.

[48]Das Territorialitätsprinzip ist auch als Ursprungsprinzip bekannt.

[49]Vgl. Brähler, 2014, S. 4; Lehner & Reimer, 2005, S. 542.

[50]Vgl. Brähler, 2014, S. 5-6; Lehner & Reimer, 2005, S. 542; Lübbehüsen & Kahle, 2016, Rz. 1.9-1.10.

[51]Vgl. Schreiber, 2017, S. 445; Prussak, 2013, S. 56-57; Miller & Oats, 2016, S. 27-28.

[52]Vgl. OECD, 2015, Tz. 27; Miller & Oats, 2016, S. 27-28.

[53]Vgl. OECD, 2015, S. 24.

[54]Vgl. OECD, 2018, Tz. 378; auch unterstützt von der Gruppe der Zwanzig

(G20); vgl. G20, 2013, S. 12-13. Bislang fehlt aber eine allgemeingültige Definition des Begriffs, vgl. Olbert & Spengel, 2017, S. 21. Für Definitionsversuche vgl. z. B. Bauer, Fritz, Schanz & Sixt, 2019, S. 6-9; Becker & Englisch, 2018, S. 1-9.

[55]Vgl. Brähler, 2014, S. 20-23.

[56]OECD (2015, Tz. 116, 136-139) ordnet digitale Werbeleistungen auch der sog. „digitalen Wirtschaft" zu. Da sich jedoch die gesamte Wirtschaft zunehmend digitalisiert und somit die digitale Wirtschaft zur Wirtschaft an sich wird (vgl. OECD, 2015, S. 11, Tz. 115; Schön, 2018, S. 280-281), erscheint eine Einteilung in digitale und analoge Wirtschaft nicht sachgerecht. Im Folgenden wird daher nur zwischen mehr und weniger digitalisierten Geschäftsmodellen unterschieden.

[57]Vgl. OECD, 2018, Tz. 130-161.

[58]Für eine Übersicht über die Entwicklung von IKT und deren Auswirkung auf Geschäftsmodelle siehe z. B. OECD, 2015, Tz. 64-108; Becker, Englisch & Schanz, 2018, S. 5-7; Bauer et al., 2019.

[59]Vgl. OECD, 2018, Tz. 131; für digitalisierte Geschäftsmodelle im Allgemeinen OECD, 2015, Tz. 246; Lück, 2018, S. 19-21.

[60]Vgl. OECD, 2018, Tz. 140-141.

[61]Vgl. OECD, 2018, Tz. 143-161 (zur Nutzerbeteiligung), Tz. 42 (zu Netzwerkeffekten).

Nutzen für die Marktseite der potenziellen Werbetreibenden, da diese ihre Werbebotschaften an mehr potenzielle Kunden richten können.[62] Je größer der Nutzen für die Werbetreibenden ist, desto höher sollte im Ergebnis auch deren Zahlungsbereitschaft für die Inanspruchnahme digitaler Werbeleistungen ausfallen. Die Bedeutung von Daten und Nutzern stellt das internationale Steuersystem also insbesondere vor die Herausforderung zu beurteilen, ob und in welchem Umfang Daten und Nutzer zur Wertschöpfung beitragen und inwiefern die Gewinne aus dieser potenziellen Wertschöpfung dem Ansässigkeits- oder Quellenstaat zuzurechnen sind.[63]

Des Weiteren stellt die Tatsache, dass digitale Geschäftsmodelle oftmals stark von immateriellen Werten wie z. B. geistigem Eigentum abhängig sind,[64] das derzeitige Steuersystem vor große Herausforderungen. Immaterielle Vermögenswerte haben mittlerweile großen Einfluss auf den Wert von Unternehmen.[65] Gleichzeitig stellt die strategische Verlagerung von immateriellem Vermögen für multinationale Unternehmen eine attraktive Möglichkeit dar, Gewinne für steuerliche Zwecke zu verlagern und empirische Untersuchungen zeigen, dass die Platzierung immaterieller Werte teilweise steuermotiviert erfolgt.[66] Dabei bleibt jedoch festzuhalten, dass die Steuervermeidung durch Gewinnverlagerung kein alleiniges Merkmal digitaler Geschäftsmodelle ist, sondern auch in „traditionellen" Geschäftsmodellen Anwendung findet.[67] Aus diesem Grund liegt der Fokus im Folgenden schwerpunktmäßig auf den beiden erstgenannten Problemen.

Daher ist im Ergebnis fraglich, inwieweit das derzeit bestehende internationale Steuerrecht überhaupt in der Lage ist, grenzüberschreitende digitale Werbeleistungen sachgerecht zu erfassen und der Besteuerung zu unterwerfen.

3. Beschränkte Steuerpflicht digitaler Werbeleistungen in Deutschland

Im Folgenden werden die bestehenden nationalen Anknüpfungspunkte vorgestellt, welche für die steuerliche Erfassung digitaler Werbeleistungen in Frage kommen könnten. Im deutschen Steuerrecht wird grundsätzlich zwischen unbeschränkter (dem Wohnsitzstaatsprinzip folgend) und beschränkter Steuerpflicht (dem Quellenstaatsprinzip fol-

gend) unterschieden.[68] Da Erbringer digitaler Werbeleistungen in Deutschland überwiegend im Ausland ansässige Körperschaften sind,[69] beschränken sich die nachfolgenden Ausführungen jedoch auf die Besteuerung von juristischen Personen in Inboundfällen, d. h. auf Fälle, in denen nicht ansässige juristische Personen digitale Werbeleistungen in Deutschland erbringen. Die Besteuerung von Inboundfällen erfolgt in Deutschland dabei im Rahmen der beschränkten Steuerpflicht. Im Folgenden wird zunächst die beschränkte von der unbeschränkten Steuerpflicht abgegrenzt (Kapitel 3.1), bevor im Anschluss die für digitale Werbeleistungen in Frage kommenden beschränkt steuerpflichtigen Einkünfte vorgestellt werden (Kapitel 3.2). Abschließend wird die Steuererhebung bei beschränkter Steuerpflicht beschrieben (Kapitel 3.3).

3.1. Abgrenzung der beschränkten von der unbeschränkten Steuerpflicht

Juristische Personen mit Geschäftsleitung (§ 10 AO) oder Sitz (§ 11 AO) im Inland gelten als unbeschränkt körperschaftsteuerpflichtig.[70] Unbeschränkt Steuerpflichtige werden gem. § 1 Abs. 2 KStG mit ihrem Welteinkommen besteuert.[71] Dem steht die beschränkte Steuerpflicht gegenüber. Juristische Personen (Körperschaften, Personenvereinigungen und Vermögensmassen) mit inländischen Einkünften sind gem. § 2 Nr. 1 KStG beschränkt körperschaftsteuerpflichtig, wenn sie weder Geschäftsleitung noch Sitz im Inland haben.[72] Beschränkt Steuerpflichtige sind im Sinne des Territorialitätsprinzip nur mit ihren inländischen Einkünften in Deutschland steuerpflichtig (§ 2 Nr. 1 KStG). § 49 EStG bestimmt abschließend den Katalog inländischer Einkünfte.[73]

3.2. Beschränkt steuerpflichtige Einkünfte

Der Einkünftekatalog des § 49 EStG knüpft an die Einkunftsarten des Einkommensteuergesetzes in § 2 Abs. 1 EStG an, sodass die Tatbestandsvoraussetzungen der §§ 13 bis 23 EStG für das Vorliegen inländischer Einkünfte erfüllt sein müssen (wie in Kapitel 1 grafisch veranschaulicht). Allerdings werden diese Tatbestandsvoraussetzungen um weitere in § 49 Abs. 1 EStG formulierte Tatbestandsmerkmale erweitert, um eine hinreichende Verbindung zum Inland herzustellen.[74] Eine weitere Besonderheit besteht in der sogenannten

---

[62]Dabei handelt es sich um indirekte Netzwerkeffekte im Kontext mehrseitiger Märkte. Bei diesen zieht eine Marktseite einen positiven Nutzen aus der Größe und den Aktivitäten der anderen Marktseite; vgl. OECD, 2018, Tz. 42.

[63]Vgl. OECD, 2018, Tz. 381.

[64]Vgl. OECD, 2018, Tz. 130, 135-138.

[65]Beispielsweise waren im Jahr 1975 immaterielle Werte für 17 % des Marktwerts von S&P 500 Unternehmen verantwortlich, im Jahr 2015 steigerte sich dieser Anteil auf 84 %. In Unternehmen des S&P 500 Europe war der Anteil immaterieller Werte im Jahr 2015 mit 71 % vergleichbar hoch; vgl. Elsten & Hill, 2017, S. 245-246.

[66]Vgl. etwa Dischinger & Riedel, 2011; Karkinsky & Riedel, 2012; Griffith, Miller & O'Connell, 2014; Skeie, Johansson, Menon & Sorbe, 2017.

[67]Vgl. OECD, 2015, Tz. 181; OECD, 2018, Tz. 257.

[68]Vgl. C. Kraft, 2018, § 49 EStG, Rn. 1; Stöber, 2018, § 49 EStG, Rn. 1, 16-17; Gosch, 2018a, § 49 EStG, Rn. 1; Schaumburg, 2017, Rn. 6.126-6.128; Brähler, 2014, 6.

[69]Vgl. Pinkernell, 2014, S. 24.

[70]Vgl. § 1 Abs. 1 KStG; Benecke, 2018, § 1 KStG, Rn. 2.

[71]Vgl. C. Kraft, 2018, § 49 EStG, Rn. 1; Benecke, 2018, § 1 KStG, Rn. 2.

[72]Neben den „Steuerausländern" des § 2 Nr. 1 KStG sind gem. § 2 Nr. 2 KStG auch bestimmte „sonstige Steuerinländer" beschränkt körperschaftsteuerpflichtig; vgl. Mohr, 2018, § 2 KStG, Rn. 2; Rengers, 2018, § 2 KStG, Rn. 1, 8. Aufgrund der geringen Relevanz für digitale Werbeleistungen wird im Folgenden auf weitere Ausführungen zu § 2 Nr. 2 KStG verzichtet.

[73]Vgl. § 2 Nr. 1, § 8 Abs. 1 Satz 1 KStG i. V. m. § 49 EStG.

[74]Vgl. Schaumburg, 2017, Rz. 6.150-6.151; Loschelder, 2018a, § 49 EStG, Rn. 11-12; Stöber, 2018, § 49 EStG, Anm. 311-312; Roth, 2018, § 49 EStG, Anm. 100; Rengers, 2018, § 2 KStG, Rz. 30.
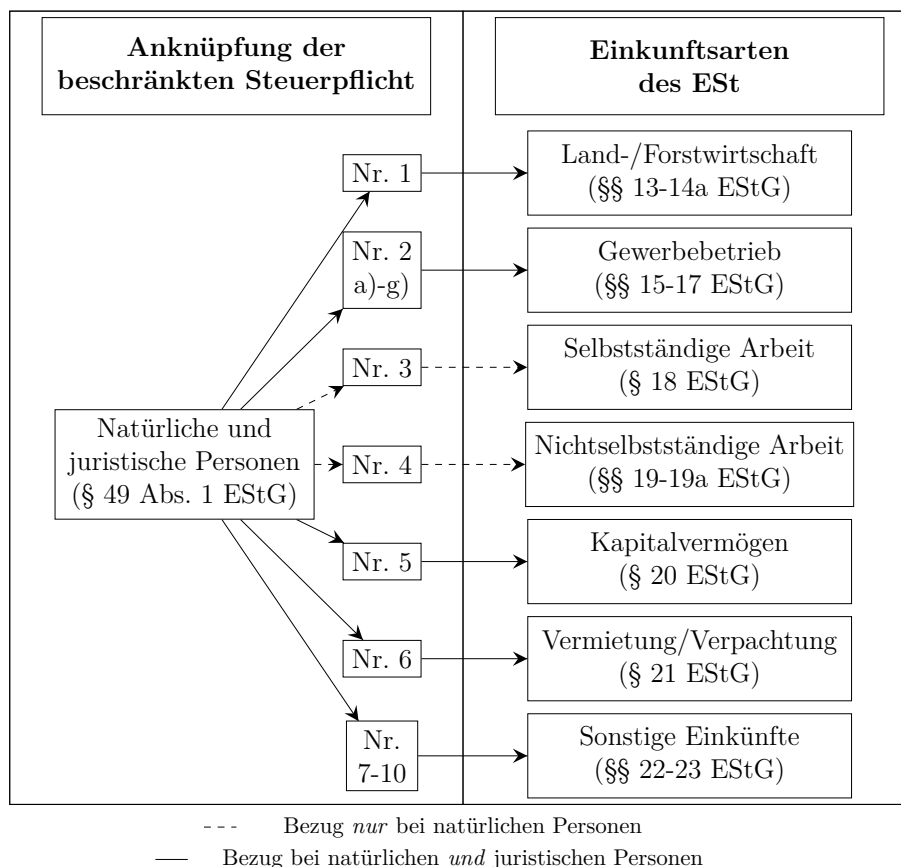
Abbildung 1: Systematik inländischer Einkünfte in Deutschland (eigene Abbildung; in Teilen angelehnt an C. Kraft, 2018, § 49 EStG, Rn. 8)

isolierenden Betrachtungsweise des § 49 Abs. 2 EStG, nach der ausländische Besteuerungsmerkmale ignoriert werden, soweit inländische Einkünfte bei deren Berücksichtigung nicht angenommen werden könnten.[75]

Im Inboundfall erbrachte digitale Werbeleistungen unterliegen in Deutschland daher nur der Besteuerung, sofern sie inländische Einkünfte i. S. d. § 49 EStG darstellen. Da die Norm jedoch keine eigenen Tatbestände für digitale Werbeleistungen enthält, müssen diese Geschäftsmodelle unter die bestehenden Tatbestände subsumiert werden.[76] Aufgrund der Eigenheiten digitaler Werbeleistungen sind insbesondere die nachfolgend dargestellten Tatbestände des § 49 EStG relevant:

- Gewerbliche Einkünfte durch Betriebsstätte oder ständigen Vertreter,
  § 49 Abs. 1 Nr. 2 Buchst. a EStG (Kapitel 3.2.1)

- Gewerbliche betriebsstättenlose Einkünfte, § 49 Abs. 1 Nr. 2 Buchst. f EStG
  (Kapitel 3.2.2)

- Sonstige Einkünfte, § 49 Abs. 1 Nr. 9 EStG (Kapitel 3.2.3)[77]

3.2.1. Gewerbliche Einkünfte durch Betriebsstätte oder ständigen Vertreter

Für das Vorliegen von inländischen Einkünften aufgrund einer Betriebsstätte oder eines ständigen Vertreters gem. § 49 Abs. 1 Nr. 2 Buchst. a Alt. 1 und 2 EStG ist zunächst erforderlich, dass es sich bei den infrage kommenden Einkünften um solche aus Gewerbebetrieb gem. § 15 EStG handelt.[78] Diese Voraussetzungen werden von Unternehmen der digitalen Wirtschaft in der Regel erfüllt.[79] Daneben muss der potenziell Steuerpflichtige im Inland entweder eine Betriebsstätte unterhalten (Alt. 1) oder einen ständigen Vertreter bestellt haben (Alt. 2).

---

[75]Vgl. § 49 Abs. 2 EStG; Frotscher, 2015, Rn. 190-191; Schreiber, 2017, S. 484.
[76]Vgl. Pinkernell, 2014, S. 25-26; Lück, 2018, S. 42-43.

[77]An dieser Stelle sei angemerkt, dass grundsätzlich auch Einkünfte aus Vermietung und Verpachtung (§ 49 Abs. 1 Nr. 6 EStG) bei digitalen Werbeleistungen relevant sein könnten. Aufgrund der Subsidiarität der Norm gegenüber § 49 Abs. 1 Nr. 2 Buchst. f EStG i. V. m. der Annahme, bei den Erbringern digitaler Werbeleistungen handle es sich um im Ausland ansässige Körperschaften und der Gewerblichkeitsfiktion des § 49 Abs. 1 Nr. 2 Buchst. f EStG (siehe Kapitel 3.2.2), ist § 49 Abs. 1 Nr. 6 EStG i. d. R. nicht einschlägig und wird daher im Folgenden nicht weiter betrachtet.
[78]Vgl. Kaligin, 2018, § 49 EStG, Anm. 703-705; Roth, 2018, § 49 EStG, Anm. 144.
[79]Vgl. Lück, 2018, S. 44; Pinkernell, 2014, 27-28.

Nach nationalem Recht ist gem. § 12 Satz 1 AO jede feste Geschäftseinrichtung oder Anlage eine Betriebsstätte, wenn sie der Tätigkeit des Unternehmens dient. Eine Betriebsstätte erfordert damit (1) eine Geschäftseinrichtung[80] mit fester Beziehung zur Erdoberfläche, die von gewisser Dauer ist, (2) der Tätigkeit des Unternehmens dient (3) und sich nicht nur vorübergehend in der Verfügungsmacht des Unternehmens befindet.[81] Die Anwesenheit von Personal innerhalb einer Betriebsstätte ist nicht erforderlich.[82] § 12 Satz 2 Nr. 1 bis 8 AO liefert eine Auflistung von Tatbeständen, die eine Betriebsstätte begründen, welche jedoch nicht abschließend ist.[83]

Im Rahmen digitaler Werbeleistungen kommen insbesondere folgende Sachverhalte zur Begründung einer Betriebsstätte in Betracht:[84]

- IT-Infrastruktur, insbesondere in Form von Internetservern

- Personal für Supportleistungen

- Webseite des Unternehmens

- Digitale Werbeflächen

Unter einem Internetserver versteht man im allgemeinen Sprachgebrauch einen Computer (sog. Host, engl. für Gastgeber) zur Speicherung von Anwendungsprogrammen und Daten, der innerhalb eines Netzwerks gegenüber anderen Computern (sog. Clients, engl. für Kunden) Leistungen erbringt bzw. Ressourcen zur Verfügung stellt.[85] Auch bei Internetservern muss es sich zunächst um eine feste Geschäftseinrichtung handeln, d. h., einen körperlichen Gegenstand, welcher der Tätigkeit des Unternehmens dient. Virtuelle (nicht körperliche) Server kommen daher nicht als Betriebsstätte in Betracht.[86] Bezüglich der erforderten Verfügungsmacht muss zwischen eigenen Servern des potenziell steuerpflichtigen ausländischen Unternehmens und Servern von Drittanbietern unterschieden werden.[87] Unterhält das besagte Unternehmen eigene physische Server im Inland, stehen diese in dessen Verfügungsmacht und begründen da-

mit eine Betriebsstätte.[88] Werden allerdings Server verwendet, die im Eigentum Dritter (sog. Webhosting-Unternehmen oder auch Internet-Service-Provider) stehen und dem ausländischen Unternehmen nur zur Nutzung überlassen werden, ist die Verfügungsmacht des ausländischen Unternehmens n. h. M. nur zu bejahen, sofern es sich um einen dedizierten Server handelt, auf den das Unternehmen direkten Zugriff hat.[89]

Inländisches Personal für Supportleistungen kann ebenfalls eine Betriebsstätte nach § 12 Satz 2 Nr. 2 AO darstellen, wenn es sich um eine juristisch unselbstständige Zweigniederlassung oder Geschäftsstelle des ausländischen Unternehmens handelt.[90] Die Webseite eines Unternehmens kann aufgrund fehlender Körperlichkeit, mangelnder Lokalisierbarkeit und fehlender Dauerhaftigkeit keine Betriebsstätte begründen.[91] Auch digitale Werbeflächen scheiden aufgrund mangelnder Körperlichkeit als Betriebsstätte aus.[92]

Neben der Betriebsstätte stellt der ständige Vertreter die weitere (subsidiäre)[93] Anknüpfungsmöglichkeit des § 49 Abs. 1 Nr. 2 Buchst. a EStG dar. Ein ständiger Vertreter gem. § 13 AO kann dabei jede natürliche oder juristische Person sein,[94] die „nachhaltig die Geschäfte eines Unternehmens besorgt und dabei dessen Sachanweisungen unterliegt"[95].

Auch bei der Frage nach ständigen Vertretern im Inland kommen hauptsächlich die oben bereits genannten Sachverhalte im Rahmen digitaler Werbeleistungen (mit Ausnahme digitaler Werbeflächen) in Betracht.[96] Sowohl eigene Internetserver als auch die Webseite des ausländischen Unternehmens scheiden jedoch als ständige Vertreter aus, da es hier schon an der Personeneigenschaft fehlt.[97] Auch inländische Webhosting-Unternehmen/Internet-Service-Provider stellen i. d. R. keine ständigen Vertreter des ausländischen Unternehmens dar, da es regelmäßig an Weisungsgebundenheit mangelt.[98] Ob inländisches Personal (z. B. für Kundensupport) des ausländischen Unternehmens, welches nicht bereits eine Betriebsstätte gem. § 12 AO begründet, als ständiger Vertreter gem. § 13 AO in Betracht kommt, ist umstritten. Eine in der Literatur vertretene Meinung setzt dafür voraus, dass der Vertreter die Geschäfte des ausländischen Unternehmens

---

[80]Als Geschäftseinrichtung kann dabei jeder körperliche Gegenstand bzw. jede Zusammenfassung körperlicher Gegenstände dienen; vgl. BFH-Urteil vom 03.02.1993 – I R 80-81/91, Tz. 46.

[81]Vgl. Drüen, 2017a, § 12 AO, Tz. 1-20; Schaumburg, 2017, Rz. 6.166-6.167; Reimer, 2018a, § 49 EStG, Rz. 89.

[82]Vgl. BFH-Urteil vom 30.10.1996 – II R 12/92, Tz. 2.

[83]Vgl. Drüen, 2017a, § 12 AO, Tz. 23.

[84]Vgl. Pinkernell, 2014, S. 27-39; Lück, 2018, S. 44-54; Hidien, 2018, § 49 EStG, Rn. D3245.

[85]Vgl. Schmidl, 2014, S. 232. In der Informatik beschreibt der Begriff Server im Grunde nur die Software, welche mit den Clients kommuniziert. Der physische Computer, auf dem dieses Programm läuft, wird als Host bezeichnet; vgl. Tappe, 2011, S. 871. Im Folgenden wird aus Gründen der Konsistenz mit der bestehenden Literatur als Server aber sowohl die Hard- als auch Softwarekomponente verstanden.

[86]Vgl. Leisner-Egensperger, 2014, S. 301; Hidien, 2018, § 49 EStG, Rn. D3281; Thiele & König, 2018, S. 378.

[87]Vgl. Lück, 2018, S. 46-47; Pinkernell, 2014, S. 30-35; Ehrmann & von Wallis, 2014, Rn. 130-132.

[88]Vgl. Kahle & Kindich, 2016, Rz. 2.151-2.153; Pinkernell, 2014, S. 30; Hecht & Lampert, 2009, S. 1129.

[89]Vgl. Lück, 2018, 46-47; Pinkernell, 2014, S. 30-31; Schaumburg, 2017, Rz. 6.166; Loschelder, 2018a, § 49 EStG, Rz. 23; Kaligin, 2018, § 49 EStG, Anm. 737; Hutter & Schmidt, 2000, S. 652-653; Kahle & Kindich, 2016, Rz. 2.150; Thiele & König, 2018, S. 379; hierzu ausführlich mit Gegenmeinungen Hidien, 2018, § 49 EStG, Rn. D3245-D3269.

[90]Vgl. Lück, 2018, S. 51-52; Pinkernell, 2014, S. 35-37.

[91]Vgl. Lück, 2018, S. 50; Kaligin, 2018, § 49 EStG, Anm. 732; Roth, 2018, § 49 EStG, Anm. 191; Hidien, 2018, § 49 EStG, Rn. D 3283; Ehrmann & von Wallis, 2014, Rn. 134.

[92]Vgl. Lück, 2018, S. 53.

[93]Vgl. Reimer, 2018a, § 49 EStG, Rz. 97; Roth, 2018, § 49 EStG, Anm. 230.

[94]Vgl. Drüen, 2017b, § 13 AO, Rz. 2.

[95]Vgl. § 13 Satz 1 AO.

[96]Vgl. Hidien, 2018, § 49 EStG, Rn. D3318; Pinkernell, 2014, S. 41; Lück, 2018, S. 55.

[97]Vgl. Hidien, 2018, § 49 EStG, Rn. D3318-D3325; Lück, 2018, S. 55-56.

[98]Vgl. Hidien, 2018, § 49 EStG, Rn. D3321; Lück, 2018, S. 56-59.

nach außen im Inland besorgt.[99]

3.2.2. Gewerbliche betriebsstättenlose Einkünfte

Neben der Anknüpfung an eine Betriebsstätte oder einen ständigen Vertreter enthält § 49 EStG weitere Sachverhalte, die inländische Einkünfte aus Gewerbebetrieb darstellen und somit zu beschränkter Körperschaftsteuerpflicht führen können. Von besonderer Bedeutung für digitale Werbeleistungen ist dabei gem. § 49 Abs. 1 Nr. 2 Buchst. f EStG die Vermietung und Verpachtung[100] oder Veräußerung von Rechten, die entweder im Inland belegen sind (Belegenheitstatbestand) oder in einer inländischen Betriebsstätte verwertet werden (Verwertungstatbestand) (vgl. Kapitel 2).[101] Dadurch werden auch ohne Anknüpfung an eine Betriebsstätte bestimmte Direktgeschäfte der nationalen Besteuerung unterworfen.[102] Der Tatbestand des Buchst. f der Norm ist dabei jedoch gegenüber der Begründung inländischer Einkünfte auf Grundlage von Buchst. a subsidiär.[103] Das Vorliegen eines Gewerbebetriebs wird für ausländische Kapitalgesellschaften fingiert,[104] sodass die isolierende Betrachtungsweise nicht greift.

Rechte i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG umfassen insbesondere, aber nicht ausschließlich, schriftstellerische, künstlerische und gewerbliche Urheberrechte, gewerbliche Erfahrungen sowie Gerechtigkeiten und Gefälle und somit den Rechtebegriff des § 21 Abs. 1 Nr. 3 EStG.[105] Der Begriff ist dabei weit auszulegen, sodass auch weitere Rechte wie Nutzungsrechte schuldrechtlicher oder dinglicher Art in Betracht kommen können.[106] Zur Beurteilung ist nicht ausländisches, sondern das deutsche Immaterialgüterrecht maßgebend.[107] Der BGH hat in einem Einzelfall entschieden, dass auch die Einräumung eines Exklusivrechts auf Werbung der Verpachtung eines Rechts gleichgestellt sein kann.[108] Bei digitalen Werbeleistungen ist daher eine Abgrenzung zwischen Nutzungsüberlassung von Rechten und Erbringung von Dienstleistungen vorzunehmen, da letztere nicht als Nutzungseinkünfte gelten.[109] Einkünfte im Rahmen von Werk-

oder Dienstverträgen fallen daher nicht unter § 49 Abs. 1 Nr. 2 Buchst. f EStG.[110] In Fällen von gemischten Verträgen, die sowohl Dienstleistungen als auch Rechteüberlassungen beinhalten, ist eine Aufteilung der Gesamtvergütung vorzunehmen, wenn keine Leistung lediglich eine unselbstständige Nebenleistung der anderen darstellt.[111]

§ 49 Abs. 1 Nr. 2 Buchst. f EStG erfasst seit 1. Januar 2009 sowohl die Vermietung und Verpachtung als auch die Veräußerung von Rechten, sodass eine Unterscheidung in befristete und unbefristete Nutzungsüberlassung von Rechten an dieser Stelle keine Relevanz mehr hat.[112] In beiden Fällen können inländische Einkünfte begründet werden.

Der Inlandsbezug wird durch den geforderten Belegenheitstatbestand bzw. Verwertungstatbestand konkretisiert. Ein Recht erfüllt den Belegenheitstatbestand des § 49 Abs. 1 Nr. 2 Buchst. f Satz 1 EStG, wenn es in ein inländisches öffentliches Buch oder Register eingetragen ist. Als inländische Register kommen dabei insbesondere das Patent- (§ 30 PatG), Design- (§ 19 DesignG), Marken- (§ 41 MarkenG) und Gebrauchsmusterregister (§ 8 GebrMG) in Betracht.[113] Die Voraussetzung ist nur erfüllt, wenn das Recht auch tatsächlich in einem solchen Register eingetragen ist; eine reine Eintragungsfähigkeit reicht nicht aus.[114] Der Inlandsbezug ist des Weiteren erfüllt, wenn das Recht innerhalb einer inländischen Betriebsstätte verwertet wird (Verwertungstatbestand).[115] Die Beschränkung auf die Verwertung in einer Betriebsstätte reduziert die Anwendbarkeit auf Nutzungsüberlassungen von Rechten im Business-to-Business Bereich (B2B), sodass bei Business-to-Consumer Geschäften (B2C) keine inländischen Einkünfte vorliegen.[116] Unter Verwerten versteht die Rechtsprechung ein „Nutzen, Benutzen oder Gebrauchen von Rechten im Rahmen eigener Tätigkeit durch eigenes Tätigwerden"[117], mit dem Ziel einen finanziellen Nutzen daraus zu ziehen.[118]

Mit Hinblick auf die Anwendung der Norm bei digitalen Werbeleistungen ist es also im Ergebnis entscheidend, ob (1) die Werbeleistung eine Rechteüberlassung i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG darstellt und das Recht entwe-

---

[99]Vgl. Pinkernell, 2014, S. 46-47; Lück, 2018, S. 60-61; Hidien, 2018, § 49 EStG, Rn. D3323.

[100]Eingeführt ab dem 1. Januar 2009; vgl. Naujok, 2018, § 49 EStG, Anm. 1409-1411.

[101]Vgl. Pinkernell, 2014, S. 51. Neben bestimmten Rechten umfasst die Norm auch inländisches unbewegliches Vermögen und Sachbegriffe; vgl. § 49 Abs. 1 Nr. 2 Buchst. f Satz 1 EStG. Da diese jedoch für digitale Werbeleistungen nicht relevant sind, fokussieren sich die nachfolgenden Ausführungen ausschließlich auf Rechte.

[102]Vgl. Schaumburg, 2017, Rn. 6.213.

[103]Vgl. § 49 Abs. 1 Nr. 2 Buchst. f Satz 1 EStG.

[104]Vgl. § 49 Abs. 1 Nr. 2 Buchst. f Satz 3 EStG.

[105]Vgl. Reimer, 2018a, § 49 EStG, Rz. 187; C. Kraft, 2018, § 49 EStG, Rn. 91; Schaumburg, 2017, Rn. 6.214; Mohr, 2018, § 2 KStG, Rn 168. Dies gilt im Ergebnis auch ohne den Verweis auf § 49 Abs. 1 Nr. 6 EStG i.V.m. § 21 EStG ab dem 1. Januar 2007 fort; vgl. Reimer, 2018a, § 49 EStG, Rz. 187; Kessler & Wald, 2015, S. 891.

[106]Vgl. Behnes, Nink & Rohde, 2016, S. 282; Kessler & Wald, 2015, S. 891; Petersen, 2013, S. 900; Ackermann, 2016, S. 259.

[107]Vgl. Ackermann, 2016, S. 259; Hecht & Lampert, 2009, S. 1130

[108]Vgl. BGH-Urteil vom 26.01.1994 – XII ZR 93/92, Tz. 15.

[109]Vgl. Pinkernell, 2014, S. 55; T. Rupp, Knies, Ott, Faust & Hüll, 2018, S. 171; von Bredow, Scheinbacher & Groß, 2017; Verband der Automobilin-

dustrie (VDA), 2018, S. 32.

[110]Vgl. T. Rupp et al., 2018, S. 171; Schlotter & Hruschka, 2018, S. 674

[111]Vgl. BFH-Urteil vom 28.01.2004 – I R 79/02, Tz. 18; BFH-Urteil vom 19.12.2007 – I R 19/06, Tz. 48; G. Kraft, 2018a, 875; G. Kraft, 2018b, 939-940.

[112]Vgl. Roth, 2018, § 49 EStG, Anm. 619-620; Kessler & Wald, 2015, S. 891.

[113]Vgl. Goebel, Wehling & Gehrmann, 2018, S. 155; Naujok, 2018, § 49 EStG, Anm. 1425; Hidien, 2018, § 49 EStG, Rn. E639; § 73a Abs. 3 EStDV.

[114]Vgl. Naujok, 2018, § 49 EStG, Anm. 1423.

[115]Vgl. § 49 Abs. 1 Nr. 2 Buchst. f Satz 1 EStG. Weiterhin reicht die Verwertung in einer anderen Einrichtung; vgl. Pinkernell, 2014, S. 54. Da dieser Sonderfall für digitale Werbeleistungen aber i. d. R. nicht relevant ist, wird er im Folgenden nicht weiter behandelt.

[116]Vgl. Lück, 2018, S. 73; Goebel et al., 2018, S. 182-183; Kessler & Wald, 2015, S. 892; Pinkernell, 2014, S. 55.

[117]BFH-Urteil vom 23.05.1973 – I R 163/71, Tz. 17.

[118]Vgl. Goebel et al., 2018, S. 155; Ackermann, 2016, S. 263; Kessler & Wald, 2015, S. 893; BMF-Schreiben vom 27.10.2017 – IV C 5 - S 2300/12/10003:004, Tz. 6.
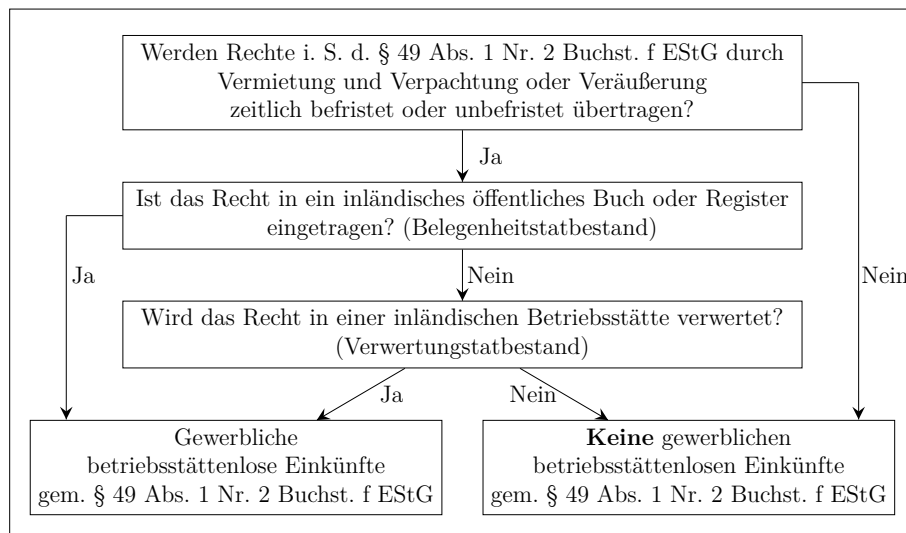
Abbildung 2: Kurzdarstellung der Tatbestandsvoraussetzungen für das Vorliegen von gewerblichen betriebsstättenlosen Einkünften (eigene Abbildung)

der (2a) den Belegenheitstatbestand oder (2b) den Verwertungstatbestand erfüllt. Dem Belegenheitstatbestand kommt dabei regelmäßig keine große Rolle zu, da die in Frage kommenden Rechte oftmals nicht in einem inländischen Register eingetragen sind.[119] Sehr wohl relevant gestaltet sich dagegen der Verwertungstatbestand. Wenn digitale Werbeleistungen den Rechtebegriff erfüllen und diese innerhalb einer inländischen Betriebsstätte des Leistungsschuldners verwertet werden, liegen auch ohne das Vorhandensein einer Betriebsstätte oder eines ständigen Vertreters seitens des Leistungserbringers inländische Einkünfte und damit eine beschränkte Steuerpflicht des Erbringers digitaler Werbeleistungen vor.

Digitale Werbeleistungen werden oftmals unter Verwendung von Software auf rein elektronischem Wege über das Internet erbracht. Sowohl Vertragsabschluss als auch Vertrieb der Leistungen erfolgen digital. Des Weiteren wird insbesondere bei personalisierter Werbung oftmals auf große Datenbestände des Erbringers der Leistung zurückgegriffen, die datenbankähnliche Strukturen aufweisen. Es stellt sich daher weiterhin die Frage, ob diese Tätigkeiten zusätzlich zur eigentlichen Erbringung von Werbeleistungen eine Überlassung von Rechten i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG darstellen und somit inländische Einkünfte des ausländischen Werbeleistungserbringers begründen können. Der Begriff Software bezeichnet Computerprogramme[120] und dazugehörige Daten, welche für den Betrieb von Datenverarbeitungsanlagen (Systemsoftware) bzw. den Gebrauch solcher Datenverarbeitungsanlagen zur Lösung von Problemen

(Anwendersoftware) dienen.[121] Datenbanken sind Sammelwerke von (meist elektronischen) Daten, die in einer gewissen Weise strukturiert sind.[122] Sowohl Software als auch Datenbanken sind urheberrechtlich geschützt.[123] Das Urheberrechtgesetz unterscheidet grundsätzlich zwischen dem Urheberrecht, Verwertungsrecht und Nutzungsrechten, wobei das Urheber- und Verwertungsrecht dem Schöpfer des Werkes zusteht und nicht übertragbar ist.[124] Der Inhaber des Urheber- und Verwertungsrechts an einer Software oder Datenbank kann aber die Verwertung des Urheberrechts an diesen Werken durch das Überlassen von Nutzungsrechten an Dritte (zeitlich begrenzt)[125] übertragen.[126]

Bei der Verwendung von Software und Datenbanken im Rahmen digitaler Werbeleistungen ist also entscheidend, ob dem Nutzer umfassende urheberrechtlich relevante Nutzungsrechte eingeräumt werden und zwar unabhängig von der Form der Überlassung (z. B. via Datenträger oder über das Internet).[127] Darunter fallen nach aktueller Ansicht der Finanzverwaltung „insbesondere Vervielfältigungs-

---

[119]Vgl. Pinkernell, 2014, S. 52-53; Lück, 2018, S. 72-73; Maßbaum, 2017, S. 44.

[120]Ein Computerprogramm ist dabei „eine Zusammenfassung von in einer bestimmten Programmiersprache geschriebenen Anweisungen zur Bewältigung bestimmter Aufgaben oder der Lösung bestimmter Probleme" (Schmidl (2014, S. 58)) und damit nichts Weiteres als ein von einem Computer umgesetzter Algorithmus; Schmidl, 2014, S. 15.

[121]Vgl. Goebel et al., 2018, S. 184-185; Schmidl, 2014, S. 242.

[122]Vgl. Goebel et al., 2018, S. 185; Schmidl, 2014, S. 67.

[123]Vgl. §§ 2 Abs. 1, 69a UrhG bei Software; §§ 4 Abs. 2, 87a UrhG bei Datenbanken; Schulze, 2018, § 2 UrhG, Rn. 126-133; Dreier, 2018, § 4 UrhG, Rn. 1; Bullinger, 2019, § 2 UrhG, Rn. 47 UrhG; Marquardt, 2019, § 4 UrhG, Rn. 8-11 UrhG; Goebel et al., 2018, S. 185; Kessler & Wald, 2015, S. 891; Ackermann, 2016, S. 259; Pinkernell, 2018, S. 140-141.

[124]Vgl. § 7 UrhG bei Urheberrechten, §§ 15 ff. UrhG bei Verwertungsrechten; § 29 Abs. 1 UrhG zur Übertragbarkeit; Hecht & Lampert, 2009, S. 1127-1128; Ackermann, 2016, S. 259.

[125]Bei urheberrechtlich geschützten Werken ist eine vollständige Übertragung ausgeschlossen; vgl. § 29 Abs. 1 UrhG; BFH-Urteil vom 01.12.1982 – I B 11/82, Tz. 18; BMF-Schreiben vom 25.11.2010 – IV C 3 - S 2303/09/10002 2010/0861549, Tz. 23; Holthaus, 2015, S. 28; Holthaus, 2017, S. 732.

[126]Vgl. §§ 29 Abs. 2, 31 ff. UrhG; Wandtke & Grunert, 2019, vor § 31 ff. UrhG, Rn. 24; Hecht & Lampert, 2009, S. 1127-1128.

[127]Vgl. BMF-Schreiben vom 27.10.2017 – IV C 5 - S 2300/12/10003:004, Tz. 3; Loschelder, 2018a, § 49 EStG, Rn. 56; Lück, 2018, S. 73-78; Goebel et al., 2018, S. 186-189; Pinkernell, 2018, S. 140. In der Vergangenheit wurde

, Bearbeitungs-, Verbreitungs- oder Veröffentlichungsrechte"[128]. So stellt z. B. die Gestattung der Fortentwicklung und des Vertriebs einer Software eine Einräumung von Nutzungsrechten dar.[129] Steht hingegen nur der bestimmungsgemäße Gebrauch einer Software im Vordergrund, liegt keine Überlassung von Nutzungsrechten vor. Die Installation und gegebenenfalls kurzfristige Speicherung von Software im Arbeitsspeicher durch den Nutzer sowie der Ablauf von Programmen gelten als bestimmungsgemäßer Gebrauch.[130] Bei Datenbanken ist die Datenbank als solche von der Software zur Darstellung der Datenbank abzugrenzen, da für Letztere die obigen Ausführungen gelten.[131] Auch hier ist bei bestimmungsgemäßem Gebrauch (z. B. bloße „Zugriffs-, Lese- und Druckfunktionen")[132] nicht von einer Nutzungsüberlassung auszugehen. Vereinzelte Meinungen in der Literatur zweifeln die Rechtmäßigkeit dieser von der Finanzverwaltung vertretenen Ansichten zu Software und Datenbanken an, sodass für Rechtssicherheit eine Klarstellung durch die Rechtsprechung erforderlich sein dürfte.[133]

3.2.3. Sonstige Einkünfte

Weiterhin stellt die lückenschließende Auffangvorschrift[134] des § 49 Abs. 1 Nr. 9 EStG einen weiteren möglichen Anknüpfungspunkt für inländische Einkünfte aus digitalen Werbeleistungen dar. Dabei ist zu beachten, dass die Regelung gem. § 49 Abs. 1 Nr. 9 2. HS EStG gegenüber allen anderen inländischen Einkünften des § 49 Abs. 1 Nr. 1 bis 8 EStG subsidiär ist.

Die Vorschrift erfasst u. a. bestimmte Einkünfte i. S. d. § 22 Nr. 3 EStG aus „der Überlassung der Nutzung oder des Rechts auf Nutzung von gewerblichen, technischen, wissenschaftlichen und ähnlichen Erfahrungen, Kenntnissen und Fertigkeiten, zum Beispiel Plänen, Mustern und Verfahren, [...] die im Inland genutzt werden oder worden sind"[135], welche sich unter dem Sammelbegriff Know-how (oder ungeschütztes Wissen) zusammenfassen lassen.[136] Know-how ist dabei „Spezialwissen als Ergebnis erfinderischer Tätigkeit, aber auch [...] Erfahrungswissen, dessen Wert darin besteht, einem Dritten, dem es vermittelt wird, Zeit und Kosten zu ersparen"[137]. Wird das Erfahrungswissen nur vom „Vermittler" selbst angewendet und nicht tatsächlich vermittelt, liegt nach Rechtsprechung des BFH keine Vermittlung von Know-how vor.[138] Gesetzlich geschütztes Know-how fällt nicht unter den Anwendungsbereich des § 49 Abs. 1 Nr. 9 EStG.[139] Die Norm erfasst n. h. M. sowohl befristete als auch unbefristete Überlassung, nicht aber die Veräußerung von Know-how.[140] Um den Inlandsbezug herzustellen, muss gem. § 49 Abs. 1 Nr. 9 EStG das überlassene Know-how auch tatsächlich im Inland genutzt werden, ohne Unterscheidung, ob es sich dabei um B2B- oder B2C-Geschäfte handelt.[141]

Für digitale Werbeleistungen bedeutet dies, dass neben einer Rechteüberlassung i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG auch die Überlassung von Know-how zu prüfen ist. § 49 Abs. 1 Nr. 9 EStG ist dabei insbesondere von Bedeutung, wenn eine Rechteüberlassung andernfalls ausscheiden würde.

3.3. Steuererhebung bei beschränkter Steuerpflicht

Bei beschränkt Steuerpflichtigen gestaltet sich die Erhebung der Körperschaftsteuer oftmals schwierig. Grund hierfür ist, dass sich beschränkt Steuerpflichtige regelmäßig nicht im Inland aufhalten und über keine Vermögensgegenstände im Inland verfügen, was eine Durchsetzung von Besteuerungsansprüchen seitens des deutschen Fiskus erschwert.[142] Dies gilt insbesondere für beschränkt steuerpflichtige Erbringer digitaler Werbeleistungen, sofern sie über keine Betriebsstätte im Inland verfügen. Um bestehende Besteuerungsansprüche dennoch unilateral durchsetzen zu können, erfolgt die Erhebung der Körperschaftsteuer für bestimmte Einkünfte im Rahmen eines Steuerabzugs an der Quelle, d. h., der Vergütungsschuldner hat die Steuer des ausländischen Steuerschuldners (beschränkt Steuerpflichtiger, Vergütungsgläubiger) einzubehalten und an die deutschen Steuerbehörden abzuführen.[143] Im Folgenden erfolgt eine

---

hierfür, basierend auf einer Verfügung der Oberfinanzdirektion OFD München, sowohl in Teilen der Literatur als auch in der Praxis der Finanzbehörden auf die zusätzliche Unterscheidung von Software in Standard- und Individualsoftware zurückgegriffen und im Ergebnis eine Überlassung von Nutzungsrechten nur bei Individualsoftware angenommen; vgl. OFD München-Verfügung vom 28.05.1998 – S 2303 - 34/11 St 41/42; G. Kraft, 2018b, S. 938-939; Goebel et al., 2018, S. 183. Diese Ansicht gilt inzwischen als überholt; vgl. Schewe, 2018, S. 102.

[128]BMF-Schreiben vom 27.10.2017 – IV C 5 - S 2300/12/10003:004, Tz. 3.

[129]Vgl. G. Kraft, 2018b, S. 939.

[130]Vgl. BMF-Schreiben vom 27.10.2017 – IV C 5 - S 2300/12/10003:004, Tz. 4; Behnes et al., 2016, 282.

[131]Vgl. BMF-Schreiben vom 27.10.2017 – IV C 5 - S 2300/12/10003:004, Tz. 33; Kessler & Wald, 2015, S. 892.

[132]Vgl. BMF-Schreiben vom 27.10.2017 – IV C 5 - S 2300/12/10003:004, Tz. 35.

