Tax Compliance and Behavioral Responses to Taxes

Von der Wirtschaftswissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades

Doktorin der Wirtschaftswissenschaften
- Dr. rer. pol. -

genehmigte Dissertation

von

M. Sc. Janine Katharina Jarzembski
geboren am 04.08.1992 in Hildesheim

2024

Referent

Prof. Dr. Kay Blaufus

Korreferent

Prof. Dr. Jens Robert Schöndube

Tag der Promotion

14. Dezember 2023

Summary

Diese Dissertation setzt sich aus drei Beiträgen zusammen, welche das Thema der Steuervermeidung aus verschiedenen Blickwinkeln untersuchen. Der erste Beitrag befasst sich mit der Wirksamkeit von Maßnahmen, mit denen die Einhaltung von Steuervorschriften verbessert werden soll. Basierend auf einer Umfrage von 10,000 individuellen Steuerzahlern wird gezeigt, dass eine verbesserte Qualität der Service-Interaktion zwischen der Finanzverwaltung und den Steuerpflichtigen mit einer erhöhten Steuerkonformität der Steuerpflichtigen assoziiert ist. Dieser Effekt ist vor allem auf Steuerpflichtige zurückzuführen, die ihre Steuererklärung selbst erstellen und über geringe Steuerkenntnisse und eine hohe Steuermoral verfügen. Weiterhin wird gezeigt, dass diese positive Assoziation vor allem von Finanzämtern mit höherer „Zwangsgewalt“ beeinflusst wird. Der zweite Beitrag befasst sich mit den Reaktionen von Unternehmen auf Vorschriften zur Eindämmung von Steuervermeidung. Seit 2016 verpflichtet die britische Finanzverwaltung Unternehmen, die bestimmte Größenschwellen überschreiten, dazu, ihre Steuerstrategie in qualitativer Form offenzulegen. Es wird untersucht, ob Firmen den Inhalt ihrer veröffentlichten Steuerstrategie strategisch beeinflussen, um z. B. potentielle Reputationskosten zu vermeiden. Gegenstand der Untersuchung sind dabei die offengelegten Steuerstrategien britischer Unternehmen, die im Financial Times Stock Exchange (FTSE) 100 und FTSE 250 gelistet sind. Die empirische Studie zeigt, dass der „Ton“, d. h. der Inhalt der offengelegten Steuerstrategie, nicht mit der tatsächlichen Steuerpolitik der Firma assoziiert ist, solange die Wahrscheinlichkeit der Aufdeckung falscher Angaben ausreichend gering ist. Der dritte Beitrag befasst sich mit den Folgen unternehmerischer Steuerplanung. Es wird untersucht, ob die Aufdeckung von Steuerplanung mit unternehmerischen Reputationskosten in Form von Umsatzeinbußen oder steigenden Werbekosten assoziiert ist sowie ob dieser Zusammenhang durch die Corporate Social Responsibility (CSR)-Leistung der Unternehmen beeinflusst wird. Zur Identifikation steuerplanender Unternehmen, werden die Daten zu Nachrichten über Steuerplanung aus der Studie von Lee et al. (2021) verwendet. Die Studie findet keinen Zusammenhang zwischen dem Umsatz oder den Werbeausgaben der Unternehmen und den Nachrichten über die Steuerplanung der Unternehmen. Die Daten enthalten keine Hinweise darauf, dass dieser Zusammenhang durch die CSR Performance der Firmen verstärkt wird.

This dissertation is composed of three articles that examine the subject of tax avoidance from different angles. The first article focuses on the effectiveness of regulations that aim to increase tax compliance. Using a survey of 10,000 individual taxpayers, it is shown that increasing the quality of service interaction between the tax administration and taxpayers is associated with increased tax compliance by taxpayers. This effect is mainly driven by self-preparing taxpayers with low tax knowledge and high tax morale. Furthermore, it is shown that this positive association is mainly influenced by tax offices with higher "coercive power". The second article deals with corporate behavioral responses to regulations that aim to curb tax avoidance. Since 2016, the U.K. tax administration requires firms that exceed specific size thresholds to disclose their tax strategy in qualitative form. It is investigated whether firms strategically influence the content of their disclosed tax strategy to avoid, for example, potential reputational costs. The investigation is based on disclosed tax strategies of U.K. companies listed on the FTSE 100 and FTSE 250. The empirical study shows that there is no association between "tone", i.e. the content of the disclosed tax strategy, and the firm's actual tax policy as long as the probability of detecting misstatements is sufficiently low. The third article addresses the consequences of corporate tax planning. It is examined whether the revelation of tax planning is associated with reputational costs in the sense of lost sales or increased advertising costs, and whether this association is influenced by firms' CSR performance. Data on news about firms' tax planning from Lee et al. (2021) is used to identify companies that engage in tax planning. The study does not find support for an association between firms' sales or advertising expenses and news about firms' tax planning. Moreover, no support is found that this link is magnified by the firms' CSR performance.