[133]Vgl. Pinkernell, 2018, 141; Maßbaum & Imhof, 2018, S. 11-12. Der Ansicht der Finanzverwaltung kritisch gegenüberstehend z. B. Holthaus (2017, S. 730).

[134]Vgl. Heß, 2018, § 49 EStG, Anm. 2041; Gosch, 2018a, § 49 EStG, Rn. 94.

[135]§ 49 Abs. 1 Nr. 9 EStG.

[136]Auch wenn der Begriff nicht explizit im Gesetzestext genannt ist, ori-

entiert sich die Aufzählung an der Definition von Know-how des BFH; vgl. BFH-Urteil vom 16.12.1970 – I R 44/67; Reimer, 2018a, § 49 EStG, Rz. 311. So auch im Gesetzesentwurf 1974; vgl. Bundestag-Drucksache 07/1509 vom 09.01.1974: Entwurf eines Zweiten Steueränderungsgesetzes 1973 (BT-Drs. 07/1509 vom 09.01.1974); Heß, 2018, § 49 EStG, Anm. 2050.

[137]BFH-Urteil vom 13.11.2002 – I R 90/01, Tz. 8.

[138]Vgl. BFH-Urteil vom 16.12.1970 – I R 44/67, Tz. 26.

[139]Vgl. BFH-Urteil vom 20.07.1988 – I R 174/85, Tz. 18; Loschelder, 2018a, § 49 EStG, Rn. 125; T. Rupp et al., 2018, S. 171; Petersen, 2013, S. 896.

[140]Vgl. Reimer, 2018a, § 49 EStG, Rz. 311; Gosch, 2018a, § 49 EStG, Rn. 94; Klein, 2018, § 49 EStG, Anm. 1111; Schaumburg, 2017, Rn. 6.263; FG München-Urteil vom 14.05.2018 – 7 K 1440/17, Tz. 17. Noch nicht höchstrichterlich entschieden vgl. BFH-Verfahren vom 20.11.2018 – I R 18/18.

[141]Vgl. Lück, 2018, S. 79; Pinkernell, 2014, 59.

[142]Vgl. Goebel et al., 2018, S. 153; Frotscher, 2015, Rn. 203; Gosch, 2018b, § 50a EStG; Rn. 1.

[143]Vgl. Goebel et al., 2018, S. 153. Häufig wird in diesem Zusammenhang auch von einer sog. „Quellensteuer" gesprochen. Dieser abkommensrechtliche Begriff bezeichnete ursprünglich nur die Zuweisung des Besteuerungsrechts an den Quellenstaat, wird inzwischen aber synonym für den Begriff des Steuerabzugs verwendet; vgl. Goebel et al., 2018, 4.

Darstellung der Steuererhebung bei beschränkter Steuerpflicht mit Fokus auf digitale Werbeleistungen.

3.3.1. Im Rahmen des Steuerabzugs

§ 50a EStG enthält die Tatbestände, bei denen die Erhebung der Körperschaftsteuer beschränkt Steuerpflichtiger im Rahmen des Steuerabzugs erfolgt.[144] Für digitale Werbeleistungen ist dabei insbesondere § 50a Abs. 1 Nr. 3 EStG relevant, nach dem Einkünfte aus der befristeten Überlassung von Rechten und Know-how (§ 49 Abs. 1 Nr. 2, 9 EStG) dem Steuerabzug unterliegen (für eine Kurzdarstellung vgl. Kapitel 3).

Voraussetzung für einen Steuerabzug gem. § 50a EStG ist grundsätzlich zunächst das Vorliegen inländischer Einkünfte nach § 49 EStG.[145] Für § 50a Abs. 1 Nr. 3 EStG bedeutet dies, dass es sich bei einer Überlassung von Rechten entweder um gewerbliche Einkünfte durch Betriebsstätte oder ständigen Vertreter (§ 49 Abs. 1 Nr. 2 Buchst. a EStG) bzw. gewerbliche betriebsstättenlose Einkünfte (§ 49 Abs. 1 Nr. 2 Buchst. f EStG) oder im Falle der Überlassung von Know-how um sonstige Einkünfte (§ 49 Abs. 1 Nr. 9 EStG) handeln muss.[146]

Jedoch unterliegen nicht alle von §§ 49 Abs. 1 Nr. 2 Buchst. a, f, Nr. 9 EStG erfassten Überlassungen von Rechten und Know-how dem Steuerabzug. § 50a Abs. 1 Nr. 3 EStG umfasst nach dem Gesetzeswortlaut insbesondere aber nicht ausschließlich Urheberrechte, gewerbliche Schutzrechte (die in § 73a Abs. 2, 3 EStDV näher konkretisiert werden) und Know-how (i. S. d. § 49 Abs. 1 Nr. 9 EStG). Urheberrechte sind dabei gem. § 73a Abs. 2 EStDV durch das UrhG geschützte Rechte. Gewerbliche Schutzrechte sind Rechte, welche durch das Design- (DesignG), Patent- (PatG), Gebrauchsmuster- (GebrMG) oder Markengesetz (MarkenG) geschützt sind.[147] Inwieweit von der nicht abschließenden Aufzählung des § 50a Abs. 1 Nr. 3 EStG auch weitere Rechte erfasst werden, die zwar keine Ausformung in einem Schutzgesetz finden, aber mit Urheberrechten bzw. gewerblichen Schutzrechten vergleichbar sind, ist bislang noch nicht höchstrichterlich entschieden.[148]

Vor dem 01.01.2007 und seit dem 01.01.2009 muss es sich zudem um eine befristete Nutzungsüberlassung von Rechten oder Know-how handeln, da die zwischenzeitliche Erweiterung um sämtliche Überlassungen inklusive der Veräußerung von Rechten und Know-how weggefallen ist.[149] Daher ist die zeitlich befristete von der unbefristeten Rechteüberlassung (auch Rechtekauf oder endgültige Übertragung

von Rechten) zu unterscheiden. Die Abgrenzung gestaltet sich aufgrund fehlender gesetzlicher Definition regelmäßig schwierig, sodass nur auf folgende in der Rechtsprechung entwickelte Kriterien für die Abgrenzung zurückgegriffen werden kann: Dauer und Exklusivität der Rechteüberlassung, Kündigungsmöglichkeiten und Rückgabeverpflichtung sowie Ausgestaltung von Verfügungsrechten.[150] Von einer zeitlich befristeten Rechteüberlassung ist daher auszugehen, wenn das Nutzungsrecht nicht endgültig bei dem durch den Vertrag Berechtigten verbleibt, der Rückfall des Rechts nicht durch Gesetz oder Vertrag ausgeschlossen ist oder eine vollständige Übertragung von vornherein nicht zulässig ist (z. B. der Fall bei urheberrechtlich geschützten Rechten).[151] Eine zeitlich unbegrenzte Überlassung von Rechten liegt hingegen vor, wenn das Nutzungsrecht endgültig bei dem durch den Vertrag Berechtigten verbleibt oder ein Rückfall des Rechts durch Gesetz oder Vertrag ausgeschlossen ist.[152] Durch Rechtsprechung des BFH steht die verbrauchende Rechteüberlassung dem Verkauf von Rechten gleich.[153] Eine Rechteüberlassung ist dann verbrauchend, wenn das wirtschaftliche Eigentum des Rechts durch die Nutzungsüberlassung endgültig an den durch den Vertrag Berechtigten übergeht und sich das Recht somit während der eingeräumten Nutzung vollständig verbraucht.[154] Bei Rechteüberlassungen im Rahmen der Erbringung digitaler Werbeleistungen unterliegen im Ergebnis also nur solche dem Steuerabzug nach § 50a EStG, welche zeitlich befristet sind.

Sind die Voraussetzungen für einen Steuerabzug nach § 50a Abs. 1 Nr. 3 EStG erfüllt, beträgt der Steuersatz für den beschränkt Steuerpflichtigen gem. § 50a Abs. 2 Satz 1 EStG 15 % der gesamten Einnahmen ohne Abzug von Betriebsausgaben zuzüglich Solidaritätszuschlag in Höhe von 5,5 % (§§ 2 Nr. 3 i. V. m. 3 Abs. 1 Nr. 6, 4 Satz 1 SolZG). Dadurch ergibt sich eine Gesamtsteuerbelastung von 15,82 %. Bei Vorliegen einer Nettovereinbarung, durch die sich der Vergütungsschuldner (derjenige, der die digitale Werbeleistung in Anspruch nimmt) zur Übernahme der Steuer verpflichtet, ist die übernommene Steuer auch Teil der Bemessungsgrundlage, was zu einem Gesamtsteuersatz von 18,79 % für den Vergütungsschuldner führt.[155] Diese Bruttobesteuerung durchbricht das objektive Nettoprinzip, nach dem nur Nettoeinkünfte besteuert werden dürfen.[156] Der EuGH hat jedoch entschieden, dass diese Bruttobesteuerung EU-rechtswidrig ist, was zur Folge hat, dass seit dem 01.01.2009 Betriebsausgaben, welche in unmittelbarem wirtschaftlichen Zusammenhang zu den Einnahmen stehen, abzugsfähig sind, sofern sie in nachprüfbarer Form nachgewiesen werden und

[144]Vgl. §§ 8, 31, 32 KStG; Helm, 2018, § 32 KStG, Rn. 9; Loschelder, 2018b, § 50a EStG, Rn. 6; Maßbaum, 2018, § 50a EStG, Anm. 6. § 50a EStG wird weiterhin durch §§ 73a - 73g EStDV ergänzt.

[145]Vgl. Gosch, 2018b, § 50a EStG; Rn. 6; Loschelder, 2018b, § 50a EStG; Rn. 1.

[146]Vgl. Goebel et al., 2018, S. 161; Rick, Gunsenheimer, Schneider & Kremer, 2017, Rn. 2694.

[147]Vgl. § 73a Abs. 3 EStDV.

[148]Vgl. Schlotter & Hruschka, 2018, S. 680-681; Diffring & Saft, 2019, S. 389.

[149]Vgl. BFH-Urteil vom 16.05.2001 – I R 64/99, Tz. 18; Holthaus, 2015, S. 27-28; Holthaus & Volkmann, 2015, S. 554.

[150]Vgl. Goebel et al., 2018, S. 178-179.

[151]Vgl. BMF-Schreiben vom 25.11.2010 – IV C 3 - S 2303/09/10002 2010/0861549, Tz. 23; BFH-Urteil vom 27.07.1988 – I R 87/85, Tz. 13.
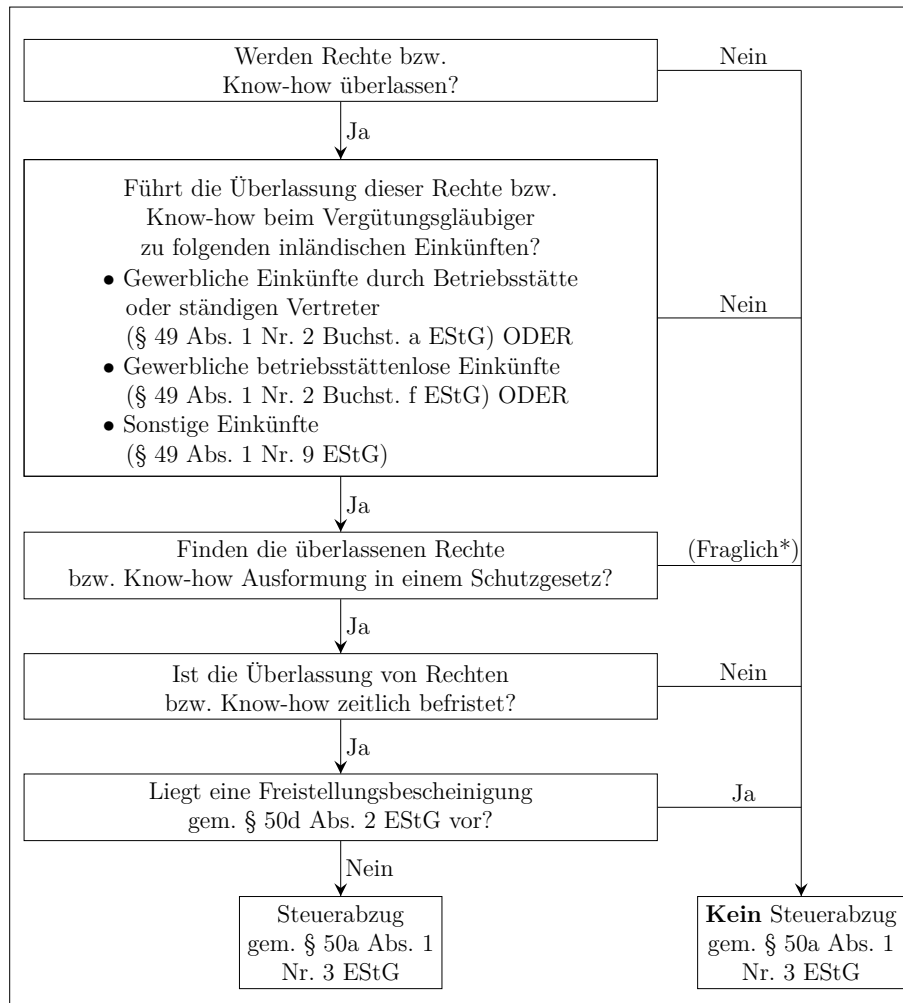
[152]Vgl. Goebel et al., 2018, S. 179-180.

[153]Vgl. BFH-Urteil vom 16.05.2001 – I R 64/99, Tz. 19; BMF-Schreiben vom 25.11.2010 – IV C 3 - S 2303/09/10002 2010/0861549, Tz. 24.

[154]Vgl. BMF-Schreiben vom 25.11.2010 – IV C 3 - S 2303/09/10002 2010/0861549, Tz. 24; BFH-Urteil vom 16.05.2001 – I R 64/99, Tz. 18-19.

[155]Vgl. Maßbaum, 2018, § 50a EStG, Anm. 10; Bundeszentralamt für Steuern, 2015.

[156]Vgl. Frotscher, 2015, Rn. 209.

Abbildung 3: Kurzdarstellung der Tatbestandsvoraussetzungen für die Steuererhebung im Rahmen des Steuerabzugs gemäß § 50a Abs. 1 Nr. 3 EStG (eigene Abbildung; in Teilen angelehnt an Goebel et al. (2018, 156))

der beschränkt Steuerpflichtige in einem Mitgliedsstaat der EU oder des Europäischen Wirtschaftsraums (EWR) ansässig ist.[157] Auch wenn dies nach dem Gesetzeswortlaut wiederum nicht bei Rechteüberlassungen gem. § 50a Abs. 1 Nr. 3 EStG gilt,[158] hat der BFH die Regelung im Rahmen einer normerhaltenden Interpretation auch auf diese ausgeweitet.[159] Die Körperschaftsteuer gilt gem. § 32 Abs. 1 KStG grundsätzlich als durch den Steuerabzug abgegolten, sofern nicht die Ausnahmen des § 32 Abs. 1, 2 KStG greifen. Hier ist vor allem § 32 Abs. 1 Nr. 2 KStG hervorzuheben, nach dem keine Abgeltungswirkung eintritt, wenn die Einkünfte innerhalb eines inländischen Betriebs, also im Falle digitaler Werbeleistungen innerhalb einer Betriebsstätte anfallen.[160]

Der Steuerabzug ist vom Vergütungsschuldner (derjenige, der die digitale Werbeleistung in Anspruch nimmt) für den Vergütungsgläubiger (der beschränkt steuerpflichtige Erbringer digitaler Werbeleistungen) vorzunehmen und an das Finanzamt abzuführen (§ 50a Abs. 5 Sätze 2, 3 EStG). Der Vergütungsschuldner haftet dabei gem. § 50a Abs. 5 Satz 5 EStG für die Abführung der Steuer.[161] Bestehen Zweifel an der Pflicht zum Steuerabzug, ist dieser dennoch vorzunehmen, es sei denn, der beschränkt Steuerpflichtige weist Gegenteiliges nach (§ 73e Satz 6 EStDV).[162] Der Steuerabzug ist zunächst grundsätzlich durchzuführen, auch wenn aufgrund eines Doppelbesteuerungsabkommens oder innerhalb der EU aufgrund der Zins- und Lizenzrichtlinie (ZiLiRL) das deutsche Besteuerungsrecht beschränkt oder vollständig aufgegeben wird.[163] Dabei handelt es sich um einen sogenann-

---

[157]Vgl. EuGH-Urteil vom 03.10.2006 – C-290/04; §§ 50a Abs. 3 Sätze 1-3 EStG i. V. m. 32 Abs. 4 KStG; Brähler, 2014, S. 83-84.

[158]Vgl. § 50a Abs. 1 Satz 1 1. HS EStG.

[159]Vgl. BFH-Urteil vom 27.07.2011 – I R 32/10; BFH-Urteil vom 25.04.2012 – I R 76/10; BMF-Schreiben vom 17.06.2014 – IV C 3-S 2303/10/10002:001; Frotscher, 2015, S. 212.

[160]Vgl. Helm, 2018, § 32 KStG, Rn. 94-99.

[161]Vgl. C. Kraft & Weigert, 2018, § 50a EStG, Rn. 68-71; Reimer, 2018b, § 50a EStG, Rz. 118-120.

[162]Vgl. Schaumburg, 2017, S. 6.301; Boochs, 2018a, § 50a EStG, Anm. 324.

[163]Vgl. § 50d Abs. 1 Satz 1 EStG. Für die deutsche Umsetzung der ZiLiRL

ten „treaty and directive overriding".[164] In einem zweiten Schritt kann der beschränkt Steuerpflichtige etwaige Erstattungsansprüche für ohne Rechtsgrund abgeführte Steuern im Rahmen des Erstattungsverfahrens geltend machen (§ 50d Abs. 1 Satz 2 EStG).[165] Nur bei Vorliegen einer sogenannten Freistellungsbescheinigung gem. § 50d Abs. 2 EStG kann die Höhe des Steuerabzugs reduziert erfolgen bzw. ganz unterbleiben.[166]

3.3.2. Im Rahmen der Veranlagung

Alle beschränkt steuerpflichtigen Einkünfte, bei der die Körperschaftsteuer nicht bereits im Rahmen des Steuerabzugs abgegolten ist, sind zu veranlagen.[167] Hierfür wird gem. § 31 Abs. 1 Satz 1 KStG auf die Regelungen des EStG verwiesen[168] und somit gelten im Grundsatz für die Veranlagung beschränkt Steuerpflichtiger dieselben Bestimmungen wie bei der unbeschränkten Steuerpflicht.[169] Besonderheiten ergeben sich dabei gem. § 50 EStG. So sind insbesondere Betriebsausgaben nur abzugsfähig, soweit sie einen wirtschaftlichen Zusammenhang mit den inländischen Einkünften aufweisen.[170] Da eine umfassende Darstellung des Veranlagungsverfahrens den Rahmen dieser Arbeit sprengen würde, wird an dieser Stelle für weitere Besonderheiten auf die Literatur verwiesen.[171]

4. Zwischenstaatliche Verteilung der Besteuerungsrechte an digitalen Werbeleistungen

In Fällen, in denen digitale Werbeleistungen, die von ausländischen Körperschaften in Deutschland erbracht werden, der beschränkten Steuerpflicht in Deutschland unterliegen, ist weiterhin zu prüfen, ob ein Doppelbesteuerungsabkommen (DBA) zwischen Deutschland und dem Ansässigkeitsstaat des beschränkt Steuerpflichtigen besteht. Dieses könnte die deutschen Besteuerungsrechte einschränken, da es den nationalen Gesetzen gem. § 2 Abs. 1 AO vorgeht. Im Folgenden werden daher zunächst die für digitale Werbeleistungen relevanten Verteilungsnormen des Musterabkommens der OECD (OECD-MA) vorgestellt, an dem sich viele deutsche

DBA orientieren (Kapitel 4.1).[172] Im Anschluss daran erfolgt eine Einordnung der inländischen Einkünfte nach deutschem Steuerrecht in die Verteilungsnormen des OECD-MA (Kapitel 4.2). Da DBA nur bestehende Steuerrechte verteilen, aber keine neuen Besteuerungsrechte begründen können, beschränken sich die Ausführungen auf die Einordnung der bereits beschriebenen potenziellen inländischen Einkünfte aus digitalen Werbeleistungen.

4.1. Verteilungsnormen des OECD-Musterabkommens

Wie schon bei der beschränkten Steuerpflicht nach nationalem Recht enthält auch das OECD-MA keine eigenständigen Regelungen für Einkünfte aus digitalen Werbeleistungen.[173] Diese müssen daher den bestehenden Verteilungsnormen zugeordnet werden. Aufgrund der gewerblichen Prägung und immateriellen Natur digitaler Werbeleistungen kommen insbesondere folgende Verteilungsnormen in Betracht:

- Unternehmensgewinne, Art. 7 OECD-MA (Kapitel 4.1.1)

- Lizenzgebühren, Art. 12 OECD-MA (Kapitel 4.1.2)

- Veräußerungsgewinne, Art. 13 OECD-MA (Kapitel 4.1.3)

4.1.1. Unternehmensgewinne

Die Verteilung der Besteuerungsrechte bei Unternehmensgewinnen wird in Art. 7 OECD-MA geregelt, sofern die Unternehmensgewinne nicht unter die spezielleren Vorschriften anderer Artikel des Musterabkommens fallen.[174] Grundsätzlich weist das OECD-MA das alleinige Besteuerungsrecht an Unternehmensgewinnen des Steuerpflichtigen dem Ansässigkeitsstaat zu, es sei denn, der Steuerpflichtige verfügt über eine Betriebsstätte im Quellenstaat.[175] Bei Vorliegen einer Betriebsstätte im Quellenstaat erhält dieser das Besteuerungsrecht an den durch die Betriebsstätte erwirtschafteten Unternehmensgewinnen.[176] In Bezug auf Deutschland als Quellenstaat bedeutet dies, dass in Deutschland nur Besteuerungsrechte an beschränkt steuerpflichtigen Einkünften aus digitalen Werbeleistungen verbleiben, welche als Unternehmensgewinne klassifiziert werden, sofern

---

vgl. § 50g EStG.

[164]Vgl. Gosch, 2018c, § 50d EStG, Rn. 1; Loschelder, 2018c, § 50d EStG, Rn. 6; kritisch hierzu Gebhardt, 2018, § 50d EStG, Rn. 32.

[165]Hierzu ausführlich Boochs, 2018b, § 50d EStG, Anm. 102-127; Klein & Hagena, 2018, § 50d EStG, Anm. 15-25.

[166]Für eine tiefergehende Betrachtung des sogenannten Freistellungsverfahrens sowie des im Rahmen dieser Arbeit aufgrund geringer Relevanz nicht betrachteten Kontrollmeldeverfahrens, des Steuererlasses bzw. der Pauschbesteuerung sei an dieser Stelle auf die einschlägige Literatur verwiesen. Siehe z. B. Boochs, 2018a, § 50d EStG, Anm. 102-186; Schaumburg, 2017, S. 6.306-6.311; Frotscher, 2015, Rn. 2016-220; Wagner, 2018, § 50d EStG, Rz. 50-63.

[167]Vgl. §§ 31, 32 Abs. 1 KStG.

[168]Vgl. Oellerich, 2018, § 31 KStG, Rn. 1; Nitzschke, 2018, § 31 KStG, Rn. 1.

[169]Vgl. Schaumburg, 2017, Rn. 6.270.

[170]Vgl. § 50 Abs. 1 Satz 1 EStG.

[171]Für weitere Besonderheiten bei beschränkter Steuerpflicht siehe z. B. Herkenroth & Striegel, 2018, § 50 EStG; Boochs, 2018c, § 50 EStG; Reimer, 2018c, § 50 EStG; Rick et al., 2017, Rn. 6.270-6.280.

---

[172]Vgl. Kahle & Kindich, 2016, Rz. 2.78; für das OECD-MA siehe OECD, 2017, 25-51.

[173]Zu diesem Ergebnis kommt auch Pinkernell (2014, S. 68) in Bezug auf Geschäftätigkeiten des E-Commerce im Allgemeinen.

[174]Vgl. Art. 7 Abs. 4 OECD-MA; OECD, 2017, Art. 7 OECD-MA, Rn. 1, 71-77.

[175]Vgl. Art. 7 Abs. 1 Satz 1 OECD-MA; OECD, 2017, Art. 7 OECD-MA, Rn. 10; Strunk & Kaminski, 2018, Art. 7 OECD-MA, Rz. 3.

[176]Vgl. Art. 7 Abs. 1 Satz 2 OECD-MA; OECD, 2017, Art. 7 OECD-MA, Rn. 12. Nach Art. 7 Abs. 1 Satz 2 i. V. m. Abs. 2 bestimmt sich der der Betriebsstätte zurechenbare Gewinn nach dem Fremdvergleichsgrundsatz. Auch wenn diese Gewinnzuteilung ebenfalls bei digitalen Werbeleistungen relevant ist, wird auf eine tiefergehende Darstellung im Folgenden verzichtet und z. B. auf die Ausführungen von OECD (2017, Art. 7 OECD-MA, Rn. 15-43) verwiesen.

sie im Rahmen einer im Inland belegenen Betriebsstätte erwirtschaftet werden.

Der Begriff der Betriebsstätte wird hierfür in Art. 5 OECD-MA definiert, was einen Rückgriff auf nationales Recht zur Begriffsdefinition ausschließt.[177] Die abkommensrechtliche Definition einer Betriebsstätte als „feste Geschäftseinrichtung, durch die die Geschäftstätigkeit eines Unternehmens ganz oder teilweise ausgeübt wird"[178], ähnelt stark dem nationalen Betriebsstättenbegriff (vgl. Kapitel 3.2.1) und fordert (1) eine Geschäftseinrichtung mit gewisser Festigkeit und Dauerhaftigkeit, (2) die der Ausübung der Unternehmenstätigkeit dient und (3) über die das Unternehmen Verfügungsmacht besitzt.[179] Art. 5 Abs. 2 OECD-MA enthält zusätzlich einen Positivkatalog betriebsstättenbegründender Sachverhalte. Der abkommensrechtliche Betriebsstättenbegriff wird allerdings durch den Negativkatalog des Art. 5 Abs. 4 OECD-MA im Vergleich zum deutschen Betriebsstättenbegriff eingeschränkt,[180] indem verschiedene Sachverhalte nicht betriebsstättenbegründend sind, wie z. B. Tätigkeiten vorbereitender Art oder Hilfstätigkeiten.

Zusätzlich kann ein abhängiger Vertreter eine Betriebsstätte begründen.[181] Dieser ist grundsätzlich mit einem ständigen Vertreter nach nationalem Recht (vgl. Kapitel 3.2.1) vergleichbar, mit dem Unterschied, dass im Abkommensfall eine Abschlussvollmacht des abhängigen Vertreters vorausgesetzt wird.[182]

4.1.2. Lizenzgebühren

Aufgrund ihres Spezialitätencharakters hat die Klassifikation von Einkünften als Lizenzgebühren gem. Art. 12 OECD-MA grundsätzlich Vorrang vor der Einordnung als Unternehmensgewinne gem. Art. 7 OECD-MA.[183] Das OECD-MA weist im Rahmen von Lizenzgebühren das alleinige Besteuerungsrecht grundsätzlich dem Ansässigkeitsstaat des Zahlungsempfängers zu.[184] Abweichend von diesen Bestimmungen ist jedoch in einigen deutschen DBA ein Besteuerungsrecht des Quellenstaates in Form eines Quellensteuerabzugs bei Lizenzgebühren vereinbart.[185]

Unter Lizenzgebühren versteht das OECD-MA „Vergütungen jeder Art, die für die Benutzung oder für das Recht auf Benutzung von Urheberrechten an literarischen, künstlerischen oder wissenschaftlichen Werken, einschließlich kinematographischer Filme, von Patenten, Marken [...], Mustern oder Modellen, Plänen, geheimen Formeln oder Verfahren oder für die Mitteilung gewerblicher, kaufmännischer oder wissenschaftlicher Erfahrungen gezahlt werden"[186]. Damit sind Lizenzgebühren sehr weit gefasst und umfassen alle Vergütungen für die Benutzung von Urheberrechten, gewerblichen Schutzrechten und Erfahrungen (Know-how).[187] Anders als im deutschen Steuerrecht ist die Aufzählung jedoch abschließend; ein Rückgriff auf innerstaatliches Recht ist aufgrund von Art. 3 Abs. 2 OECD-MA für abkommensrechtlich definierte Begriffe untersagt.[188] Weiterhin ist es unerheblich, ob die bezeichneten Rechte in einem öffentlichen Buch oder Register eingetragen sind.[189]

Art. 12 OECD-MA findet nur auf die zeitlich befristete Überlassung von Rechten i. S. d. Norm Anwendung.[190] Liegt hingegen eine vollständige Übertragung des Rechts vor (Rechtekauf), fallen die Einkünfte unter Art. 7 oder 13 OECD-MA.[191] Des Weiteren werden Lizenzzahlungen, die in Verbindung zu einer Betriebsstätte des Lizenzgebers stehen, nicht Art. 12, sondern Art. 7 OECD-MA (Unternehmensgewinne) zugerechnet.[192]

In Bezug auf digitale Werbeleistungen kommt insbesondere die zeitlich befristete Überlassung von Werberechten zur Begründung von Lizenzgebühren in Betracht.

4.1.3. Veräußerungsgewinne

Art. 13 OECD-MA regelt die Veräußerung von Vermögen

---

[177]Vgl. Art. 3 Abs. 2 OECD-MA; hierzu ausführlich Wassermeyer & Kaeser, 2018b, Art. 3 OECD-MA, Rz. 71-74; Hruschka, 2013, Art. 5 OECD-MA, Rz. 32.

[178]Beck-online, 2017, Art. 5 Abs. 1 OECD-MA. Hierbei handelt es sich um eine nichtamtliche Übersetzung des Beck Verlags. Der offizielle Gesamttext in englischer Sprache definiert eine Betriebsstätte als „fixed place of business through which the business of an enterprise is wholly or partly carried on"; Art. 5 Abs. 1 OECD-MA. Für Zwecke der Einheitlichkeit der Sprache werden wörtliche Zitate im Rahmen des OECD-MA im Folgenden dieser nichtamtlichen Übersetzung entnommen und in den jeweiligen Fußnoten der englischsprachige Text wiedergegeben.

[179]Vgl. OECD, 2017, Art. 5 OECD-MA, Rn. 6; Kahle & Kindich, 2016, Rz. 2.88.

[180]Vgl. Geutebrück, 2016, S. 44; Thiele & König, 2018, S. 378.

[181]Vgl. Art. 5 Abs. 5 OECD-MA.

[182]Vgl. Mohr, 2018, § 2 KStG, Rn. 142. Für weitere Ausführungen siehe z. B. OECD, 2017, Art. 5 OECD-MA, Rn. 82-101.

[183]Vgl. OECD, 2017, Art. 7 OECD-MA, Rn. 73-74.

[184]Vgl. Art. 12 Abs. 1 OECD-MA.

[185]Mit Hinblick auf die G20 Staaten hat Deutschland beispielsweise mit folgenden Staaten die Möglichkeit zum Quellensteuerabzug in DBA vereinbart:

Argentinien (15 %), Australien (5 %), China (6-10 %), Indien (10 %), Indonesien (7,5-15 %), Italien (0-5 %), Kanada (0-10 %), Mexiko (10 %), Türkei (10 %); (vgl. Bundeszentralamt für Steuern, 2018). Für einen weiteren Überblick über deutsche DBA inklusive der vereinbarten Quellensteuersätze sei an dieser Stelle auf die Übersicht des Bundeszentralamt für Steuern (2018) verwiesen. Die Möglichkeit der Quellenbesteuerung läuft bei Lizenzzahlungen zwischen verbundenen Unternehmen in deutschen DBA mit Mitgliedsstaaten der EU aber dennoch aufgrund der Zins- und Lizenzrichtlinie ins Leere, da diese zum Verzicht auf etwaige Besteuerungsrechte des Quellenstaates verpflichtet; vgl. 200349; Bozza-Bodden, 2013, Art. 12 OECD-MA, Rz. 23.

[186]Beck-online, 2017, Art. 12 Abs. 2 OECD-MA (nichtamtliche Übersetzung); englischsprachiger Originaltext: „payments of any kind received as a consideration for the use of, or the right to use, any copyright of literary, artistic or scientific work including cinematograph films, any patent, trade mark, design or model, plan, secret formula or process, or for information concerning industrial, commercial or scientific experience"; Art. 12 Abs. 2 OECD-MA.

[187]Vgl. Bozza-Bodden, 2013, Art. 12 OECD-MA, Rz. 78; Käbisch & Strunk, 2018, Art. 12 OECD-MA, Rz. 26.

[188]Vgl. Bozza-Bodden, 2013, Art. 12 OECD-MA, Rz. 79-80. Ein Rückgriff auf nationales Recht für abkommensrechtlich nicht definierte Begriffe ist hingegen möglich; vgl. Art. 3 Abs. 2 OECD-MA; Käbisch & Strunk, 2018, Art. 12 OECD-MA, Rz. 24.

[189]Vgl. OECD, 2017, Art. 12 OECD-MA, Rn. 8.

[190]Vgl. OECD, 2017, Art. 12 OECD-MA, Rn. 8.2; Bozza-Bodden, 2013, Art. 12 OECD-MA, Rz. 83; Wassermeyer & Kaeser, 2018a, Art. 12 OECD-MA, Rz. 58.

[191]Vgl. OECD, 2017, Art. 12 OECD-MA, Rn. 8.2; ausführlich hierzu Käbisch & Strunk, 2018, Art. 12 OECD-MA, Rz. 32-79.1.

[192]Vgl. Art. 12 Abs. 3 OECD-MA.

und findet bei digitalen Werbeleistungen hauptsächlich bei der vollständigen Übertragung von Rechten i. S. d. Art 12 OECD-MA Anwendung.[193] Hierbei ist für die Verteilung der Besteuerungsrechte grundsätzlich relevant, ob die veräußerten Rechte in Verbindung mit einer Betriebsstätte des Veräußernden (d. h. des Erbringers digitaler Werbeleistungen) im Quellenstaat stehen.[194] Falls dies der Fall ist, wird das alleinige Besteuerungsrecht an den Veräußerungsgewinnen dem Quellenstaat zugewiesen.[195] Steht die Veräußerung von Rechten i. S. d. Art. 12 OECD-MA nicht mit einer solchen Betriebsstätte in Verbindung, wird das alleinige Besteuerungsrecht gem. Art. 13 Abs. 5 OECD-MA dem Ansässigkeitsstaat des Veräußernden (d. h. des Erbringers digitaler Werbeleistungen) zugewiesen.

4.2. Zuordnung beschränkt steuerpflichtiger digitaler Werbeleistungen auf Abkommensebene

Die Einordnung inländischer Einkünfte von in Deutschland beschränkt steuerpflichtigen Erbringern digitaler Werbeleistungen in die Verteilungsnormen des OECD-MA spielt insgesamt nur eine stark untergeordnete Rolle. Deutschland als Quellenstaat gibt unabhängig von der Einordnung als Unternehmensgewinne, Lizenzgebühren oder Veräußerungsgewinne i. d. R. alle Besteuerungsrechte an Einkünften aus digitalen Werbeleistungen zugunsten des Ansässigkeitsstaates auf, es sei denn, diese Einkünfte fallen im Rahmen einer inländischen Betriebsstätte an. Mit Hinblick auf die in Kapitel 3 erarbeiteten nationalen Anknüpfungspunkte an gewerbliche Einkünfte durch Betriebsstätte oder ständigen Vertreter (vgl. Kapitel 3.2.1), gewerbliche betriebsstättenlose Einkünfte (vgl. Kapitel 3.2.2) und sonstige Einkünfte (vgl. Kapitel 3.2.3) kann Folgendes festgehalten werden:

- Sofern gewerbliche Einkünfte durch Betriebsstätte oder ständigen Vertreter gem. § 49 Abs. 1 Nr. 2 Buchst. a EStG auch die engeren abkommensrechtlichen Vorgaben der Begriffe der Betriebsstätte oder des abhängigen Vertreters erfüllen, sind diese als Unternehmensgewinne einzuordnen. Damit verbleibt das Besteuerungsrecht an diesen Einkünften in Deutschland (Art. 7 Abs. 1 Satz 2 OECD-MA).

- Bei gewerblichen betriebsstättenlosen Einkünften durch Vermietung und Verpachtung von Rechten i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG kommt es für die Einordnung als Lizenzgebühren darauf an, ob die zeitlich befristete Rechteüberlassung nach nationalem Recht auch den Lizenzbegriff des Art. 12 Abs. 2 OECD-MA erfüllt. Ist dies der Fall, verbleibt in Deutschland als Quellenstaat allenfalls nur bei gesonderter Vereinbarung eines Quellensteuerabzugs im jeweils geltenden DBA ein Besteuerungsrecht (abweichend von Art. 12

Abs. 1 OECD-MA). Erfüllen die inländischen Einkünfte schon den Lizenzbegriff nicht, sind die Einkünfte als (betriebsstättenlose) Unternehmensgewinne einzuordnen und damit in Deutschland nicht zu besteuern (Art. 12 Abs. 3 OECD-MA).

- Gewerbliche betriebsstättenlose Einkünfte aus der Veräußerung von Rechten i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG unterliegen i. d. R. Art. 7 oder 13 OECD-MA. Somit gibt Deutschland im Ergebnis alle Besteuerungsrechte an diesen (betriebsstättenlosen) Einkünften auf (Art. 7 Abs. 1, 13 Abs. 5 OECD-MA).

- Für die Überlassung von Know-how i. S. d. § 49 Abs. 1 Nr. 9 EStG (sonstige Einkünfte) gilt dasselbe wie für gewerbliche betriebsstättenlose Einkünfte.

5. Besteuerung digitaler Werbeleistungen am Beispiel von Google Ads

Im Folgenden wird die Wirksamkeit der in Kapitel 3 und 4 erarbeiteten steuerlichen Rahmenbedingungen bei der Erfassung von digitalen Werbeleistungen anhand eines konkreten Beispiels geprüft. Dieser Fokus auf einen Einzelfall ist notwendig, da die steuerliche Einordnung digitaler Werbeleistungen immer vom konkreten Geschäftsmodell und der vertraglichen Ausgestaltung abhängt. Um trotz der Fokussierung auf einen Einzelfall die maximale Abdeckung des deutschen Marktes für digitale Werbeleistungen zu erreichen, soll als exemplarisches Geschäftsmodell die Erbringung von Suchmaschinenwerbung (engl. search engine advertisement, kurz SEA) durch das Unternehmen Google dienen. Dies rechtfertigt sich dadurch, dass Suchmaschinenwerbung die umsatzstärkste Werbeform digitaler Werbeleistungen darstellt (vgl. Kapitel 2.1) und Google innerhalb dieses Marktes sowohl als weltweiter als auch als deutscher Marktführer gilt.[196]

Im Folgenden wird daher zunächst das Geschäftsmodell Googles im Zusammenhang mit der Erbringung von Suchmaschinenwerbung vorgestellt (Kapitel 5.1). Im Anschluss erfolgt eine steuerrechtliche Würdigung des Sachverhalts, um festzustellen, ob Deutschland nach aktueller Rechtslage im Inboundfall ein Besteuerungsrecht an im Inland erbrachten digitalen Werbeleistungen von Google zusteht (Kapitel 5.2).

5.1. Digitale Werbeleistungen als Geschäftsmodell

Das Geschäftsmodell des US-amerikanischen Unternehmens Google Inc. (im Folgenden Google), welches eine Tochtergesellschaft der ebenfalls US-amerikanischen Alphabet Inc. (im Folgenden Alphabet) ist, basiert im Wesentlichen auf der Erbringung von digitalen Werbeleistungen im Rahmen

---

[193]Vgl. OECD, 2017, Art. 12 OECD-MA, Rn. 8.2; Bozza-Bodden, 2013, Art. 12 OECD-MA, Rn. 83.
[194]Vgl. Gradel & Klaeren, 2018, Art. 13 OECD-MA, Rz. 18.
[195]Vgl. Art. 13 Abs. 2 OECD-MA.

[196]Auch wenn der tatsächliche Umsatz durch Suchmaschinenwerbung von Google in Deutschland unbekannt ist, liegt dieser Schluss nahe, da Google in den Jahren 2009 bis 2019 einen durchschnittlichen Marktanteil von 95 % am deutschen Suchmaschinenmarkt hatte (vgl. statcounter, 2019) und eine hohe Korrelation zwischen Marktanteil am Suchmaschinenmarkt und Umsatz im Bereich Suchmaschinenwerbung zu erwarten ist; so in seiner Argumentation wohl auch Bundeskartellamt, 2018, S. 5.

der gleichnamigen Google Suchmaschine. Der weltweite Umsatz von Google betrug im Jahr 2018 ca. 137 Mrd. USD, wovon ca. 116 Mrd. USD oder 85 % auf Umsätze aus digitalen Werbeleistungen entfielen.[197] Nach Schätzungen entfallen davon wiederum ca. 90 bis 95 % auf Suchmaschinenwerbung im Rahmen des sogenannten Google Ads Programms (vormals Google Adwords).[198]

5.1.1. Suchmaschinenwerbung mit Google Ads

Zentraler Bestandteil und Ausgangspunkt des Werbemodells von Google ist die gleichnamige Suchmaschine. Unter einer (Web-)Suchmaschine wie der Google Suchmaschine versteht man im Allgemeinen „ein Computersystem, das Inhalte aus dem *World Wide Web* mittels *Crawling* erfasst und über eine Benutzerschnittstelle durchsuchbar macht, wobei die Ergebnisse in einer nach systemseitig angenommener Relevanz geordneten Darstellung aufgeführt werden"[199]. Die Google Suchmaschine besteht also aus mehreren Komponenten, wie in Kapitel 4 dargestellt wird.

Der sogenannte Crawler[200] durchsucht vereinfacht gesprochen fortlaufend weite Teile des Internets nach Webseiten. Diese vom Crawler erfassten Webseiten werden vom Indexer aufbereitet, mit passenden Schlagworten versehen und innerhalb einer Datenbank, im sogenannten Index, abgelegt.[201] Zudem erfolgt eine Bewertung (Ranking) der im Index abgelegten Webseiten bezüglich der Qualität des Webseiteninhalts in Bezug auf die vergebenen Schlagworte.[202] Gibt ein Suchmaschinennutzer einen oder mehrere Suchbegriffe in die Suchmaske von Google ein, gleicht der Searcher die Suchanfrage mit dem aktuellen Datenbestand des Index ab und gibt die (organischen) Suchergebnisse nach Relevanz sortiert in der Suchergebnisseite (engl. search engine result page, kurz SERP) aus, wobei auch das beobachtete Suchverhalten des Nutzers berücksichtigt wird.[203] Eine beispielhafte Google Suchergebnisseite findet sich im Anhang.

Suchmaschinen wie Google sind Informationsintermediäre, welche die Informationssuche im Internet oftmals überhaupt erst ermöglichen.[204] Ihnen kommt insoweit gewissermaßen eine Gatekeeper-Funktion[205] zu, da sie den Nutzern nicht nur die Informationssuche ermöglichen, sondern auch deren Wahrnehmung von Internetinhalten, insbesondere durch die Auswahl und Anordnung der Suchergebnisse, beeinflussen können. Dies wird durch das Suchverhalten der Nutzer weiter verstärkt, die in der Regel nur die obersten Treffer einer Suche auswerten (sog. Primacy-

Effekt).[206] Insbesondere für Unternehmen, die ihre Produkte oder Dienstleistungen über das Internet vertreiben, ist es daher essenziell, bei Suchanfragen ihrer potenziellen Kunden in den obersten Suchtreffern zu erscheinen, um von diesen überhaupt wahrgenommen zu werden. Hier setzt das Geschäftsmodell der Suchmaschinenwerbung durch das Google Ads Programm an, indem Google gegen Entgelt bei jeder Suchanfrage eine begrenzte Anzahl gesondert gekennzeichneter „Suchtreffer" über, unter oder neben den organischen Suchtreffern der Suchmaschine platziert.[207]

Im Rahmen von Google Ads werden diese kostenpflichtigen „Suchtreffer" (im Folgenden Anzeigen) an Webseitenbetreiber (im Folgenden Werbetreibende) für jede einzelne Suchanfrage versteigert.[208] Werbetreibende können dabei im Vorhinein ihre maximalen Gebote auf von ihnen selbst gewählte Suchbegriffe (Keywords) in der webbasierten Google Ads Software abgeben.[209] Immer wenn diese Suchbegriffe in die Google Suchmaschine eingegeben werden, findet eine vollautomatische Auktion[210] statt, in der die zur Verfügung stehenden Anzeigenplätze an die bietenden Werbetreibenden versteigert werden.[211] Eine Besonderheit des Auktionsmechanismus liegt darin, dass nicht nur die maximale Zahlungsbereitschaft der Bietenden über das Auktionsergebnis entscheidet, sondern Google auch die Qualität der Anzeigen in Bezug auf die Relevanz für den Suchmaschinennutzer nach geheimen Verfahren bewertet (Qualitätsfaktor). Gewinner der jeweiligen Auktion sind die Werbetreibenden, die die Kombination aus maximaler Zahlungsbereitschaft des Gebots und Qualitätsfaktor der Anzeige maximieren.[212] Die Anzeigen der Auktionsgewinner werden dem Suchmaschinennutzer dann auf der Suchergebnisseite angezeigt. Eine weitere Besonderheit des Auktionsmechanismus von Google ist, dass die Werbetreibenden nur dann für die Anzeigenplätze bezahlen, wenn der Suchmaschinennutzer auch tatsächlich auf die Anzeige klickt, d. h., die Werbemaßnahme zum Erfolg führt und der Suchmaschinennutzer auf die Webseite des Werbetreibenden weitergeleitet wird; dies wird auch als Cost-per-Klick (CPC) Verfahren bezeichnet.[213] Im Erfolgsfall bezahlen die Werbetreibenden weiterhin nicht das eigene Maximalgebot, sondern nur das Gebot, welches nötig ist, um den nächst

---

[197]Vgl. Alphabet Inc., 2019, S. 27.

[198]Vgl. Lammenett, 2017, S. 154.

[199]Lewandowski, 2013, S. 495.

[200]Auch Spider oder Robot genannt; vgl. Müller-Terpitz & Knodel, 2017, Rn. 7; S. Rupp, 2010, S. 26.

[201]Vgl. Lewandowski, 2015, S. 31-57; Pelzer, Sommeregger & Linnenbrink, 2015, S. 31.

[202]Vgl. Müller-Terpitz & Knodel, 2017, Rn. 13-15.

[203]Vgl. Müller-Terpitz & Knodel, 2017, Rn. 7; Lewandowski, 2015, S. 58-61; Pelzer et al., 2015, S. 32.

[204]Vgl. Müller-Terpitz & Knodel, 2017, Rn. 1; Paal, 2012, S. 25.

[205]Vgl. Müller-Terpitz & Knodel, 2017, Rn. 2.

[206]Vgl. Rath, 2005, S. 826; für empirische Nachweise vgl. anstatt vieler z. B. Kim, Thomas, Sankaranarayana, Gedeon & Yoon, 2015.

[207]Vgl. Lammenett, 2017, S. 135; neben der Möglichkeit der kostenpflichtigen Suchmaschinenwerbung verfolgen viele Unternehmen die Strategie, im Rahmen der Suchmaschinenoptimierung (engl. search engine optimization kurz SEO) die eigene Webseite so zu optimieren, dass diese von der Google Suchmaschine als möglichst relevant eingestuft wird und somit unter den ersten organischen Suchtreffern erscheint.

[208]Vgl. Pelzer et al., 2015, S. 60.

[209]Vgl. Pelzer et al., 2015, S. 60.

[210]Dabei handelt es sich um eine allgemeine Zweit-Preis-Auktion mit versiegelten Geboten, engl. generalized sealed-bid second-price auction. Für weitere Informationen zu dieser Auktionsform siehe z. B. Varian (2016, S. 365-377).

[211]Vgl. Pelzer et al., 2015, S. 60.

[212]Vgl. Pelzer et al., 2015, S. 61-63; Ciomek, 2016, S. 225.

[213]Vgl. Pelzer et al., 2015, S. 59; Müller-Terpitz & Knodel, 2017, Rn. 72; Lammenett, 2017, S. 146.
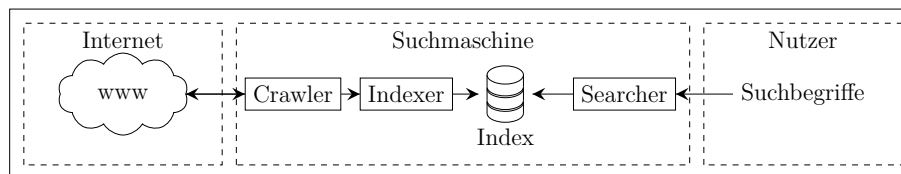
Abbildung 4: Komponenten einer Suchmaschine (eigene Abbildung; in Teilen angelehnt an Risvik und Michelsen (2002, 290))

tiefer platzierten Konkurrenten zu überbieten.[214]

5.1.2. Europäische Vertriebsstruktur im Rahmen von Google Ads

Die digitalen Werbeleistungen im Rahmen von Google Ads werden in Europa von der irischen Tochtergesellschaft Google Ireland Ltd. (im Folgenden Google Irland) erbracht, d. h., Vertragspartner der deutschen Werbetreibenden ist Google Irland (vgl. Kapitel 5).[215] Zwar unterhält Google auch eine Tochtergesellschaft in Deutschland (Google Germany GmbH, im Folgenden Google Deutschland), jedoch ist diese nicht in die Erbringung digitaler Werbeleistungen involviert, sondern erbringt als Servicegesellschaft lediglich Dienstleistungen im Bereich Marketing, Web-Hosting sowie Forschung und Entwicklung gegenüber der Muttergesellschaft und Google Irland.[216]

5.2. Beschränkte Steuerpflicht der Suchmaschinenwerbung in Deutschland

Da Google seine Werbeleistungen im europäischen Markt durch die Tochtergesellschaft Google Irland vertreibt, ist zu prüfen, inwiefern diese Tochtergesellschaft in Deutschland steuerpflichtig ist. Google Irland mit Geschäftsleitung und Sitz in Irland[217] unterliegt in Deutschland nicht der unbeschränkten Körperschaftsteuerpflicht gem. § 1 Abs. 2 KStG. Zu prüfen bleibt aber, ob eine beschränkte Körperschaftsteuerpflicht gem. § 2 Nr. 1 KStG besteht. Hierfür ist das Vorliegen inländischer Einkünfte gem. § 49 EStG erforderlich, wobei bei digitalen Werbeleistungen insbesondere gewerbliche Einkünfte durch Betriebsstätte oder ständigen Vertreter (Nr. 2 Buchst. a), gewerbliche betriebsstättenlose Einkünfte (Nr. 2 Buchst. f) oder sonstige Einkünfte (Nr. 9) in Betracht kommen.

5.2.1. Gewerbliche Einkünfte durch Betriebsstätte oder ständigen Vertreter

Wie in Kapitel 3.2.1 dargestellt, müssen für inländische Einkünfte gem. § 49 Abs. 1 Nr. 2 Buchst. a EStG (1) Einkünfte aus Gewerbebetrieb vorliegen und (2a) eine Betriebsstätte oder (2b) ein ständiger Vertreter in Deutschland unterhalten werden. Die Gewerblichkeit der Einkünfte gem. § 15 EStG

könnte im vorliegenden Fall wohl angenommen werden, jedoch scheitert es am Vorliegen einer Betriebsstätte oder eines ständigen Vertreters.

Aus Kapitel im Anhang geht zwar hervor, dass die amerikanische Google Inc. eine Tochtergesellschaft in Deutschland unterhält. Dort sind ca. 1.065 Mitarbeiter beschäftigt.[218] Diese erbringt auch Supportleistungen gegenüber Google Irland. Dabei handelt es sich jedoch um eine selbstständige Kapitalgesellschaft, die keine juristisch unselbstständige Zweigniederlassung oder Geschäftsstelle von Google Irland darstellt und somit keine Betriebsstätte gem. § 12 Satz 2 Nr. 2, 3 EStG begründet.[219] Weiterhin werden in Deutschland auch keine Internetserver unterhalten, die eine Betriebsstätte für Google Irland begründen könnten. Google Irland tritt einzig durch die Unternehmenswebseite mit darauf integrierten Werbeflächen in Deutschland auf. Wie bereits beschrieben, kann diese jedoch schon allein aufgrund mangelnder Körperlichkeit keine Betriebsstätte begründen.

Nach aktuellem Kenntnisstand ist kein inländisches Personal von Google Irland in Deutschland tätig, welches als ständiger Vertreter gem. § 13 AO anzusehen ist. Weiterhin ist aufgrund mangelnder Internetserver im Inland nicht anzunehmen, dass in Deutschland weisungsgebundene Webhosting-Unternehmen/Internet-Service-Provider für Google Irland tätig sind. Somit sind insgesamt keine Anhaltspunkte für das Vorhandensein eines ständigen Vertreters gem. § 13 AO gegeben.

Im Ergebnis begründen damit digitale Werbeleistungen, die von Google Irland in Deutschland erbracht werden, keine inländischen gewerblichen Einkünfte durch Betriebsstätte oder ständigen Vertreter gem. § 49 Abs. 1 Nr. 2 Buchst. a EStG.

5.2.2. Gewerbliche betriebsstättenlose Einkünfte

Ob Google Irland bei der Erbringung von Suchmaschinenwerbung im Rahmen des Google Ads Programms in Deutschland gewerbliche betriebsstättenlose Einkünfte aus der Vermietung/Verpachtung oder Veräußerung von Rechten i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG erzielt, ist in der jüngsten Literatur umstritten. Der Diskussion liegen insbesondere Vorstöße aus Teilen der Finanzverwaltung zugrunde, die das Vorliegen solcher inländischer Einkünfte bei Google Irland als gegeben sehen.[220] Aus den Ausführungen zu Kapitel 3.2.2

---

[214]Vgl. Pelzer et al., 2015, S. 61.
[215]Vgl. Google Inc., 2019; Pinkernell, 2012, S. 370-371.
[216]Vgl. Google Germany GmbH, 2019; Pinkernell, 2012, S. 370-371.
[217]Vgl. Google Ireland Ltd., 2018, S. 15.

[218]Die Mitarbeiterzahl bezieht sich auf den Stichtag 31.12.2017; vgl. Google Germany GmbH, 2019.
[219]So auch Lück, 2018, S. 51.
[220]Von diesen Vorstößen berichten etwa Schlotter & Hruschka, 2018, S. 711-712; Diffring, 2019; Diffring & Saft, 2019, S. 387; Linn, 2019, S. 419;
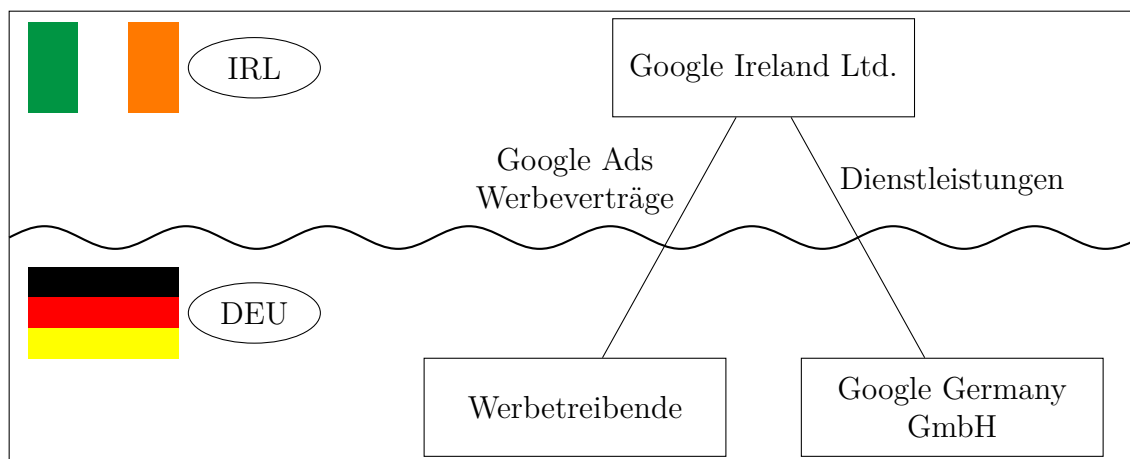
Abbildung 5: Europäische Vertriebsstruktur von Google Ads (eigene Abbildung)

geht hervor, dass für das Vorliegen gewerblicher betriebsstättenloser Einkünfte aus der Erbringung von Suchmaschinenwerbung (1) Rechte i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG zeitlich befristet oder unbefristet überlassen werden müssen, welche entweder (2a) den Belegenheits- oder (2b) Verwertungstatbestand erfüllen.

Damit ist in einem ersten Schritt zunächst zu beurteilen, ob es sich bei der Erbringung von Suchmaschinenwerbung im Rahmen des Google Ads Programms um eine Überlassung von Rechten oder um eine bloße Erbringung einer Dienstleistung handelt, da Letztere keine inländischen Einkünfte gem. § 49 Abs. 1 Nr. 2 Buchst. f EStG begründet. Nach hier vertretener Auffassung ist dabei die eigentliche Platzierung der Werbeanzeige in der SERP (Kapitel 5.2.2) gesondert von weiteren Teilleistungen zu beurteilen, die im Rahmen des Google Ads Programms erbracht werden (Kapitel 5.2.2). Weitere Teilleistungen sind z. B. die Überlassung der internetbasierten Google Ads Software. Dieses Vorgehen rechtfertigt sich daraus, dass bei gemischten Verträgen steuerlich unterschiedlich zu bewertende Teilleistungen grundsätzlich auch unterschiedlich behandelt werden.[221] Weiterhin wird die Ansicht vertreten, dass zur Abgrenzung der Überlassung von Rechten von der Erbringung von Dienstleistungen auf die zivilrechtliche Einordnung des zugrundeliegenden Vertrages abzustellen ist. Eine steuerspezifisch abweichende Einordnung lässt sich nicht, wie von Hruschka (2019) vertreten, aus § 4 Abs. 1 Satz 2 EStG ableiten[222] und ist darüber hinaus auch nicht in der Rechtsprechung des BFH erkenntlich[223]. In einem zweiten Schritt sind der Belegenheits- und Verwertungstatbestand zu prüfen (Kapitel 5.2.2).

Rechteüberlassung aus der Platzierung von Werbeanzeigen
Von verschiedenen Seiten wird berichtet, dass von der Fi-

nanzverwaltung in Betriebsprüfungen vermehrt die Ansicht vertreten wird, die Inanspruchnahme digitaler Werbeleistungen wie Suchmaschinenwerbung im Rahmen von Google Ads sei nicht als Inanspruchnahme von Dienstleistungen, sondern als Überlassung von Werberechten anzusehen.[224] Dieser Argumentation liegt wohl ein Urteil des BGH zugrunde, welches die Einräumung eines Exklusivrechts auf Werbung als Verpachtung von Rechten klassifiziert.[225] Der Ansicht der Betriebsprüfung ist jedoch nicht zu folgen. Die Platzierung der Anzeigen von Werbetreibenden durch Google Irland in der SERP von Suchanfragen stellt keine Rechteüberlassung, sondern eine Dienstleistung dar. Der BFH hat sich bereits mit der Abgrenzung von Rechteüberlassungen und Dienstleistungen (im Rahmen der Gewerbesteuer) beschäftigt.[226] Er versteht unter Rechten „Immaterialgüterrechte, d. h. subjektive Rechte an unkörperlichen Gütern mit selbständigem Vermögenswert, die eine Nutzungsbefugnis enthalten und an denen [...] ein Abwehrrecht [...] besteht"[227]. Dienstleistungen grenzen sich nach Ansicht des BFH insofern von Rechten ab, als dass sie keine Nutzungs- und Abwehrbefugnis an unkörperlichen Gütern überlassen.[228]

Im vorliegenden Fall wird den Werbetreibenden von Google Irland jedoch keine Nutzungsbefugnis an bestimmten Werbeflächen überlassen, womit eine Rechteüberlassung ausscheidet.[229] Die Werbeflächen werden für jede Suchanfrage neu im Rahmen einer Auktion versteigert. Werbetreibende erhalten keinerlei Garantie, dass ihre Anzeige bei Suchanfragen überhaupt bzw. bei mehreren identischen Suchanfragen immer an gleicher Stelle platziert wird, selbst wenn sie das höchste Gebot innerhalb der Auktion abgeben. Grund hier-

---

Reichl & Wiedmann, 2019, S. 202.

[221]Vgl. BFH-Urteil vom 28.01.2004 – I R 79/02, Tz. 18; BFH-Urteil vom 19.12.2007 – I R 19/06, Tz. 48.

[222]Vgl. Diffring & Saft, 2019, S. 390. Für weitere Ausführungen hierzu wird auf vorstehenden Aufsatz verwiesen.

[223]Vgl. etwa BFH-Urteil vom 26.04.2018 – III R 25/16.

[224]So etwa Schlotter & Hruschka, 2018, S. 711-712; Diffring, 2019; Linn, 2019, S. 419; Reichl & Wiedmann, 2019, S. 202; Haselrieder & Pontzen, 2019; Wieduwilt, 2019.

[225]Vgl. BGH-Urteil vom 26.01.1994 – XII ZR 93/92.

[226]Vgl. BFH-Urteil vom 26.04.2018 – III R 25/16.

[227]BFH-Urteil vom 26.04.2018 – III R 25/16, Tz. 29.

[228]Vgl. BFH-Urteil vom 26.04.2018 – III R 25/16, Tz. 31.

[229]So etwa Reichl & Wiedmann, 2019, S. 204; Diffring & Saft, 2019, S. 390-391; Linn, 2019, S. 420.

für ist, dass sich der Qualitätsfaktor der Anzeige bei jeder einzelnen Suchanfrage und für jeden einzelnen Suchmaschinennutzer unterscheiden kann und somit das für die Auktion entscheidende Produkt aus maximaler Zahlungsbereitschaft und Qualität der Anzeige variiert. Die Werbetreibenden können somit nur ein Gebot für eine Platzierung auf der SERP abgeben, woraus sich jedoch keine Nutzungsbefugnis für eine bestimmte Werbefläche ableiten lässt. Google Irland entscheidet für jede Suchanfrage neu, ob und welche Werbeanzeigen wo auf der Webseite platziert werden. Weiterhin sind keine Abwehrrechte gegenüber Dritten aus dem Vertragsverhältnis zu erkennen.[230]

Im Fall von Google Ads werden den Werbetreibenden somit weder durchsetzbare Nutzungsbefugnisse gegenüber Google Irland noch Abwehrrechte gegenüber Dritten überlassen, wodurch nach zutreffender Ansicht des BFH keine Rechteüberlassung vorliegt. Weiterhin ist der Sachverhalt damit auch nicht mit dem vom BGH entschiedenen Fall vergleichbar,[231] bei dem die Überlassung eines Exklusivrechts auf Werbung als Verpachtung eines Rechts klassifiziert wurde.[232] In Letzterem wurde dem Werbetreibenden die exklusive Möglichkeit geboten, auf für diesen Zweck aufgestellten Werbetafeln für einen festgelegten Zeitraum von zehn Jahren zu einem festgelegten Entgelt auf einem Golfplatz Werbung zu treiben.[233] Dem Werbetreibenden wurden insofern Nutzungsbefugnisse und Abwehrrechte an einer bestimmten Werbefläche überlassen.[234] Genau darin unterscheidet sich aber gerade die Erbringung von Suchmaschinenwerbung.

Bei den digitalen Werbeleistungen von Google Irland im Rahmen der Suchmaschinenwerbung mit Google Ads handelt es sich vielmehr um die Erbringung von Dienstleistungen.[235] Dabei ist grundsätzlich unerheblich, dass die Dienstleistung rein maschinell und ohne Mitwirkung von Personen erbracht wird.[236] Nach Rechtsprechung des BFH spricht eine erfolgsbezogene Vergütung für einen bestimmten Vermittlungserfolg bei digitalen Werbeleistungen dafür, dass es sich um eine Dienstleistung handelt.[237] Bei Google Ads erfolgt eine Vergütung durch den Werbetreibenden immer nur dann, wenn ein Suchmaschinennutzer auch tatsächlich auf die geschaltete Anzeige klickt und so auf die Webseite des Werbetreibenden weitergeleitet wird, d. h., wenn die Werbemaßnahme zum Erfolg führt. Die Vergütung ist damit erfolgsabhängig und wird für den Vermittlungserfolg und nicht für die Platzierung der Werbeanzeige bezahlt. Der BFH würdigt solche erfolgsabhängigen Vergütungen ähnlich einer Provision eines Handelsvertreters (§ 84 Abs. 1 HGB) bzw. Handels-

maklers (§ 93 Abs. 1 HGB) und spricht in seiner Entscheidung von einer Klassifikation als Dienst- (§ 611 BGB) bzw. Werkvertrag (§ 631 BGB).[238] Daher sind die Werbeleistungen von Google Irland als Dienstleistungen anzusehen. Dies deckt sich auch mit der Rechtsprechung des BGH, der in der zeitlich befristeten Platzierung einer Werbeanzeige auf einer Webseite selbst bei einem festgelegten monatlichen Entgelt von einem Werkvertrag i. S. d. § 631 BGB ausgeht.[239] Ob im Fall von Google Ads letztendlich ein Werkvertrag oder aufgrund der konkret geschuldeten Werbeleistung ein Dienstvertrag zugrunde liegt, kann letztendlich dahingestellt bleiben, da in beiden Fällen eine Dienstleistung und somit keine Rechteüberlassung vorliegt und damit eine beschränkte Steuerpflicht gem. § 49 Abs. 1 Nr. 2 Buchst. f EStG ausscheidet.[240]

Im Ergebnis werden durch das Platzieren von Werbeanzeigen keine Rechte i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG überlassen und keine inländischen Einkünfte begründet.

Rechteüberlassung aus weiteren Teilleistungen

Neben der bloßen Platzierung von Werbeanzeigen kommen weitere Teilleistungen der Werbeleistungserbringung in Betracht, die eine Rechteüberlassung beinhalten könnten.

Google Irland greift zur Erbringung der Suchmaschinenwerbung im Rahmen von Google Ads auf eine webbasierte Anwendung, das sogenannte Google Ads Konto zurück, in dem sich alle Werbetreibende zur Teilnahme am Google Ads Programm anmelden müssen.[241] Innerhalb dieser Webanwendung können die Werbetreibenden Schlagworte auswählen, ihre Werbeanzeigen gestalten und ihre maximalen Gebote für die ausgewählten Schlagworte abgeben.[242] Sofern diese Webanwendung als urheberrechtlich geschützte Software (§§ 2 Abs. 1, 69a UrhG) zu klassifizieren ist, stellt sich die Frage, ob den Werbetreibenden umfassende Nutzungsrechte an dieser Software (zeitlich begrenzt) übertragen werden, die als Rechteüberlassung i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG gelten (vgl. Ausführungen zu Kapitel 3.2.2). Da diese Software den Werbetreibenden aber nur zum bestimmungsgemäßen Gebrauch überlassen wird und in diesem Zuge den Werbetreibenden auch keine „Vervielfältigungs-, Bearbeitungs-, Verbreitungs- oder Veröffentlichungsrechte"[243] zur wirtschaftlichen Weiterverwertung überlassen werden,[244] dürfte der bisherigen Argumentation der Finanzverwaltung folgend keine Rechteüberlassung vorliegen.[245]

[230]Vgl. Reichl & Wiedmann, 2019, S. 204.

[231]So auch Reichl & Wiedmann, 2019, S. 204.

[232]Vgl. BGH-Urteil vom 26.01.1994 – XII ZR 93/92, Tz. 15.

[233]Vgl. BGH-Urteil vom 26.01.1994 – XII ZR 93/92, Tz. 2, 4, 10.

[234]Vgl. Reichl & Wiedmann, 2019, S. 204; so wohl auch Linn, 2019, S. 420.

[235]Zu diesem Ergebnis kommen auch Pinkernell, 2014, S. 55; Schlotter & Hruschka, 2018, S. 711; Diffring, 2019; Diffring & Saft, 2019, S. 390-391; Linn, 2019, S. 420-421; Reichl & Wiedmann, 2019, S. 203-204; des Weiteren im Kontext der Gewerbesteuer auch Rapp, 2017, S. 571-572.

[236]Vgl. BFH-Urteil vom 26.04.2018 – III R 25/16, Tz. 37.

[237]Vgl. BFH-Urteil vom 26.04.2018 – III R 25/16, Tz. 38.

[238]Vgl. BFH-Urteil vom 26.04.2018 – III R 25/16, Tz. 38, 40.

[239]Vgl. BGH-Urteil vom 22.03.2018 – VII ZR 71/17, Tz. 1-4, 12. Auch in der Literatur werden Verträge über Werbemaßnahmen überwiegend als Werkverträge angesehen, vgl. etwa Busche, 2018, § 631 BGB, Rn. 175; Schlotter & Hruschka, 2018, S. 711.

[240]Vgl. Reichl & Wiedmann, 2019, S. 203.

[241]Vgl. Pelzer & Düssel, 2013, S. 166-167.

[242]Für eine umfassende Beschreibung der Funktionalität des Google Ads Kontos siehe Pelzer & Düssel, 2013, S. 166-206.

[243]BMF-Schreiben vom 27.10.2017 – IV C 5 - S 2300/12/10003:004, Tz. 3.

[244]Es finden sich keine Anhaltspunkte für solche umfassenden Nutzungsrechte in den allgemeinen Nutzungsbedingungen; vgl. Google Inc., 2019.

[245]Vgl. BMF-Schreiben vom 27.10.2017 – IV C 5 - S 2300/12/10003:004,

Da sowohl beim Betreiben der Suchmaschine als auch bei der Erbringung der Suchmaschinenwerbung große Datensätze Verwendung finden, bleibt weiterhin zu prüfen, ob den Werbetreibenden von Google Irland evtl. umfassende Nutzungsrechte an Datenbankinhalten eingeräumt werden, die als Rechteüberlassung klassifiziert werden könnten. So greift Google z. B. bei der Bereitstellung der Suchmaschine auf den sogenannten Index zurück, der alle von der Suchmaschine erfassten Webseiten beinhaltet (vgl. Kapitel 5.1.1). Weiterhin speichert Google umfassende Daten zum Suchverhalten der Nutzer.[246] Jedoch ergeben sich aus den Nutzungsbedingungen von Google Ads keinerlei Anhaltspunkte, dass Werbetreibenden überhaupt Zugriff auf diese datenbankähnlichen Strukturen gewährt wird, oder diese sogar umfassende Nutzungsrechte an diesen erhalten.[247]

Auch durch weitere Teilleistungen werden im Ergebnis keine Rechte i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG überlassen und damit keine inländischen Einkünfte begründet.

Belegenheits- und Verwertungstatbestand

Da es wie beschrieben bereits an Rechten i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG mangelt, die überlassen werden könnten, ist eine weitere Prüfung des Belegenheits- bzw. Verwertungstatbestands hinfällig. Unter der hypothetischen Annahme, es lägen Rechte i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG vor, würden diese aber ebenfalls nicht den Belegenheitstatbestand erfüllen, da die hypothetischen Werberechte keine Ausformung in einem Schutzgesetz fänden und auch in keinem inländischen Buch oder Register eingetragen wären.[248] Der Verwertungstatbestand hingegen könnte aber wohl regelmäßig im Falle von Vertragsschlüssen im gewerblichen B2B-Bereich als erfüllt angesehen werden.[249]

Sonstige Einkünfte

Derzeit wird von Hruschka (2019) vertreten, dass den Werbetreibenden bei der Erbringung von Suchmaschinenwerbung Know-how i. S. d. § 49 Abs. 1 Nr. 9 EStG in Form von bestimmten Algorithmen[250] zur Nutzung überlassen wird.[251] Nach dieser Ansicht würden auch im Falle der Suchmaschinenwerbung durch Google Ads inländische Einkünfte aus der Überlassung von Know-how vorliegen. Voraussetzungen für das Vorliegen sonstiger Einkünfte i. S. d. § 49 Abs. 1 Nr. 9 EStG sind (1) die Überlassung von Know-how (Kapitel 5.2.2) (2) zur Nutzung im Inland (Kapitel 5.2.2).[252] Da diese Voraussetzungen wie nachfolgend dargelegt jedoch nicht

erfüllt sind, ist der Ansicht von Hruschka (2019) nicht zu folgen.[253] Im Ergebnis liegen keine inländischen Einkünfte Google Irlands aufgrund sonstiger Einkünfte i. S. d. § 49 Abs. 1 Nr. 9 EStG vor.

Überlassung von Know-how

Es ist nicht ersichtlich, dass Google Irland den Werbetreibenden im Rahmen von Google Ads in irgendeiner Weise Know-how vermittelt.[254] Die Werbetreibenden erhalten zum einen keine gesonderten Kenntnisse über die Funktionsweise bzw. die verwendeten Algorithmen der Google Suchmaschine an sich. So gewinnen sie z. B. keinerlei Kenntnisse über die genaue Funktionsweise des Crawlers, Indexers oder Searchers der Suchmaschine. Zum anderen überlässt Google Irland den Werbetreibenden auch kein gesondertes Know-how bezüglich der Algorithmen, die über die Auswahl der Werbeanzeigen und deren Platzierung entscheiden, welches über das öffentlich bereits Bekannte hinausgeht. Google wird diese betriebsnotwendigen Geschäftsgeheimnisse auch nicht aus der Hand geben wollen.[255]

Das zum Betreiben der Suchmaschine und zum Erbringen der Werbedienstleistung nötige Know-how wird hingegen von Google Irland selbst angewandt.[256] Google stellt den Nutzern bzw. Werbetreibenden nur die fertigen Produkte (Suchmaschine, Werbeleistung) zur Verfügung, ohne dass ein Wissenstransfer am zugrundeliegenden Know-how stattfindet. Nach Rechtsprechung des BFH liegt jedoch genau dann keine Überlassung von Know-how vor, da dieses nicht tatsächlich vermittelt, sondern nur vom „Vermittler" (im vorliegenden Fall Google Irland) selbst angewendet wird.[257] Diese Würdigung unterstreicht auch das FG München, welches in der Durchführung einer klinischen Studie auf Grundlage ähnlicher Argumente keine Know-how-Überlassung, sondern die Erbringung einer konkreten Dienstleistung sieht.[258] Auch in besagtem Fall haben nach Ansicht des FG München die die Studie durchführenden ausländischen Ärzte ihr Know-how nur selbst angewandt und dem Auftraggeber lediglich im Rahmen einer technischen Dienstleistungserbringung das Ergebnis dieser Know-how Anwendung in Form einer entsprechenden Ergebnisdokumentation und nicht das Know-how selbst zur Nutzung überlassen.[259]

Würde Google Irland den Werbetreibenden tatsächlich die der Werbeleistungserbringung zugrunde liegenden Algorithmen überlassen, wären diese höchstwahrscheinlich in einer Software eingebettet.[260] Software wiederum ist jedoch

---

Tz. 3, 4.

[246]Vgl. Müller-Terpitz & Knodel, 2017, Rn. 72.

[247]Vgl. Google Inc., 2019; Lück, 2018, S. 81-82; Pinkernell, 2014, S. 64-65.

[248]Diesem stimmt selbst Hruschka, 2019, S. 89 zu.

[249]Vgl. Kapitel 3.2.2. So in ihrer Argumentation wohl auch Diffring und Saft (2019, S. 388).

[250]Unter einem Algorithmus versteht man eine „Steueranweisung, um ein bestimmtes Problem zu bewältigen"; Schmidl, 2014, 15. Algorithmen sind dabei auch zentraler Bestandteil von Computerprogrammen; vgl. hierzu auch die Definitionen zu Computerprogrammen und Software in footnote 120 auf page 486 (Kapitel 3.2.2).

[251]Vgl. Hruschka, 2019, S. 90, 92.

[252]Vgl. Kapitel 3.2.3 in Kapitel 3.

[253]Eine Steuerpflicht auf dieser Grundlage verwerfen auch Diffring & Saft, 2019, S. 388; Linn, 2019, S. 419-420; Reichl & Wiedmann, 2019, S. 205-206.

[254]So auch Diffring & Saft, 2019, S. 388; Reichl & Wiedmann, 2019, S. 205.

[255]Vgl. Pinkernell, 2014, S. 65.

[256]So in ihrer Argumentation auch Linn, 2019, S. 421; Reichl & Wiedmann, 2019, S. 205.

[257]Vgl. BFH-Urteil vom 16.12.1970 – I R 44/67, Tz. 26; zu diesem Ergebnis kommen auch Lück, 2018, S. 81; Pinkernell, 2014, S. 63-64.

[258]Vgl. FG München-Urteil vom 27.05.2013 – 7 K 3552/10.

[259]Vgl. FG München-Urteil vom 27.05.2013 – 7 K 3552/10, Tz. 25.

[260]So auch Diffring & Saft, 2019, S. 391

urheberrechtlich geschützt (§§ 2 Abs. 1, 69a UrhG) und fällt damit, wie bereits in Kapitel 3.2.3 beschrieben, nicht unter den Anwendungsbereich des § 49 Abs. 1 Nr. 9 EStG. Diese Überlassung von Software fiele vielmehr unter § 49 Abs. 1 Nr. 2 Buchst. f EStG. Eine Pflicht zur getrennten Beurteilung einer Software von den darin enthaltenen Algorithmen lässt sich nicht im BMF-Schreiben vom 27.10.2017 erkennen, an das die Finanzverwaltung gebunden ist.[261] Zur steuerrechtlichen Einordnung würde daher das in Kapitel 3.2.2 und 5.2.2 Dargestellte gelten und inländische Einkünfte würden im Ergebnis ausscheiden.

Nutzung von Know-how im Inland

Selbst unter der fälschlichen Annahme, dass Google Irland den Werbetreibenden im Rahmen von Google Ads Know-how in Form von (reinen) Algorithmen überlässt, würde die Begründung inländischer Einkünfte gem. § 49 Abs. 1 Nr. 9 EStG ausscheiden. Dieses Know-how würde nämlich von den Werbetreibenden nicht selbst genutzt werden.[262] Genutzt würde das Know-how von den Werbetreibenden nur dann, wenn sie das überlassene Know-how auch tatsächlich durch bestimmungsgemäße Verwendung selbst zur Lösung von technischen oder betriebswirtschaftlichen Fragen einsetzen würden bzw. ihnen dies zumindest mittelbar ermöglicht würde.[263] Dies ist aber gerade nicht der Fall. Die Werbetreibenden werden durch die (einmalige) Nutzung von Google Ads nicht in die Lage versetzt, in Zukunft selbstständig hochgradig personalisierte Werbung an ihre potenziellen Kunden durch Nutzung der „überlassenen Algorithmen" zu erbringen. Dies könnte von Google Irland zum Schutz des eigenen Geschäftsmodells auch gar nicht gewollt sein.

Steuererhebung im Rahmen der beschränkten Steuerpflicht

Wie aus den vorherigen Ausführungen hervorgegangen ist, erzielt Google Irland durch die Erbringung von Suchmaschinenwerbung im Rahmen von Google Ads keine inländischen Einkünfte i. S. d. § 49 EStG und unterliegt somit nicht der beschränkten Körperschaftsteuerpflicht gem. § 2 Nr. 1 KStG. Damit hat auch grundsätzlich keine Steuererhebung (weder im Rahmen des Steuerabzugs noch im Rahmen der Veranlagung) zu erfolgen. Aber auch bei fälschlicher Annahme des Vorliegens inländischer Einkünfte aus der Überlassung von Rechten bzw. Know-how scheidet zumindest eine Steuererhebung im Rahmen des Steuerabzugs (§ 50a EStG) aus, da es an den dafür notwendigen Tatbestandsvoraussetzungen mangelt. Zu Argumentationszwecken soll im Folgenden dennoch von einer (hypothetischen) Überlassung von Rechten bzw. Know-how bei der Erbringung von Suchmaschinenwerbung durch Google Irland ausgegangen werden.

Liegt eine Rechteüberlassung i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG vor, handelt es sich bei den Werberechten um Rechte, die keine Ausformung in einem Schutzgesetz finden.[264] Somit ist allein schon fraglich, ob diese Rechte überhaupt von der Steuerabzugsnorm des § 50a Abs. 1 Nr. 3 EStG erfasst werden, welche in ihrer beispielhaften Aufzählung nur Rechte nennt, die bestimmten Schutzgesetzen unterfallen.[265] Das BMF vertritt die Ansicht, dass auch weitere vergleichbare, jedoch nicht explizit genannte Rechte der Steuerabzugsnorm unterliegen können, insbesondere dann, wenn sie Ausformung in einem (weiteren nicht von der Norm genannten) Schutzgesetz erfahren.[266] Die von Google Irland überlassenen Werberechte unterliegen aber nicht nur keinem Schutzgesetz i. S. d. § 50a Abs. 1 Nr. 3 EStG, sondern auch keinem anderen Schutzgesetz. Damit ist die Anwendbarkeit des § 50a Abs. 3 EStG insgesamt zweifelhaft.

Eine Rechteüberlassung i. S. d. § 49 Abs. 1 Nr. 2 Buchst. f EStG durch Suchmaschinenwerbung von Google Irland ist aber in jedem Fall verbrauchend und damit als zeitlich unbegrenzte Rechteüberlassung bzw. als Verkauf von Rechten anzusehen.[267] Die von Google vergebenen Werberechte werden für jede einzelne getätigte Suchanfrage jedes einzelnen Suchmaschinennutzers neu vergeben, d. h., die Dauer der Rechteüberlassung beschränkt sich jeweils auf eine einzelne getätigte Suchanfrage. Weiterhin kann Google Irland jede Werbefläche einer Suchanfrage nur einmal vergeben, d. h., dieselbe Werbefläche kann pro Suchanfrage nur an einen Werbetreibenden überlassen werden. Das vergebene Werberecht endet, sobald der Suchmaschinennutzer die Suchanfrage verlässt oder auf die Anzeige klickt. Google Irland überlässt den Werbetreibenden das jeweilige Werberecht für die gesamte Dauer der Suchanfrage und damit für die gesamte Dauer, in der das Recht überhaupt besteht. Nach Beenden der Suchanfrage durch den Suchmaschinennutzer hat sich das Werberecht verbraucht. Google kann das Werberecht einer bereits vergangenen Suchanfrage nicht erneut vergeben. Dem steht auch nicht entgegen, dass Google Irland bei einer neuen Suchanfrage erneut Werbeflächen bereitstellen und neue Werberechte vergeben kann.[268] Dabei handelt es sich jeweils um neue Werbeflächen und -rechte und nicht um die „verbrauchten" Werberechte an Werbeflächen vergangener Suchanfragen.[269] Da § 50a Abs. 1 Nr. 3 EStG aber keine verbrauchenden Rechteüberlassung erfasst, besteht keine Pflicht zum Steuerabzug seitens der Vergütungsschuldner (Werbetreibende).[270] Die Steuererhebung hat damit im Rahmen der Veranlagung zu erfolgen.[271]

---

[261]Vgl. BMF-Schreiben vom 27.10.2017 – IV C 5 - S 2300/12/10003:004.

[262]So auch Linn, 2019, S. 421-422; Reichl & Wiedmann, 2019, S. 205.

[263]Vgl. BFH-Urteil vom 16.12.1970 – I R 44/67, Tz. 26; BFH-Urteil vom 29.09.1987 – X R 17/82, Tz. 13; BFH-Urteil vom 10.04.2013 – I R 22/12, Tz. 12; FG München-Urteil vom 27.05.2013 – 7 K 3552/10, Tz. 24.

[264]Vgl. Diffring & Saft, 2019, S. 389.

[265]Vgl. § 50a Abs. 1 Nr. 3 EStG i. V. m. § 73a Abs. 2-3 EStDV; Schlotter & Hruschka, 2018, S. 680-681; Diffring & Saft, 2019, S. 389; anderer Ansicht Hruschka, 2019, S. 89-90.

[266]Vgl. BMF-Schreiben vom 25.11.2010 – IV C 3 - S 2303/09/10002 2010/0861549, Tz. 22.

[267]So auch Schlotter & Hruschka, 2018, S. 712; Linn, 2019, S. 421; Reichl & Wiedmann, 2019, S. 206-207. Zur Gleichstellung einer verbrauchenden Rechteüberlassung mit einem Rechteverkauf vgl. BFH-Urteil vom 16.05.2001 – I R 64/99, Tz. 18.

[268]Anderer Ansicht Hruschka, 2019, S. 90.

[269]Vgl. Reichl & Wiedmann, 2019, S. 206.

[270]Vgl. Holthaus, 2015, S. 27-28.

[271]Vgl. §§ 31, 32 Abs. 1 KStG.

Wird annahmegemäß durch die Erbringung von Suchmaschinenwerbung Know-how i. S. d. § 49 Abs. 1 Nr. 9 EStG (in Form der Werbeleistung zugrunde liegender Algorithmen) an die Werbetreibenden überlassen, kann dies nur zeitlich unbefristet erfolgen. Eine zeitlich befristete Überlassung von Know-how ist nach Ansicht von Pinkernell überhaupt nur bei körperlich manifestierten Gegenständen wie z. B. bei Plänen und Mustern, nicht hingegen bei der Überlassung von Kenntnissen denkbar.[272] Auch der BFH hat wiederholt angezweifelt, dass die zeitlich befristete Überlassung von Kenntnissen und Erfahrungen aufgrund ihrer Natur überhaupt möglich ist.[273] Es ist nicht ersichtlich, warum das dem Werbetreibenden überlassene Know-how von diesem nur zeitlich begrenzt, etwa für die Dauer des Vertragsverhältnisses, genutzt werden kann. Somit scheidet ein Steuerabzug i. S. d. § 50a Abs. 1 Nr. 3 EStG aus; es hat eine Veranlagung zu erfolgen.

5.3. Zwischenstaatliche Verteilung der Besteuerungsrechte

Aus dem vorherigen Abschnitt geht hervor, dass Deutschland bereits gar kein nationales Besteuerungsrecht an den Umsätzen von Google Irland aus der Erbringung von Suchmaschinenwerbung hat. Aber selbst unter der fälschlichen Annahme, dass doch deutsche Besteuerungsrechte an den Umsätzen von Google Irland nach nationalem Recht bestehen, gibt Deutschland in fast allen Fällen sämtliche Besteuerungsrechte aufgrund des mit Irland geschlossenen Doppelbesteuerungsabkommens (DBA DEU-IRL) auf.

Unabhängig davon, unter welche Verteilungsnorm die hypothetischen Umsätze fallen, bleiben deutsche Besteuerungsrechte an diesen nur dann erhalten, wenn sie im Rahmen einer deutschen Betriebsstätte von Google Irland erwirtschaftet werden. Das DBA DEU-IRL weicht innerhalb der relevanten Verteilungsnormen (Art. 7, 12, 13 DBA DEU-IRL) nicht von den Bestimmungen des OECD-MA ab.[274] Daher gilt das in Kapitel 4.2 Beschriebene analog auch für das DBA DEU-IRL. Aufgrund einer fehlenden Betriebsstätte[275] gibt Deutschland gem. Art. 7 Abs. 1 DBA DEU-IRL die Besteuerungsrechte an Unternehmensgewinnen auf. Deutschland wird zudem gem. Art. 12 DBA DEU-IRL keine Möglichkeit zur Quellenbesteuerung von Lizenzgebühren ermöglicht. Weiterhin erhält Deutschland gem. Art. 13 Abs. 5 DBA DEU-IRL ohne Vorliegen einer Betriebsstätte keine Besteuerungsrechte an Veräußerungsgewinnen.

Im Ergebnis scheitern deutsche Besteuerungsrechte selbst bei hypothetisch vorliegendem nationalen Besteuerungsrecht unabhängig von der tatsächlichen Zuordnung inländischer Einkünfte zu den verschiedenen Verteilungsnormen an den Regelungen des DBA DEU-IRL. Deutschland verbleibt im Ergebnis kein Besteuerungsrecht an den Umsätzen, die

Google Irland durch Suchmaschinenwerbung in Deutschland erwirtschaftet.

6. Zwischenfazit: Status quo der Besteuerung digitaler Werbeleistungen in Deutschland

Die vorhergehenden Ausführungen haben den Status quo der Besteuerung digitaler Werbeleistungen in Deutschland im Inboundfall dargestellt. Dabei ist festzustellen, dass die derzeitigen nationalen Steuergesetze und bilateralen Doppelbesteuerungsabkommen in Ermangelung einer inländischen Betriebsstätte nicht in der Lage sind, digitale Werbeleistungen, welche von im Ausland ansässigen Körperschaften mithilfe des Internets in Deutschland erbracht werden, steuerlich zu erfassen.

Wie in Kapitel 3 deutlich wurde, sind die derzeit bestehenden deutschen Regelungen des nationalen Steuerrechts nur in speziellen Fällen in der Lage, überhaupt einen hinreichenden steuerlichen Anknüpfungspunkt (steuerlicher Nexus, genuine link) und damit ein deutsches Besteuerungsrecht an digitalen Werbeleistungen im Inboundfall herzustellen. Dies liegt daran, dass das deutsche Steuerrecht keine eigenen Tatbestände zur Erfassung digitaler Werbeleistungen innerhalb der inländischen Einkünfte des § 49 EStG enthält. Somit müssen diese unter die bestehenden Tatbestände, insbesondere gewerbliche Einkünfte durch Betriebsstätte oder ständigen Vertreter (§ 49 Abs. 1 Nr. 2 Buchst. a EStG), gewerbliche betriebsstättenlose Einkünfte (§ 49 Abs. 1 Nr. 2 Buchst. f EStG) und sonstige Einkünfte (§ 49 Abs. 1 Nr. 9 EStG), subsumiert werden.

Darüber hinaus hat Kapitel 4 gezeigt, dass Deutschland in DBA-Fällen in den meisten Fällen ohnehin alle bestehenden Besteuerungsrechte an digitalen Werbeleistungen aufgibt. Somit ist nationalen Besteuerungsrechten insgesamt nur wenig Bedeutung beizumessen. Diese bleiben nur für Gewinne aus der Erbringung digitaler Werbeleistungen erhalten, die im Rahmen einer inländischen Betriebsstätte erwirtschaftet werden.

Kapitel 5 hat für den marktführenden Anbieter der umsatzstärksten Werbeform digitaler Werbeleistungen in Deutschland, die Suchmaschinenwerbung durch Google im Rahmen des Google Ads Programms, weiterhin gezeigt, dass bereits die nationalen Tatbestände nicht in der Lage sind digitale Werbeleistungen in Deutschland der Besteuerung zu unterwerfen. Google Irland, welche die digitalen Werbeleistungen Googles in Europa vertreibt, unterhält weder eine Betriebsstätte noch einen ständigen Vertreter im Inland, sodass gewerbliche Einkünfte durch Betriebsstätte oder ständigen Vertreter ausscheiden. Suchmaschinenwerbung kann auch nicht als Überlassung von Rechten oder Know-how angesehen werden, wodurch auch die anderen beiden möglichen Anknüpfungspunkte ausscheiden. Im Ergebnis besteht also schon keine Besteuerungsgrundlage nach nationalem Recht. Selbst im hypothetischen Fall, dass Google Irland in Deutschland doch beschränkt steuerpflichtige Einkünfte erwirtschaftet, könnte Deutschland die Besteuerungsrechte an

---

[272]Vgl. Pinkernell, 2014, S. 63.

[273]Vgl. etwa BFH-Urteil vom 04.03.1970 – I R 140/66, Tz. 23; BFH-Urteil vom 04.03.1970 – I R 86/69, Tz. 12; BFH-Urteil vom 27.04.1977 – I R 211/74, Tz. 21. Diese Ansicht vertreten auch Reimer, 2018a, § 49 EStG, Rz. 311; Gosch, 2018a, § 49 EStG, Rn. 94.

[274]Vgl. Art. 7, 12, 13 DBA DEU-IRL; Art. 7, 12, 13 OECD-MA.

[275]Vgl. hierzu die Abbildung im Anhang.

diesen, in Ermangelung einer inländischen Betriebsstätte, aufgrund des mit Irland bestehenden DBA nicht behaupten.

An dieser Stelle sei angemerkt, dass durch die Betrachtung eines Einzelfalls nur begrenzt Rückschlüsse auf die steuerliche Beurteilung von Werbeleistungen (1) anderer Anbieter, (2) anderer Werbeformen bzw. (3) einer Kombination beider Vorhergenannten möglich sind. Dennoch sollten die Ergebnisse des Kapitel 5 zumindest auf die Leistungen anderer Anbieter von Suchmaschinenwerbung (d. h. gleicher Werbeform) übertragbar sein, da hier die Leistungserbringung ähnlich funktioniert. Auch hier sollten bei fehlender Betriebsstätte im Inland keine inländischen Einkünfte und damit keine Besteuerungsrechte in Deutschland bestehen. Weiterhin sollten die vorgebrachten Argumente auch in ähnlicher Weise bei der Erbringung von Bannerwerbung gelten, da der BGH die der Platzierung von Werbebannern zugrunde liegenden Verträge als Werkverträge einordnet.[276] So verneinen z. B. Diffring und Saft (2019, S. 392-394) auf Grundlage von Argumenten, welche denen in Kapitel 5 sehr ähnlich sind, auch eine beschränkte Steuerpflicht aufgrund der Erbringung von Bannerwerbung.

Dem Vorstoß aus Teilen der Finanzverwaltung ist damit im Fall von Suchmaschinenwerbung (und wohl auch Bannerwerbung) nicht zu folgen. Die bestehenden steuerlichen Rahmenbedingungen sind in Ermangelung einer Betriebsstätte nicht in der Lage, im Inboundfall erbrachte digitale Werbeleistungen in Deutschland der Besteuerung zu unterwerfen. Damit bestehen auch keine Steuerabzugsverpflichtungen für Unternehmen, die solche Werbeleistungen in Anspruch genommen haben bzw. nehmen. Diese Unternehmen sind keinen Haftungsrisiken ausgesetzt. Mittlerweile hat dies auch das Bayerische Staatsministerium der Finanzen und für Heimat bestätigt, nach dessen Aussage keine Steuerabzugsverpflichtungen im Rahmen von digitalen Werbeleistungen (im Allgemeinen) bestehen.[277]

Unter der Annahme, dass durch die Erbringung digitaler Werbeleistungen jedoch Wertschöpfung im Inland stattfindet, scheint damit auch in Deutschland Handlungsbedarf gegeben die bestehenden Regelungen anzupassen, um die Besteuerung grenzüberschreitender digitaler Werbeleistungen besser mit der Wertschöpfung im Inland in Einklang zu bringen und die in Kapitel 2 vorgestellte steuerliche Grundproblematik zu lösen.

7. Mögliche Handlungsalternativen auf Ebene der Europäischen Union zur Sicherstellung deutscher Besteuerungsrechte bei digitalen Werbeleistungen

Nachdem, wie vorhergehend dargestellt, die Besteuerung grenzüberschreitender digitaler Werbeleistungen im Inboundfall in Deutschland derzeit nicht sichergestellt werden kann, sollen nachfolgend die Richtlinienentwürfe der EU zur Sicherstellung der sachgerechten Besteuerung grenzüberschreitender digitaler Transaktionen aus digitalen Geschäftsmodellen vorgestellt werden. Im Hinblick auf die Forschungsfrage wird geprüft, inwieweit diese Handlungsalternativen geeignet erscheinen, deutsche Besteuerungsrechte an digitalen Werbeleistungen sicherzustellen, welche von im Ausland ansässigen Körperschaften mithilfe des Internets im Inland erbracht werden.

Hierzu werden die Handlungsalternativen zunächst vorgestellt. Anschließend wird geprüft, ob durch diese Handlungsalternativen auch digitale Werbeleistungen steuerlich erfasst werden könnten. Nachfolgend wird auf deren wesentliche Kritikpunkte eingegangen. Da es hierbei nicht zielführend erscheint, die Kritikpunkte allein auf digitale Werbeleistungen zu beziehen, werden die Handlungsalternativen in ihrer Gänze, d. h., allgemein im Kontext digitaler Transaktionen, betrachtet. Abschließend wird auf der Basis dieser Kritikpunkte eine Bewertung darüber abgegeben, ob die Handlungsalternativen die in Kapitel 2 beschriebenen steuerlichen Probleme lösen und damit deren Implementierung empfohlen werden kann.

Auf Ebene der EU stehen derzeit zwei verschiedene Richtlinienentwürfe im Mittelpunkt der Diskussion. Zum einen schlägt die EU eine Ausgleichsbesteuerung in Form einer Digitalsteuer vor (Kapitel 7.1).[278] Diese soll als Interimsmaßnahme bestimmte digitale Transaktionen der Besteuerung im Quellenstaat unterwerfen, bis eine langfristige Lösung gefunden wird. Diese langfristige Lösung zur Sicherstellung von Besteuerungsrechten liegt nach Ansicht der EU in der Schaffung eines neuen steuerlichen Nexus aufgrund einer signifikanten digitalen Präsenz (Kapitel 7.2).[279]

7.1. Richtlinienentwurf zur Digitalsteuer

Die europäische Digitalsteuer (engl. digital services tax, kurz DST) basiert auf der Idee einer Ausgleichsbesteuerung (engl. equalisation levy), welche im Jahr 2015 von der OECD im Rahmen des Abschlussberichts zu BEPS Aktionspunkt 1 vorgestellt wurde.[280] Diese Ausgleichsbesteuerung der OECD zielt darauf ab, bestimmte Sachverhalte, meist Dienstleistungen, die von nicht ansässigen Unternehmen auf digitalem Wege erbracht werden, der Besteuerung im Quellenstaat zu unterwerfen.[281] Dieser Grundgedanke wurde von der EU aufgegriffen und bildet die Grundlage für einen Richtlinienentwurf zur kurzfristigen Einführung einer Digitalsteuer auf „Erträge aus der Erbringung bestimmter digitaler Dienstleistungen"[282] (DST-RL-E). Die Digitalsteuer soll Erbringer digita-

---

[276]Vgl. BGH-Urteil vom 22.03.2018 – VII ZR 71/17, Tz. 1-4, 12.

[277]Vgl. Bayerisches Staatsministerium der Finanzen und für Heimat, 2019. Allerdings wurde hierbei offengelassen, ob es bereits an der beschränkten Steuerpflicht der Erbringer digitaler Werbeleistungen in Deutschland oder nur an der Verpflichtung zum Steuerabzug scheitert.

[278]Vgl. DST-RL-E; Rat der Europäischen Union, 2018; Rat Wirtschaft und Finanzen, 2018.

[279]Vgl. SDP-RL-E.

[280]Für die ursprüngliche Idee einer Ausgleichsbesteuerung vgl. OECD, 2015, Tz. 302.

[281]Vgl. OECD, 2015, Tz. 302-303. Die Einführung einer solchen Ausgleichsbesteuerung wurde von der OECD jedoch nicht empfohlen, vgl. OECD, 2015, Tz. 357.

[282]DST-RL-E, Titelblatt.

ler Dienstleistungen treffen, die in einem Mitgliedsstaat der EU über keine physische Präsenz verfügen, jedoch in diesem Staat wertschöpfend tätig werden.[283] Als Zwischenlösung angedacht soll die Digitalsteuer eine Besteuerung solcher Dienstleistungen sicherstellen, bis eine langfristige Lösung implementiert wird.[284]

7.1.1. Beschreibung des Richtlinienentwurfs

Der Richtlinienentwurf der EU zur Digitalsteuer vom 21.03.2018 erfasst juristische Personen und Konstrukte, welche Erträge aus der Erbringung bestimmter digitaler Dienstleistungen gegenüber Nutzern[285] erzielen.[286] Diese Dienstleistungen umfassen (1) die Platzierung digitaler Werbung auf digitalen Schnittstellen, (2) die Bereitstellung mehrseitiger digitaler Schnittstellen und (3) die Übermittlung von Nutzerdaten, welche auf digitalen Schnittstellen gesammelt werden.[287] Unter einer digitalen Schnittstelle wird dabei jede Art von Software verstanden, was auch Webseiten und sonstige Anwendungen mit einschließt.[288]

Von der DST werden nur solche Erbringer digitaler Dienstleistungen erfasst, deren jährliche Gesamterträge weltweit 750 Mio. EUR und deren Erträge aus diesen Dienstleistungen innerhalb der EU 50 Mio. EUR übersteigen.[289] Weiterhin werden nur Erträge aus der Erbringung solcher Dienstleistungen steuerlich erfasst, welche innerhalb der EU erwirtschaftet werden. Als in der EU erwirtschaftet gelten die Erträge immer dann, wenn der jeweilige Nutzer der Dienstleistung in der EU ansässig ist.[290] Ein Nutzer gilt für Zwecke der Digitalsteuer dabei in dem Staat als ansässig, von dem aus er auf die digitale Schnittstelle zugreift.[291]

Sind die Voraussetzungen erfüllt, unterliegen die in der EU erzielten Bruttoerträge aus diesen Dienstleistungen einem Steuersatz von 3 %.[292] Die Steuer wird anteilig in den Mitgliedsstaaten der EU fällig, in denen die steuerbaren Erträge erwirtschaftet werden.[293]

Der Richtlinienentwurf vom 21.03.2018 soll ab 01.01.2020 in Kraft treten.[294]

Ein französisch-deutscher Kompromissvorschlag vom 04.12.2018 schlägt eine reduzierte Form der Digitalsteuer vor, deren Anwendbarkeit alleinig auf digitale Werbeleistungen begrenzt ist.[295] Diese reduzierte Digitalsteuer soll ab 01.01.2021 in Kraft treten und im Jahr 2025 außer Kraft gesetzt werden. Weiterhin soll diese Steuer nicht in Kraft treten, wenn bis dahin eine internationale Lösung (auf Ebene der OECD) gefunden wurde.[296]

7.1.2. Möglichkeit zur Erfassung digitaler Werbeleistungen

Sowohl der Richtlinienentwurf vom 21.03.2018 als auch der französisch-deutsche Vorschlag vom 04.12.2018 erfassen per Definition digitale Werbeleistungen. Durch die Implementierung des Richtlinienentwurfs könnte Deutschland damit die Besteuerung von im Inland erbrachten digitalen Werbeleistungen nicht ansässiger Körperschaften sicherstellen, sofern diese die Ertragsschwellen überschreiten.

Im Hinblick auf den deutschen Gesamtmarkt für digitale Werbeleistungen in Höhe von ca. 7,2 Mrd. EUR[297] im Jahr 2018 könnte die Digitalsteuer damit zu einem Steueraufkommen von maximal 216 Mio. EUR führen. Das tatsächliche Steueraufkommen sollte realisticherweise aber wohl niedriger ausfallen, da nicht davon ausgegangen werden kann, dass alle Umsätze aus digitalen Werbeleistungen zur Gänze von nicht ansässigen Unternehmen erwirtschaftet werden, welche die Voraussetzungen der Digitalsteuer erfüllen.[298]

7.1.3. Kritik am Richtlinienentwurf

Die Einführung einer solchen Digitalsteuer ist jedoch problematisch und wird vielfach kritisiert.[299] Neben rechtlichen Fragen wie der Vereinbarkeit einer Digitalsteuer in der vorgeschlagenen Ausgestaltung mit EU-Recht[300] und deren tatsächlichem Anwendungsbereich[301] ist eine Digitalsteuer vor allem mit ökonomischen Problemen verbunden.

Zunächst ist zweifelhaft, inwiefern die Digitalsteuer am Prinzip der Besteuerung von Gewinnen am Ort der Wertschöpfung ausgerichtet ist. Es wird pauschal die Vermutung aufgestellt, dass durch bestimmte digitale Dienstleistungen Wertschöpfung in dem Staat stattfindet, in dem die Nutzer der Dienstleistung ansässig sind.[302] Dabei findet auch keine Bewertung der Höhe dieser angenommenen Wertschöpfung statt. So wird beispielsweise keine Bewertung des Beitrages von Daten und Nutzern (vgl. Kapitel 2.3) zur Wertschöpfung für die konkrete Dienstleistungserbringung vorgenommen. Auch die Höhe der Wertschöpfung wird vielmehr pauschal

---

[283]Vgl. DST-RL-E, Begründung S. 2-3, 8.

[284]Vgl. DST-RL-E, Begründung S. 4.

[285]Nutzer kann dabei eine natürliche Person oder ein Unternehmen sein, vgl. Art. 2 Nr. 4 DST-RL-E.

[286]Vgl. DST-RL-E, Begründung S. 8.

[287]Vgl. Art. 3 Abs. 1 DST-RL-E.

[288]Vgl. Art. 2 Abs. 3 DST-RL-E.

[289]Vgl. Art. 3, 4 DST-RL-E.

[290]Vgl. Art. 5 Abs. 1 DST-RL-E.

[291]Vgl. Art. 5 Abs. 2 DST-RL-E.

[292]Vgl. Art. 8 DST-RL-E.

[293]Vgl. Art. 6 DST-RL-E.

[294]Vgl. Art. 25 DST-RL-E.

[295]Vgl. Rat Wirtschaft und Finanzen, 2018.

[296]Vgl. Rat Wirtschaft und Finanzen, 2018.

[297]Vgl. Statista, 2018.

[298]Im Vergleich hierzu schätzen z. B. Fuest, Meier und Neumeier (2018, S. 23-27) das deutsche Steueraufkommen der Digitalsteuer aus allen erfassten Dienstleistungen auf maximal ca. 836 Mio. EUR und damit insgesamt nur auf ca. 0,1 % des deutschen Gesamtsteueraufkommens.

[299]So z. B. einer Ausgleichsbesteuerung in Form einer Digitalsteuer kritisch gegenüberstehend Kofler, Mayr & Schlager, 2017; Becker, 2018; Dorenkamp, 2018; Eilers & Oppel, 2018; Fuest et al., 2018; Mason & Parada, 2018; Pinkernell, 2018; Schanz & Sixt, 2018; Schön, 2018; Sheppard, 2018; Spengel, 2018; Valta, 2018; Zöller, 2018; Wissenschaftlicher Beirat beim Bundesministerium der Finanzen, 2018; Kokott, 2019; Wünnemann, 2019; für einen umfassenden Überblick über das Stimmungsbild zur einer Ausgleichsbesteuerung im Allgemeinen siehe auch Fritz, Schanz & Siegel, 2018.

[300]Hierzu kritisch z. B. Mason & Parada, 2018; Valta, 2018; Sheppard, 2018, S. 2.

[301]Hierzu ausführlich Schanz & Sixt, 2018.

[302]Vgl. DST-RL-E, Begründung S. 8; hierzu kritisch z. B. Roderburg (2018, S. 255).

angenommen und mit 3 % des Bruttoertrags der Besteuerung unterworfen.

Darüber hinaus sind die Ertragsschwellen zur Begründung der Steuerpflicht zu kritisieren. Diese erscheinen zum einen willkürlich gewählt[303] und begünstigen „kleinere" gegenüber „größeren" Unternehmen.[304] Durch die Ausgestaltung als Freigrenze führen sie zum anderen zu Belastungssprüngen an der Besteuerungsgrenze.[305] So beträgt die Steuerbelastung eines hypothetischen Unternehmens A, welches in der EU potenziell steuerpflichtige Erträge i. H. v. 49.999.999,99 EUR erwirtschaftet, 0 EUR. Dem gegenüber wird ein hypothetisches Unternehmen B, welches in der EU potenziell steuerpflichtige Erträge i. H. v. 50.000.000,01 EUR erwirtschaftet, mit 1,5 Mio. EUR Digitalsteuer belastet. Dies lässt Ausweichreaktionen von Unternehmen erwarten, welche die Ertragsschwellen der Digitalsteuer knapp über- bzw. unterschreiten.[306]

Weiterhin ist die ertragsunabhängige Struktur der vorgeschlagenen Digitalsteuer problematisch.[307] Da die Steuer vom Bruttoertrag und nicht vom Gewinn berechnet wird, durchbricht sie das objektive Nettoprinzip.[308] Damit unterliegen Unternehmen mit höherer Rendite im Vergleich zu Unternehmen mit niedrigerer Rendite relativ gesehen einer geringeren effektiven Steuerbelastung.[309] So liegt die effektive Gewinnsteuerbelastung aus der Digitalsteuer bei einer Umsatzrendite von 5 % bei 60 %, wohingegen diese bei einer Umsatzrendite von 20 % nur 15 % beträgt.[310] Ob dies ein gewünschter Effekt ist, darf angezweifelt werden. Des Weiteren steigt dadurch die effektive Steuerbelastung in Zeiten sinkender Profitabilität (z. B. in konjunkturellen Schwächephasen), was im Extremfall sogar dazu führen kann, dass die Digitalsteuer wie eine Substanzsteuer wirkt.[311] Dies wäre immer dann der Fall, wenn die Umsatzrendite kleiner als der Digitalsteuersatz ausfällt.

Des Weiteren erhöht sich durch die geplante Digitalsteuer das Risiko einer Doppelbesteuerung.[312] Die erfassten digitalen Dienstleistungen unterliegen neben der Digitalsteuer regelmäßig auch im Ansässigkeitsstaat der betroffenen Unternehmen der Ertragsbesteuerung. Damit würden die Gewinne eines Unternehmens aus der Erbringung digitaler Dienstleistungen mehrfach besteuert. Gleichzeitig ist jedoch fraglich, ob die Digitalsteuer von derzeit geltenden DBA erfasst wird, die eine solche Doppelbesteuerung vermeiden bzw. vermindern könnten.[313]

Nicht zuletzt könnte die DST auch zu wirtschaftspolitischen Konflikten beitragen.[314] Von der Steuer würden hauptsächlich US-amerikanische Unternehmen betroffen sein.[315] Die Digitalsteuer könnte damit als Importzoll auf Dienstleistungen US-amerikanischer Unternehmen interpretiert werden und zu vergleichbaren Gegenmaßnahmen der USA gegenüber der EU führen.[316] So wurden vonseiten der USA z. B. bereits entsprechende Reaktionen auf die Einführung einer unilateralen Digitalsteuer in Frankreich diskutiert.[317]

Schlussendlich wäre die Durchsetzung von aus der DST resultierenden Steueransprüchen schwierig.[318] Dem Quellenstaat, der die DST erhebt, dürfte der Zugriff auf den Steuerpflichtigen in Ermangelung physischer Verbindungen im Inland regelmäßig schwerfallen. Der Richtlinienentwurf äußert sich jedoch nicht dazu, wie diesem Problem des Steuervollzugs begegnet werden soll.

7.1.4. Bewertung des Richtlinienentwurfs

Abschließend lässt sich festhalten, dass die Einführung einer Digitalsteuer zwar grundsätzlich dazu in der Lage wäre, digitale Werbeleistungen in Deutschland der Besteuerung zu unterwerfen. Insofern würde sie mit Hinblick auf die in Kapitel 2.3 erarbeitete Grundproblematik bei der Besteuerung digitaler Werbeleistungen das Problem des fehlenden steuerlichen Nexus in Ermangelung einer physischen Betriebsstätte lösen.

Aufgrund der vorangegangenen Kritikpunkte kann nach hier vertretener Ansicht die Einführung einer Digitalsteuer jedoch weder in Form des Richtlinienentwurfs vom 21.03.2018 noch in Form des französisch-deutschen Kompromissvorschlags vom 04.12.2018 empfohlen werden. Beide Vorschläge würden bei einem geringen Steueraufkommen zu erheblichen rechtlichen und ökonomischen Problemen führen. Durch die Digitalsteuer wäre im Ergebnis auch weder eine sachgerechte Bewertung des Wertschöpfungsbeitrags durch Daten und Nutzer gewährleistet (vgl. Kapitel 2.3) noch würde sie sich ohne Weiteres in das bestehende internationale Steuerrecht (vgl. Kapitel 2.2) eingliedern lassen.

---

[303]Vgl. Fuest et al., 2018, S. 17.

[304]Vgl. Kokott, 2019, S. 127.

[305]Vgl. Fuest et al., 2018, S. 17.

[306]Vgl. Fuest et al., 2018, S. 17.

[307]Vgl. Fuest et al., 2018, IV; Kokott, 2019, S. 128.

[308]Vgl. hierzu auch die Ausführungen in Kapitel 3.3.1.

[309]Vgl. Dorenkamp, 2018, S. 641; Fuest et al., 2018, IV.

[310]Berücksichtigt man zusätzlich dazu noch die steuerliche Belastung durch weitere Ertragsteuern, liegt die gesamte effektive Gewinnsteuerbelastung wesentlich höher. Schanz und Sixt (2018) errechnen für Unternehmen mit einer Umsatzrendite von 5 % bei einer zusätzlichen 30 %igen Ertragsbesteuerung eine gesamte effektive Gewinnsteuerbelastung von 90 %, sofern die Digitalsteuer nicht von der ertragsteuerlichen Bemessungsgrundlage abziehbar ist und 72 % falls eine solche Abzugsmöglichkeit besteht (S. 1989). Bei einer Umsatzrendite von 20 % läge die gesamte effektive Gewinnsteuerbelastung ceteris paribus bei 45 % bzw. respektive 40,5 %.

[311]Vgl. Kokott, 2019, S. 128; Dorenkamp, 2018, S. 641; Fuest et al., 2018, S. 18; Roderburg, 2018, S. 254.

[312]Vgl. Spengel, 2018, S. M4; Kokott, 2019, S. 129; Wissenschaftlicher Beirat beim Bundesministerium der Finanzen, 2018, S. 2-3.

[313]Vgl. Wissenschaftlicher Beirat beim Bundesministerium der Finanzen, 2018, S. 2-3; Kokott, 2019, S. 129.

[314]Vgl. Fuest et al., 2018, S. 28-30; Kokott, 2019, S. 129; Roderburg, 2018, S. 256; Wissenschaftlicher Beirat beim Bundesministerium der Finanzen, 2018, S. 5; Zöller, 2018, S. 2905.

[315]Fuest et al. (2018, S. 28-30) schätzen etwa, dass 50 % der digitalsteuerpflichtigen Erträge von US-amerikanischen Unternehmen erwirtschaftet würden. Somit würden diese Unternehmen auch den Großteil der Steuerlast tragen.

[316]Vgl. Fuest et al., 2018, S. 28.

[317]Vgl. United States House Committee on Ways and Means, 2019.

[318]Vgl. Kokott, 2019, S. 130; Kelm & Müller, 2018, S. 594; Roderburg, 2018, S. 255. Zu den Problemen des Steuervollzugs bei fehlender physischer Präsenz sei an dieser Stelle auch auf Kapitel 3.3 zurückverwiesen.

Auch auf politischer Ebene konnte die Digitalsteuer bisher nicht überzeugen. Mehrere EU-Mitgliedsstaaten, darunter Schweden, Irland, Dänemark und Finnland, lehnen die Einführung der Digitalsteuer ab.[319] Da für steuerliche Maßnahmen auf Ebene der EU jedoch Einstimmigkeit der Mitgliedsstaaten vorausgesetzt wird,[320] erscheint die Einführung einer Digitalsteuer in der EU zweifelhaft. Sowohl der Richtlinienentwurf vom 21.03.2018 als auch der französisch-deutsche Vorschlag vom 04.12.2018 wurde auf Ebene der EU bisher abgelehnt.[321]

7.2. Richtlinienentwurf zur signifikanten digitalen Präsenz

Am 21.03.2018 hat die EU weiterhin eine langfristige Lösung zur Besteuerung digitaler Geschäftsmodelle in Form eines neuen steuerlichen Anknüpfungspunktes auf Grundlage einer signifikanten digitalen Präsenz (engl. significant digital presence, kurz SDP) vorgestellt.[322] Diese Idee basiert dabei auf dem Konzept der signifikanten ökonomischen Präsenz (engl. significant economic presence, kurz SEP), welches von der OECD in ihrem Abschlussbericht zu BEPS Aktionspunkt 1 vorgestellt wurde.[323] Das Konzept der SEP der OECD würde einen steuerlichen Anknüpfungspunkt im Quellenstaat schaffen, wenn ein nicht ansässiges Unternehmen durch eine signifikante ökonomische Präsenz mithilfe von Informations- und Kommunikationstechnologien oder automatisierten Verfahren zielgerichtet und nachhaltig in diesem Staat tätig wird.[324] An diesen Gedanken knüpft die EU an und verfolgt mit dem Richtlinienentwurf zur SDP das Ziel, einen steuerlichen Anknüpfungspunkt für „grenzüberschreitend tätige digitale Unternehmen ohne physische Präsenz"[325] zu schaffen. Diesen neuen steuerlichen Anknüpfungspunkt versteht die EU als Ergänzung des bestehenden Betriebsstättenbegriffs.[326] Damit sollen nicht ansässige Unternehmen in Mitgliedsstaaten der EU der Besteuerung unterworfen werden, wenn sie in diesen (ohne physische Präsenz) auf digitalem Wege wirtschaftlich tätig werden und gleichzeitig Wertschöpfung in diesen Mitgliedsstaaten stattfindet.[327]

7.2.1. Beschreibung des Richtlinienentwurfs

Der Richtlinienentwurf erfasst die Erbringer (automatisierter) digitaler Dienstleistungen.[328] Darunter werden Dienstleistungen verstanden, „[...] die über das Internet oder ein elektronisches Netzwerk erbracht werden, [und] deren Erbringung aufgrund ihrer Art im Wesentlichen automatisiert und nur mit minimaler menschlicher Beteiligung erfolgt

[...]"[329]. Zur Abgrenzung werden in einem nicht abschließenden Positivkatalog beispielhaft 31 von der Richtlinie zu erfassende digitale Dienstleistungen aufgelistet.[330] Weiterhin werden in einem Negativkatalog verschiedene Dienstleistungen genannt, die dem Richtlinienentwurf nicht unterliegen sollen.[331]

Der Geltungsbereich des Richtlinienentwurfs ist weiterhin auf bestimmte Erbringer digitaler Dienstleistungen begrenzt. Zum einen soll sie nur Erbringer digitaler Dienstleistungen erfassen, die entweder (1) in einem Mitgliedsstaat der EU ansässig sind oder (2) eine SDP in einem Mitgliedsstaat der EU unterhalten und in einem Staat ansässig sind, mit dem dieser EU-Mitgliedsstaat kein DBA unterhält.[332] Zum anderen wird eine signifikante digitale Präsenz in einem Quellenstaat nur dann bejaht, wenn der Erbringer digitaler Dienstleistungen dort im jeweiligen Steuerjahr mindestens eine der drei nachfolgenden Bedingungen erfüllt:

- Gesamterträge übersteigen 7 Mio. EUR ODER
- Anzahl der Nutzer übersteigt 100 Tsd. ODER
- Anzahl der abgeschlossenen Verträge übersteigt 3.000

Liegt in einem Quellenstaat eine SDP vor, werden dieser alle mit ihr im Zusammenhang stehenden Gewinne unter Beachtung des Fremdvergleichsgrundsatzes anhand von Verrechnungspreisen zugerechnet.[333] Die Gewinnzuordnung soll sich dabei am sog. „Authorised OECD Approach" (AOA) orientieren und der signifikanten digitalen Präsenz unter Berücksichtigung der übernommenen (Personal-)Funktionen und Risiken einen Gewinn zuordnen, den ein separates, unabhängiges Unternehmen unter gleichen Umständen erwirtschaftet hätte.[334] Hierfür sollen laut Richtlinienentwurf zusätzlich auch die ausgeübten Funktionen der SDP (z. B. Erhebung von Daten und nutzergenerierten Inhalten, Verkauf von Online-Werbeflächen, Bereitstellung von Inhalten Dritter) Beachtung finden.[335] Standardmäßig sollen die Erbringer digitaler Dienstleistungen zur Gewinnzuordnung die Gewinnaufteilungsmethode anwenden, es sei denn, eine andere anerkannte Verrechnungspreismethode ist nachweislich geeigneter.[336] Die der SDP so zugeordneten Gewinne sind ausschließlich in dem Quellenstaat zu besteuern, in dem die SDP besteht.[337]

Die Richtlinie soll ab dem 01.01.2020 in Kraft treten.[338]

---

[319]Vgl. Ueberbach, 2019.
[320]Vgl. Paternoster, 2018.
[321]Vgl. Zeit Online, 2018; Süddeutsche Zeitung, 2019; Ueberbach, 2019.
[322]Vgl. SDP-RL-E.
[323]Vgl. OECD, 2015, Tz. 277-291.
[324]Vgl. OECD, 2015, Tz. 277; OECD, 2019a, Tz. 51.
[325]SDP-RL-E, Begründung S. 3.
[326]Vgl. Art. 1 SDP-RL-E.
[327]Vgl. SDP-RL-E, Begründung S. 2, 3.
[328]Vgl. Art. 3 Nr. 5 SDP-RL-E.

[329]Art. 3 Nr. 5 SDP-RL-E.
[330]Vgl. Art. 3 Nr. 5 SDP-RL-E.
[331]Vgl. Art. 3 Nr. 5 i. V. m. Anhang III zum Richtlinienentwurf (vgl. Europäische Kommission, 2018a, S. 4).
[332]Vgl. SDP-RL-E, Begründung S. 8.
[333]Vgl. Art. 5 Abs. 1, 2 SDP-RL-E; SDP-RL-E, Begründung, S. 10-11.
[334]Vgl. SDP-RL-E, Begründung S. 10; zum AOA sei an dieser Stelle auch auf OECD (2010) verwiesen.
[335]Vgl. Art. 5 Abs. 5 SDP-RL-E; SDP-RL-E, Begründung S. 10-11.
[336]Vgl. Art. 5 Abs. 6 SDP-RL-E.
[337]Vgl. Art. 5 Abs. 1 SDP-RL-E.
[338]Vgl. Art. 9 Abs. 1 SDP-RL-E.

7.2.2. Möglichkeit zur Erfassung digitaler Werbeleistungen

Der Richtlinienentwurf zur signifikanten digitalen Präsenz vom 21.03.2018 erfasst auch die Erbringung digitaler Werbeleistungen. Die Bereitstellung von digitalen Werbeplätzen ist explizit im Positivkatalog genannt.[339] Durch die Implementierung könnte Deutschland damit Besteuerungsrechte an im Inland erbrachten digitalen Werbeleistungen bestimmter nicht ansässiger Körperschaften sicherstellen. Diese Körperschaften müssten dafür aber entweder in der EU oder in einem Drittstaat, mit dem Deutschland kein DBA unterhält, ansässig sein und die weiteren Größenbedingungen erfüllen.

7.2.3. Kritik am Richtlinienentwurf

Allerdings ist auch das Konzept der signifikanten digitalen Präsenz mit verschiedenen Kritikpunkten verbunden, welche die Einführung eines solchen Konzeptes und deren Umsetzung schwierig gestalten.

Zunächst ist die Begrenzung einer signifikanten digitalen Präsenz auf digitalisierte Geschäftsmodelle problematisch. Die SDP der EU ist eindeutig auf digitale Dienstleistungen begrenzt.[340] Dies steht in klarem Widerspruch zur Ansicht der OECD, dass eine Abgrenzung der digitalen Wirtschaft vom Rest der Wirtschaft für steuerliche Zwecke nicht möglich ist.[341] Dem Grundgedanken der OECD folgend sollte eine SDP demnach nicht zwischen digitalen und analogen Geschäftstätigkeiten unterscheiden. Eine getrennte Betrachtung von digitalisierten und analogen Geschäftsmodellen würde zudem zu Zuordnungsschwierigkeiten führen, wenn Unternehmen sowohl vom bestehenden Konzept der physischen Betriebsstätte als auch vom Konzept der SDP erfasst werden.[342]

Weiterhin sind die gewählten Schwellenwerte zur Begründung der SDP zu kritisieren. Vorausgesetzt man folgt dem Prinzip, dass Gewinne am Ort der Wertschöpfung besteuert werden sollten, müssen die Tatbestände so gewählt sein, dass sie überhaupt nur Geschäftstätigkeiten nicht ansässiger Unternehmen erfassen, bei denen auch tatsächlich Wertschöpfung im Quellenstaat stattfindet. Wenn keine Wertschöpfung im Quellenstaat stattfindet, sollte diesem auch kein Besteuerungsrecht zustehen.[343] Allerdings ist zweifelhaft, ob schon bei Überschreiten eines einzelnen Schwellenwerts, wie es im derzeitigen Richtlinienentwurf gefordert wird, davon ausgegangen werden kann, dass auch tatsächlich Wertschöpfung im Quellenstaat stattfindet. Es ist z. B. nicht ersichtlich, warum allein durch das Überschreiten der Ertragsschwelle (Gesamterträge > 7 Mio. EUR) naheliegt, dass Wertschöpfung im Quellenstaat stattfindet. So wurde bereits in der ursprünglichen Idee der SEP, auf dem die

SDP basiert, von der OECD angemerkt, dass das alleinige Abstellen auf Erträge kein ausreichender Hinweis auf Wertschöpfung im Quellenstaat ist.[344] Erwirtschaftete Erträge sollten vielmehr nur als erster Indikator dienen und mit weiteren Bedingungen verknüpft werden.[345] Da es im Richtlinienentwurf zur SDP jedoch schon ausreicht, wenn eine der genannten Bedingungen erfüllt ist, findet eine solche Kombination mit weiteren Faktoren nicht statt. Im Endeffekt begründen damit bereits alle Erbringer digitaler Dienstleistungen in einem Quellenstaat eine SDP, sofern sie allein die Ertragsschwelle überschreiten. Selbiges Argument kann für die anderen beiden Schwellenwerte angebracht werden. Allein das Vorhandensein von mehr als 100 Tsd. Nutzern oder der Abschluss von 3.000 Verträgen sagt per se nichts über vorhandene Wertschöpfung im Quellenstaat aus. Damit werden von der SDP unter Umständen auch digitale Geschäftsmodelle erfasst, bei denen überhaupt keine Wertschöpfung im Quellenstaat stattfindet.

Das weit größere Problem liegt jedoch nicht darin, einen steuerlichen Nexus auf Grundlage einer signifikanten digitalen Präsenz zu definieren, sondern diesem in einem zweiten Schritt einen sachgerechten Gewinn zuzuweisen.[346] Hierzu ist es notwendig, nicht nur zu bestimmen, ob Wertschöpfung im Quellenstaat stattfindet, sondern auch wie hoch diese ausfällt. Die EU hat richtigerweise erkannt, dass dem Quellenstaat bei Anwendung des bestehenden AOA regelmäßig keine Gewinne zugeteilt werden könnten, da es dort i. d. R. an eigenen Personalfunktionen fehlt.[347] Um dieses Problem zu umgehen, sollen daher bei der Gewinnzuteilung auch die oben genannten Funktionen (z. B. Erhebung von Daten und nutzergenerierten Inhalten, Verkauf von Online-Werbeflächen, Bereitstellung von Inhalten Dritter) berücksichtigt werden. Die Bewertung, ob diese Funktionen aber überhaupt zur Wertschöpfung beitragen bzw. wie hoch dieser Beitrag ausfällt, ist aber alles andere als trivial.[348] Der Richtlinienentwurf äußert sich jedoch gerade nicht dazu, wie diese zusätzlich zu beachtenden Funktionen konkret bei der Gewinnzuordnung bewertet werden sollen.[349] Ohne weitere Handlungsvorschriften dürfte eine konkrete Gewinnzuordnung zu einer SDP daher schwerfallen.

Darüber hinaus könnten potenziell steuerpflichtige Unternehmen der Begründung einer SDP relativ einfach ausweichen. Die Richtlinie ist nur auf Erbringer digitaler Dienstleistungen anzuwenden, die entweder (1) in einem Mitgliedsstaat der EU ansässig sind oder (2) eine SDP in einem Mitgliedsstaat der EU unterhalten und in einem Staat ansässig sind, mit dem dieser EU-Mitgliedsstaat kein DBA unter-

---

[339]Vgl. Art. 3 Nr. 5 Buchst. f SDP-RL-E i. V. m. Anhang II Buchst. r des Richtlinienentwurfs (vgl. Europäische Kommission, 2018a, S. 2).

[340]Vgl. Art. 4 Abs. 3 SDP-RL-E.

[341]Vgl. Geutebrück, 2016, S. 124-125; Olbert & Spengel, 2017, S. 16; Bendlinger, 2018, S. 50; Lück, 2018, S. 460; Haase, 2018, S. 264; vgl. zur Grundaussage der OECD OECD, 2015, Tz. 115.

[342]Olbert & Spengel, 2017, S. 16.

[343]So in ihrer Argumentation Becker et al., 2018, S. 7; Farruggia-Weber, 2019, S. 639; Eilers & Oppel, 2018, S. 370; Fuest, 2018, S. 33-34.

[344]Vgl. OECD, 2015, Tz. 278.

[345]Vgl. OECD, 2015, Tz. 278.

[346]Vgl. Devereux & Vella, 2018, S. 392; Dorenkamp, 2018, S. 642; Haase, 2018, S. 262.

[347]Vgl. SDP-RL-E, Begründung S. 10-11.

[348]So stellt sich z. B. die Frage, ob und unter welchen Voraussetzungen das Sammeln von Daten überhaupt wertschöpfend ist, vgl. hierzu ausführlich Becker et al., 2018; Becker, Englisch & Schanz, 2019; Schneemelcher & Dittrich, 2019; Mayr & Schlager, 2018, 29; Fehling, 2015, S. 801.

[349]Vgl. Art. 5 SDP-RL-E, SDP-RL-E, Begründung S. 10-11.

hält.[350] Somit könnten potenziell steuerpflichtige Unternehmen der Begründung einer SDP entgehen, indem sie als in einem Staat ansässig gelten, der (1) kein Mitgliedsstaat der EU ist und gleichzeitig (2) ein DBA mit allen (oder möglichst vielen) EU-Mitgliedsstaaten unterhält, in dem das Unternehmen seine digitalen Dienstleistungen anbieten will. Somit wäre die Richtlinie zur signifikanten digitalen Präsenz in ihrer Gänze nicht anwendbar.

Abschließend sei festgehalten, dass sich auch der Steuervollzug (wie bereits bei der DST in Kapitel 7.1.3 beschrieben) in Ermangelung physischer Verbindungen zum Inland als schwierig gestaltet.

7.2.4. Bewertung des Richtlinienentwurfs

Wie die vorangegangenen Ausführungen gezeigt haben, würde der Richtlinienentwurf zur signifikanten digitalen Präsenz grundsätzlich auch digitale Werbeleistungen erfassen. Somit könnte Deutschland durch die Implementierung der Richtlinie Besteuerungsrechte (dem Grunde nach) an digitalen Werbeleistungen sicherstellen, welche von im Ausland ansässigen Körperschaften mithilfe des Internets im Inland erbracht werden. Das Problem eines fehlenden steuerlichen Nexus aufgrund mangelnder physischer Präsenz (vgl. hierzu Kapitel 2.3) könnte damit durch die SDP gelöst werden.

Aufgrund der vorangehenden Kritikpunkte ist jedoch fraglich, inwieweit die Einführung des Richtlinienentwurfs zur signifikanten digitalen Präsenz in der derzeitigen Form zu empfehlen ist. In der aktuellen Form ist fraglich, inwieweit die SDP in der Lage ist, zu beurteilen, ob durch Daten und Nutzer tatsächlich Wertschöpfung im Quellenstaat stattfindet, die eine Besteuerung rechtfertigt (vgl. hierzu Kapitel 2.2), und diesen Wertschöpfungsbeitrag zu bewerten (vgl. hierzu Kapitel 2.3). Um dies sicherzustellen, sollten die Schwellenwerte zur Begründung einer SDP besser aufeinander abgestimmt werden, um zu gewährleisten, dass nur Geschäftstätigkeiten erfasst werden, bei denen auch tatsächlich hinreichende Indizien bestehen, dass Wertschöpfung im Quellenstaat stattfindet. Weiterhin sollte die EU vor Einführung einer SDP genauere Vorgaben erstellen, die regeln, wie dieser ein sachgerechter Gewinn zuzurechnen ist. Andernfalls besteht die Gefahr, dass eine Besteuerung zwar dem Grunde nach gegeben ist, dem Umfang nach aber ins Leere läuft.

8. Fazit

Die vorliegende Arbeit hat das Ziel, die Frage zu beantworten, wie deutsche Besteuerungsrechte an digitalen Werbeleistungen sichergestellt werden können, welche von im Ausland ansässigen Körperschaften mithilfe des Internets im Inland erbracht werden.

Dazu wurden zunächst die Problemfelder dargestellt, welche die steuerliche Erfassung digitaler Werbeleistungen

im derzeit bestehenden internationalen Steuerrecht erschweren. Im Anschluss daran wurden die aktuellen nationalen steuerlichen Anknüpfungspunkte vorgestellt, die in Deutschland im Rahmen der beschränkten Steuerpflicht für die Erfassung digitaler Werbeleistungen in Frage kommen und wie diese potenziellen Steuerrechte im Kontext von bilateralen Doppelbesteuerungsabkommen eingeschränkt werden. Anhand eines konkreten Beispiels, der Suchmaschinenwerbung durch Google, wurde weiterhin untersucht, inwiefern diese bestehenden nationalen und bilateralen Steuergesetze im Inboundfall in der Lage sind, in Deutschland erbrachte digitale Werbeleistungen der Besteuerung zu unterwerfen. Abschließend wurden die von der EU vorgeschlagenen Handlungsalternativen in Form einer Digitalsteuer und eines neuen steuerlichen Tatbestands in Form einer signifikanten digitalen Präsenz vorgestellt. Dabei wurde analysiert, inwiefern diese geeignet erscheinen, digitale Werbeleistungen in Deutschland der Besteuerung zu unterwerfen.

Die vorangegangenen Ausführungen haben gezeigt, dass das deutsche Steuerrecht im Rahmen der beschränkten Steuerpflicht über keinen eigenen Tatbestand zur Erfassung digitaler Werbeleistungen verfügt und diese damit unter die bestehenden Tatbestände subsumiert werden müssen. Dafür kommen inländische gewerbliche Einkünfte durch Betriebsstätte oder ständigen Vertreter (§ 49 Abs. 1 Nr. 2 Buchst. a EStG), gewerbliche betriebsstättenlose Einkünfte (§ 49 Abs. 1 Nr. 2 Buchst. f EStG) und sonstige Einkünfte (§ 49 Abs. 1 Nr. 9 EStG) in Betracht. Aber selbst wenn digitale Werbeleistungen unter einen dieser drei Tatbestände subsumiert werden können, verbleiben in Deutschland im Fall eines bestehenden Doppelbesteuerungsabkommens i. d. R. nur dann Besteuerungsrechte an digitalen Werbeleistungen, wenn diese im Rahmen einer Betriebsstätte erwirtschaftet werden. In anderen Fällen gibt Deutschland im Normalfall alle Besteuerungsrechte an digitalen Werbeleistungen auf.

Die Anwendung der bestehenden nationalen und bilateralen Steuergesetze auf den Fall der Suchmaschinenwerbung durch Google hat weiterhin gezeigt, dass in Ermangelung einer physischen Betriebsstätte in Deutschland schon gar kein Besteuerungsrecht nach nationalem Recht besteht. Im konkreten Fall können die erbrachten digitalen Werbeleistungen nicht als Überlassung von Rechten oder Know-how angesehen werden. Somit liegen keine inländischen Einkünfte und damit auch keine Besteuerungsrechte aufgrund gewerblicher betriebsstättenloser oder sonstiger Einkünfte vor. Folglich sind Unternehmen, welche diese digitalen Werbeleistungen in Anspruch nehmen, auch nicht zum Einbehalt etwaiger Steueransprüche des deutschen Staates gegenüber den Erbringern digitaler Werbeleistungen verpflichtet. Insofern ist dem Vorstoß aus Teilen der Finanzverwaltung nicht zu folgen. Die derzeit bestehenden nationalen Steuergesetze sind in Ermangelung einer Betriebsstätte nicht in der Lage, die Besteuerung digitaler Werbeleistungen sicherzustellen. Vor diesem Hintergrund hat sich mittlerweile auch das Bayerische Staatsministerium der Finanzen und für Heimat in begrüßenswerter Weise dafür ausgesprochen, dass mit der Inanspruchnahme digitaler Werbeleistungen keine Steuerab-

---

[350]Vgl. SDP-RL-E, Begründung S. 8.

zugsverpflichtungen einhergehen.[351]

In Anbetracht dieser Ergebnisse scheint es unter der Annahme, dass Erbringer digitaler Werbeleistungen Wertschöpfung in Deutschland betreiben, gerechtfertigt, Maßnahmen in Betracht zu ziehen, welche die Besteuerung digitaler Werbeleistungen in Deutschland sicherstellen. Die beiden Richtlinienentwürfe auf Ebene der EU (Digitalsteuer und signifikante digitale Präsenz) wären grundsätzlich dazu in der Lage, digitale Werbeleistungen in Deutschland der Besteuerung zu unterwerfen. Allerdings sind beide Konzepte in ihrer derzeitigen Ausgestaltung zu kritisieren, sodass nach hier vertretener Ansicht die Einführung beider Konzepte nicht empfohlen werden kann.

An dieser Stelle sei angemerkt, dass dieser Arbeit verschiedene entscheidende Annahmen zugrunde liegen. Zum einen wurde davon ausgegangen, dass es sich bei Erbringern digitaler Werbeleistungen um juristische Personen in Form von Körperschaften handelt. Diese Annahme scheint in vielen Fällen zutreffend zu sein. Dennoch erfassen die vorangegangenen Ausführungen keine Fälle, in denen natürliche Personen digitale Werbeleistungen erbringen. Da die Regelungen zur beschränkten Steuerpflicht jedoch bei juristischen und natürlichen Personen nahezu identisch sind, sollten die Aussagen zu großen Teilen auch auf Erbringer digitaler Werbeleistungen anzuwenden sein, die als natürliche Personen zu klassifizieren sind.

Zum anderen basiert die Analyse der Wirksamkeit der bestehenden steuerlichen Regelungen in Kapitel 5 auf einem Einzelfall. Daher können die Ergebnisse nur bedingt dazu dienen generalisierte Aussagen zu treffen. Da sich die Geschäftsmodelle bei Erbringern digitaler Werbeleistungen teilweise stark voneinander unterscheiden, mag durchaus in Einzelfällen doch eine Steuerpflicht nach deutschem Steuerrecht zu bejahen sein. Bei der Übertragung der Ergebnisse auf andere Sachverhalte ist daher Vorsicht geboten. Wie dargestellt, lassen sich die Ergebnisse aber wohl auf ähnliche Sachverhalte (z. B. Suchmaschinenwerbung anderer Anbieter oder ähnliche Werbeformen wie z. B. Bannerwerbung) übertragen.

Die wohl kritischste Annahme der Arbeit liegt darin, dass durch die Erbringung digitaler Werbeleistungen im Inboundfall in jedem Fall Wertschöpfung in Quellenstaaten wie Deutschland stattfindet. Diese Annahme mag wohl für viele Erbringer digitaler Werbeleistungen durchaus vertretbar sein, insbesondere dann, wenn die Werbeleistungserbringung durch das exzessive Sammeln und Auswerten von Daten stark personalisiert erfolgt. Allerdings sollte im Einzelfall dennoch geprüft werden, ob diese angenommene Wertschöpfung im Quellenstaat auch tatsächlich stattfindet. Andernfalls fällt es schwer, einen Besteuerungsanspruch zu rechtfertigen.

Im Ergebnis kann damit die Frage, wie deutsche Besteuerungsrechte an digitalen Werbeleistungen sichergestellt werden können, welche von im Ausland ansässigen Körperschaften mithilfe des Internets im Inland erbracht werden, nicht abschließend beantwortet werden. Im Zuge der Arbeit konnte vielmehr festgehalten werden, wie solche Besteuerungsrechte in Deutschland nicht begründet werden können bzw. sollten. Daher besteht auch weiterhin Bedarf, angemessene Handlungsempfehlungen zu erarbeiten, welche die sachgerechte Besteuerung digitaler Werbeleistungen und im Allgemeinen digitaler Geschäftsmodelle sicherstellen. Da digitale Geschäftsmodelle eine globale Herausforderung für das internationale Steuersystem darstellen, scheint es auch angemessen, diesen Herausforderungen auf globaler Ebene zu begegnen. Es wäre daher wünschenswert, wenn auf Ebene der OECD eine international anerkannte konsensfähige Lösung erarbeitet werden würde. Vor diesem Hintergrund ist es begrüßenswert, dass die OECD seit Januar 2019 neue Handlungsalternativen diskutiert, um den Herausforderungen digitaler Geschäftsmodelle für das internationale Steuersystem zu begegnen.[352] Zukünftige Forschung sollte diesen Prozess unterstützen, indem sie untersucht, ob die neuen Vorschläge dazu geeignet wären, digitale Geschäftsmodelle im Allgemeinen und digitale Werbeleistungen im Speziellen einer sachgerechten Besteuerung zu unterwerfen.

---

[351]Vgl. Bayerisches Staatsministerium der Finanzen und für Heimat, 2019.

[352]Vgl. OECD, 2019a, 2019b. An diesen Diskussionen beteiligen sich mittlerweile 129 Staaten weltweit, vgl. OECD, 2019c.

Literatur

Ackermann, M. (2016). Beschränkte Steuerpflicht bei der grenzüberschreitenden Überlassung von Software. *Internationale Steuer-Rundschau*, *5* (7), 258–264.

Alphabet Inc. (2019). *Annual report for the fiscal year 2018.* Zugriff am 2019-02-22 auf https://abc.xyz/investor/static/pdf/20180204_alphabet_10K.pdf?cache=11336e3

Bauer, G., Fritz, J., Schanz, D. & Sixt, M. (2019). *Corporate income tax challenges arising from digitalised business models.* Zugriff am 2019-03-08 auf https://www.steuern.bwl.uni-muenchen.de/aktuelles/veroeffentlichung_digi/working-paper.pdf

Bayerisches Staatsministerium der Finanzen und für Heimat. (2019). *Füracker: Kein Steuerabzug bei Onlinewerbung.* Zugriff am 2019-03-14 auf http://www.stmfh.bayern.de/internet/stmf/aktuelles/pressemitteilungen/23847/index.htm

Becker, J. (2018). Besteuerung einer digitalen Präsenz. *Internationales Steuerrecht*, *27* (16), 634–638.

Becker, J. & Englisch, J. (2018). *Taxing where value is created.* Zugriff am 2019-03-25 auf https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3258387

Becker, J., Englisch, J. & Schanz, D. (2018). *How Data Should (Not) Be Taxed.* Zugriff am 2018-11-23 auf https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3289036

Becker, J., Englisch, J. & Schanz, D. (2019). A SURE Way of Taxing the Digital Economy. *Tax Notes International*, *93* (3), 309–312.

Beck-online. (2017). *OECD-Musterabkommen 2017 zur Beseitigung der Doppelbesteuerung sowie der Steuerverkürzung und -umgehung auf dem Gebiet der Steuern vom Einkommen und vom Vermögen.* Zugriff am 2019-02-13 auf https://beck-online.beck.de/Dokument/Gesamtversion?vpath=bibdata%2Fges%2FOECD_MUSTERABK%2Fcont%2FOECD_MUSTERABK.A5.htm&isAktuellGueltigeGesamtversion=True#lawanchor_oecd_musterabk.amtabschnitt_0

Behnes, R., Nink, J. & Rohde, K. (2016). Nutzung internetbasierter Datenbankanwendungen - Haftung des Lizenznehmers für Quellensteuern des ausländischen Anbieters. *Computer und Recht*, *2016* (5), 281–285.

Bendlinger, S. (2018). Die digitale Betriebsstätte. In S. Kirchmayr, G. Mayr & K. Hirschler (Hrsg.), *Digitalisierung im Konzernsteuerrecht* (S. 33–55).

Benecke, A. (2018). § 1 KStG. In A. Schnitger & O. Fehrenbacher (Hrsg.), *Kommentar Körperschaftsteuer KStG* (2. Aufl.). Wiesbaden.

Berschens, R. (2017). *Der Fiskus im digitalen Zeitalter.* Zugriff am 2019-04-04 auf https://www.handelsblatt.com/politik/international/steuervermeidung-der-fiskus-im-digitalen-zeitalter/20302648.html?ticket=ST-577433-UzeE2LKFXjDnWheKE0fF-ap6

Boochs, W. (2018a). § 50a EStG. In F. Lademann, W. Anissimov, W. Boochs & I. Heß (Hrsg.), *Beschränkte Steuerpflicht in Deutschland.* Stuttgart.

Boochs, W. (2018b). § 50d EStG. In F. Lademann, W. Anissimov, W. Boochs & I. Heß (Hrsg.), *Beschränkte Steuerpflicht in Deutschland.* Stuttgart.

Boochs, W. (2018c). § 50 EStG. In F. Lademann, W. Anissimov, W. Boochs & I. Heß (Hrsg.), *Beschränkte Steuerpflicht in Deutschland.* Stuttgart.

Bozza-Bodden, N. (2013). Art. 12 OECD-MA. In J. Schönfeld & X. Ditz (Hrsg.), *Doppelbesteuerungsabkommen.* Köln.

Brähler, G. (2014). *Internationales Steuerrecht* (8. Aufl.).

British Broadcasting Corporation (BBC). (2018). *Budget 2018.* Zugriff am 2019-04-02 auf https://www.bbc.com/news/business-46023450

Bruhn, M., Burmann, C., Esch, F.-R., Hennig, A., Kirchgeorg, M., Markgraf, D. & Schneider, W. (2013). *222 Keywords Marketing.*

Bruins, G. W. J., Einaudi, L., Seligman, E. R. A. & Stamp, J. (1923). *Report on Double Taxation.* Zugriff am 2019-01-03 auf http://adc.library.usyd.edu.au/view?docId=split/law/xml-main-texts/brulegi-source-bibl-1.xml&chunk.id=item-1&toc.id=item-1&database=&collection=&brand=default

Bullinger, W. (2019). § 2 UrhG. In A.-A. Wandtke & W. Bullinger (Hrsg.), *Praxiskommentar zum Urheberrecht* (5. Aufl.). München.

Bundeskartellamt. (2018). *Online-Werbung.* Zugriff am 2019-02-08 auf https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Schriftenreihe_Digitales/Schriftenreihe_Digitales_3.pdf?__blob=publicationFile&v=5

Bundeszentralamt für Steuern. (2015). *Was ist eine Nettovereinbarung?* Zugriff am 2019-04-15 auf https://www.bzst.de/DE/Steuern_International/Abzugsteuern_gem_%2050_50a_EStG/Fragen_&_Antworten/FAQ_Texte_Bemessungsgrundlage/faq_02.html?nn=31082

Bundeszentralamt für Steuern. (2018). *Übersicht über das deutsche Besteuerungsrecht an Lizenzgebühren nach Doppelbesteuerungsabkommen.* Zugriff am 2019-02-19 auf http://www.bzst.de/DE/Steuern_International/Abzugsteuerentlastung/Freistellung_Erstattung/Merkblaetter/Reststeuersatzliste_Lizenzen.pdf

Bunn, D. (2019). *France's Digital Services Tax.* Zugriff am 2019-04-02 auf https://taxfoundation.org/france-digital-services-tax/

Busche, J. (2018). § 631 BGB. In M. Henssler (Hrsg.), *Münchener Kommentar zum Bürgerlichen Gesetzbuch* (7. Aufl.).

Ciomek, N. (2016). Internet advertising and google adwords. In N. Lee (Hrsg.), *Google it* (S. 211–236).

Dams, J. & Heuzeroth, T. (2017). *Warum die EU Amazon und Co. nicht zu fassen bekommt.* Zugriff am 2019-04-04 auf https://www.welt.de/wirtschaft/article169441495/Warum-die-EU-Amazon-Co-nicht-zu-fassen-bekommt.html

Devereux, M. & Vella, J. (2014). Are We Heading towards a Corporate Tax System Fit for the 21st Century? *Fiscal Studies*, *35* (4), 449–475.

Devereux, M. & Vella, J. (2018). Taxing the Digitalised Economy. *British Tax Review*, *2018* (4), 387–406.

Diffring, P. (2019). *Online-Werbung im Visier der Finanzverwaltung.* Zugriff am 2019-02-15 auf https://taxtech.blog/2019/02/15/online-werbung-im-visier-der-finanzverwaltung/

Diffring, P. & Saft, A. (2019). Der Steuerabzug bei Online-Werbung am Beispiel von Google Ads. *Der Betrieb*, *2019* (8), 387–394.

Dischinger, M. & Riedel, N. (2011). Corporate taxes and the location of intangible assets within multinational firms. *Journal of Public Economics*, *95* (7-8), 691–707.

Dorenkamp, C. (2018). EU-Vorschläge zur Besteuerung der Digitalwirtschaft sowie EU-Beihilfenrecht und Verrechnungspreise. *Internationales Steuerrecht*, *27* (16), 640–642.

Dreier, T. (2018). § 4 UrhG. In T. Dreier & G. Schulze (Hrsg.), *Urheberrechtsgesetz* (6. Aufl.). München.

Drüen, K.-D. (2017a). § 12 AO. In K. Tipke & H. W. Kruse (Hrsg.), *Abgabenordnung, Finanzgerichtsordnung* (Stand: 153. EL Aufl.). Köln.

Drüen, K.-D. (2017b). § 13 AO. In K. Tipke & H. W. Kruse (Hrsg.), *Abgabenordnung, Finanzgerichtsordnung* (Stand: 153. EL Aufl.). Köln.

Ehrmann, O. C. & von Wallis, G. (2014). Steuerrechtliche Aspekte von Online-Transaktionen. In B. Holznagel, T. Hoeren & U. Sieber (Hrsg.), *Handbuch Multimedia-Recht* (Stand: 46. EL Aufl.).

Eilers, S. & Oppel, F. (2018). Die Besteuerung der digitalen Wirtschaft: Trends und Diskussionen Überblick über die Arbeiten der OECD und EU mit kritischer Einordnung. *Internationales Steuerrecht*, *27* (10), 361–370.

Elsten, C. & Hill, N. (2017). *Intangible Asset Market Value Study?* Zugriff am 2019-03-17 auf https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3009783

Ernst and Young (EY). (2018). *Turkey introduces new taxes on electronic commerce.* Zugriff am 2019-04-02 auf https://www.ey.com/Publication/vwLUAssets/Turkey_introduces_new_taxes_on_electronic_commerce/$FILE/2018G_012576-18Gbl_Turkey%20introduces%20new%20taxes%20on%20electronic%20commerce.pdf

Ernst and Young (EY). (2019a). *Austria announces new digital tax.* Zugriff am 2019-04-02 auf https://www.ey.com/Publication/vwLUAssets/Austria_announces_new_digital_tax/$FILE/2019G_012825-18Gbl_Indirect_Austria%20announces%20new%20digital%20tax.pdf

Ernst and Young (EY). (2019b). *Italy introduces new digital services tax.* Zugriff am 2019-04-02 auf https://www.ey.com/Publication/vwLUAssets/Italy_introduces_new_digital_services_tax/$FILE/2019G_012828-18Gbl_Indirect_Italy%20introduces%20new%20digital%20services%20tax.pdf

Ernst and Young (EY). (2019c). *Spain sends bill on Digital Services Tax to Parliament for approval.* Zugriff am 2019-04-02 auf https://www.ey.com/Publication/vwLUAssets/Spain_sends_bill_on_Digital_Services_Tax_to_Parliament_for_approval/$FILE/2019G_012876-18Gbl_Indirect_Spain%20sends%20DST%20bill%20to%20Parliament%20for%20approval.pdf

Europäische Kommission. (2018a). *Anhänge des Vorschlags für eine Richtlinie des Rates zur Festlegung von Vorschriften für die Unternehmensbesteuerung einer signifikanten digitalen Präsenz.* Zugriff am 2019-03-26 auf https://ec.europa.eu/taxation_customs/sites/taxation/files/proposal_significant_digital_presence_annex_21032018_de.pdf

Europäische Kommission. (2018b). *Faire Besteuerung der digitalen Wirtschaft.* Zugriff am 2019-04-04 auf https://ec.europa.eu/taxation_customs/business/company-tax/fair-taxation-digital-economy_de

Facebook Germany GmbH. (2018). *Jahresabschluss zum Geschäftsjahr vom 01.01.2017 bis zum 31.12.2017.* Zugriff am 2019-04-04 auf https://www.bundesanzeiger.de/ebanzwww/wexsservlet

Facebook Inc. (2019). *Annual Report for the fiscal year ended December 31, 2018.* Zugriff am 2019-04-02 auf https://investor.fb.com/financials/sec-filings-details/default.aspx?FilingId=13183451

Farruggia-Weber, F. (2019). Signifikante digitale Präsenz und Digitalsteuer. *Deutsches Steuerrecht, 57* (12), 638–644.

Fehling, D. (2015). Neues zu den Herausforderungen für die Besteuerung der Digitalen Wirtschaft - Der Abschlussbericht zu Maßnahme 1 des BEPS-Aktionsplans liegt vor. *Internationales Steuerrecht, 24* (21), 797–802.

Forschungsunion & acatech. (2013). *Umsetzunsempfehlungen für das Zukunftsprojekt Industrie 4.0.* Zugriff am 2019-04-02 auf https://www.bmbf.de/files/Umsetzungsempfehlungen_Industrie4_0.pdf

Frankfurter Allgemeine Zeitung. (2018). Briten preschen mit Digitalsteuer vor. In *Frankfurter Allgemeine Zeitung* (S. 30.10.2018, S. 21).

Französisches Ministerium für Wirtschaft und Finanzen. (2019). *Projet de loi relatif a la taxation des grandes entreprises du numerique.*

Fritz, J., Schanz, D. & Siegel, F. (2018). Stimmungsbild zur Equalisation Levy - Eine Analyse der Kommentare zum Abschlussbericht von BEPS Aktionspunkt 1. *Internationales Steuerrecht, 27* (22), 853–860.

Frotscher, G. (2015). *Internationales Steuerrecht* (4. Aufl.).

Fuest, C. (2018). *Digitalisierung und Steuerpolitik.* Zugriff am 2018-11-14 auf https://www.cesifo-group.de/DocDL/sd-2018-14-fuest-digitalisierung-steuerpolitik-2018-07-26.pdf

Fuest, C., Meier, V. & Neumeier, F. (2018). *Die Besteuerung der Digitalwirtschaft.* Zugriff am 2018-12-10 auf http://www.cesifo-group.de/DocDL/Studie-Digitalsteuer-2018.pdf

G20. (2013). *Russia G20.* Zugriff am 2019-01-14 auf http://www.g20.utoronto.ca/2013/Saint_Petersburg_Declaration_ENG.pdf

Gebhardt, R. (2018). § 50d EStG. In H.-J. Kanzler, G. Kraft, S. Bäuml, J. Marx & F. Hechtner (Hrsg.), *Einkommensteuergesetz* (3. Aufl.). Herne.

Geutebrück, G. (2016). *Aktuelle Entwicklungen in der Besteuerung von grenzüberschreitenden Lieferungen digitaler Güter und Leistungen zur Bekämpfung von Gewinnverkürzung und Gewinnverlagerung (BEPS)* (Dissertation). Zugriff am 2018-11-27 auf https://www.netidee.at/sites/default/files/2017-08/Dissertation_Geutebrueck_2016.pdf

Goebel, S., Wehling, C. & Gehrmann, S. (2018). *Handbuch der Quellenbesteuerung.*

Google Germany GmbH. (2019). *Jahresabschluss zum Geschäftsjahr vom 01.01.2017 bis zum 31.12.2017.* Zugriff am 2019-04-04 auf https://www.bundesanzeiger.de/ebanzwww/wexsservlet

Google Inc. (2019). *Nutzungsbedingungen für Werbeprogramme der Google Ireland Ltd.* Zugriff am 2019-02-16 auf https://payments.google.com/payments/paymentsinfofinder?hostOrigin=aHR0cHM6Ly9wYXltZW50cy5nb29nbGUuY29tOjQ0NDM.&sri=-21

Google Ireland Ltd. (2018). *Directors' report and financial statements for the year ended 31 december 2017.* Zugriff am 2019-02-27 auf https://search.cro.ie/company/ListSubDocs.aspx?id=368047&type=C

Gosch, D. (2018a). § 49 EStG. In P. Kirchhof (Hrsg.), *Einkommensteuergesetz Kommentar* (17. Aufl.). Köln.

Gosch, D. (2018b). § 50a EStG. In P. Kirchhof (Hrsg.), *Einkommensteuergesetz Kommentar* (17. Aufl.). Köln.

Gosch, D. (2018c). § 50d EStG. In P. Kirchhof (Hrsg.), *Einkommensteuergesetz Kommentar* (17. Aufl.). Köln.

Gradel, C. & Klaeren, J. (2018). Art. 13 OECD-MA. In G. Strunk, B. Kaminski & S. Köhler (Hrsg.), *Außensteuergesetz, Doppelbesteuerungsabkommen* (Stand: 51. EL Aufl.). Bonn.

Griffith, R., Miller, H. & O'Connell, M. (2014). Ownership of intellectual property and corporate taxation. *Journal of Public Economics, 112,* 12–23.

Haase, F. (2018). Der EU-Richtlinienvorschlag zur Besteuerung einer signifikanten digitalen Präsenz. *Die Unternehmensbesteuerung, 11* (5), 259–264.

Haselrieder, M. & Pontzen, D. (2019). *Steuereintreiber bei Google & Co.* Zugriff am 2019-02-21 auf https://www.zdf.de/nachrichten/heute/wegen-google-und-co-deutschen-mittelstaendlern-droht-insolvenz-102.html

Hecht, S. A. & Lampert, S. (2009). Die einkommensteuerrechtliche Behandlung der Überlassung von Software (Teil I). *Finanz-Rundschau, 2009* (24), 1127–1132.

Heil, S. & Pupeter, A. (2019). Google-Tax: Gibt es sie schon? Zur Abzugsteuer auf Online-Werbung nach § 50a EStG. *Betriebs-Berater, 2019* (11), 604–611.

Helm, M. (2018). § 32 KStG. In A. Schnitger & O. Fehrenbacher (Hrsg.), *Kommentar Körperschaftsteuer KStG* (2. Aufl.). Wiesbaden.

Herkenroth, K. E. & Striegel, A. (2018). § 50 EStG. In C. Herrmann et al. (Hrsg.), *Einkommensteuer- und Körperschaftsteuergesetz* (Stand: 288. EL Aufl.). Köln.

Heß, I. (2018). § 49 estg. In F. Lademann, W. Anissimov, W. Boochs & I. Heß (Hrsg.), *Beschränkte Steuerpflicht in Deutschland.* Stuttgart.

Hidien, J. (2018). § 49 EStG. In P. Kirchhof, H. Söhn & R. Mellinghoff (Hrsg.), *Einkommensteuergesetz* (Stand: 292. EL Aufl.). Heidelberg.

Hoke, W. (2019). Germany studying potential royalty withholding tax on online ads. *Tax Notes International, 93* (8), 869.

Holthaus, J. (2015). *Ausländische Künstler und Sportler im Steuerrecht* (2. Aufl.).

Holthaus, J. (2017). Besteuerung von Rechteüberlassungen ins Inland und Steuerabzug gemäß § 50a Abs. 1 Nr. 3 EStG vor dem Hintergrund des geplanten BMF-Schreibens. *Internationales Steuerrecht, 26* (17), 729–735.

Holthaus, J. & Volkmann, T. (2015). Besteuerung von Vergütungen nach § 13b UStG sowie nach § 50a EStG bei elektronischer Datennutzung, insbesondere von ausländischen Datenbanken. *Deutsche Steuer-Zeitung, 103* (14), 550–558.

Hruschka, F. (2013). Art. 5 OECD-MA. In J. Schönfeld & X. Ditz (Hrsg.), *Doppelbesteuerungsabkommen.* Köln.

Hruschka, F. (2019). Onlinewerbung im Fokus der Betriebsprüfung. *Deutsches Steuerrecht, 57* (3), 88–92.

Hutter, H. & Schmidt, J. (2000). Aktueller Überblick zum Thema e-commerce - unter besonderer Berücksichtigung der Frage: Betriebsstätteneigenschaft eines Internetservers. *Internationales Steuerrecht, 9* (21), 650–654.

Ittermann, P. & Niehaus, J. (2015). Industrie 4.0 und Wandel von Industriearbeit. In H. Hirsch-Kreinsen, P. Ittermann & J. Falkenberg (Hrsg.), *Digitalisierung industrieller Arbeit* (S. 32–51).

Käbisch, V. & Strunk, G. (2018). Art. 12 OECD-MA. In G. Strunk, B. Kaminski & S. Köhler (Hrsg.), *Außensteuergesetz, Doppelbesteuerungsabkommen* (Stand: 51. EL Aufl.). Bonn.

Kagermann, H., Lukas, W.-D. & Wahlster, W. (2011). *Industrie 4.0.* Zugriff am 2019-04-02 auf http://www.wolfgang-wahlster.de/wordpress/wp-content/uploads/Industrie_4_0_Mit_dem_Internet_der_Dinge_auf_dem_Weg_zur_vierten_industriellen_Revolution_2.pdf

Kahle, H. & Kindich, A. (2016). Begriff der Betriebsstätte nach nationalem Recht und DBA-Recht. In T. Lübbehüsen & H. Kahle (Hrsg.), *Brennpunkte der Besteuerung von Betriebsstätten* (Rechtsstand vom 1.9.2015 Aufl., S. 39–159).

Kaligin, T. (2018). § 49 EStG. In F. Lademann, W. Anissimov, W. Boochs & I. Heß (Hrsg.), *Beschränkte Steuerpflicht in Deutschland.* Stuttgart.

Karkinsky, T. & Riedel, N. (2012). Corporate taxation and the choice of patent location within multinational firms. *Journal of International Economics*, *88* (1), 176–185.

Kelm, D. & Müller, J. P. (2018). Die Vorschläge der EU-Kommission zur Besteuerung der digitalen Wirtschaft. *Die Wirtschaftsprüfung*, *2018* (9), 587–594.

Kessler, W. & Wald, R. (2015). Datenbankanwendungen – Quellensteuerabzug aufgrund automatischer Rechteverwertung oder nicht steuerbares Direktgeschäft? *Internationales Steuerrecht*, *24* (23), 889–895.

Kim, J., Thomas, P., Sankaranarayana, R., Gedeon, T. & Yoon, H.-J. (2015). Eye-tracking analysis of user behavior and performance in web search on large and small screens. *Journal of the Association for Information Science and Technology*, *66* (3), 526–544.

Klein, M. (2018). § 49 EStG. In C. Herrmann et al. (Hrsg.), *Einkommensteuer- und Körperschaftsteuergesetz* (Stand: 288. EL Aufl.). Köln.

Klein, M. & Hagena, A. (2018). § 50d EStG. In C. Herrmann et al. (Hrsg.), *Einkommensteuer- und Körperschaftsteuergesetz* (Stand: 288. EL Aufl.). Köln.

Kofler, G., Mayr, G. & Schlager, C. (2017). Digitalisierung und Betriebsstättenkonzept. *Betriebs-Berater*, *2017* (32), 1815–1819.

Kokott, J. (2019). Herausforderungen einer Digitalsteuer. *Internationales Steuerrecht*, *28* (4), 123–134.

Korts, S. (2016). *Grundzüge im internationalen Steuerrecht* (2. Aufl.).

KPMG. (2019). *Führt Bayern die Quellensteuer für digitale Werbeleistungen ein?* Zugriff am 2019-04-03 auf https://home.kpmg/de/de/home/themen/2019/02/quellensteuer-fuer-digitale-werbeleistungen.html

Kraft, C. (2018). § 49 EStG. In H.-J. Kanzler, G. Kraft, S. Bäuml, J. Marx & F. Hechtner (Hrsg.), *Einkommensteuergesetz* (3. Aufl.). Herne.

Kraft, C. & Weigert, K. (2018). § 50a EStG. In H.-J. Kanzler, G. Kraft, S. Bäuml, J. Marx & F. Hechtner (Hrsg.), *Einkommensteuergesetz* (3. Aufl.). Herne.

Kraft, G. (2018a). Softwareüberlassung und Gestattung von Datenbanknutzung durch ausländische Digitalanbieter (Teil 1). *Steuer- und Wirtschaftsrecht*, *2018* (13), 868–876.

Kraft, G. (2018b). Softwareüberlassung und Gestattung von Datenbanknutzung durch ausländische Digitalanbieter (Teil 2). *Steuer- und Wirtschaftsrecht*, *2018* (14), 937–944.

Lammenett, E. (2017). *Praxiswissen Online-Marketing* (6. Aufl.).

Le Maire, B., Schäuble, W., Padoan, P.-C. & De Guindos, L. (2017). *Joint initiative on the taxation of companies operating in the digital economy.* Zugriff am 2019-04-02 auf https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Themen/Europa/ECOFIN_und_Eurogruppe/2017-09-18-eurogruppe-informeller-ecofin-nachbericht-september-anl3.pdf?__blob=publicationFile&v=1

Lehner, M. & Reimer, E. (2005). Generalthema I: Quelle Versus Ansässigkeit - Wie sind die grundlegenden Verteilungsprinzipien des Internationalen Steuerrechts austariert? *Internationales Steuerrecht*, *14* (16), 542–551.

Leisner-Egensperger, A. (2014). Betriebsstättenbegriff und Besteuerung des E-Commerce. *Steuer und Wirtschaft*, *91* (4), 298–304.

Lewandowski, D. (2013). Suchmaschinen. In R. Kuhlen, W. Semar & D. Strauch (Hrsg.), *Grundlagen der praktischen Information und Dokumentation* (6. Aufl., S. 495–508).

Lewandowski, D. (2015). *Suchmaschinen verstehen*.

Linn, A. (2019). Keine Steuerabzugsverpflichtung für Aufwendungen für Online-Marketing. *Deutsches Steuerrecht*, *57* (9), 418–422.

Loschelder, F. (2018a). § 49 EStG. In H. Weber-Grellet (Hrsg.), *Einkommensteuergesetz* (37. Aufl.). München.

Loschelder, F. (2018b). § 50a EStG. In H. Weber-Grellet (Hrsg.), *Einkommensteuergesetz* (37. Aufl.). München.

Loschelder, F. (2018c). § 50d EStG. In H. Weber-Grellet (Hrsg.), *Einkommensteuergesetz* (37. Aufl., S. 1–70). München.

Lübbehüsen, T. & Kahle, H. (2016). Besteuerung von Betriebsstätten und Grundfragen der internationalen Unternehmensbesteuerung. In T. Lübbehüsen & H. Kahle (Hrsg.), *Brennpunkte der Besteuerung von Betriebsstätten* (Rechtsstand vom 1.9.2015 Aufl., S. 5–25).

Lück, K. (2018). *Steuerrecht und digital economy*.

Marquardt, M. (2019). § 4 UrhG. In A.-A. Wandtke & W. Bullinger (Hrsg.), *Praxiskommentar zum Urheberrecht* (5. Aufl.). München.

Mason, R. & Parada, L. (2018). Digital Battlefront in the Tax Wars. *Tax Notes International*, *92* (17. Dezember), 1183–1197.

Maßbaum, M. (2017). Quellensteuerabzug bei Rechteüberlassungen in der Automobilindustrie. *Recht, Automobil, Wirtschaft*, *11* (1), 42–50.

Maßbaum, M. (2018). § 50a EStG. In C. Herrmann et al. (Hrsg.), *Einkommensteuer- und Körperschaftsteuergesetz* (Stand: 288. EL Aufl.). Köln.

Maßbaum, M. & Imhof, D. (2018). Beschränkte Steuerpflicht und Steuerabzug bei grenzüberschreitender Überlassung von Software und Datenbanken. *Finanz-Rundschau*, *100* (1), 6–12.

Mayr, G. & Schlager, C. (2018). Digitale Geschäftsmodelle und Betriebsstättenkonzept. In S. Kirchmayr, G. Mayr & K. Hirschler (Hrsg.), *Digitalisierung im Konzernsteuerrecht* (S. 21–31).

Miller, A. & Oats, L. (2016). *Principles of international taxation* (5. Aufl.).

Mohr, D. (2018). § 2 KStG. In A. Schnitger & O. Fehrenbacher (Hrsg.), *Kommentar Körperschaftsteuer KStG* (2. Aufl.). Wiesbaden.

Müller-Terpitz, R. & Knodel, A. (2017). Search engines. In P. Bräutigam, D. Rücker & A. Berger (Hrsg.), *E-commerce* (S. 671–715).

Naujok, J.-P. (2018). § 49 EStG. In F. Lademann, W. Anissimov, W. Boochs & I. Heß (Hrsg.), *Beschränkte Steuerpflicht in Deutschland.* Stuttgart.

Nitzschke, D. (2018). § 31 KStG. In A. Schnitger & O. Fehrenbacher (Hrsg.), *Kommentar Körperschaftsteuer KStG* (2. Aufl.). Wiesbaden.

OECD. (2010). *Bericht über die Zurechnung von Gewinnen zu Betriebsstätten von 2010.* Zugriff am 2019-03-27 auf http://www.oecd.org/ctp/transfer-pricing/attributes-of-profits-permanent-establishments-german.pdf

OECD. (2011). *Oecd guide to measuring the information society 2011.* Zugriff am 2019-02-17 auf http://www.oecd-ilibrary.org/science-and-technology/oecd-guide-to-measuring-the-information-society-2011_9789264113541-en

OECD. (2014). *Aktionsplan zur Bekämpfung der Gewinnverkürzung und Gewinnverlagerung.* Zugriff am 2018-12-03 auf http://www.oecd.org/tax/beps/action-plan-on-base-erosion-and-profit-shifting-9789264202719-en.htm

OECD. (2015). *Addressing the Tax Challenges of the Digital Economy.* Zugriff am 2018-11-27 auf http://www.oecd.org/ctp/addressing-the-tax-challenges-of-the-digital-economy-action-1-2015-final-report-9789264241046-en.htm

OECD. (2017). *Model tax convention on income and on capital.*

OECD. (2018). *Steuerliche Herausforderungen der Digitalisierung - Zwischenbericht 2018.* Zugriff am 2018-11-27 auf https://www.oecd-ilibrary.org/taxation/steuerliche-herausforderungen-der-digitalisierung-zwischenbericht-2018_9789264310438-de

OECD. (2019a). *Addressing the Tax Challenges of the Digitalisation of the Economy.* Zugriff am 2019-02-15 auf http://www.oecd.org/tax/beps/public-consultation-document-addressing-the-tax-challenges-of-the-digitalisation-of-the-economy.pdf

OECD. (2019b). *Addressing the Tax Challenges of the Digitalisation of the Economy.* Zugriff am 2019-01-30 auf http://www.oecd.org/tax/beps/policy-note-beps-inclusive-framework-addressing-tax-challenges-digitalisation.pdf

OECD. (2019c). *Members of the Inclusive Framework on BEPS.* Zugriff am 2019-03-26 auf http://www.oecd.org/tax/beps/inclusive-framework-on-beps-composition.pdf

Oellerich, I. (2018). § 31 KStG. In W. Blümich & B. Heuermann (Hrsg.), *Einkommensteuergesetz, Körperschaftsteuergesetz, Gewerbesteuergesetz* (Stand: 142. EL Aufl.). München.

Olbert, M. & Spengel, C. (2017). International Taxation in the Digital Economy. *World Tax Journal*, *2017* (Februar 2017), 3–45.

Paal, B. P. (2012). *Suchmaschinen, Marktmacht und Meinungsbildung.*

Paperlein, J. (2018). Ein Plus, aber nicht für TV. *Horizont*, *2018* (50), 8.

Paternoster, D. (2018). *Kurzdarstellungen zur Europäischen Union.* Zugriff am 2019-03-29 auf http://www.europarl.europa.eu/factsheets/de/sheet/92/allgemeine-steuerpolitik

Pelzer, G. & Düssel, M. (2013). *Online-Marketing mit Google.*

Pelzer, G., Sommeregger, T. & Linnenbrink, R. (2015). *Google AdWords.*

Petersen, S. (2013). Quellensteuer bei Überlassung von Standardsoftware. *Internationales Steuerrecht*, *22* (23), 896–904.

Pinkernell, R. (2012). Ein Musterfall zur internationalen Steuerminimierung

durch US-Konzerne. *Steuer und Wirtschaft*, *2012* (4), 369–374.

Pinkernell, R. (2014). *Internationale Steuergestaltung im Electronic Commerce.* Zugriff am 2019-01-27 auf https://www.ifst.de/wp-content/uploads/2013/01/494.pdf

Pinkernell, R. (2018). Digitale Wirtschaft. *Die Unternehmensbesteuerung*, *11* (3), 139–147.

Prussak, A. Y. (2013). The income of the twenty-first century. *Tulane Journal of Technology and Intellectual Property*, *16* (39), 39–77.

Rapp, B. (2017). Die gewerbesteuerliche Hinzurechnung in Zusammenhang mit digitalen Services und Produkten. *Finanz-Rundschau*, *2017* (12), 563–572.

Rat der Europäischen Union. (2017). *Bewältigung der Problematik der Besteuerung der Gewinne der digitalen Wirtschaft.* Zugriff am 2019-04-02 auf http://data.consilium.europa.eu/doc/document/ST-15445-2017-INIT/de/pdf

Rat der Europäischen Union. (2018). *Vorschlag für eine Richtlinie des Rates zum gemeinsamen System einer Digitalsteuer auf Erträge aus der Erbringung bestimmter digitaler Dienstleistungen.* Zugriff am 2018-12-04 auf http://data.consilium.europa.eu/doc/document/ST-14886-2018-INIT/de/pdf

Rat Wirtschaft und Finanzen. (2018). *Franco-german joint declaration on the taxation of digital companies and minimum taxation.* Zugriff am 2018-12-04 auf https://www.consilium.europa.eu/media/37276/fr-de-joint-declaration-on-the-taxation-of-digital-companies-final.pdf

Rath, M. (2005). Suchmaschinen sind auch nicht mehr das, was sie einmal waren. *Wettbewerb in Recht und Praxis*, *2005* (7), 826–833.

Reichl, A. & Wiedmann, B. (2019). Besteuerung von Online-Marketingleistungen mit dem Steuerabzug nach § 50a EStG. *Finanz-Rundschau*, *2019* (5), 201–207.

Reimer, E. (2018a). § 49 EStG. In W. Blümich & B. Heuermann (Hrsg.), *Einkommensteuergesetz, Körperschaftsteuergesetz, Gewerbesteuergesetz* (Stand: 142. EL Aufl.). München.

Reimer, E. (2018b). § 50a EStG. In W. Blümich & B. Heuermann (Hrsg.), *Einkommensteuergesetz, Körperschaftsteuergesetz, Gewerbesteuergesetz* (Stand: 142. EL Aufl.). München.

Reimer, E. (2018c). § 50 EStG. In W. Blümich & B. Heuermann (Hrsg.), *Einkommensteuergesetz, Körperschaftsteuergesetz, Gewerbesteuergesetz* (Stand: 142. EL Aufl.). München.

Rengers, J. (2018). § 2 KStG. In W. Blümich & B. Heuermann (Hrsg.), *Einkommensteuergesetz, Körperschaftsteuergesetz, Gewerbesteuergesetz* (Stand: 142. EL Aufl.). München.

Rick, E., Gunsenheimer, G., Schneider, J. & Kremer, T. (2017). *Lehrbuch Einkommensteuer* (23. Aufl.).

Risvik, K. M. & Michelsen, R. (2002). Search engines and web dynamics. *Computer Networks*, *39* (3), 289–302.

Roderburg, G. (2018). Neues zur Besteuerung der digitalen Wirtschaft. *Die Unternehmensbesteuerung*, *11* (5), 249–258.

Roth, A. (2018). § 49 EStG. In C. Herrmann et al. (Hrsg.), *Einkommensteuer- und Körperschaftsteuergesetz* (Stand: 288. EL Aufl.). Köln.

Rupp, S. (2010). *Google Marketing*.

Rupp, T., Knies, J. T., Ott, J.-P., Faust, T. & Hüll, M. (2018). *Internationales Steuerrecht* (4. Aufl.).

Schanz, D. & Sixt, M. (2018). Betroffene Geschäftsmodelle des EU-Richtlinienvorschlags zur Digitalsteuer. *Deutsches Steuerrecht*, *56* (38), 1985–1990.

Schaumburg, H. (2017). Einkommensteuer. In H. Schaumburg (Hrsg.), *Internationales Steuerrecht* (4. Aufl., S. 129–318).

Schewe, M. (2018). Beschränkte Steuerpflicht ausländischer Softwareanbieter. *Internationales Steuer- und Wirtschaftsrecht*, *2018* (3), 98–107.

Schlotter, C. & Hruschka, F. (2018). Aktuelle Fragen des Steuerabzugs nach § 50a EStG aus Unternehmenssicht. In T. Rödder & R. Hüttemann (Hrsg.), *Steuerberater-Jahrbuch 2017/2018* (S. 667–712).

Schmidl, M. (2014). *IT-Recht von A - Z* (2. Aufl.).

Schmidt, L., Sigloch, J. & Henselmann, K. (2005). *Internationale Steuerlehre*.

Schneemelcher, P. & Dittrich, P.-J. (2019). *Nutzer, Daten, Netzwerke.* Zugriff am 2019-03-07 auf https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/nutzer-daten-netzwerke/

Schön, W. (2018). Ten Questions about Why and How to Tax the Digitalized Economy. *Bulletin for International Taxation*, *2018* (4/5), 278–292.

Schreiber, U. (2017). *Besteuerung der Unternehmen* (4. Aufl.).

Schulze, G. (2018). § 2 UrhG. In T. Dreier & G. Schulze (Hrsg.), *Urheberrechtsgesetz* (6. Aufl.). München.

Sheppard, L. (2018). Digital Permanent Establishment and Digital Equalization Taxes. *Bulletin for International Taxation*, *72* (4a/Special Issue), 1–6.

Skeie, Ø. B., Johansson, Å., Menon, C. & Sorbe, S. (2017). *Innovation, patent location and tax planning by multinationals.* Zugriff am 2019-03-21 auf http://www.oecd.org/eco/Innovation-patent-location-and-tax-planning-by-multinationals.pdf

Spengel, C. (2018). Besteuerung der digitalen Wirtschaft - Europa einmal mehr auf dem falschen Weg. *Der Betrieb*, *2018* (15), M4-M5.

statcounter. (2019). *Search Engine Market Share Worldwide - January 2019.* Zugriff am 2019-02-17 auf http://gs.statcounter.com/search-engine-market-share

Statista. (2018). *Umsatz im Markt für Digitale Werbung nach Segmenten in Deutschland im Jahr 2017 sowie eine Prognose bis 2023.* Zugriff am 2019-02-10 auf https://de-statista-com.emedien.ub.uni-muenchen.de/statistik/daten/studie/456157/umfrage/umsaetze-im-markt-fuer-digitale-werbung-in-deutschland/

Stelzer, D. (2004). Produktion digitaler Güter. In A. Braßler, H. Corsten & T. Blecker (Hrsg.), *Entwicklungen im Produktionsmanagement* (S. 233–250).

Stöber, M. (2018). § 49 EStG. In F. Lademann, W. Anissimov, W. Boochs & I. Heß (Hrsg.), *Beschränkte Steuerpflicht in Deutschland.* Stuttgart.

Strunk, G. & Kaminski, B. (2018). Art. 7 OECD-MA. In G. Strunk, B. Kaminski & S. Köhler (Hrsg.), *Außensteuergesetz, Doppelbesteuerungsabkommen* (Stand: 51. EL Aufl.). Bonn.

Süddeutsche Zeitung. (2019). *Einführung von Digitalsteuer in Europa gescheitert.* Zugriff am 2019-03-13 auf https://www.sueddeutsche.de/news/wirtschaft/finanzen-einfuehrung-von-digitalsteuer-in-europa-gescheitert-dpa.urn-newsml-dpa-com-20090101-190312-99-344663

Tappe, H. (2011). Steuerliche Betriebsstätten in der Cloud. *Internationales Steuerrecht*, *20* (22), 870–874.

Thiele, P. & König, I. (2018). Die steuerliche Behandlung von Serverbetriebsstätten. *Internationales Steuer- und Wirtschaftsrecht*, *2018* (10), 376–385.

Ueberbach, S. (2019). *EU-Kompromiss gescheitert.* Zugriff am 2019-03-22 auf https://www.tagesschau.de/wirtschaft/digitalsteuer-107.html

United States House Committee on Ways and Means. (2019). *Brady laments france digital tax proposal.* Zugriff am 2019-03-12 auf https://gop-waysandmeans.house.gov/brady-laments-france-digital-tax-proposal/

Valta, M. (2018). Verfassungs- und Abkommensrechtsfragen des Richtlinienentwurfs für eine Steuer auf digitale Dienstleistungen. *Internationales Steuerrecht*, *27* (19), 765–772.

Varian, H. R. (2016). *Grundzüge der Mikroökonomik* (9. Aufl.).

Verband der Automobilindustrie (VDA). (2018). *Steuerfragen digitaler Geschäftsmodelle in der Automobilindustrie.* Zugriff am 2018-11-16 auf https://www.ey.com/de/de/services/tax/ey-steuerfragen-digitaler-geschaeftsmodelle-in-der-automobilindustrie

von Bredow, A., Scheinbacher, A. & Groß, S. (2017). *Cloud Services im Bann der Abzugsteuer.* Zugriff am 2018-12-10 auf https://www.psp.eu/artikel/55/cloud-services-im-bann-der-abzugsteuer-die-digitale-transformation-fordert-den-fiskus-zum-handeln/

Wagner, K. J. (2018). § 50d EStG. In W. Blümich & B. Heuermann (Hrsg.), *Einkommensteuergesetz, Körperschaftsteuergesetz, Gewerbesteuergesetz* (Stand: 142. EL Aufl.). München.

Wandtke, A.-A. & Grunert, W. (2019). vor § 31 ff. UrhG. In A.-A. Wandtke & W. Bullinger (Hrsg.), *Praxiskommentar zum Urheberrecht* (5. Aufl.). München.

Wassermeyer, F. & Kaeser, C. (2018a). Art. 12 OECD-MA. In F. Wassermeyer, C. Kaesner, M. Schwenke & K.-D. Drüen (Hrsg.), *Doppelbesteuerung* (Stand: 143. EL Aufl.). München.

Wassermeyer, F. & Kaeser, C. (2018b). Art. 3 OECD-MA. In F. Wassermeyer, C. Kaesner, M. Schwenke & K.-D. Drüen (Hrsg.), *Doppelbesteuerung*

(Stand: 143. EL Aufl.). München.

Wieduwilt, H. (2019). *Bayern nutzt die Hintertür zur Digitalsteuer.* Zugriff am 2019-03-05 auf https://www.faz.net/aktuell/wirtschaft/diginomics/bayern-nutzt-die-hintertuer-zur-digitalsteuer-16071968.html?service=printPreview

Wissenschaftlicher Beirat beim Bundesministerium der Finanzen. (2018). *Stellungnahme zu den EU-Vorschlägen für eine Besteuerung der digitalen Wirtschaft.* Zugriff am 2019-03-18 auf https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Ministerium/Geschaeftsbereich/Wissenschaftlicher_Beirat/Gutachten_und_Stellungnahmen/Ausgewaehlte_Texte/2018-09-27-digitale-Wirtschaft-anl.pdf?__blob=publicationFile&v=3

Wünnemann, M. (2019). Herausforderungen der Besteuerung der digitalen Wirtschaft im Jahr 2019. *Internationales Steuerrecht*, *28* (4), 134–140.

Zeit Online. (2018). *Deutsch-französischer Vorschlag für Digitalsteuer ist gestoppt.* Zugriff am 2018-12-04 auf https://www.zeit.de/wirtschaft/2018-12/eu-finanzminister-digitalsteuer-facebook-google?print

Zöller, D. (2018). Die Besteuerung der digitalen Wirtschaft. *Betriebs-Berater*, *72* (49), 2903–2909.

Junior Management Science

journal homepage: www.jums.academy

Economics of Hydrogen: Scenario-based Evaluation of the Power-to-Gas Technology

Ubald Bauer

Technische Universität München

Abstract

Power-to-gas (PtG) facilities apply the chemical process of water electrolysis to produce hydrogen and represent a low-carbon alternative to conventional hydrogen production methods when coupled with renewable energy sources. This thesis aims to evaluate the economic potential of the PtG technology and explore how policy changes can affect its profitability, measured by the break-even price of hydrogen. For the derivation of the break-even price, I rely on a net present value model that considers cost and revenue components as levelized terms, which I adapt by incorporating energy policy instruments. I develop an algorithm for the investment analysis of PtG projects, which considers both the capacity of the PtG facility and the renewable energy source as variables and optimizes their ratio for profitability. My analysis shows that large-scale PtG facilities can already compete on the market for medium-scale hydrogen supply at a price of 3.55 € /kg. However, profitable operations of small-scale PtG plants still depend on the implementation of policy changes. I find that small systems could produce pure renewable hydrogen at a break-even price below 3.00 € /kg and thus more than halve their costs, if supportive policy measures were adopted.

Keywords: Hydrogen economics; power-to-gas; renewable energy; capacity optimization.

1. Introduction

In view of climate change, the German government set the goal to reduce the emission of greenhouse gases drastically and transform the electricity mix toward using renewable sources to an extent of 80% of gross electricity consumption by 2050.[1] In addition to the current consumption, increased power generation will be necessary to cover the added demand from sectors being electrified by then, such as mobility. This will require a massive deployment of renewable energy sources in the coming decades, particularly of onshore and offshore wind power and solar PV. Due to the intermittency of these technologies, long-term energy storage will be required at a large scale, to balance the seasonally diverging supply and demand of electricity.[2] Stored energy can then be reconverted to supply power during calm period, i.e., an interval of several days without wind and solar power generation due to meteorological conditions occurring in winter. Currently no technologies other than chemical energy storage methods, mainly power-to-gas (PtG), are capable of providing the required long-term storage capacities.[3] This has

also been acknowledged by the German government, which plans for the announcement of a strategy concerning hydrogen, the energy carrier produced by the PtG process, toward the end of 2019. The main component of a power-to-gas facility is the electrolyzer, which performs the chemical process of water electrolysis. This process consumes water and electricity as input materials to generate hydrogen and oxygen. In addition to long-term storage, PtG has the potential to provide ancillary services rapidly due to electrolyzer flexibility, when preferential short-term storage capacity is already utilized at full capacity.

Hydrogen cannot only be reconverted to power but could also replace fossil feedstock to cut carbon emissions in hardly electrifiable industries, such as steel production, aviation or ammonia production for fertilizers. For instance, steelmaking, which is among the largest industrial carbon emitters, could become fossil-free through the application of renewable hydrogen in the reduction of iron ore.[4] Due to its potential to satisfy the needs of various different energy consumers, PtG is considered to be a key element when it comes to "coupling of the electricity sector with other energy sec-

---

[1]Cf. BMWi (2010), p. 5.
[2]Cf. Sterner et al. (2019b), p. 110.
[3]Cf. Sterner et al. (2019a), p. 327 and Schenuit et al. (2016), p. 28.

[4]Cf. Vogl et al. (2018), p. 736.

tors".[5]

Considering the current electricity mix of Germany, which still exhibits a high share of fossil fuels, the production of renewable hydrogen requires the immediate integration of PtG technology with renewable energy sources, e.g., wind turbines or solar PV. In this thesis, I refer to such systems, where the electrolyzer directly relies on a renewable energy source, as vertically integrated energy systems.

The focus of my thesis lies on the examination of the economics of vertically integrated energy systems based under current legislation and considering potential policy changes. In particular, I will identify suitable instruments to incentivize investment in vertically integrated energy systems in favor of renewable hydrogen production. The profitability of such systems is measured by the break-even price of hydrogen generated by the PtG subsystem, which is derived based on a net present value (NPV) model.

One aim of the thesis is the implementation of a tool, that facilitates the evaluation of vertically integrated systems under various scenarios covering both current legislation and a range of potential future policy designs. Such a tool could serve policymakers to understand the implications of their decisions on power-to-gas profitability and help investors when assessing the economic potential of vertically integrated systems. It also allows the simulation of a wide range of system compositions in order to optimize the ratio of the PtG and renewable energy subsystems.

Section 2 describes the analyzed scenario and introduces the relevant regulatory aspects, followed by the methodology based on an adopted net present value model and some model alterations. Section 3 of my thesis concerns the implementation of the mentioned analysis tool and addresses some issues resolved throughout the development of the algorithms for the break-even price detection and capacity optimization. In section 4 I describe the selected model input variables and present the results of various scenario simulations, followed by an interpretation of the results and an examination of the sensitivity of the hydrogen price with respect to key model input parameters. Finally, section 5 gives a conclusion of the thesis.

2. Methodology

2.1. Scenario description

The objective of this thesis is the evaluation of the power-to-gas technology for a series of scenarios and under consideration of current regulation and potential future modifications of the regulatory framework. All underlying scenarios are based on a vertically integrated energy system, composed of a power-to-gas facility and a renewable energy source (RES). The RES itself can be composed of stand-alone wind turbines or solar PV technology for power generation, or of a hybrid RES comprising both technologies. Thus, the PtG plant can either utilize renewable power or electricity

sourced from the power grid to produce hydrogen by application of water electrolysis. With regard to grid power, access to wholesale electricity prices is assumed. The PtG facility is able to run flexibly and can be switched on and off at any time without technological restrictions, thus enabling an optimization of the power conversion to hydrogen in reaction to the development of real-time electricity prices. The gained value from converting renewable electricity to hydrogen is traded off against the earnings that would result from feeding electricity into the grid at any point in time.

In some of the analyzed scenarios the capture of oxygen, a by-product of water electrolysis, is considered for the evaluation, in order to find out if the commercialization of oxygen could help electrolytic hydrogen to become more cost-competitive.

For the profitability analysis of the vertically integrated system, I take the perspective of an investor in search of the ideal sizing of the PtG and RES subsystems, which yields a minimum break-even price of hydrogen.

2.2. Literature review

The economics of hydrogen production based on water electrolysis have been extensively investigated in the past decades and several research projects have been established around the power-to-gas technology, such as "FCH JU"[6], an EU-funded consortium composed of fuel cell and hydrogen industries and academia. Many research papers regarding power-to-gas refer to the use of renewable power but generally do not examine power-to-gas directly coupled with a renewable energy source. Furthermore, the focus of research seems to lie on PtG as an energy storage technology and profitability many times is not assessed in terms of a hydrogen price. However, many articles and reports exist which analyze integrated energy systems, as defined in this thesis, and compute comparable profitability measures. A recent study[7] provides an insightful literature review concerning clean hydrogen production, including based on water electrolysis with solar or wind power, and offers a cost-based comparison of the identified literature. The review reports costs of renewable hydrogen between 5 - 8 $/kg in the case of wind power and finds electrolytic hydrogen from solar power at a cost of 8 - 9 $/kg.[8]

Glenk and Reichelstein[9] examined the economic potential of PtG facilities integrated with wind power and derived a net present value model for the profitability assessment of vertically integrated energy systems, based on time-variant electricity prices and renewable energy capacity factors, while leveraging the concept of unit cost, e.g., by relying on the common measure of levelized cost of electricity (LCOE).

None of the studies, I have found, directly compare the effects of legislation on PtG economics. Therefore, I set the

---

[5]Sterner et al. (2019c), p. 758.

[6]Fuel Cells and Hydrogen Joint Undertaking.
[7]Cf. El-Emam and Özcan (2019).
[8]Cf. El-Emam and Özcan (2019), p. 603-604.
[9]Cf. Glenk and Reichelstein (2019b).

focus of my thesis on the profitability analysis of vertically integrated energy systems under the legal framework effective in Germany and highlight the implications of policy changes on PtG profitability. In particular, I will identify instruments necessary to incentivize the production of renewable hydrogen, in contrast to grid-supplied electrolysis.

For this purpose, I build on the model framework developed by Glenk and Reichelstein and add components to represent the policy options and the possibility to consider wind turbines and solar PV as a hybrid renewable energy system integrated with a PtG facility. Additionally, I accounted for the capture of the by-product oxygen and included a scenario for consideration of a rising price of carbon emission allowances, which affects the price of competing fossil hydrogen, as well as the cost associated to grid electricity consumption. Both the analysis of a hybrid renewable energy system and the consideration of the oxygen by-product can rarely be observed in literature.

2.3. Legal environment

The described scenario setup is mainly subject to regulations under the Renewable Energy Sources Act (EEG[10]), which was initially introduced in 2000. The latest amendment took effect in January 2017. The EEG guarantees a subsidy for renewable energy systems during a period of 20 years. The subsidy is granted in the form of a feed-in premium, which compensates for the difference between the guaranteed tariff and the market price of electricity. The premium is only paid if power is fed into the public power grid. For renewable energy systems above 750 kW the tariff is determined by auctions, in which future system operators need to submit bids. Below the threshold of 750 kW, the subsidy value is calculated based on the accepted bid values from past auctions. While the subsidy is only granted to solar PV systems with a size of up to 10 MW[11], subsidization of wind turbines does not depend on the installed capacity volume.

The entitlement to EEG subsidy comes along with some restrictions. The consumption of own electricity from a renewable energy system is only permitted to owners of small systems with a size not exceeding 750 kW.[12] This threshold applies to wind turbines and solar PV separately[13], therefore a combined system of up to 1500 kW could benefit from this permission. The concept of own consumption is very narrowly defined. Own consumption – in the legal sense – is only given, when the generated electricity is consumed by the same legal entity, meaning that the consumption by a subsidy is not considered as own consumption, legally. Moreover, consumption must occur within immediate vicinity to the renewable energy source and cannot pass through the public

power grid.[14] As a consequence of the consumption of self-generated power, operators can avoid various levies or even benefit from complete exemptions, e.g., from fees incurred by utilization of the public grid.

However, operators of renewable energy systems with a size above 750 kW can still utilize electricity from their own facility when consumption occurs through another legal entity, which can be a subsidiary of the system operator. From a legal standpoint, this is not considered own consumption, but a direct sale[15] of electricity, which applies when electricity is sold to a third party within immediate vicinity without feeding the power through the public grid.[16]

The proposed scenario is subject to various statutory levies, charges and taxes, which can incur at different rates depending on the type of electricity consumption.

Grid-supplied electricity generally is burdened with the EEG levy[17], which allocates the cost of the EEG subsidy to all electricity consumers, and various levies and charges for grid access and utilization. Those grid-related burdens are composed of the transmission charge[18], which is the main fee for grid utilization, and various levies coupled to the transmission charge. Those levies include the concession charge, which results from fees payable by grid operators to local authorities for installation of the electricity infrastructure on public property, the CHP levy[19], which covers the cost of a subsidy granted to promote combined heat and power (CHP) plants, the offshore grid levy, charged to cover the costs related to the network link between offshore electricity generation facilities and the public grid, and the so-called §19 StromNEV levy, which covers any missing amount for transmission grid expenses resulting from reductions or exemptions from the transmission charge. In addition, there is a levy for interruptible loads for allocation of the cost accrued due to reduction of the grid load to ensure power grid stability at times of electricity undersupply.[20]

Besides those levies and charges, electricity tax, value-added tax and income tax are imposed. The value-added tax is not considered in the calculation, since it is reimbursed.

There are several reliefs on taxes and levies applicable to the described scenario, primarily due to the use of a power-to-gas facility. Electricity applied by industrial companies in an electrolysis process is exempt from the electricity tax[21]. Another exemption is provided for the transmission charge, which is not imposed on facilities for the production of hydrogen by water electrolysis.[22] After a law change entered into effect in May 2019 the current wording of the relevant law requires re-electrification of the generated hydrogen for eligibility of the exemption. This requirement was

---

[10]German term: „Erneuerbare-Energien-Gesetz".

[11]Cf. § 37 (3) EEG.

[12]Cf. § 27a EEG.

[13]Since wind turbines and solar PV are different renewable technologies, no aggregation of the two systems is required for legal purposes in accordance with § 24 (1) Nr. 2 EEG.

[14]Cf. § 3 Nr. 19 EEG.

[15]German term: „Direktlieferung".

[16]Cf. § 21b (4) Nr. 2 EEG.

[17]German term: „EEG-Umlage".

[18]German term: „Netzentgelte".

[19]German term: „KWKG-Umlage".

[20]Cf. Bundesnetzagentur (2019d).

[21]Cf. § 9a (1) Nr. 1 StromStG (Electricity Tax Act).

[22]Cf. § 118 (1) EnWG (Energy Industry Act).

recently revoked retroactively and therefore the exemption is still valid irrespective of the hydrogen utilization further down the value chain.[23]

In addition, companies with high electricity cost burdens can benefit from reduced rates with regard to almost all of the other levies. For instance, electricity consumers with a yearly consumption above 30 MWh and a load exceeding 30 kW in at least two months of a year receive a rebate on the concession charge. There is an enormous potential for reduction of the EEG levy for companies operating within specific industrial branches, provided their electricity cost intensity lies above a certain threshold. Producers of industrial gases, such as hydrogen, are generally eligible for a reduction of the EEG levy, as regulated under the special compensation scheme.[24] In addition to activity within one of the specified branches[25], eligibility for the cost relief requires a high electricity cost burden, measured as the proportion of electricity cost to the company's gross value added. The applied electricity rate is standardized and depends on the company's consumed full load hours per year and the total yearly electricity consumption. The lowest rate for 2019 lies slightly above 11 € ct/kWh.[26] For a rough estimate, to explore if PtG facilities are eligible for a reduction of the EEG levy, the gross value added per unit is approximated based on a medium-scale supply price of industrial hydrogen of 3.0 € /kg[27] and assuming an electricity consumption of 50 kWh/kgH$_2$. The application of the least beneficial electricity cost rate of 11 € ct/kWh results in an electricity cost per kilogram of hydrogen of 5.50 € /kg. This value compared to the revenue per kilogram, which for calculation of the gross value added would still be reduced by variable and fixed costs, lies well above the threshold required for eligibility. This is even true, when a price much higher than the industrial price of hydrogen applies. Hence, PtG facilities satisfy the requirements and can benefit from a reduced EEG levy under the special compensation scheme. The reduced levy does not apply on the entire electricity consumption. The first gigawatt hour is always non-exempt and subject to the full EEG levy. Electricity consumption beyond the non-exempt volume is burdened with a levy rate reduced to 15% of the base rate. However, the total cost of the levy for consumption above 1 GWh is limited to a maximum of 0.5% of the company's gross value added, while the resulting levy can never fall below the absolute minimum of 0.1 € ct/kWh.[28] Again, using the industrial hydrogen price as upper bound of the gross value added the maximum levy cost would result in a total of 0.015 € [29], which must be allocated to the 50 kWh required per kgH$_2$. Since the levy per kWh results in a value much lower than the defined minimum of 0.1 € ct/kWh, the minimum rate is the applicable levy rate for electricity consumption above 1

GWh. This does not change, even when calculating with a hydrogen price of 10 € .

The same scheme – but with another minimum rate – applies to the CHP levy[30] and the offshore grid levy[31], which refer to the section on the special compensation scheme of the EEG. Applying the same logic as above to determine the applicable levy rates, shows that the specified minimum rates also are the determining rates, which is 0.03 € ct / kWh in both cases.

Although, the §19 StromNEV levy does not refer to the EEG provision of the special compensation scheme, it follows a similar approach. The full base rate is charged for the first gigawatt hour, above that a reduced rate exists for industrial companies with an electricity cost burden above 4%, measured as the proportion of electricity cost to sales. This is also given, in accordance with the estimate shown above.

All grid-related burdens can be avoided by consumption of self-produced electricity from the renewable energy system due to the lack of grid utilization. Hence, in case of own power consumption only the EEG levy is charged and the described scheme for reduction of the levy also applies. In addition, when own consumption – in the legal sense and in contrast to a direct sale – is given, the EEG levy is reduced to 40% of the base rate.[32] Remember, this discount is only available to operators of renewable energy systems not exceeding 750 kW, with wind and solar power being considered separately for this threshold.

Table 1 gives an overview of the relevant levies, charges and taxes and indicates which fees are imposed depending on the source of power supply:

It should be noted that the EEG subsidy is only granted when the electricity generated by the renewable energy source is fed into the public power grid – in the form of a feed-in premium. Hence, under the current legislation electricity absorbed by the PtG facility is not entitled to the EEG subsidy, which creates a huge barrier for the conversion of electricity from the RES and thus the production of renewable hydrogen.

In a very limited scope, the EEG provides ways for an entire relief from the EEG levy payable on the consumption of own electricity in a few, rather unfavorable, cases. The only case, compatible with the described setup, would require the operator of the vertically integrated energy system to entirely forgo the EEG subsidy and utilize exclusively renewable energy for hydrogen conversion, no grid power, to be eligible for a complete waiver of the EEG levy.[34] Excess power, which cannot be absorbed by the electrolyzer could still be sold at the power exchange, but without a subsidy granted.

This scenario has a huge disadvantage, since the loss of the subsidy needs to be compensated for by the profits on hydrogen sale. However, this scenario will also be analyzed

[23]Cf. Resolution 383/19 (30 August 2019).

[24]Cf. §63 - 69a EEG ("Besondere Ausgleichsregelung").

[25]Cf. Appendix 4 EEG in connection with § 64 (2) EEG.

[26]Cf. BAFA (2019), p. 14.

[27]Cf. Glenk and Reichelstein (2019b), p. 218.

[28]Cf. § 64 (2) Nr. 2 a) and Nr. 3 a) and Nr. 4 b) EEG.

[29]0.5% of 3.0 € /kg (medium-scale supply industrial hydrogen price).

[30]Cf. § 27 (1) KWKG.

[31]Cf. § 17f (5) EnWG in connection with § 27 (1) KWKG.

[32]Cf. § 61b EEG.

[33]Own research.

[34]Cf. § 61a Nr. 3 EEG.

Table 1: Statutory levies, charges and taxes imposed on electricity consumption.[33]

	Own consumption	Direct sale	Grid supply
EEG levy	40%	100%	
Transmission charge	Exemption for PtG facilities		
Concession charge	No		Yes
CHP levy	No		Yes
Offshore grid levy	No		Yes
§ 19 StromNEV levy	No		Yes
Levy for interruptible loads	No		Yes
Electricity tax	Exemption for PtG facilities		

in one case in section 4 of the thesis.

2.4. Model framework

2.4.1. Requirements to the model framework

In addition to the implementation of the regulatory framework presented in the previous section, the effects of conceivable policy changes shall be analyzed. Therefore, the proposed model framework should be designed to support a range of policy changes, explained briefly below. In addition, some other features are presented, which need to be considered within the model framework.

One analyzed policy measure is the decoupling of the EEG subsidy from the feed-in requirement and the payment of a production premium instead, which the operator of the renewable energy system receives irrespective of a power feed-in. This measure would remove the barrier to utilize electricity from the renewable source, especially for sizes above 750 kW, where the benefit of a reduced EEG levy is not provided. Thus, such a measure could promote the generation of renewable hydrogen, compared to a grid-power hydrogen production.

Another supportive policy involves an extension of the permission of own consumption in the legal sense, which under current legislation is only granted to systems sized up to 750 kW. As a consequence, the addressed systems would benefit from the reduction of the EEG levy to 40% of the full rate. The NPV framework should also allow for a waiver of the entire EEG levy and other statutory fees on self-produced renewable energy and/or grid-supplied power absorbed by a power-to-gas facility. Not only the system operators would benefit from such a policy, but it could also help to balance demand and supply and relieve the power grid, if the right incentives are set. Furthermore, a scenario based on (rising) carbon emission prices needs to be implemented within the model framework.

Beyond the representation of the regulatory framework and the described policy changes, the developed tool shall support several other features.

The framework must allow for either the use of a stand-alone wind turbine or solar PV system, or a hybrid system combining both electricity generation technologies. It should also be feasible to switch between different operating modes:

an integrated mode, which allows power consumption from the renewable energy source and the public grid, a grid-only mode and a res-only mode, which consider electricity consumption from one of the two sources only. Regarding the EEG subsidy, there must be the possibility to compare a scenario with and without granting of the subsidy. There should exist a mode considering the capture of oxygen, a by-product from water electrolysis, to allow for an assessment of the effect on profitability resulting from oxygen commercialization.

2.4.2. Base model

The methodology of my thesis builds on the model framework developed by Glenk and Reichelstein[35], which can be applied to calculate the net present value of a vertically integrated system. I will present the main characteristics of the model in this section based on its derivation in the referred article. In the subsequent section I will show the model adjustments, which I conducted in order to satisfy the defined requirements. For a detailed insight into the model construction, I recommend the referenced article for a read.

The authors derived a formula for the calculation of the net present value of vertically integrated systems, which can be separated in three terms: the stand-alone NPVs of the two subsystems renewable energy system and power-to-gas facility, and the synergistic value resulting from their integration. The NPV derivation is based on an optimized management of the PtG operations, which is adjusted to the time-variant electricity prices and power output from the intermittent renewable energy source in order to identify the best scheduling strategy for power conversion.[36] Therefore, the selling price of electricity from the renewable energy source, denoted by $p^s(t)$, is compared with the buying price of electricity from the grid $p^b(t)$, and the conversion value of hydrogen CV_h. The model assumes constant yearly prices and an unchanging distribution of the yearly power generation during the system lifetime, denoted by T. While the price of grid-supplied electricity can fall below zero, own electricity is never sold at a negative price, but the renewable energy system is rather

---

[35]Glenk and Reichelstein (2019a).
[36]Cf.Glenk and Reichelstein (2019a), p. 6.

switched off or the generated power is consumed internally, provided that grid power is not available at a lower rate.

Formally, the model builds on the assumption:

$$p^s(t) \begin{cases} \leq p^b(t) & \text{if } p^b(t) \geq 0 \\ = 0 & \text{if } p^b(t) < 0 \end{cases} \qquad (1)[37]$$

In accordance with equation (2), the conversion value is computed as the product of the contribution margin per kilogram (kg) of produced hydrogen (H2) and the conversion rate of the PtG facility, denoted by η with the unit kg/kWh. The contribution margin in € /kg equals the difference between the time-invariant selling price of hydrogen, denoted by p_h in € /kg, and the variable cost of hydrogen conversion, denoted by w_h in € /kg, excluding the cost of electricity.

Thus, the conversion value of hydrogen is:

$$CV_h = \eta \cdot (p_h - w_h). \qquad (2)$$

Whenever CV_h exceeds $p^s(t)$ or $p^b(t)$, running the PtG facility yields a positive contribution margin, because the value of each kWh converted to hydrogen exceeds the cost of electricity. Thus, the PtG facility would be switched on.

When $p^s(t)$ falls below $p^b(t)$ and CV_h, synergies arise from the system integration. At that point in time, a stand-alone PtG facility would either use the more costly grid power or would be idle, while in the integrated scenario the cheaper power from the renewable energy source can be leveraged.

There are four phases with varying constellations of the electricity prices and the conversion value, as displayed in Figure 1.

The optimized contribution margin representing all phases of the integrated system equals:

$$CM(t \mid k_e, k_h) = p^s(t) \cdot CF(t) \cdot k_e + \left[p^{b+}(t) - p^b(t) \right] \cdot k_h$$
$$+ \left[p^+(t) - p^s(t) \right] \cdot z(t \mid k_e, k_h)$$
$$(3)[39]$$

where $p^{b+}(t) = \max \left\{ p^b(t), CV_h \right\}$
and $p^+(t) = \max \left\{ \min \left\{ p^b(t), CV_h \right\}, p^s(t) \right\}$
and $z(t \mid k_e, k_h) = \min \left\{ CF(t) \cdot k_e, k_h \right\}$

The variables k_e and k_h refer to the system sizes of the RES and PtG facility in kilowatt (kW) with the subscripts e and h, respectively. $CF(t)$ represents the time-variant capacity factors of the RES. The term $z(t|k_e, k_h)$ equals the maximum amount of electricity from the renewable source, which can be utilized by the PtG facility at a point in time. It is determined by the lower value of the power output from the RES or the installed capacity of the PtG plant.

The authors leveraged a general NPV formula, composed of the sum of all yearly discounted cash flows minus the total cost of the investment as shown in equation (4), and transformed it into an NPV term based on unit cost expressions and average prices.

$$NPV(k_e, k_h) = \sum_{i=0}^{T} CFL_i(k_e, k_h) \cdot \gamma^i - \text{Investment} . \quad (4)$$

The yearly cash flows are determined by the sum of the hourly optimized contribution margin reduced by the yearly fixed operating costs. The investment cost is determined based on system prices and installed capacities of the subsystems. A convenient transformation of equation (4) results in the net present value term of vertically integrated systems:

$$NPV(k_e, k_h) = (1 - \alpha) \cdot L \cdot [(\Gamma^s \cdot p^s - LCOE) \cdot CF \cdot k_e$$
$$+ (p^{b+} - p^b - LFCH) \cdot k_h$$
$$+ (p^+ - p^s) \cdot z(k_e, k_h)].$$
$$(5)[40]$$

The first of the three terms, $(1 - \alpha) \cdot L \cdot (\Gamma^s \cdot p^s - LCOE) \cdot CF \cdot k_e$, equals the stand-alone NPV of a RES, the second term, $(1 - \alpha) \cdot L \cdot (p^{b+} - p^b - LFCH) \cdot k_h$, is the stand-alone NPV of a PtG facility and the third term, $(1 - \alpha) \cdot L \cdot (p^+ - p^s) \cdot z(k_e, k_h)$, represents the synergistic value attributed to the integrated system.

α denotes the applying income tax rate. $L = m \cdot \sum_{i=0}^{T} x^{i-1} \cdot \gamma^i$ is the levelization factor, which "expresses the discounted number of hours that are available from the facility over its entire lifetime".[41] m is the number of hours per year and equals $365 \cdot 24 = 8760$. The yearly degradation is denoted by the factor x, which leads to a reduction of the capacity of both the renewable energy system and PtG facility. The discount factor is expressed by $\gamma = \frac{1}{1+r}$, where r is the weighted average cost of capital (WACC). The variables CF, p^s, p^b, p^{b+} and p^+ denote the average value of the underlying time-variant data. For instance, the average capacity factor (CF) is computed by $CF = \frac{1}{m} \int_0^m CF(t) dt$. The other variables are computed accordingly.

The levelized cost of electricity (LCOE) allocates the investment cost and fixed operating costs of the renewable energy source to its lifetime electricity production and thus expresses the discounted average cost per kWh of the produced energy:

$$LCOE = w_e + f_e + \Delta \cdot c_e \qquad (6)$$

Since a renewable energy system has insignificant variable operating costs: $w_e = 0$.

---

[37] All formulas in section 2.4.2 are cited from Glenk and Reichelstein (2019b).

[38] Source: Glenk and Reichelstein (2019a), p. 7.

[39] Refer for the proof to "Proof of Lemma 1" in the Appendix of Glenk and Reichelstein (2019a).

[40] For a mathematical proof refer to "Proof of Proposition 1" in the Appendix of Glenk and Reichelstein (2019b).

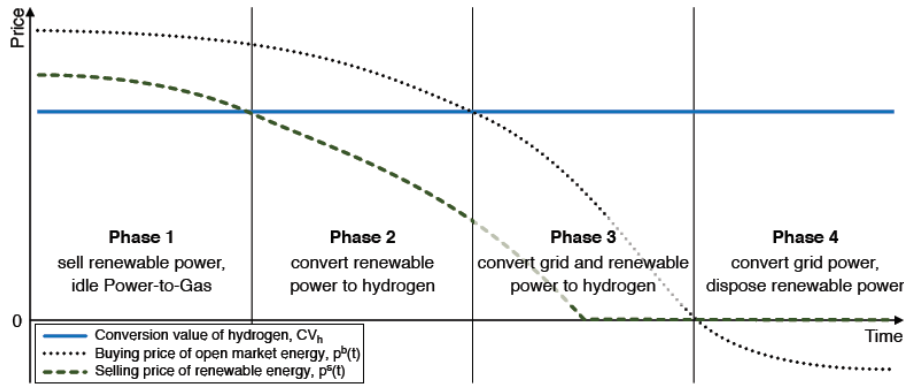[41] Glenk and Reichelstein (2019b), p. 10-11.

Figure 1: Phase diagram.[38]

The levelized fixed operating costs $f_e = \frac{\sum_{i=1}^{T} F_{ei}\gamma^i}{CF \cdot L}$ are based on the yearly fixed cost F_{ei}, where i denotes the year. The levelized capacity cost $c_e = \frac{SP_e}{CF \cdot L}$ is based on the system price SP_e, which indicates the cost per installed kW of the RES. Both costs are levelized on a per kWh basis.

The tax factor is expressed by $\Delta = \frac{1-\alpha \cdot \sum_{i=1}^{T} d_i \cdot \gamma^i}{1-\alpha}$ and is based on the income tax rate α, the yearly tax depreciation rate d_i and the discount factor γ.

The construction of the levelized fixed cost of hydrogen (LFCH) follows the same logic, but does not consider the variable operating cost of hydrogen production, which are already accounted for in the conversion value of hydrogen CV_h in equation (2):

$$LFCH = f_h + \Delta \cdot c_h \qquad (7)$$

where $f_h = \frac{\sum_{i=0}^{T} F_{hi}\gamma^i}{L}$ and $c_h = \frac{SP_h}{L}$.

The denominators disregard a capacity factor, since PtG is a dispatchable technology. LFCH therefore expresses the investment and fixed cost of the PtG facility allocated to every kWh, which could be converted to hydrogen assuming a full load of the electrolyzer during its lifetime. Thus, at times when the PtG facility is idle, LFCH represents the loss resulting from the investment in PtG, which must be compensated for by the contribution margin gained during production hours for overall profitability of the PtG subsystem.

The term for the stand-alone net present value of a renewable energy system, applies the co-variation coefficient Γ^s to the average revenues p^s per kWh produced. This factor adjusts for the variation between electricity price and power generation output based on their average values.[42] This adjustment allows that the NPV term is computed based on the average values of the electricity selling price, p^s, and the capacity factor, CF.

The co-variant coefficient is defined as:

$$\Gamma^s = \frac{1}{m} \cdot \int_0^m \frac{CF(t)}{CF} \cdot \frac{p^s(t)}{p^s} dt. \qquad (8)$$

Similarly, the third component of the NPV term in equation (5), needs to be adjusted for the variation between the price premium of hydrogen, $p^+(t) - p^s(t)$, and the hydrogen output at every point in time.[43] This adjustment is carried out by the term:

$$z(k_e, k_h) = \frac{1}{m} \cdot \int_0^m z(t \mid k_e, k_h) \cdot \frac{p^+(t) - p^s(t)}{p^+ - p^s} dt \qquad (9)$$

2.4.3. Model alterations

In order to align the model framework of Glenk and Reichelstein with the effective legislation, described in section 2.3, and provide the flexibility to switch between different scenarios, some adaptations of the original model are necessary.

The model shall support the current legislation and conceivable policy changes, such as the waiver of the feed-in requirement and supplementary reliefs regarding the statutory fees on electricity imposed on PtG operators. It shall also account for the possibility of a carbon price and the commercialization of the oxygen by-product. Furthermore, the model framework should allow for hybrid renewable energy systems composed of solar PV and wind turbines and be compatible with different operating modes.

To facilitate an easy implementation of the net present value model, I introduce Boolean parameters, which can take the value 0 or 1 to specify the analyzed scenario (see Table 2).

For instance, when the renewable energy system is entitled to the EEG subsidy ($sub = 1$) in the form of a production premium: $fit = 0$; and in the form of a feed-in premium: $fit = 1$.

In case the parameter $tax^{res} = 0$, the electricity absorbed directly from the renewable energy source is assumed to be exempt from statutory levies, charges and taxes. The Boolean parameter tax^{grid} controls the fees applicable to grid-supplied electricity, accordingly.

The selling price of electricity from the renewable energy source and the buying price of grid-supplied electricity is

---

[42]Cf. Glenk and Reichelstein (2019b), p. 12.

[43]Cf. Glenk and Reichelstein (2019b), p. 13.

Table 2: Boolean parameters for scenario specification.

Description	Boolean	Value
Eligibility to EEG subsidy	sub	
Feed-in requirement	fit	
Onsite consumption of hydrogen	$onsite$	Yes = 1 /
Liability to pay statutory fees on grid-supplied power	tax^{grid}	No = 0
Liability to pay statutory fees on renewable power	tax^{res}	
Consideration of an emission price	em	
Consideration of the oxygen by-product	$o2$	

based on the wholesale electricity price at the electric power exchange, in the case of Germany mainly at the European Power Exchange (EPEX). Therefore, I denote the time-variant electricity price by $epex(t)$. It is assumed that renewable energy systems, unlike other power generation technologies, have the flexibility to curtail power production immediately when prices become negative. Therefore, own electricity is not sold at a negative price, but at the rate expressed by $epex^+(t) = \max{epex(t), 0}$. In contrast, the price of grid-supplied electricity can fall below zero and is based on the unadjusted $epex(t)$.

Under the current legal framework both own electricity and grid-supplied electricity consumption are subject to statutory fees. As explained in section 2.3, some of fees are based on a fixed rate, while others have a varying average rate, that changes depending on the consumption quantity. Among the latter the EEG levy is the one with the greatest impact. It is convenient to group the statutory fees in two different variables based on their varying design, called $fees_{fix}$ and $fees_{var,i}$. They will be completed by the superscripts res and grid to indicate the source of power they refer to. $fees_{var,i}$, which expresses the average fee allocated per kWh, can vary along the years of system lifetime, since the system degradation leads to a decrease of the yearly power consumption, which again influences the average fee rate. Section 3.3 covers the derivation of the fee rates of $fees_{var,i}$, which must be determined in an iterative approach.

In addition to the wholesale electricity price and the statutory fees, grid-supplied power is also subject to a price markup resulting from the cost of electricity trading, and the cost of emissions in case carbon emissions are considered in the analysis ($em = 1$).

Overall, the buying price of electricity can be expressed by:

$$p^b(t) = epex(t) + markup + tax^{grid} \cdot \left(fees_{fix}^{grid} + fees_{var,i}^{grid} \right)$$
$$+ em \cdot \Delta \, cost_{em}^{el}(t) \quad (10)$$

$\Delta \, cost_{em}^{el}(t)$ represents the hourly emission cost per kWh of electricity at time t. Its derivation will be shown later in this section.

Beyond statutory fees, the price of own electricity consumption includes the opportunity cost that results from the

lost profit due to the internal power absorption instead of a sale at the power exchange. This cost does not accrue in reality but needs to be considered as part of the optimization of the electrolyzer operating strategy. It is composed of the cash flows, which would be obtained only when renewable electricity was sold at the market. These comprise the non-negative electricity price $epex^+(t)$ and the subsidy premium, provided the subsidy is tied to the power feed-in. If the feed-in requirement is waived, then the premium is not considered as an opportunity cost, since it is obtained even when own power is absorbed internally.

Note, the resulting price does not represent the electricity selling price, which is defined by $epex^+(t)$ only, but rather the imputed price for consumption of own electricity. Still, I remain with the original variable $p^s(t)$ and define it by:

$$p^s(t) = epex^+(t) + tax^{res} \cdot \left(fees_{fix}^{res} + fees_{var,i}^{res} \right)$$
$$+ fit \cdot premium(t). \quad (11)$$

The conversion value of electrolysis is initially based on the conversion value of hydrogen. As soon as the oxygen by-product is also captured and commercialized, the conversion value can increase, provided that the price of oxygen covers the additional variable operating costs.

During the process of electrolysis, water – chemically denoted by H_2O – is split into hydrogen (H_2) and oxygen (O_2). Since the water molecule is composed of two hydrogen and one oxygen atom, the resulting quantity of hydrogen molecules is twice the quantity of oxygen molecules. However, hydrogen and oxygen possess different standard atomic weights of 1.008 and 15.999, respectively.[44] To obtain the oxygen output per kg produced from 1 kWh electricity, it is necessary to relate the conversion rate of hydrogen (in kgH_2/kWh) to both the ratio of the molecule numbers and atomic weights.

The resulting conversion rate of oxygen (in kgO_2/kWh) is:

$$\eta_o = \frac{15.999}{2 \cdot 1.008} \cdot \eta_h \approx 8 \cdot \eta_h.$$

In other words, each kg of hydrogen captured through electrolysis, has a by-product of around 8 kg of oxygen, which

---

[44]CIAAW: Atomic weights of the elements 2017. Available online at www.ciaaw.org.

is generally vented as a waste product. The resulting total conversion value is composed of the conversion value of hydrogen and oxygen:

$$CV = CV_h + CV_o \cdot o2 \tag{12}$$

where $CV_o = \frac{15.999}{2 \cdot 1.008} \cdot \eta_h \cdot (p_o - w_o)$.

p_o and w_o denote the price of oxygen and the variable operating costs of oxygen capture, respectively. $o2$ is the Boolean operator, which determines whether oxygen is considered for the calculation of the NPV.

In addition to the variable operating costs, the capture of oxygen might require further investments and/or regular maintenance and thus can cause fixed costs.

Hence, the term levelized fixed cost of oxygen (LFCO) captures the capacity cost and yearly fixed operating costs related to oxygen operations, analogous to the construction of LFCH:

$$LFCO = (f_o + \Delta \cdot c_o) \cdot o2 \tag{13}$$

$$\text{where} \quad f_0 = \frac{\sum_{i=1}^{T} F_{oi}\gamma^i}{L}$$
$$\text{and} \quad c_o = \frac{SP_o}{L}$$

An adaptation of the conversion value of hydrogen $CV_h = \eta_h \cdot (p_h - w_h)$ is also necessary due to the distinction between onsite utilization of hydrogen and the sale of hydrogen. Furthermore, the price of carbon emissions, imposed on fossil hydrogen, needs to be incorporated into the equation. Both adjustments concern the price of hydrogen:

$$p_h = \text{price}_{H_2} + em \cdot \Delta \, \text{cost}_{em,i}^{\text{fossil}} + \text{onsite} \cdot \text{cost}_{\text{transport}} \tag{14}$$

price_{H_2} represents the base price of hydrogen, which generally already includes a carbon price, since hydrogen production is subject to the EU Emissions Trading System (EU ETS).

When putting into effect a rising carbon price, the emission price markup on fossil hydrogen, expressed by $\Delta \cos t_{em,i}^{fossil}$, raises the price of hydrogen. The consideration of emission prices is explained in the subsequent paragraph. The cost of transportation needs to be included in the hydrogen price in case the produced hydrogen is consumed onsite. It reflects the price advantage gained from onsite production from the avoided cost of transportation, which are payable for hydrogen obtained externally.

I assume that an emission price is entirely added to the price of hydrogen or grid power, respectively. $\Delta \text{cost}_{em,i}^{fossil}$ expresses the cost of emission certificates required to cover up for the carbon emissions per kg of fossil hydrogen, when considering a yearly varying emission price, denoted by $\sigma_{\text{target},i}$ in $\text{\euro}/tCO_2eq$. σ_h expresses the emission price relevant for the production of fossil hydrogen in the base year and thus needs to be deducted from the target price. The emission factor of fossil hydrogen ε_h (in $gCO2eq/kgH_2$) must be multiplied with the offset of the emission prices to obtain the emission cost in $\text{\euro}/kg$:

$$\Delta \, \text{cost}_{em,i}^{\text{fossil}} = \left(\sigma_{\text{target},i} - \sigma_h\right) \cdot \varepsilon_h \cdot 10^{-6} \tag{15}$$

The buying price of grid electricity also has an emission cost component, which was mentioned in equation (10). $\Delta \text{cost}_{em}^{el}(t)$ represents the time-variant emission cost assigned per kWh at any point in time t based on the hourly national electricity mix:

$$\Delta \, \text{cost}_{em}^{el}(t) = \left(\sigma_{\text{target},i} - \sigma_{el}\right) \cdot \frac{\sum_{type=1}^{n} E_{type} \cdot G_{type}(t)}{\sum_{type=1}^{n} G_{type}(t)} \cdot 10^{-6} \tag{16}$$

Again, it is necessary to adjust the target emission price, $\sigma_{\text{target},i}$, by the actual emission price, σ_{el}, valid in the year from which the price data originates in order to account for absolute emission prices. Therefore, the difference between the two emission prices is calculated and then multiplied with the respective emission factor at time t, expressed by the fraction in equation (16). The emission factor is computed as the weighted average carbon intensity of the time-variant electricity mix, which is derived based on the carbon intensity E_{type} of each power generation type measured in gCO_2eq/kWh and the time-variant output $G_{type}(t)$ of each source of power generation in kWh.

This approach assumes a yearly constant electricity mix during system lifetime. Since a rising emission price would affect the composition of the electricity mix, the emission scenario can only suggest tendencies caused by a rising emission price and should not be interpreted in detail.

The original model was based on the distinction of four different price phases displayed in Figure 1. The four phases need to be extended to six phases in order to represent all price constellations potentially occurring in the altered model. This is due to the fact that some of the underlying assumptions of the original model are violated. In fact, the dependence of $p^s(t)$ on $p^b(t)$ does not follow the rule defined in (1), when statutory fees are imposed. In that case $p^s(t)$ never obtains the value zero due to the regulatory burden, while the original model is designed for $p^s(t)$ to be zero in case of a negative spot market price.

Therefore, the phases need to be extended as can be seen in Table 3

Based on these phases the term expressing the optimized contribution margin, presented by equation (3), needs to be updated to represent the changed definitions of the electricity prices and the conversion value, and its applicability needs to be validated.

In phase 1, the contribution margin of the vertically integrated system only comprises the revenues from the renewable energy system, since the PtG facility is idle. The contribution margin thus includes the gained electricity price and the subsidy premium, if granted:

$$CM_1(t \mid k_e, k_h) = \left(epex^+(t) + sub \cdot premium(t)\right) \cdot CF(t) \cdot k_e \tag{17}$$

In phase 2, as much power from the renewable energy system is utilized, as the PtG facility can absorb. Any re-

Table 3: Distinction of price phases in the adapted model.

Phase	Original model	Adapted model	PtG operating mode
1	$p^b(t) \geq p^s(t) \geq CV_h \geq 0$	$p^b(t) \geq p^s(t) \geq CV \geq 0$	PtG facility idle
2	$p^b(t) \geq CV_h > p^s(t) \geq 0$	$p^b(t) \geq CV > p^s(t) \geq 0$	Electrolysis from RES power
3	$CV_h > p^b(t) \geq p^s(t) \geq 0$	$CV > p^b(t) \geq p^s(t) \geq 0$	Electrolysis from RES & grid power
4	$CV_h \geq p^s(t) = 0 > p^b(t)$	$CV \geq p^s(t) > p^b(t)$	Electrolysis from grid power
5	n/a	$p^s(t) > p^b(t) \geq CV \geq 0$	PtG facility idle
6	n/a	$p^s(t) > CV > p^b(t)$	Electrolysis from grid power

maining power is fed into the grid. The contribution margin equals:

$$CM_2(t \mid k_e, k_h) = \left(epex^+(t) + sub \cdot premium(t)\right) \cdot CF(t) \cdot k_e$$
$$+ \left[CV - p^s(t)\right] \cdot z(t \mid k_e, k_h)$$
(18)

In phase 3, electricity from the renewable source is absorbed with priority. In case the PtG facility disposes any excess capacity beyond that, grid power is utilized to run the electrolyzer at full capacity:

$$CM_3(t \mid k_e, k_h) = \left(epex^+(t) + sub \cdot premium(t)\right) \cdot CF(t) \cdot k_e$$
$$+ \left[CV - p^b(t)\right] \cdot k_h$$
$$+ \left[p^b(t) - p^s(t)\right] \cdot z(t \mid k_e, k_h)$$
(19)

In phase 4, the entire renewable electricity is fed into the grid, or curtailed in case of a negative wholesale electricity price, and only grid-supplied electricity, which is cheaper at those times, is converted in the PtG facility, leading to the contribution margin:

$$CM_4(t \mid k_e, k_h) = \left(epex^+(t) + sub \cdot premium(t)\right) \cdot CF(t) \cdot k_e$$
$$+ \left[CV - p^b(t)\right] \cdot k_h$$
(20)

Phase 5 exhibits the same operation mode as phase 1 and the operations in phase 6 equal those in phase 4. This means that
$$CM_5(t \mid k_e, k_h) = CM_1(t \mid k_e, k_h)$$
$$\text{and } CM_6(t \mid k_e, k_h) = CM_4(t \mid k_e, k_h)$$
The contribution margin terms of all phases are aggregated in one term using the auxiliary price variables $p^{b+}(t)$ and $p^+(t)$ from the original model, defined under equation (3), such that:

$$CM(t \mid k_e, k_h) = \left(epex^+(t) + sub \cdot premium(t)\right) \cdot CF(t) \cdot k_e$$
$$+ \left[p^{b+}(t) - p^b(t)\right] \cdot k_h$$
$$+ \left[p^+(t) - p^s(t)\right] \cdot z(t \mid k_e, k_h)$$
(21)[45]

Based on the adapted optimized contribution margin the net present value term can be updated, following the approach of the original model.

The net present value of a vertically integrated system is defined as:

$$NPV(k_e, k_h) = (1 - \alpha) \cdot L \cdot$$
$$\left[\left(\Gamma^{ep^+} \cdot \overline{epex^+}_m + sub \cdot premium - LCOE\right) \cdot \overline{CF}_m \cdot k_e\right.$$
$$+ \left(\Gamma^{b+b}_{x\gamma} \cdot \left(\overline{p^{b+}}_{Tm} - \overline{p^b}_{Tm}\right) - LFCH - LFCO\right) \cdot k_h$$
$$\left. + \left(\overline{p^+}_{Tm} - \overline{p^s}_{Tm}\right) \cdot z_{x\gamma}(k_e, k_h)\right]$$
(22)[46]

where $\Gamma^{ep^+} = \frac{1}{m} \cdot \int_0^m \frac{epex^+(t)}{\overline{epex^+}_m} \cdot \frac{CF(t)}{\overline{CF}_m} dt$

and $premium = \frac{1}{m} \cdot \int_0^m premium(t) \cdot \frac{CF(t)}{\overline{CF}_m} dt \cdot \frac{\sum_{i=1}^{T_{eeg}} \gamma^i \cdot x^{i-1}}{\sum_{i=1}^{T} 1 \gamma^i \cdot x^{i-1}}$

and $\Gamma^{b+b}_{x\gamma} = \frac{1}{T \cdot m} \cdot \int_0^{T \cdot m} \frac{p^{b+}(t) - p^b(t)}{\overline{p^{b+}}_{Tm} - \overline{p^b}_{Tm}} \cdot \frac{xy(t)}{\overline{xy}_{Tm}} dt$

and $z_{x\gamma}(k_e, k_h) = \frac{1}{T \cdot m} \cdot \int_0^{T \cdot m} z(t \mid k_e, k_h) \cdot \frac{p^+(t) - p^s(t)}{\overline{p^+}_{Tm} - \overline{p^s}_{Tm}} \cdot \frac{xy(t)}{\overline{xy}_{Tm}} dt$

and $xy(t) = x^{\lfloor \frac{t-1}{m} \rfloor} \cdot \gamma^{\lfloor \frac{t-1}{m} \rfloor + 1}$

Again, the NPV term is composed of the stand-alone NPVs of the renewable energy and power-to-gas subsystems, and the synergistic value obtained by system integration.

Similarly, as done by Glenk and Reichelstein[47], I incorporate the levelized subsidy premium into the stand-alone NPV of a renewable energy system, if a subsidy is granted. Thus, the subsidy is always considered in the NPV of the renewable energy source, independent from the electricity utilization. In case the feed-in requirement is effective, and thus no subsidy is granted for electricity absorbed by the PtG facility, the subsidy is still attributed to the NPV of RES, but at the same time the synergistic value is reduced by the amount of the subsidy premium through $p^s(t)$. The levelization of the premium is necessary, since it is time-variant and the subsidy lifetime, denoted by T_{eeg}, might be shorter than the system lifetime T.

---

[45]The derivation of all contribution margin terms is shown in the Appendix A.1.

[46]The derivation of the NPV term is based on the updated contribution margin and follows the procedure shown in the Appendix of Glenk and Reichelstein (2019b) under "Proof of Proposition 1" (p. 28-30).

[47]Glenk and Reichelstein (2019a), p. 221.

Furthermore, additional co-variation coefficients were introduced, since the assumption of constant prices during system lifetime is violated by a yearly variation of the prices $p^b(t)$ and $p^s(t)$ in the new model framework, which can result from increasing emission prices or from varying statutory fees. Those price variations are compensated by co-variation coefficients, similarly as in the base model. The new coefficients are computed based on time-variant data over the entire lifetime in the interval $[0, T \cdot m]$, instead of originally over the data of one year and thus in the interval $[0, m]$. This changed approach expresses itself in the mean values of the price variables, present in the second and third NPV term. Those are now computed based on the entire lifetime. For clarity, I denote the mean value of a data series $f(t)$ over the time period $[0, T]$ by $\bar{f}_T$, which equals:

$$\bar{f}_T = \frac{1}{T} \cdot \int_0^T f(t) dt$$

Since the calculation of the stand-alone NPV of the renewable energy source, reflected by the first term of (22), still satisfies the assumption of constant prices over the system lifetime, the respective coefficient Γ^{ep^+} can still be calculated on a one year's basis. Consequently, the relevant mean values are constructed over the time span $[0, m]$ and are denoted by $\overline{epex^+}_m$ and $\overline{CF}_m$.

Finally, the term $xy(t)$ is the product of the yearly degradation factor and the discount factor, scaled down to an hourly granularity. Thus, it corresponds to the term "$\gamma^i \quad x^{i1}$", which expresses the same product, but on a yearly basis. All other terms not defined explicitly, such as $p^{b+}(t)$ and $p^+(t)$, are equally defined as in the base model.

The construction of the NPV framework, based on Boolean parameters throughout all layers of the model, allows for the flexibility to switch between scenarios conveniently, such as changing the subsidy form between production and feed-in premium or waiving the imposed statutory fees on renewable energy consumption. Due to its design the scheme can easily be implemented in programming code.

In order to make the NPV term applicable to hybrid renewable energy systems, further adjustments need to be made.

First of all, the total renewable energy system capacity is defined as the sum of wind turbine and solar PV capacity:

$$k_e = k_{wind} + k_{pv} \tag{23}$$

Now, the combined capacity factor, system price and yearly fixed operating costs are constructed from the respective values of both systems, weighted with the ratio of the installed capacity to the total system capacity, such that:

$$CF(t) = \frac{k_{wind}}{k_e} \cdot CF_{wind}(t) + \frac{k_{pv}}{k_e} \cdot CF_{pv}(t) \tag{24}$$

$$SP_e = \frac{k_{wind}}{k_e} \cdot SP_{wind} + \frac{k_{pv}}{k_e} \cdot SP_{pv} \tag{25}$$

$$F_{ei} = \frac{k_{wind}}{k_e} \cdot F_{wind,t} + \frac{k_{pv}}{k_e} \cdot F_{pv,i} \tag{26}$$

The subsidy premium also needs to be adjusted to the hybrid system by computing the average over the premiums received for each renewable technology, weighted by the time-variant power generation of each renewable system as a proportion of total power generation:

$$\begin{aligned} premium(t) = & \frac{CF_{wind}(t) \cdot k_{wind}}{CF(t) \cdot k_e} \cdot premium_{wind}(t) \\ & + \frac{CF_{pv}(t) \cdot k_{pv}}{CF(t) \cdot k_e} \cdot premium_{pv}(t) \end{aligned} \tag{27}$$

This assignment is convenient, since it does not require any further adaptation of the model, however it is subject to some limitations, which are caused by the design of the regulatory framework.

The two renewable subsystems always need to be scaled either both larger than 750 kW or not exceeding the threshold of 750 kW. This results from the policy framework, which favors smaller systems and therefore burdens consumption of self-produced electricity from small renewable systems not exceeding 750 kW with lower statutory fees. If the wind turbine component was sized above the threshold and solar PV below it, or vice versa, the imputed price of own electricity consumption, $p^s(t)$, would be different for the wind turbine and the solar PV system. This would add a fourth price component to the phase distinction shown in Table 3. Consequently, the number of distinguished phases could increase sixfold and thus complicate the model framework significantly.

Finally, the produced hydrogen quantity is measured at each point in time depending on the phase distinction and broken down to hydrogen production from renewable and grid-supplied electricity (see Table 4).

The variable $z(t|k_e, k_h)$ expresses the amount of renewable power absorbed by the PtG facility at any point in time. The degradation factor $x^{(i-1)}$ accounts for decreasing systems performances of the renewable power generation and the electrolyzer. The conversion rate of hydrogen, η_h, is applied to convert the absorbed power volume to the produced hydrogen volume in kg. The produced hydrogen quantity from grid power in phase 3 includes the term "$[k_h - z(t \mid k_e, k_h)]$", which expresses the excess capacity, that the PtG plant can still provide after all renewable electricity has been absorbed.

Beyond the presented characteristics of the adapted model framework, some of the features, mentioned in the specification of the model requirements in section 2.4.1, are not considered in the model design. Instead they are realized programmatically. This concerns the extension of the permission of own consumption, which would benefit systems exceeding 750 kW by a reduction of the EEG levy, and the option to switch between different operating modes of the PtG facility.

Table 4: Production quantity of hydrogen from the different energy sources.

Phase	Power source of PtG operations	Total hydrogen production	Hydrogen production from RES power	Hydrogen production from grid power
1	n/a	0	0	0
2	RES power	$\eta_h \cdot x^{i-1} \cdot z(t \mid k_e, k_h)$	$\eta_h \cdot x^{i-1} \cdot z(t \mid k_e, k_h)$	0
3	RES & grid power	$\eta_h \cdot x^{i-1} \cdot k_h$	$\eta_h \cdot x^{i-1} \cdot z(t \mid k_e, k_h)$	$\eta_h \cdot x^{i-1} \cdot [k_h - z(t \mid k_e, k_h)]$
4	Grid power	$\eta_h \cdot x^{i-1} \cdot k_h$	0	$\eta_h \cdot x^{i-1} \cdot k_h$
5	n/a	0	0	0
6	Grid power	$\eta_h \cdot x^{i-1} \cdot k_h$	0	$\eta_h \cdot x^{i-1} \cdot k_h$

3. Model implementation

3.1. Tool functionalities

Based on the presented model framework, I developed an easy-to-use application using the programming language Python to facilitate the profitability evaluation of vertically integrated energy systems. The tool provides the functionality to run an optimization analysis for detection of the ideally sized renewable energy system for a fixed PtG facility size, i.e., the capacity ratio which minimizes the break-even price of hydrogen.

Due to the efficient simulation of numerous scenarios along a wide range of capacity combinations the implemented tool has a great advantage compared to conventional analysis tools, like Microsoft Excel. It is better suited for the consumption of the large amounts of data and computations, necessary to solve the described optimization problem. The model input can be easily set by the user through a graphical user interface (GUI) and thus does not require knowledge of any programming language, which makes the tool accessible to businesspeople and policymakers without computer science background.

In addition to the capacity optimization mode, which covers the macro perspective of the profitability analysis, a scenario analysis mode enables a closer inspection of the details of a single case, where both renewable energy source and PtG facility have fixed capacities and a projected hydrogen price is set. This mode is helpful to determine the feasibility of an investment in a specific project.

The transformation of the model framework into the implementation of programming code and, in particular, the optimization problems posed some challenges. Later in this chapter, I will discuss the developed approaches to solve these issues.

Moreover, issues related to the cost allocation of the variable statutory fees and the circular dependency between the electricity consumption volume and the average rate of the variable statutory fees needed to be solved. I will describe the applied solution techniques to these problems and show the implementation of the operating modes of the PtG facility.

Selected screenshots of the implemented tool are shown in the Appendix A.6.

3.2. Allocation of statutory fees

PtG facility operators are granted various exemptions or reliefs from fees on the electricity price. Some of the reliefs are structured in a way that a non-exempt volume of electricity is charged with the base rate of the respective fee, while only beyond this volume a lower rate applies. The amount in excess of the reduced rate, which is charged on the non-exempt volume, must then be allocated adequately to the entire consumption volume.

Generally, the entire cost related to a fee is added up and equally allocated to all units of electricity consumed. The resulting mean value of this fee then directly adds to $fees_{var,i}$. This approach can create anomalies and wrong incentives[48], when one source of electricity is subject to a reduced rate, while electricity from another source is burdened with the full rate. This is the case with regard to the EEG levy, when the size of the renewable energy source does not exceed 750 kW[49]. In this case own consumption is burdened with an EEG levy reduced to 40% of the base rate, while power supply from the grid is subject to the entire EEG levy. As long as total consumption lies below the non-exempt threshold of 1 GWh, intuitively renewable energy is assigned its real fee amounting to 40% of the base rate, and the full rate is assigned to grid power. As soon as consumption exceeds the threshold of 1 GWh, the entire cost from the statutory fee is still allocated with the ratio 40:100 in order to avoid jumps in the assigned levy value and facilitate a continuous development of the assigned cost for varying power consumption scenarios. This results in a lower allocation of the burden to power consumption from the RES for calculative purposes. The mathematical expressions of the resulting EEG levy rates are provided in the Appendix A.2.

3.3. Circularity problem of the variable statutory fees

In addition to the EEG levy, there are other reliefs, which only apply for electricity consumption exceeding a non-exempt volume. Those fees, described in section 2.3, are grouped in the variables $fees_{var,i}^{grid}$ and $fees_{var,i}^{res}$, which express the assigned average fee rates and are part of $p^b(t)$ and $p^s(t)$, based on which the operating hours of the PtG facility are determined in accordance with the distinction of the six presented phases.

This means, that the fee rates have a direct effect on the amount of power absorbed by the PtG plant. At the same

---

[48]Explanation is shown in the Appendix A.2.

[49]This threshold applies to the wind turbines and solar PV system separately.

time, both rates are determined as an average value based on the entire power consumption. Thus, there is a circular dependency, where the average fee rate affects the consumed electricity volume, while the volume serves as a basis to determine the average rate. This issue can only be solved by the application of an iterative approach, to determine the rates for which both components, electricity consumption and statutory fees, keep the balance.

Therefore, in the first iteration, the consumption volume is determined for the case of $fees_{var,i}^{grid} = fees_{var,i}^{res} = 0$. Based on the resulting consumption volume, now an initial value can be assigned to the statutory fees. Consequently, the prices $p^b(t)$ and $p^s(t)$ increase, which leads to a decrease of the consumption volume, since conversion to hydrogen becomes less profitable. Again, based on the new consumption volume, $fees_{var,i}^{grid}$ and $fees_{var,i}^{res}$ need to be updated. The updated rates will exceed the previous rates, since the fixed cost associated to the non-exempt consumption volume is allocated to less kWh. This, again, increases the prices for renewable and grid electricity, leading to yet another decrease of electricity consumption. This procedure is continued until the electricity consumption and the statutory fees are not exposed to further value changes, which implies that the final values of $fees_{var,i}^{grid}$ and $fees_{var,i}^{res}$ have been identified.

3.4. Implementation of the PtG operating modes

The model framework allows to calculate the NPV of vertically integrated energy systems, which operate in an integrated mode, using grid-supplied power and electricity generated by the own renewable energy system. This integrated mode is expanded by two additional modes, a "RES-only" mode, which only allows the use of power from the renewable energy source and the "Grid-only" mode, which deploys solely grid electricity for electrolysis. Those operating modes are only programmatically realized following a straightforward approach.

To implement the "RES-only" mode, I set the value of $p^b(t) > CV$ at all times. This guarantees that electricity supplied by the public grid is not utilized for power conversion at any point in time. In reference to the phase distinction in Table 3, it would have the effect that the phases 3, 4 and 6 never occurred. The same logic applies for the "Grid-only" mode based on the value assignment $p^s(t) > CV$, which prevents the phases 2 - 4 from happening.

This simple approach is beneficial, since it does not require any adaptation of the NPV framework.

3.5. Break-even price detection

As defined by Glenk and Reichelstein, a vertically integrated system breaks even, if its net present value equals the sum of the non-negative stand-alone NPVs of its subsystems, such that:

$$NPV(k_e, k_h) = \max\{NPV(k_e, 0), 0\} + \max\{NPV(0, k_h), 0\}$$

$$(28)^{[50]}$$

Consequently, an integrated system must compensate for the negative stand-alone NPVs of its components, while their positive NPVs are not attributed to the NPV of the integrated system for the purpose of determining its break-even point. Thus, a system's hydrogen price is considered its break-even price, if above-listed equation holds.

The break-even price of a vertically integrated energy system is detected by iteratively calculating its NPV while changing the value assigned to the hydrogen price until the break-even condition of Equation (28) is met.

The NPV term can be expressed as the stand-alone NPVs of its components – renewable energy system and PtG facility – plus the synergistic term resulting from their (vertical) integration:

$$NPV(k_e, k_h) = NPV(k_e, 0) + NPV(0, k_h) + NPV(synergies).$$

$$(29)$$

From now on, I denote the NPV components as: $NPV_{RES} = NPV(k_e, 0)$ and $NPV_{PtG} = NPV(0, k_h)$ and $NPV_{syn} = NPV(synergies)$.

The general behavior of the different NPV components for a varying hydrogen price is illustrated in Figure 2.

When the hydrogen price is modified, both the NPV of the PtG facility (NPV_{PtG}) and the value of synergies (NPV_{RES}) are affected, while the NPV of the renewable energy system (NPV_{RES}) remains unchanged, since it is unrelated to the price of hydrogen. Both NPV_{PtG} and NPV_{RES} can evaluate to negative or positive values.

For the break-even detection algorithm, it is convenient to differentiate between scenarios with a positive or negative stand-alone net present value of the RES, since NPV_{RES} remains constant.

Suppose the renewable energy system is profitable on its own ($NPV_{RES} \geq 0$).

1. Combining equation (28) and (29), we obtain for the case of $NPV_{PtG} < 0$:

$$NPV_{PtG} + NPV_{syn} = 0 \qquad (30)$$

2. For the case of $NPV_{PtG} \geq 0$, we obtain:

$$NPV_{syn} = 0 \qquad (31)$$

The second case can only occur for hydrogen prices close to zero, so that it is never favorable to convert power from the renewable energy source to hydrogen and thus no synergies of an integrated system exist.

For the low values of hydrogen prices where this applies, a PtG facility is never profitable considering current and projected future system prices. Therefore, ii) can be neglected and in case of a positive net present value of the RES, a break-even price is defined where equation (30) holds.

Now suppose the renewable energy system on its own shows a negative NPV ($NPV_{RES} < 0$).

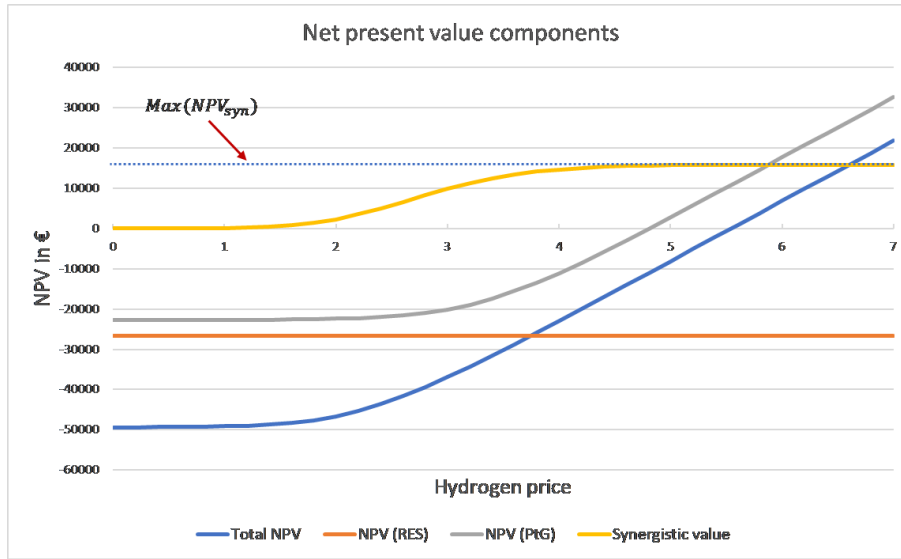[50]Glenk and Reichelstein (2019b), p. 14.
[51]Own figure.

Figure 2: Development of the NPV components at a variation of the hydrogen price.[51]

1. For the case of $NPV_{PtG} < 0$ we obtain as break-even condition:

$$NPV(k_e, k_h) = 0 \qquad (32)$$

2. For the case of $NPV_{PtG} \geq 0$ the condition evaluates to:

$$NPV_{RES} + NPV_{syn} = 0 \qquad (33)$$

In a nutshell, the value of synergies in any case must make up for a negative NPV of at least one of the subsystems. The value of synergies only increases – at an increasing hydrogen price – as long as additional renewable power can be absorbed by the PtG facility. At some point either the entire power generation is absorbed or the electrolyzer runs on full load, which means that no additional power can be converted to hydrogen and thus no further improvement of NPV_{syn} can be realized. Thus, the synergistic value does not scale infinitely, unlike NPV_{PtG}, but possesses an upper bound and a maximum value.[52] If the maximum value lies below the absolute value of a negative NPV_{RES}, which remains constant, then the conditions 32 and 33 can never hold and a break-even price is not defined. In such a case the vertically integrated system cannot be cost-competitive compared to a stand-alone PtG facility.

Consequently, a vertically integrated system can only break even if:

$$NPV_{syn}^{max} \geq |\min\{NPV_{RES}; 0\}| \qquad (34)$$

The algorithm for the detection of the break-even price of a vertically integrated system is based on the preceding insights. First of all, it distinguishes between a profitable and a loss-making renewable energy subsystem.

---
[52]Proof is shown in the Appendix A.3 and A.4.

For the former a break-even price always exists and is present for the hydrogen price, that satisfies the condition defined in equation (30).

In case of a loss-making RES, first of all it needs to be verified whether a break-even price is defined according to equation (34). If that is the case, initially the break-even price is searched using the condition stated in equation (32), which requires the entire NPV to be zero. Any identified break-even price candidate is only valid if $NPV_{PtG} < 0$, which is a prerequisite for the applicability of equation (32). In case of $NPV_{PtG} \geq 0$, search is continued using equation (33) until eventually the break-even price is found for the case $NPV_{RES} < 0$ and $NPV_{PtG} \geq 0$.

Hence, there are three different break-even terms, which must evaluate to zero for detection of a break-even price and vary depending on the NPV constellation of the subsystems (see Table 5).

From now on, I refer to the term to the left of the equal sign, which needs to be zeroed to detect the break-even price of a system, as the break-even term or simply term. All terms follow the same approach, in order to detect the break-even price of hydrogen. First of all, the term value is computed for an arbitrary start value of the hydrogen price. Depending on the sign (+/-) of the resulting value of the break-even term, in the next step another arbitrary value is selected above or below the start value. The actual algorithm can only unfold after these initial steps have been conducted. Now the approximation can be treated as a linear interpolation (or extrapolation) problem with the goal to identify the price at which the break-even term equals zero – i.e., the break-even price. For that purpose, in each iteration the break-even price is guessed based on the last two result values of the relevant break-even term and the tested prices. Based on those two points a linear curve can be constructed for a coordinate system with the vertical axis "Term value" and the horizontal axis "Price of hydrogen". Based on the curve function, which

Table 5: Break-even conditions of a vertically integrated system.

NPVs of subsystems	Break-even condition
$NPV_{RES} \geq 0$	$NPV_{PtG} + NPV_{syn} = 0$
$NPV_{RES} < 0$ and $NPV_{PtG} < 0$	$NPV\,(k_e, k_h) = 0$
$NPV_{RES} < 0$ and $NPV_{PtG} \geq 0$	$NPV_{RES} + NPV_{syn} = 0$

is defined by its slope and vertical intercept, the term value is zeroed in order to calculate the next break-even price candidate. The real result of the relevant term value is computed by applying the price candidate, thus adding another data point. Based on the new data point a more precise break-even price candidate can be computed, again based on the last two points in the described coordinate system. This procedure is continued until the result value of the break-even term lies within a threshold close enough to zero, which is the stopping condition of the algorithm. The break-even price candidate meeting this condition first, is the sufficiently approximated break-even price of a vertically integrated system. The general procedure of the algorithm is illustrated in Figure 3.

3.6. Capacity optimization algorithm

The NPV model framework allows to calculate the net present value of a vertically integrated system, with fixed renewable energy and power-to-gas nameplate capacities and for a proposed hydrogen price. In section 3.5, I described how the framework can be utilized in order to derive the break-even price of hydrogen for a specified system with a fixed subsystem capacity ratio. As a next step, the break-even price can be leveraged to compare differently sized systems regarding their profitability and identify the ideal capacity ratio of the renewable energy system and the power-to-gas facility.

Generally, the search for an optimal capacity ratio could be easily solved by calculating the break-even price for a wide range of capacity combinations. This simple approach, though, would be very time-consuming, since the break-even price of hydrogen – itself already evoking multiple NPV calculations – would need to be calculated for numerous capacity combinations. Therefore, the derivation of the price minimizing capacity ratio of just one scenario could take several hours due to the inefficient approach.

In order to speed up this process and to create an agile tool, a capacity optimization algorithm is proposed, which can detect the price minimizing capacity of the renewable source for a given PtG facility size. This algorithm should also offer an adequate reconstruction of the break-even price curve.

From now on, the PtG system size is taken as a fixed value, while only the RES capacity is altered for optimization of the subsystem ratio.

The proposed algorithm is divided into the following subtasks: First, the whole capacity range is divided into its largest subintervals, in which the break-even price function

is continuous. Next, for each subinterval the range is identified, where the break-even price function is defined. This range is denoted as the domain of the break-even price function. Then each identified domain is checked for extreme points and the general curve properties are sampled. The data points, collected in this process, will be enriched, wherever more detail is necessary in order to reduce the error of the resulting break-even price curve. In a final step, linear interpolation is applied on all collected data points in order to retrieve the final break-even price curve. Based on the derived price curve the profitability of the selected PtG facility can be analyzed for an integration with differently sized renewable energy systems. The price curve also allows to identify the ideal capacity ratio of the subsystems.

3.6.1. Division into subintervals

First of all, it is necessary to split up the whole capacity range of the renewable energy system into subintervals, within which the break-even price curve satisfies the condition of a continuous function, i.e., where it does not exhibit jumps in the function value. Those jumps can arise from the design of the regulatory framework. Considering the presented regulations, jumps can appear at the thresholds of 750 kW and 10 MW.

The jump at 750 kW, primarily, results from the increased EEG levy charged for the consumption of renewable power when the renewable energy source exceeds 750 kW. Secondly, for systems exceeding 750 kW the EEG subsidy value is determined by auctions instead of relying on the legally defined amount and thus, can be based on different values.

Another jump exists for systems including solar PV at 10 MW, since no subsidy is granted to open space solar PV installations above that threshold. Vertically integrated systems based on wind power only are not subject to the jump at 10 MW.

For a vertically integrated system with a hybrid renewable energy source the resulting subintervals would be:

[0 kW; 750 kW] and] 750 kW; 10 MW] and] 10 MW; 100 MW]

If no solar PV system is included the subintervals reduce to:

[0 kW; 750 kW] and] 750 kW; 100 MW]

The threshold of 100 MW was arbitrarily selected as an upper bound of the analyzed capacities.

3.6.2. Detection of the break-even price domain

Next, for each of the resulting subintervals the domain of the break-even price function needs to be determined by excluding any capacity ranges, where a break-even price is not
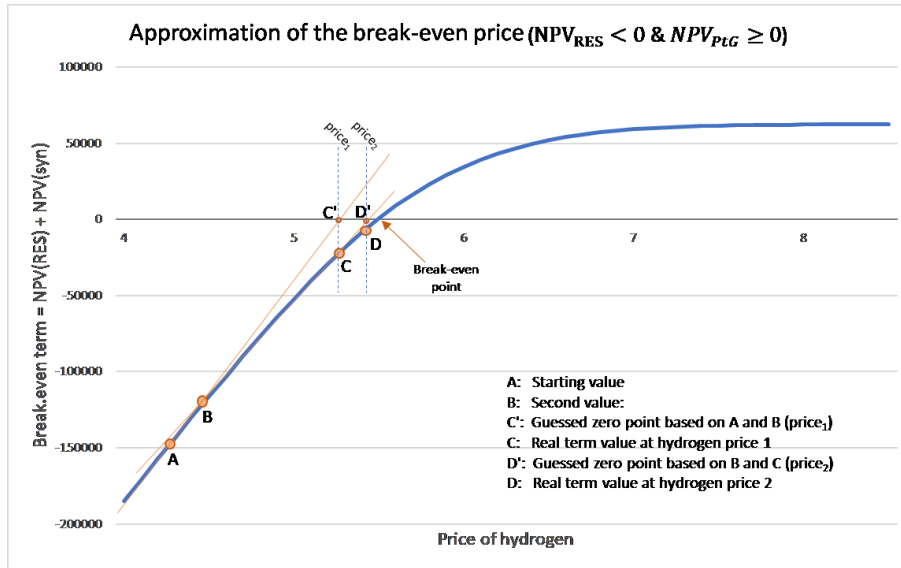
Figure 3: Procedure of the break-even price detection algorithm.

defined according to equation (34). This is the case where the absolute value of a negative NPV of the renewable energy system exceeds the maximum attainable synergistic value.

In order to verify which range satisfies the break-even condition, again, the distinction between stand-alone profitable and loss-making renewable energy systems is made. Remember, a break-even price is defined in the entire subinterval if the renewable energy source exhibits a positive stand-alone NPV, while the synergistic value might not be able to compensate for the loss of a highly negative NPV of the renewable source.

Therefore, I analyzed scenarios with a negative NPV_{RES} regarding their maximum value of synergies $NPV_{syn}^{\max}$[53] for an increasing RES capacity k_e. While NPV_{RES} always scales with factor k_e and thus proportionally to the sizing of the renewable source, $NPV_{syn}^{\max}$ can scale with a factor slightly above k_e for low values of k_e. For larger renewable energy systems, the growth factor of $NPV_{syn}^{\max}$ decreases toward zero.[54]

As an example, the development of the ratio of $NPV_{syn}^{\max}$ to the absolute NPV_{RES} is displayed in Figure 4. It shows to which degree the negative NPV of the renewable energy system can be compensated for by the maximum attainable synergistic value. A break-even price is only defined in the area, where the degree of compensation lies above 100%. The ratio curve does not in every case cross the threshold of 1, it might also lie entirely above or below the threshold.

In order to retrieve the bounds of the defined area, denoted as domain, the intersections of the ratio curve and the threshold line must be detected. They occur where:

$$\frac{NPV_{Syn}}{|NPV_{RES}|} = 1 \qquad (35)$$

The algorithm to identify the boundaries of the domain proceeds in the following way: First the ratio is calculated at the borders of the subinterval. If the ratio already lies above 1 at both borders, any point in between also exhibits a ratio above 100% and therefore the break-even price can be computed in the entire range of the subinterval. If only one of the borders exhibits a ratio above 1, the other border must be searched for by application of a search algorithm described below. If none of the borders shows a ratio above the threshold, there might still be a range for which a break-even price is defined, if the ratio value is increasing at the left border of the subinterval, as shown in Figure 4. If the ratio is decreasing right from the subinterval's left border, the entire subinterval can be excluded from further analysis since a break-even price is not defined for any k_e. This is owed to the construction of the NPV of synergies, whose value development depends on two terms with reverse tendencies. One of them has an increasing and the other a reducing effect on the synergistic value. In some scenarios the increasing effect is stronger for small values of k_e resulting in an over proportional increase of NPV_{syn}^{max} compared to NPV_{RES} and thus the potential to exceed the negative NPV of the renewable source, even when the ratio at the left border lies below 1.[56] If this increasing tendency is not present at the left border, than it will not appear with an increasing k_e either and the respective subinterval can be excluded from the analysis.

The search algorithm for the domain borders follows an easy approach and does not require an optimization, since the performed calculations are not computationally expensive, and the algorithm is only applied rarely throughout program run.

Starting from the subinterval borders, the next compensation ratio is computed at the center of the subinterval. If

---

[53]Compare findings in the Appendix A.3 and A.4.
[54]Explanation is shown in the Appendix A.5.
[55]Own figure.
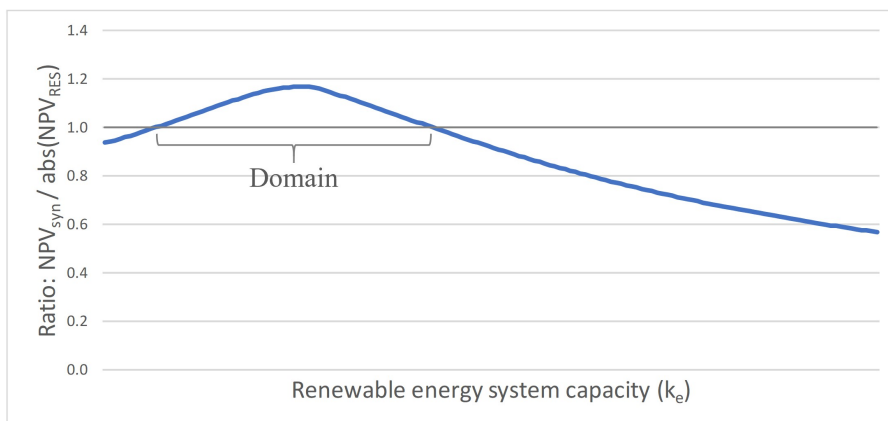
[56]Explanation is shown in the Appendix A.5.

Figure 4: Degree of compensation of a negative stand-alone NPV of the renewable source.[55]

the resulting ratio is above 1, it lies within the domain and search must continue to the left or right of the current center, depending on whether the domain's left or right bound is searched for. Assuming that we are trying to locate the left border of the defined area, then the current center point would be set as the new right border of the search range and the next ratio would be computed for the center between the original left border and the new right border. If the center point exhibits a ratio below 1, meaning it is located outside of the defined area, it needs to be checked if the center point lies to the left or right of the domain. This can be detected by computing the ratio for a second point next to the current center to find out if the ratio is increasing or decreasing at the center point. Now, if the center lies to the left/right of the domain, it would replace the left/right border of the search range, respectively.

This procedure is continued until the searched domain border is approximated sufficiently. In the event no defined range was detected, the algorithm terminates, when the distance between the borders of the search range undercuts a threshold. For subintervals, where the domain lies somewhere in the middle of the range, this procedure is conducted twice, for the left and right domain border. The procedure is illustrated in Figure 5.

3.6.3. Detection of the minimum break-even price and re-construction of the price curve

After the derivation of the area where a break-even price can be computed, the curve needs to be sampled in the domain in order to retrieve the curve characteristics and identify the extreme points. Particularly, the capacity of the renewable source k_e, where the break-even price of hydrogen reaches its minimum value, shall be detected.

Based on sample data, generated for a wide range of scenarios, it became evident, that the derivative of the break-even price curve with respect to k_e has at most one sign change (+/-) indicating the existence of a local extreme

point. In some cases the price curve is even monotone decreasing or increasing and thus the derivative does not exhibit any sign change. The existence of a local minimum seems to be the prevalent characteristic of scenarios, for integrated systems with a loss-making renewable energy subsystem, while in case of a positive NPV_{RES} the price curve generally does not exhibit an extremum within the range. Instead the minimum price would lie at one of the domain borders. Compare Figure 6 and Figure 7 for typical break-even price curves of the respective vertically integrated systems.

The algorithm for detection of the minimum break-even price builds on the described insights. It applies varying approaches depending on the monotonicity of the price curve and therefore distinguishes between curves with a sign change of the derivative and those without.

Similar to the domain search algorithm, the capacity optimization algorithm starts by computing the function gradients, i.e., the values of the derivative, at the left and right border. The gradients are derived based on two neighboring points with a minimal offset in k_e and thus indicate the curve development in their proximate neighborhood. They are computed by subtracting the break-even price at one point from the break-even price at a neighboring point and dividing by the offset, such that:

$$\text{gradient}(k_e) = \frac{\text{price}_{k_e + \text{offset}} - \text{price}_{k_e}}{\text{offset}} \quad (36)$$

Next, the gradient is computed at the center point of the interval. In the following the left or right interval border is replaced by the center in a way that the critical point is located between the new borders. This manner of approximating the critical point corresponds to the procedure of the domain search algorithm. The critical point denotes the extreme point, where the gradient value is zero, in case of a sign change of the derivative. If no sign change can be observed,

---
[57]Own figure.
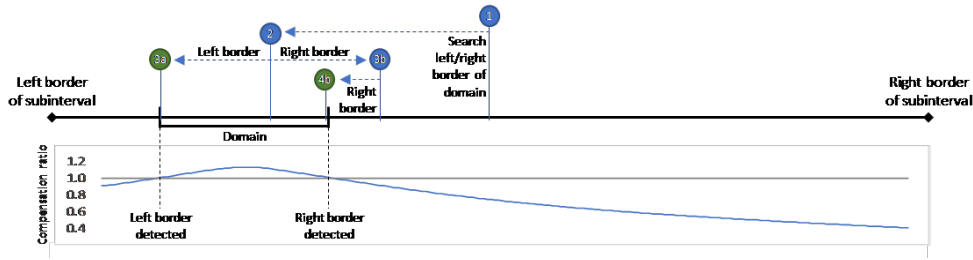
[58]Own figure.
[59]Own figure.

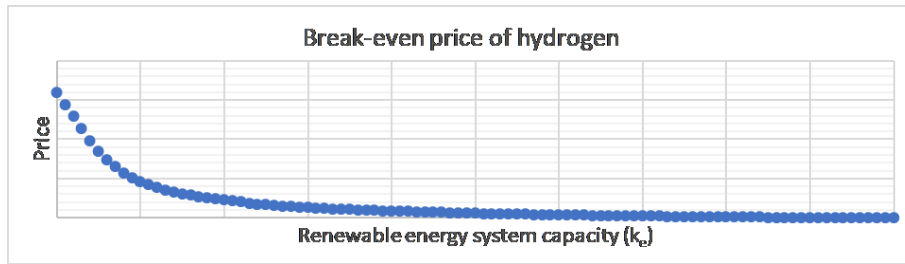**Figure 5:** Algorithm of the detection of the domain borders.[57]

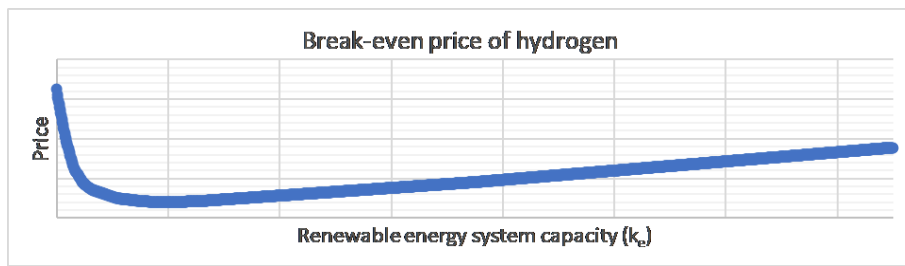**Figure 6:** Typical break-even price curve for a system with a positive NPV_{RES}.[58]

Figure 7: Typical break-even price curve for a system with a negative NPV_{RES}.[59]

the critical point refers to the area exhibiting the highest variation of the derivative, which is explored, in order to enable an adequate reconstruction of the price curve.

The application of the algorithm in this manner can lead to high inefficiencies, since each derivation of a break-even price is computationally expensive and the calculation of a gradient is based on two break-even prices and thus doubles the cost of computation. Several price curves from the produced sample data showed a high variation of the gradients only in a very limited range of the domain, while in the remaining parts the gradient only changes marginally. In such a case, the approach of a plain selection of the center point to approximate the critical point produces many data points with low significance. To avoid that, I followed a method, which enables approaching the critical point more quickly, in a partly expontential way. Therefore, the gradients at the three computed points – left/right border and intermediate point – are compared. When there is only a marginal change between the gradients at one border and the mid point, then it is expected that the critical area is considerably closer to the opposite border. I refer to this constellation as a border-sided extremum. Under those circumstances it is beneficial to move quickly toward the respective border and not simply

select the center point of the interval. Whenever a border-sided extremum appears, in the first step the mid point is selected at that tertile of the current range, which is closest to the respective border, in the next steps at the quartile, quintile and so forth, thus moving quicker in each iteration toward the respective border, where the critical point seems to be located.

Provided that there is no sign change between the left and mid gradient, the condition to initiate the exponential approximation in case of a right border-sided extremum is:

$$\frac{\left|\text{gradient}_{\text{mid}} - \text{gradient}_{\text{left}}\right|}{|\text{ gradient}_{\text{mid}} - \text{gradient}_{\text{right}}|} \leq 0.2 \tag{37}$$

Conversely, a left border-sided extremum is expected if the right and mid gradient have the same sign, and when:

$$\frac{\left|\text{gradient}_{\text{mid}} - \text{gradient}_{\text{right}}\right|}{|\text{ gradient}_{\text{mid}} - \text{gradient}_{\text{left}}|} \leq 0.2 \tag{38}$$

In a nutshell, if the difference between the gradient at the mid point and at one of the borders is considerably lower than the difference between the gradient at the mid point and at the opposite border, the critical point is searched by exponentially approaching the opposite border.

The conditions 37 and 38 are verified after each selection of a new mid point. When none of the two criteria holds, the mid point is selected at the center of the current interval, following the conventional method. Each time the condition for a border-sided extremum is entered from another category, the selection of the intermediate point starts again at the tertile.

The algorithm has two stopping conditions, one of them treating the case of a local extreme point, the other taking effect when the underlying price curve is monotone.

In case of the former, the algorithm stops, when the distance between the current left and right border lies within a precision threshold. The latter can terminate on two criteria, but always presumes that all gradients show the same sign. Either, again, the distance between the current left and right border lies within a precision threshold, or the difference between the gradients at the left border, mid point and right border is only marginal, which implies that the underlying curve is practically linear.

The applied precision threshold is computed in each iteration as a proportion of the mean value of the current range and thus is adjusted to the system scale. This leads to a higher accuracy of the critical point, in absolute terms, for small systems compared to large-scale renewable energy subsystems and increases computation efficiency when deriving the ideal capacity ratio of a large-scale vertically integrated system, where a variation of the detected ideal capacity by a few kW does not have a high impact.

3.6.4. Curve enrichment

After termination of the algorithm, which provides a rough reconstruction of the break-even price curve, the collected data points are checked for major gaps, which might lead to inaccuracies of the result data. For that purpose, each pair of neighboring gradient data is compared, and the maximum potential error between the two points is calculated.

The approach for the derivation of the maximum error is illustrated in Figure 8. At each capacity of RES k_e, for which a gradient has been computed, the tangent to the break-even price curve can be constructed from the gradient, which equals the slope of the tangent, and the data point itself.

Then, for each neighboring data pair (A,B) the intersection I of their tangents is calculated. As a second step, the break-even price is calculated for the capacity k_e, where I is located, based on a direct line between the points A and B. The difference between the break-even prices of the points C and I (red line) represents the maximum potential error of the break-even price between the neighboring data points A and B, when applying linear interpolation. It should be noted that the calculated maximum error generally overestimates the real error, while it always constitutes an upper bound for the real error. Areas with a high break-even price, compared to the computed minimum price, are disregarded for the curve enrichment, since they have a low significance and their computation can be avoided.

3.6.5. Linear interpolation

As the last step, all collected data points are interpolated linearly to produce a curve plot, which allows to visually analyze the profitability of various system combinations and identify the price-minimizing capacity ratio of the renewable energy system and the power-to-gas facility.

4. Scenario evaluation

I now use the adapted model framework and apply the developed optimization algorithm on a vertically integrated system assuming a range of scenarios with different policy designs.

4.1. Model input selection

The calculations are conducted for an investment in a vertically integrated energy system in Germany with an expected project lifetime of 30 years and a guaranteed EEG subsidy during the first 20 years of operation. The projected investment starts in 2020, but I use electricity price data and power generation data for simulation of the renewable energy sources from the year 2015. I assume costs of capital of 4%[60] and an effective corporate income tax rate of 30%.[61] Concerning the depreciation of the fixed assets, for simplification I apply the depreciation schedule valid for wind turbines to the entire system. Thus, the investment cost is depreciated linearly over a period of 16 years.[62]

My calculations are based on the use of a polymer electrolyte membrane (PEM) electrolyzer, since PEM electrolyzers can operate dynamically at a wide range of loads and react quickly due to a short system response time within milliseconds[63], while technically more mature Alkaline electrolyzers are only capable of operating at low partial load ranges and can lack in hydrogen purity depending on the load. They are also inferior to PEM electrolyzers in terms of dynamic operations.[64]

For the system prices of the power-to-gas facility I rely on the price and cost regression of electrolyzers derived by Glenk and Reichelstein[65]. I only adjust the cost of the electrolyzer, which is the main component of the power-to-gas facility, since it has been and will be subject to significant price drops in the recent and coming years. This price adjustment is based on an average annual price drop rate of 4.77%, which was found by the authors through regression of the price development of PEM electrolyzers between the years 2003 and 2016.

Additionally, the fixed operating costs, which are generally expressed as a percentage of the system price, must also

---

[60]Data sources: Noothout et al. (2016), p.40 and ZSW and Bosch & Partner GmbH (2019), p. 43.
[61]Data source: OECD (2018).
[62]Bundesfinanzhof (2011). Ruling of the Federal Fiscal Court of Germany.
[63]Cf. Schmidt et al. (2017), p. 30472.
[64]Cf. Carmo et al. (2013), p. 4903-4904.
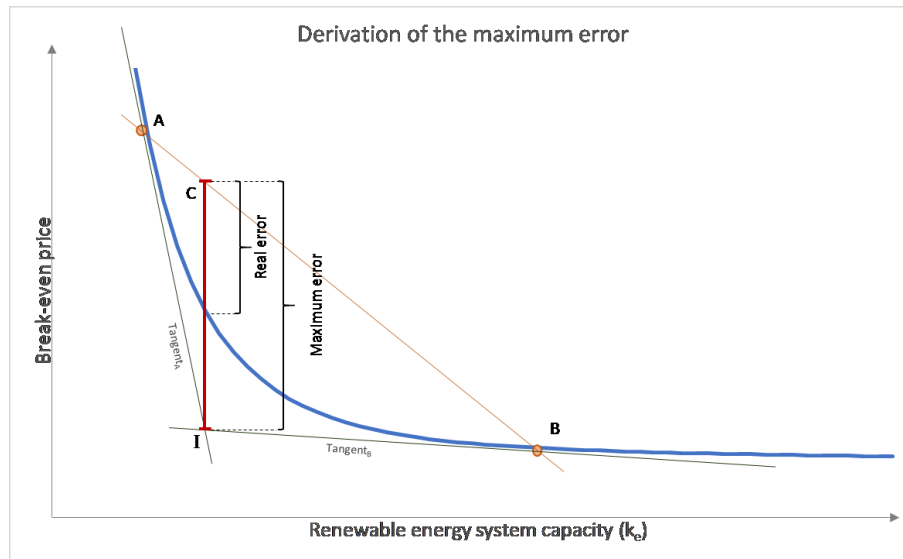[65]Data source: Glenk and Reichelstein (2019b), p. 217-218.

Figure 8: Approach to calculate the maximum price error between two neighboring points.

be adjusted. For my calculation I assume operational expenditure (OpEx) amounting to 2%[66] of the capital expenditure (CapEx), which coincides with the cost ratio present in the work of Glenk and Reichelstein.

Thus, the system price of the electrolyzer reduces from 1863 € /kW[67] to 1459 € /kW in 2020. For the remaining PtG cost factors, such as foundation, access, power connection, compression, which are valued at 424 € /kW, I do not assume price reductions. The system price of the entire PtG facility totals 1883 € /kW and the respective fixed operating costs amount to 37.66 € /kW.

Regarding the wind turbine, I assume a system price of 1180 € /kW and fixed operating costs of 38 € /kW.[68]

For the current system price and fixed operating costs of solar PV, I rely on a report prepared for the Federal Ministry for Economic Affairs and Energy of Germany, which found the current cost of solar PV systems – including inverter, mounting system, grid access, solar PV modules, planning and installation – at 770 € /kWp for open space installations with a nameplate capacity of at least 750 kW. The fixed operating costs are estimated 15 € /kW with a yearly growth factor of 1.5%.[69] Based on formulas of mathematical finance the growing operating costs are converted to an annuity over the system lifetime using the cost of capital as interest rate. The calculated yearly fixed operating costs result in a constant amount of 18 € /kW.

It should be noted that these are only average prices, since the CapEx and OpEx generally reduce with higher system scales.[70]

The yearly degradation rate of solar PV systems lies in the range of 0.4 – 0.7%.[71] Wind turbines are also impacted by degradation over their lifetime and an annual rate of 0.6% is assumed, which can be lowered by increased operations and maintenance procedures, such as blade erosion repair.[72]

The system lifetime and degradation of PEM electrolyzers has not received enough attention in research and is still changing significantly due to technological enhancements. Today many researchers expect a lifetime of up to 60,000 hours.[73] Those lifetime measures are generally not put into perspective by indicating the remaining electrolyzer capacity after a system reaches its lifetime. Though, Proton OnSite, a producer of electrolyzers, has "stated that it has achieved 60,000 hours lifetime in its commercial stacks without any detected voltage decay."[74] This implies that electrolyzers might already reach operating hours well above the often cited 60,000 hours. Degradation is at least to some degree reversible, e.g., by treatment with dilute sulfur acid.[75] Therefore, I account for potential degradation of the PtG system by applying additional maintenance fees at the cost of 1% of the system price, to recover electrolyzer efficiency. This requires an adjustment of OpEx of the PtG facility to the amount of 56.50 € /kW. Apart from that correction, I apply a unique degradation rate to the entire vertically integrated system amounting to 0.6%, which is consistent with the wind turbine and solar PV system components.

The summarized cost and degradation measures of the vertically integrated system are displayed in Table 6.

For simulation of the power generation by the wind turbine and solar PV system, I rely on the same wind turbine

---

[66]Data source: Bertuccioli et al. (2014), p. 14.

[67]System price of PEM electrolyzers in 2015.

[68]Data source: Glenk and Reichelstein (2019a), p. 32.

[69]Data source: ZSW and Bosch & Partner GmbH (2019), p. 41-44.

[70]Cf. Bertuccioli et al. (2014), p. 14.

[71]Data source: ZSW and Bosch & Partner GmbH (2019), p.44.

[72]Cf. Rubert et al. (2017), p. 12-13.

[73]Cf. Schmidt et al. (2017), p. 30472.

[74]Price (2017), p.50.

[75]Cf. Sun et al. (2014), Abstract.

Table 6: Technological system specifications.

	Wind turbine	Solar PV	Power-to-gas
System price	1180.00 € /kW	770.00 € /kW	1883.00 € /kW
Fixed operating costs	38.00 € /kW	18.00 € /kW	56.50 € /kW
Degradation factor		0.994	

capacity factors as Glenk and Reichelstein.[76] I obtained solar PV production data for Germany in 2015 from a tool provided by Pfenninger and Staffell. The average capacity factors are 30.31% for the wind turbine data and 12.85% for the solar PV data.

The electricity prices are based on the day-ahead auction price at the EPEX spot market from 2015. The mean electricity price of the respective year amounted to 3.16 € ct/kWh. In addition, a markup on grid-supplied electricity is necessary to account for the cost of electricity trading, which is estimated at 1.00 € ct/kWh.[77] Furthermore, a range of statutory fees is charged on grid-supplied and own electricity consumption, as described in section 2.3.

The statutory fee rates effective in 2020 are shown in Table 7.

There are no fixed fees charged on the use of renewable power, the fixed fees for grid-supplied power are composed of 1) and 2). The variable fees are determined in an iterative approach, which I described in section 3.3. The variable fees of renewable energy are only based on the EEG levy, while variable fees imposed on grid power are composed of items 3) to 6).

The conversion value of hydrogen depends on the conversion rate of hydrogen and the variable cost of hydrogen production. Additionally, in the event of onsite utilization, the avoided transportation cost is attributed to the conversion value.

The conversion rate is defined as the hydrogen output in kg per kWh of electricity utilized. For the calculation of the conversion rate I assume an electricity input of 48 kWh/kg, which corresponds to an electrolyzer efficiency of 68%[79], and an electricity intensity of the hydrogen compression amounting to 2-4 kWh/kg.[80] This amount considers the compression from the output pressure of the electrolyzer, assumed at ambient pressure, to 350 bar, a pressure level suitable for hydrogen storage or sale.

When combining the electricity consumption of conversion and compression, the conversion rate amounts to:

$$\eta_h = \frac{1\,kg}{52kWh} \approx 0.01923 \frac{kg}{kWh}. \tag{39}$$

For the variable operating costs of the PtG facility, I rely on

the value assigned by Glenk and Reichelstein and therefore define $w_h = 0.10€\,/kg$.[81]

The cost of transportation depend very much on the specific use case. For the scenario of onsite hydrogen utilization, I consider an avoided distribution cost accounting for transportation by compressed gas trailer trucks for a distance of 300 km amounting to 1.05 € /kg.[82]

Furthermore, the conversion value could be increased by the contribution margin resulting from the sale or use of oxygen, if the oxygen by-product is captured. In that case the price of oxygen and any additional variable and fixed cost components need to be considered.

There is a market for industrial and medical oxygen. While medical oxygen might yield a higher contribution margin, it could also be subject to stricter regulation. Since only few and almost no recent research articles seem to exist on the economic potential of oxygen, I rely on a price range between 24 and 40 $/t, which is the cost for industrial on-site oxygen production based on vacuum pressure swing adsorption.[83] I assume the average value, which corresponds to a price of 0.03 € /kg[84]. Higher prices might be achievable, since the produced oxygen possesses a high purity and is therefore suitable for the sale as medical oxygen.

Due to a lack of details about the subsequent processing of oxygen and which additional investments might be required, I account for no further costs. I consider this approach sufficient for an initial evaluation of the economic potential of the oxygen by-product. However, it remains unclear if additional compression capacity needs be acquired for the compression of oxygen, or if the existing compressor can be shared between hydrogen and oxygen.

The EEG subsidy premium is calculated as the difference between the guaranteed subsidy value[85] and the observed market prices of electricity, which are published on a monthly basis.[86] Renewable energy systems with a size up to 750 kW can rely on legally defined subsidy values, while systems above that threshold need to participate in auctions to get a subsidy granted. Thus, for systems exceeding 750 kW, I assume the weighted average value of the accepted auction bids[87] in 2019. As a result, the corresponding subsidy val-

---

[76]Data source: Pfenninger and Staffell (2016). Tool accessible at https://www.renewables.ninja/.

[77]Data source: Glenk and Reichelstein (2019a), p. 33.

[78]Refer for legal sources to section 2.3.

[79]Data source: Bertuccioli et al. (2014), p. 62. Electricity input and LHV efficiency estimates for 2020.

[80]Data source: Gardiner (2009), p. 3.

[81]Data source: Glenk and Reichelstein (2019a), p. 32.

[82]Data source: International Energy Agency (2019), p. 80: 1.20 $/kg; Conversion with the avg. exchange rate of Q1-Q3 2019: 1.1238 $/€ (European Central Bank (2019), p. 73).

[83]Data source: Dorris et al. (2016), p. 19.

[84]Conversion with the avg. exchange rate of Q1-Q3 2019: 1.1238 $/€ (European Central Bank (2019), p. 73).

[85]German term: "Anzulegender Wert".

[86]Cf. Netztransparenz (2019).

[87]Data source: Bundesnetzagentur (2019b).

Table 7: Current rates of statutory fees (in € ct/kWh).[78]

	Base rate	Reduced rate (above 1 GWh)	Own consumption	Grid consumption
1) Concession charge	0.110	-	-	X
2) Levy for interruptible loads	0.007	-	-	X
3) EEG levy (base rate)	6.756	0.100	40% / 100%	100%
4) CHP levy	0.226	0.030	-	X
5) Offshore grid levy	0.416	0.030	-	X
6) § 19 StromNEV levy	0.358	0.025	-	X
7) Transmission charge	Exempt			
8) Electricity tax	Exempt			

ues for wind turbines and solar PV systems amount to 6.17 € ct/kWh and 5.84 € ct/kWh, respectively. The subsidy values for small systems are determined based on the results of past auctions and are published on the website of the Federal Network Agency. Systems installed in January 2020 are supported with a subsidy amounting to 6.04 € ct/kWh for wind turbines and 7.20 € ct/kWh for open space solar PV installations.[88]

The subsidy value granted to wind turbines is adjusted to account for performance differences due to the turbine location. According to § 36h EEG, the proportion of the real electricity output to a reference output, determined by independent experts for each wind turbine type, needs to be calculated. Based on the computed proportion, the subsidy value from the auction bid is increased or decreased in accordance with a factor defined in the mentioned EEG section.

In this case, the reference output, which refers to a wind turbine of type Enercon E101 with a hub height of 149 m and a nameplate capacity of 3050 kW, is specified for a five years period at 49,221,048 kWh.[89]

The real output of the wind turbine in its first five years, based on the average capacity factor, the degradation rate and m as the number of hours per year, is determined by:

$$output_{Y1-Y5} = \sum_{i=1}^{5} 3050kW \cdot CF \cdot m \cdot x^{1-i}. \qquad (40)$$

The resulting electricity output adds up to 40,982,915 kWh, which corresponds to a proportion of 83.3% and leads to an adjustment with the factor 1.1303.

Consequently, the subsidy values granted to wind turbine operators increase to 6.97 € ct/kWh for turbines exceeding 750 kW and 6.83 € ct/kWh for smaller turbines. Eventually, the monthly subsidy premium is derived by deducting the gained market price from the respective subsidy value. The monthly premium is expressed on an hourly basis by the variable premium(t).

Finally, the input variables for the scenario accounting for emission prices need to be defined. The EU ETS regulates the allocation and trading of emission allowances. In general, both electricity generators and industrial installations of specific branches, including the production of hydrogen by reforming or partial oxidation, are subject to EU ETS regulations and therefore need to acquire emission allowances.

While power suppliers do not receive free allowances since 2013, allowances are allocated to industrial installations free of charge for a gradually decreasing proportion of their total emissions.[90] Furthermore, some activities receive 100% free allowances, if they belong to a sector with exposure to carbon leakage. Since such a waiver was effective for the production of fossil hydrogen in 2015[91], the emission price of hydrogen in the base year is zero. On the contrary, I assume that the cost of emission allowances, which amounted to 7.60 € /tCO$_2$eq in 2015[92], was entirely added to electricity prices.

For the emission scenario, a consideration of the entire emissions without free allocation of allowances is assumed during the projected system lifetime. I assume a linearly growing emission price between 20 € /tCO$_2$eq in 2020 and 100 € /tCO$_2$eq in 2049. As a result, the entire target emission price is applied to emissions from fossil hydrogen production. In the case of electricity, the carbon price of 2015 needs to be deducted from the target price. The adjusted target emission price is then applied on the emissions per kWh of electricity.

The emission factor of fossil hydrogen lies at 7330 gCO$_2$/kgH$_2$ in the case of steam methane reforming, which is the most frequent source of hydrogen production.[93] The hourly emission factor of electricity is determined by computing the average emission factor of all electricity generation sources weighted by their generation output for each point in time. Thus, an hourly emission factor of the German electricity mix is derived.

The emission factors displayed in Table 8 are considered for the different sources of electricity.

All factors are multiplied with the hourly electricity generation by their source. Their sum is then divided by the total electricity production to derive the hourly emission factor of the electricity mix. I retrieved the electricity generation

---

[88]Data source: Bundesnetzagentur (2019a).
[89]Data source: Wind-FGW.

[90]Cf. European Commission.
[91]Cf. Statute 2014/746/EU.
[92]Data source: DEHSt (Umweltbundesamt) (2016), p. 5.
[93]Data source: Kothari et al. (2008), p. 554-558.
[94]Data source: Agora Energiewende (2019), p. 13.

Table 8: Emission factors of fossil electricity sources (in gCO_2eq/kWh).[94]

	Brown coal	Hard coal	Natural gas	Others (incl. fossil waste, oil)
Emission factor	1100	850	370	1590

data from a platform provided by the European Network of Transmission System Operators for Electricity (ENTSO-E) on a quarter-hourly basis.[95]

4.2. Evaluation of results

For the calculations I considered a range of differently sized vertically integrated systems based on PtG facility sizes of 100 kW, 500 kW, 1 MW, 5 MW, 20 MW and 100 MW. These facilities are capable of producing between 46.15 kg/day and 46.15 t/day at full load. I computed the break-even price curves for various cases accounting for the current legislation and potential future policies and assuming vertically integrated systems, where the renewable energy system can be composed of wind turbines and/or a solar PV system. When analyzing hybrid systems, I base my calculations on equal nameplate capacities of wind and solar PV. The optimization of the renewable energy system composition would go beyond the scope.

Eventually, I compare the derived break-even prices with the prices of industrial hydrogen. The market for industrial hydrogen divides into three segments: supply on a large-scale, where prices between 1.5 - 2.0 € /kg are observed, a market for medium-scale supply exhibiting prices in the range 3.0 – 4.0 € /kg and the small-scale market with prices above 4.0 € /kg.[96]

Initially, it will be interesting to analyze the economic potential of electrolytic hydrogen under the effective legislation. Figure 9 shows the break-even price curves of a vertically integrated system for the defined scales of the PtG plant. The horizontal axis shows the combined installed capacity of the wind turbines and solar PV system. The figure reveals, what currently is the challenge for profitability of PtG on a small-scale. The power consumption of a PtG facility with an installed capacity of 100 kW does not exceed the volume of 1 GWh, which is not exempt from a reduction of statutory fees, and thus cannot benefit from the reliefs granted to large industrial electricity consumers under the EEG framework. This results in excessive break-even prices of electrolytic hydrogen of up to 7.28 € /kg above the 750 kW threshold. The minimum price of 6.13 € /kg can be observed for an installation of each wind turbines and solar PV sized 750 kW, which in the case of wind turbines would be a fairly small system. This result is a consequence of the reduction of the EEG levy to 40%, charged on own power consumption from small renewable energy systems. When considering investments in PtG capacity on a larger scale, the reduction of statutory fees, applying to electricity consumption above 1 GWh, can unfold

its effect of reducing the cost of power supply. This leads to a substantial decrease of the break-even price level and facilitates that minimum prices of 4.42 € /kg, 3.93 € /kg or 3.55 € /kg can be accomplished by vertically integrated systems with a PtG component scaled 500 kW, 1 MW or 5-100 MW, respectively. However, the effect from scaling almost disappears for plants above 5 MW.

At 20 MW, where the hybrid system is composed of wind turbines and a solar PV system each sized 10 MW, the price curves "jump", which is caused by the non-eligibility to the EEG subsidy of solar PV systems above 10 MW. In the case of a feed-in requirement, the opportunity cost of the lost feed-in premium disappears for solar power above 10 MW. This leads to a lower barrier for absorption of own power by the PtG facility resulting in a higher load and decreased break-even price. At the same time, the NPV will decrease due to the lost subsidy. This shows a limitation of the sole consideration of the break-even price when comparing the prices of different subintervals[98] of the RES capacity range. Hence, investors need to look at the break-even price coupled with the NPV, when comparing different investment options.

Based on Figure 9 it also becomes evident that large-scale PtG facilities are favored by the design of the regulatory framework. In the next step, I will compare the effects of various conceivable policy changes on the profitability of PtG focusing primarily on small-scale installations. The proposed policy changes comprise, firstly, the replacement of the feed-in premium by a production premium, which is paid irrespective of the feed-in of the produced renewable power. This instrument would remove the barrier of a lost subsidy, when converting electricity to hydrogen. Secondly, an extension of the reduced EEG levy to renewable energy systems of any capacity is considered, which is currently only granted to operators of renewable systems of up to 750 kW. Lastly, a complete waiver of the statutory fees is proposed for renewable electricity consumption and for both renewable and grid electricity consumption. The price curves resulting from implementation of these policies are illustrated in Figure 10 for PtG facilities with nameplate capacities of 100 kW and 1 MW, respectively. All cases are implemented based on the current legislation, while only changing the parameters representing the described policy change. This ensures the comparability of the proposed policy measures regarding their effectiveness.

The figure shows that especially small power-to-gas

---

[95]Data source: ENTSO-E.
[96]Cf. Glenk and Reichelstein (2019b), p. 218.

[97]Own figure. Case: Feed-in premium | all statutory fees included | wind & solar combined.
[98]Refers to the subintervals of the RES capacity range as part of the optimization algorithm in section 3.6.
[99]Own figure. Case: PtG = 100 kW (left) / 1 MW (right) | wind & solar combined.
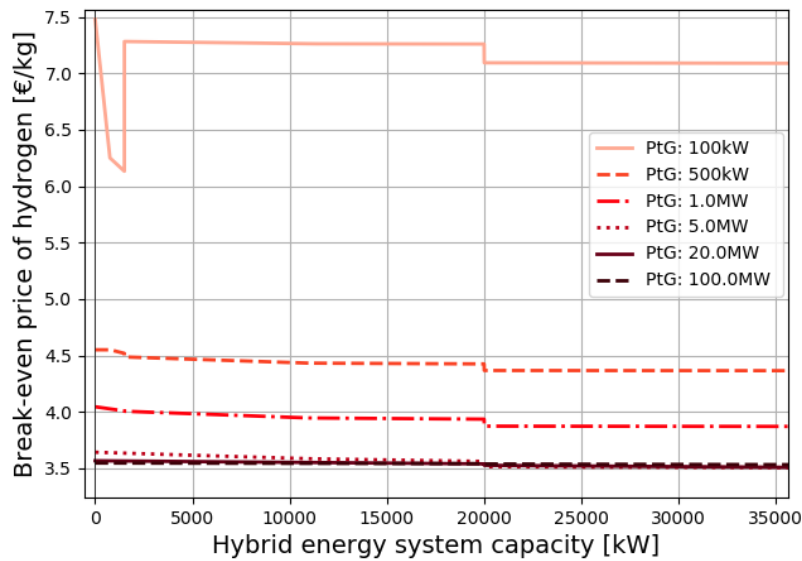
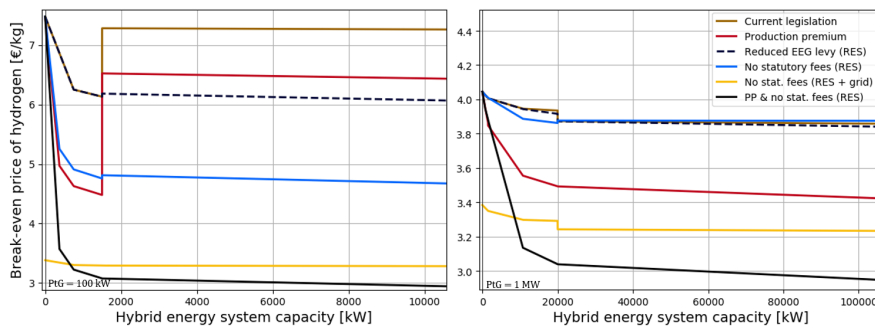Figure 9: Break-even price of hydrogen under the current regulatory framework.[97]

Figure 10: Effects of policy changes on the break-even price of hydrogen.[99]

plants depend on changes of the regulatory framework in order to be able to compete with industrial hydrogen, while large PtG facilities can already accomplish industrial hydrogen prices for a medium-scale supply. When waiving the feed-in requirement (red line), a break-even price of 4.50 € /kg can be reached for the optimal capacity ratio and electrolytic hydrogen could participate in the small-scale hydrogen market. Compared to the other policy measures, none can accomplish a price as low in the lower capacity range up to 1500 kW. However, above a combined capacity of 1500 kW, the waiver of the feed-in requirement loses much of its potential due to the increased EEG levy imposed on renewable power. In that range a waiver of the statutory fees on renewable electricity (blue line), i.e., the waiver of the EEG levy, performs much better and is able to lower the price to 4.80 € /kg for hybrid systems above the 1500 kW-threshold and down to 4.65 € /kg for a 20 MW renewable energy system. The extension of the reduced EEG levy to renewable energy systems above 750 kW (dashed line) is not able to shift the hydrogen price below 6.00 € /kg and thus will

not be sufficient to make electrolytic hydrogen economical. Only a waiver of all statutory fees, imposed on renewable and grid power, or the introduction of a production premium combined with a waiver of the EEG levy payable on renewable power can shift the break-even price below 3.30 € /kg and in the case of the latter even under 3.00 € /kg and thus can help the profitability of small-scale PtG facilities substantially. It should not remain unmentioned that stand-alone PtG plants would also benefit extremely from a waiver of all statutory fees and could operate at a break-even price of 3.40 € /kg and thus at a price slightly below the price of 3.65 € /kg at which stand-alone PtG facilities with a capacity of at least 5 MW can already produce under the current framework. At the same time, the policy suggesting the waiver of the feed-in requirement and EEG levy for renewable power would not be to the benefit of stand-alone PtG facilities, thus setting an incentive for coupling of the PtG plant with renewable energy sources and promoting the production of renewable hydrogen.

Under limited circumstances, electricity consumption can

already be exempt from the EEG levy under the current regulatory framework. The EEG levy does not apply when an electricity consumer is not connected to the grid and satisfy their power demand exclusively from an adjacent renewable energy source. The renewable source can still be connected to the grid and sell power on the wholesale market but is not entitled to the EEG subsidy.[100] This policy could benefit a vertically integrated system, where the PtG facility only absorbs power from the renewable source, while a subsidy is not received on the feed-in of any excess power.

The resulting break-even price curves are depicted in Figure 11. It becomes evident that all break-even price curves have a pronounced minimum point which indicates that off-grid operating power-to-gas facilities always have an ideal renewable energy system capacity. It is also visible that differently sized systems always exhibit the same minimum price, which lies for the hybrid, wind-only and PV-only scenario at 4.92 € /kg, 5.03 € /kg and 7.58 € /kg, respectively. The fact that the combined system shows the lowest price, proofs that a combination of wind turbines and solar PV in a hybrid system can have synergistic value. Although the price improvement of a hybrid system is only minor, compared to the wind-only scenario, an optimization of the ratio of wind turbine to solar PV capacity might lead to increased synergies. It should be noted, though, that good conditions for wind turbines and solar PV systems might not easily be found at the same geographical locations. The three scenario variations also show equal capacity ratios. In case of the hybrid system, the ratio of the PtG facility size to the size of the renewable energy system amounts to 26% and for the wind and PV scenario the ratio equals 35% and 30%, respectively (see Figure 11).

In particular, the consideration of this scenario could be relevant for solar PV open space installations exceeding the threshold of 10 MW and thus losing the eligibility for a subsidy in any case. The installation of such a system with a nameplate capacity of 175 MW in Germany was recently announced and start of construction is scheduled for early 2020.[102] However, the computed value for renewable hydrogen production from solar PV in this work is not applicable in that context, since such a project can enormously benefit from economies of scale and a much lower system price of the solar PV system must be assumed. The high intermittency of solar PV stations could be detrimental to such a project, though.

Finally, I ran analyses of the scenarios considering a rising emission price and the commercialization of the oxygen by-product, both based on the current legislation. The results, shown in Figure 12, indicate that an emission price, especially for large-scale PtG facilities, generates increased costs for electricity supply, since large PtG systems absorb high amounts of grid power, when not coupled with adequately sized renewable energy systems. The costs of emissions related to fossil hydrogen production, which raise the conversion value of hydrogen, appear to have a minor effect compared to the additional costs of emission allowances required for consumed grid power. Thus, for all PtG scales a rising emission price results in higher break-even prices of hydrogen almost along the entire vertical axis. However, emission prices can also serve as an incentive to couple the PtG facility with a renewable energy source, since the penalized grid power could be replaced by cheaper renewable power. It should be noted, that the analysis of the emission price is based on a constant electricity mix over the entire lifetime of the investment and should therefore be interpreted with caution.

The scenarios considering the commercialization of the oxygen by-product possess break-even price curves, which run almost perfectly parallel to the break-even price curves corresponding to the scenario under current legislation. This is not only valid for the scenario displayed in the figure, but for all PtG scales and hybrid, wind-only and PV-only renewable energy system. The price advantage of the oxygen scenario compared to the break-even price under the current legislation constantly lies around 0.24 € /kg. It is questionable, though, if and to what extent additional investment and operational costs could eat up this small margin. Also, the market price of oxygen must be explored in more detail, to find out if oxygen can really improve the economics of power-to-gas technology.

4.3. Sensitivity analysis

In order to evaluate the risks of an investment in a vertically integrated energy system accordingly and account for existing uncertainties, the sensitivity of the break-even price of hydrogen concerning various model input parameters needs to be analyzed. Therefore, I computed the sensitivity of the derived break-even price for a hybrid system composition of solar PV and wind turbines each 750 kW and a PtG facility sized 100 kW and for a larger system with both renewable components having a size of 10 MW and a 1 MW sized electrolyzer. The analysis results for the small-scale system are presented in Figure 13 for a range of discussed scenarios. The analysis of the large-scale scenario is available in the Appendix A.7. Both results show that the respective sensitivities exhibit almost equal characteristics among the different cases.

The break-even price exhibits the most pronounced sensitivity with respect to a variation of the conversion rate. The conversion rate depends on the efficiency of the installed electrolyzer and constitutes one of its key performance indicators. However, particularly, PEM electrolyzers are still being researched extensively and more long-term trials are required to validate their continuity in performance. Hence, the high sensitivity of the conversion rate currently expresses a high risk for investments in PtG technology. A decrease of

---

[100]Cf. § 61a Nr. 3 EEG.
[101]Own figure. Case: No subsidy | no statutory fees on RES power.
[102]pv magazine pv magazine Deutschland (2019).

[103]Own figure. Case: PtG = 20.0 MW | wind turbine only.
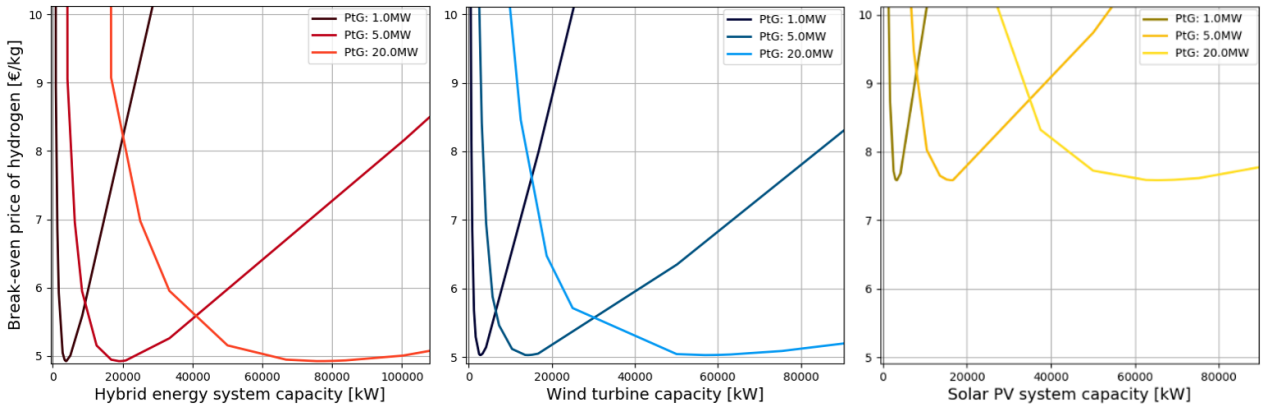[104]Own figure.

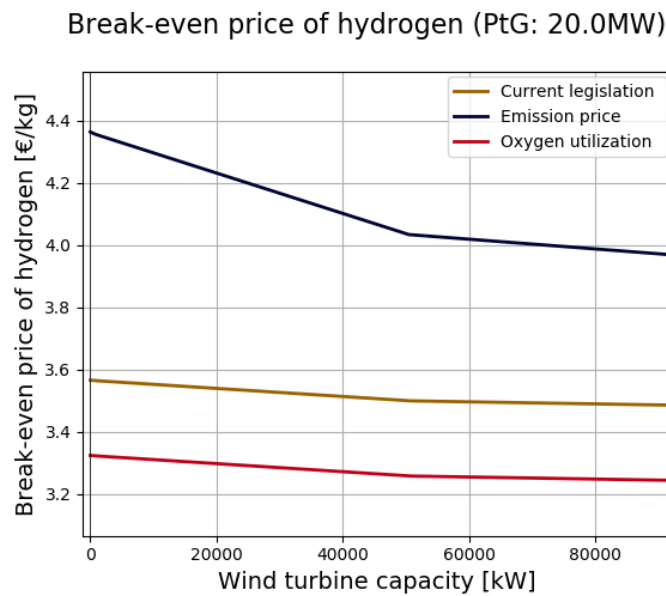**Figure 11:** Break-even price of hydrogen of hybrid, wind-only and PV-only systems.[101]

Figure 12: Effects of the emission price and oxygen commercialization scenarios.[103]
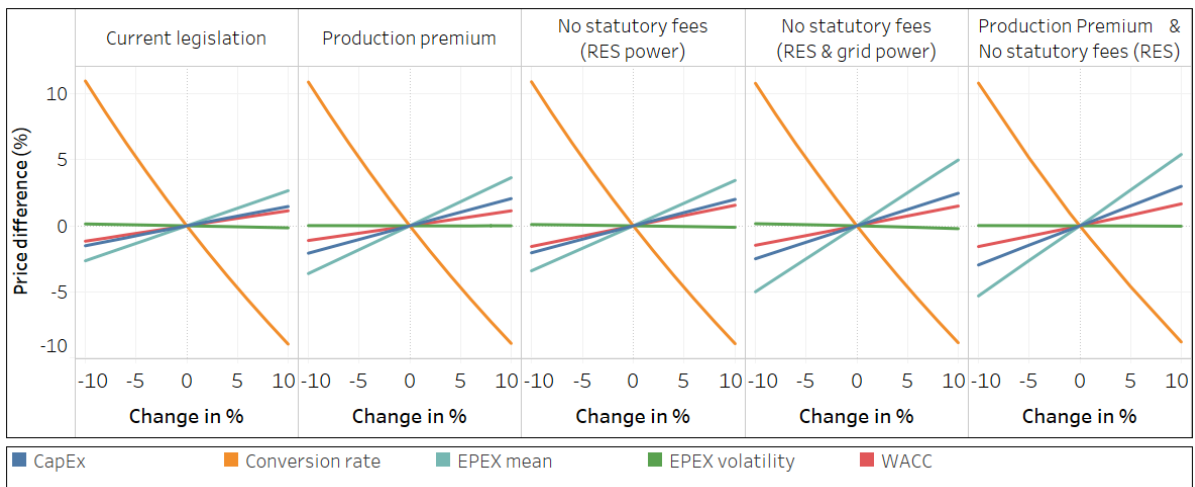

Figure 13: Sensitivity of the break-even price of hydrogen.[104]

the conversion rate by 10% would result in a significant increase of the break-even price slightly above 10%. The other analyzed input parameters have lower impacts on the break-even price.

I defined the volatility of the electricity price by its standard deviation. The volatility analysis is implemented such that only the volatility is changed, while the average electricity price remains constant.

To facilitate the interpretation, I calculated the mean electricity prices and price volatilities of the years of 2015 through 2018 as displayed in Table 9.

While the effect of price volatility seems to be negligible, a change of the electricity price level has a higher impact. This is, particularly, the case for scenarios, which are based on a higher consumption of grid electricity, such as the two scenarios on the right of Figure 13.

In 2018 the electricity price was extraordinarily high compared to the other years and 40% higher than in 2015. When running the analysis based on electricity prices of 2018, results would have shown a significant 10% and 16.6% increase of the break-even price for small and large system under current legislation, respectively. This shows, that the development of the electricity price can have a huge impact on the economics of electrolytic hydrogen.

In case of small-scale systems, an increased electricity price adversely affects the price of hydrogen, since the opportunity costs for the conversion of renewable power rise, due to higher yields of a feed-in resulting from higher electricity prices. Thus, the electricity price also has a significant impact on systems with a high internal power absorption, but still to a lower extent than on large systems that rely to a higher degree on power supply from the public grid.

Besides, the sensitivity of the WACC and CapEx are displayed as additional information but will not be discussed. The CapEx does not change during lifetime and thus only presents a risk prior to the installation of a vertically integrated system.

4.4. Implications for policymakers and investors

To conclude this section, I will sum up the implications of the analysis results and draw conclusions relevant for policymakers and investors.

The break-even price of hydrogen computed for large vertically integrated energy systems with a 5 MW sized power-to-gas facility gets as low as 3.55 € /kg and thus lies above the price of large-scale hydrogen (2.0€ /kg), but can already compete on the market for medium-scale hydrogen supply, where prices lie in the range 3.0 – 4.0 € /kg.[106] This result confirms the findings of Glenk and Reichelstein.

The relatively low prices of large-scale PtG can only be accomplished due to the reliefs from statutory burdens granted to large industrial electricity consumers with a consumption

above the threshold of 1 GWh. Based on those reliefs the regulatory framework incentivizes large-scale PtG facilities. However, large systems involve substantial risks, due to high capital expenditure, technological uncertainties and a lacking market for electrolytic hydrogen. In addition, large-scale PtG facilities base their low break-even price to a certain extent on the absorption of grid electricity and have a higher dependence on the development of the wholesale electricity price.

Furthermore, large power-to-gas systems do not make much sense when producing hydrogen from grid power under the current German electricity mix. The resulting hydrogen product is not only more expensive than hydrogen produced by steam methane reforming but would also cause carbon emissions more than three times as high as the emissions resulting from the conventional technology for hydrogen production.[107]

Therefore, I consider it necessary to help accelerate the economy for renewable hydrogen by improving the economics of small-scale PtG facilities[108] coupled with renewable energy sources. The analysis in section 4.2 has shown that the waiver of the feed-in requirement, in favor of a production premium, coupled with a waiver of the EEG levy payable on self-generated renewable electricity would be the most effective instrument to improve the economics of small-scale vertically integrated systems. At the same time this measure would provide an incentive to couple power-to-gas technology with renewable power sources. When additionally adding an exemption from the statutory fees charged on grid electricity during times of negative or low electricity prices, such systems could also help to balance the electricity market and the power grid.

The introduction of a production premium would substantially lower the opportunity cost of converting electricity from an adjacent renewable energy system to hydrogen, since the subsidy premium would not need to be compensated for by the profits from hydrogen sale. As a consequence, electrolysis based on power from a renewable energy source would be incentivized and renewable hydrogen could be produced at a lower cost. The analysis results have shown that renewable hydrogen could be produced at a price below 3.00 € /kg in this case and already at 4.80 € /kg if only an exemption from the EEG levy for consumption of self-produced renewable energy was granted. These prices compare to the current minimum cost of production of a small-scale vertically integrated system of 6.13 € /kg.

The simulation of an integrated system with a 100 kW sized PtG facility and a hybrid energy system composed of wind turbines and solar PV, each with a nameplate capacity of 5 MW, has shown that such a system would utilize exclusively

---

[105]Data source: Bundesnetzagentur (2019c). Day-ahead auction prices at the EPEX spot market.

[106]Cf. Glenk and Reichelstein (2019b), p. 218.

[107]The average emission factor of the German electricity mix in 2018 of 474 gCO$_2$/kWh (Agora Energiewende (2019), p. 13) at a power consumption of 52 kWh/kg results in total emissions of 24648 gCO$_2$/kgH$_2$ compared to the emissions of 7330 gCO$_2$/kgH$_2$ for hydrogen produced by steam methane reforming.

[108]"Small-scale PtG facilities" here refers to facilities with a nameplate capacity in the range of 100 kW.

Table 9: Mean values and price volatilities of the wholesale electricity prices (2015 - 2018).[105]

Year	2015	2016	2017	2018
Mean price	3.16	2.90	3.42	4.45
% of 2015 mean value	0.0%	−8.4%	8.1%	40.6%
Price volatility	1.27	1.25	1.77	1.78
% of 2015 volatility	0.0%	−1.5%	39.4%	40.3%

power from the renewable energy source. This confirms that a small-scale PtG facility under the described policy change would be suitable for production of pure renewable hydrogen at an economical price. A production based on grid power is avoided due to the high statutory fees imposed on grid-supplied electricity, since a small PtG facility does not exceed the non-exempt volume of 1 GWh.

Small-scale hydrogen applications could substantially benefit from break-even prices below 5 € /kg, such as hydrogen filling stations with onsite hydrogen production, which do not face high daily demand volumes and therefore would not consider an investment in a large-scale PtG facility.

Small systems might also be more likely to attract investments since they have a lower capital expenditure and moderate risks associated to the price development of electricity.

The utilization of the oxygen by-product showed a small price advantage of 0.24 € /kg, compared to the break-even price under current legislation. This could improve the profitability of investments to a small degree if an oxygen consumer is located nearby. However, further analysis is necessary whether additional investment costs accrue and if the sale of medical instead of industrial oxygen could additionally improve the economics of power-to-gas.

5. Conclusion

This thesis has analyzed the economics of vertically integrated energy systems under the current regulatory framework and for potential policy changes to determine the factors which inhibit an economical production of renewable hydrogen and find out what action is required for its improvement.

For this purpose, I applied a net present value model that considers investment costs in the form of the levelized cost of electricity and levelized fixed cost of hydrogen. These costs are offset by levelized terms of the contribution margin, which originate from the sale of hydrogen and renewable electricity. The model is structured in a way that the stand-alone net present values of both subsystems are separately computed and a third term representing the synergistic value of the system integration is accounted for. I leveraged this model to develop an algorithm for the derivation of the break-even price of hydrogen and built an optimization algorithm to facilitate the analysis of a wide range of system compositions and identify the price-minimizing capacity combination of the renewable energy system and the power-to-gas facility. Based on the developed algorithms, I implemented

a tool that allows for a relatively quick computation of the break-even price curve of a power-to-gas facility.

I find that small-scale power-to-gas facilities cannot be profitable stand-alone or in combination with a renewable energy source under current regulations, while large-scale facilities can compete on the market for medium-scale hydrogen supply due to the policies in favor of large industrial electricity consumers. However, hydrogen produced at large facilities would often be based on the use of grid electricity to a large share and thus cannot be considered renewable. The resulting hydrogen product would therefore be inferior to hydrogen from conventional production from an economic and ecological perspective, when comparing to hydrogen sourced on the market for large-scale supply.

Small-scale PtG facilities could be made profitable by two straightforward policy changes, which at the same time are suitable to incentivize the coupling of power-to-gas technology with a renewable energy source. These policies concern an exemption from the EEG levy payable on electricity consumed from the renewable energy source and the waiver of the EEG's feed-in requirement in favor of a production premium. When implemented, these measures would enable the production of truly renewable hydrogen at a break-even price below 3 € /kg, which could compete on the market for small- and medium-scale hydrogen supply and thus is suitable for deployment in some hydrogen applications. At the same time those PtG systems could be leveraged to balance the electricity market, if the statutory fees would be waived at times of negative electricity prices.

I also found that hybrid renewable energy sources, composed of wind turbines and solar PV, yield a lower break-even price of hydrogen, compared to systems only based on one renewable technology. However, further research is necessary, to find the ideal ratio of combination of solar PV and wind turbines. It must also be verified if suitable meteorological conditions for a joint operation exist.

My analysis has shown that the capture of oxygen as a by-product from water electrolysis, initially, has a low potential to improve the cost-effectiveness of electrolytic hydrogen. It remains open to explore the attainable oxygen price and whether the sale of medical oxygen would have a stronger effect on hydrogen economics.

Under real conditions, the described scenario is subject to uncertainties of various kinds, such as technological aspects concerning the system performance, meteorological conditions regarding the electricity output from the renewable energy source and uncertainties with respect to the electricity price movement at the power exchange, which might justify

further model adjustments. For a more grounded simulation of the power-to-gas facility, additional technical parameters should be taken into consideration. For instance, PEM electrolyzers have the potential to operate at an overload, which could be utilized to some extent. But varying loads could also adversely affect the long-term electrolyzer performance. The data for power generation has been regarded as preexisting. A calculation of the break-even price should also be conducted based on forecasts of the capacity factors of the renewable energy source in order to validate the price of hydrogen under assumption of the meteorological uncertainties present for intermittent renewable energy sources. Instead of only relying on the day-ahead auction price of electricity, several other electricity products, such as intraday trading, could be included in future research to capitalize on the flexibility of PEM electrolyzers.

References

Agora Energiewende. Agorameter - Dokumentation. Technical report, 2019. URL www.agora-energiewende.de.

BAFA. Merkblatt stromkostenintensive Unternehmen 2019. Technical report, Bundesamt für Wirtschaft und Ausfuhrkontrolle, 2019. URL https://www.bafa.de/SharedDocs/Downloads/DE/Energie/bar_merkblatt_unternehmen.html.

Bertuccioli, L., Chan, A., Hart, D., Lehner, F., Madden, B., and Standen, E. Study on development of water electrolysis in the EU - Final Report. *Fuel Cells and hydrogen Joint Undertaking*, (February):83, 2014. URL https://www.fch.europa.eu/sites/default/files/FCHJUElectrolysisStudy_FullReport%20(ID%20199214).pdf.

BMWi. Energiekonzept für eine umweltschonende, zuverlässige und bezahlbare Energieversorgung. Technical report, Bundesministerium für Wirtschaft und Technologie, 2010.

Bundesfinanzhof. BFH v. 14.04.2011 - IV R 46/09, 2011. URL https://datenbank.nwb.de/Dokument/Anzeigen/409262/.

Bundesnetzagentur. Ausschreibungen für EE- und KWK-Anlagen, 2019a. URL https://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen_Institutionen/Ausschreibungen/Ausschreibungen_node.html.

Bundesnetzagentur. Veröffentlichung von EEG-Registerdaten und -Fördersätzen, 2019b. URL https://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen_Institutionen/ErneuerbareEnergien/ZahlenDatenInformationen/EEG_Registerdaten/EEG_Registerdaten_node.html.

Bundesnetzagentur. SMARD | Download market data, 2019c. URL https://www.smard.de/en/downloadcenter/download-market-data/.

Bundesnetzagentur. Verbraucherthemen Energie: Preise und Rechnungen, 2019d. URL https://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Verbraucher/PreiseRechnTarife/preiseundRechnungen-node.html.

Carmo, M., Fritz, D. L., Mergel, J., and Stolten, D. A comprehensive review on PEM water electrolysis, apr 2013. ISSN 03603199.

DEHSt (Umweltbundesamt). Deutsche Versteigerungen von Emissionsberechtigungen - Jahresbericht 2015. Technical report, 2016. URL www.dehst.de.

Dorris, C. C., Lu, E., Park, S., and Toro, F. H. High-Purity Oxygen Production Using Mixed Ionic-Electronic Conducting Sorbents. *Senior Design Reports (CBE)*, 2016. URL http://repository.upenn.edu/cbe_sdr/78.

El-Emam, R. S. and Özcan, H. Comprehensive review on the techno-economics of sustainable large-scale clean hydrogen production, may 2019. ISSN 09596526.

ENTSO-E. Actual Generation per Production Type. URL https://transparency.entsoe.eu/generation/r2/actualGenerationPerProductionType/show.

European Central Bank. Statistics Bulletin, 2019. URL http://sdw.ecb.europa.eu.

European Commission. EU ETS - Free allocation of allowances. URL https://ec.europa.eu/clima/policies/ets/allowances.

Gardiner, M. Energy requirements for hydrogen gas compression and liquefaction as related to vehicle storage needs. *U.S Department Of Energy*, 25:1–6, 2009. URL https://www.hydrogen.energy.gov/pdfs/9013_energy_requirements_for_hydrogen_gas_compression.pdf.

Glenk, G. and Reichelstein, S. Economics of converting renewable power to hydrogen. *Nature Energy*, 4(3):216–222, mar 2019a. ISSN 20587546. doi: 10.1038/s41560-019-0326-1.

Glenk, G. and Reichelstein, S. Synergistic Value in Vertically Integrated Power-to-Gas Energy Systems. *Production and Operations Management*, oct 2019b. ISSN 1059-1478. doi: 10.1111/poms.13116. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/poms.13116.

International Energy Agency. *The Future of Hydrogen*. 2019. doi: 10.1787/1e0514c4-en.

Kothari, R., Buddhi, D., and Sawhney, R. L. Comparison of environmental and economic aspects of various hydrogen production methods, feb 2008. ISSN 13640321.

Netztransparenz. EEG Marktwertübersicht, 2019. URL https://www.netztransparenz.de/EEG/Marktpraemie/Marktwerte.

Noothout, P., De Jager, D., Tesnière, L., Van Rooijen, S., Karypidis, N., Brückmann, R., Jirouš, F., Breitschopf, B., Angelopoulos, D., Doukas, H., Konstantinavičiūtė, I., Resch, G., and Wien, T. U. The impact of risks in re-newable energy investments and the role of smart policies Final report Work Package 3 Leader Organisation: ECOFYS. Technical report, 2016.

OECD. Statutory Corporate Income Tax Rates, 2018. URL https://stats.oecd.org/Index.aspx?DataSetCode=CTS_CIT.

Pfenninger, S. and Staffell, I. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy*, 114:1251–1265, nov 2016. ISSN 03605442. doi: 10.1016/j.energy.2016.08.060.

Price, E. Durability and degradation issues in PEM electrolysis cells and its components: The second international workshop organised by the NOVEL EU-funded programme, jan 2017. ISSN 20565135.

pv magazine Deutschland. EnBW macht Weg frei für Baustart des 175 Megawatt Solarparks ohne EEG-Förderung, 2019. URL https://www.pv-magazine.de/2019/10/16/enbw-macht-weg-frei-fuer-baustart-des-175-megawatt-solarparks-ohne-eeg-foerderung/.

Rubert, T., McMillan, D., and Niewczas, P. The Effect of Upscaling and Performance Degradation on Onshore Wind Turbine Lifetime Extension Decision Making. In *Journal of Physics: Conference Series*, volume 926. Institute of Physics Publishing, nov 2017. doi: 10.1088/1742-6596/926/1/012013.

Schenuit, C., Heuke, R., and Paschke, J. Potenzialatlas Power to Gas - Eine Studie der Deutschen Energie-Agentur (dena). *Deutsche Energie-Agentur GmbH (dena)*, 2016.

Schmidt, O., Gambhir, A., Staffell, I., Hawkes, A., Nelson, J., and Few, S. Future cost and performance of water electrolysis: An expert elicitation study. *International Journal of Hydrogen Energy*, 42(52):30470–30492, dec 2017. ISSN 03603199. doi: 10.1016/j.ijhydene.2017.10.045.

Sterner, M., Bauer, F., Crotogino, F., Eckert, F., von Olshausen, C., Teichmann, D., and Thema, M. Chemical Energy Storage. In *Handbook of Energy Storage*, pages 325–482. Springer Berlin Heidelberg, Berlin, Heidelberg, 2019a. doi: 10.1007/978-3-662-55504-0_8. URL http://link.springer.com/10.1007/978-3-662-55504-0_8.

Sterner, M., Breuer, C., Drees, T., Eckert, F., Maaz, A., Pape, C., Rotering, N., and Thema, M. Storage Demand in Power Supply. In *Handbook of Energy Storage*, pages 51–136. Springer Berlin Heidelberg, 2019b. doi: 10.1007/978-3-662-55504-0_3.

Sterner, M., Stadler, I., Eckert, F., Gerhardt, N., von Olshausen, C., Thema, M., and Trost, T. Storage Integration for Coupling Different Energy Sectors. In *Handbook of Energy Storage*, pages 757–803. Springer Berlin Heidelberg, 2019c. doi: 10.1007/978-3-662-55504-0_14.

Sun, S., Shao, Z., Yu, H., Li, G., and Yi, B. Investigations on degradation of the long-term proton exchange membrane water electrolysis stack. *Journal of Power Sources*, 267:515–520, dec 2014. ISSN 03787753. doi: 10.1016/j.jpowsour.2014.05.117.

Tuson, G., Kobayashi, H., and Campbell, M. Oxygen enriched combustion system performance study. Phase 2: 100 percent oxygen enriched combustion in regenerative glass melters, Final report. Technical report, aug 1994. URL http://www.osti.gov/servlets/purl/10196777-oNIvYD/webviewable/.

Vogl, V., Åhman, M., and Nilsson, L. J. Assessment of hydrogen direct reduction for fossil-free steelmaking. *Journal of Cleaner Production*, 203:736–745, dec 2018. ISSN 09596526. doi: 10.1016/j.jclepro.2018.08.279.

Wind-FGW. Referenzerträge EEG 2017 - Anlage 2 (Enercon). URL https://wind-fgw.de/themen/referenzertraege.

ZSW and Bosch & Partner GmbH. Erfahrungsbericht gemäß § 97 EEG zur Solaren Strahlungsenergie. Technical report, 2019.