Schlagwörter: Compliance · Qualität des öffentlichen Dienstes · Umfrage · Bürger-Staat-Interaktionen · Durchsetzung · Steuertransparenz · Steuerliche Offenlegung · Steuerliche Falschangaben · Externe Überwachung · Steuerplanung · Steuervermeidung · Steuerhinterziehung · Steueroasen · Corporate Social Responsibility · Reputationskosten · Steuernachrichten

Keywords: Compliance · Public Service Quality · Survey · Citizen-State Interactions · Enforcement · Tax Transparency · Tax Disclosure · Tax Misreporting · External Monitoring · Tax Planning · Tax Avoidance · Tax Evasion · Tax Haven · Corporate Social Responsibility · Reputational Costs · Tax News

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Taxes impose costs on firms and individual taxpayers that can significantly reduce their after-tax profits. As a consequence, tax planning is pervasive in all tax regimes around the world. According to the Organisation for Economic Co-operation and Development (OECD), the global annual revenue loss due to base erosion and profit shifting is estimated to be between USD 100 billion to USD 240 billion (OECD 2017). Research shows that corporate tax rates significantly reduced over the last decades as firms face a high pressure to reduce their costs (Dyreng et al. 2017). The increased tax competition leads firms to apply tax planning that can range from illegal actions (i.e., tax evasion) to legal tax avoidance techniques that may or may not be in line with the spirit of the law. Not only can large corporations be tax aggressive, but individual taxpayers can also intentionally or unintentionally minimize their tax burden. The resulting gap between the statutory tax rate and the effective tax rates of corporations and individual taxpayers leads to undesirable consequences for the government, for which taxes are one of the most important sources of revenue (Dyreng et al. 2017). Consequently, effectively curbing tax planning is one of the most important issues for tax administrations around the world.

A modern approach by tax administrations to curb tax planning and increase tax compliance is to improve service provisions to be perceived as more ‘customer friendly’ (OECD 2007, 2013, 2020). Unlike tax audits, which are intended to detect misconduct, this service approach aims to prevent tax non-compliance ex ante by assisting taxpayers and affecting their attitude towards the tax authority. The service provisions aim not only to prevent unintentional mistakes, but also to build trust in the tax administration to increase voluntary compliance. In contrast to many other anti-avoidance regulations,

this approach also addresses tax compliance of individual taxpayers, which are often no tax-experts (Slemrod 2010). Since the expansion of services provisions implies significant investments for tax authorities, it is essential to determine whether they have the desired effect on taxpayers' tax compliance.

Another approach to combat tax planning is to increase tax transparency, forcing firms to disclose the relevant information by themselves. By providing information on the firms' tax policies, the tax authorities can improve their tax audits and detect legal loopholes (Müller et al. 2020). Numerous countries have enacted regulations to improve corporate tax transparency in recent years, which shows the relevance of this topic (see Müller et al. 2020 for an overview). Thereby, tax disclosures can be private or public. Public tax disclosures additionally discipline firms by holding them accountable to the public (Müller et al. 2020). If firms disclose that they are being tax aggressive and not paying their 'fair share' of taxes, they risk being labeled as 'poor corporate citizen'. The fear of public shaming and potential reputational costs are expected to reduce tax planning and encourage voluntary compliance. Nevertheless, firms adapt to new regulations and may find ways to circumvent the purpose of a law. These reactions must be taken into account when enacting or adjusting tax transparency regulations to achieve the desired guidance of corporate behavior and maximize the benefits. Regulations that require disclosure of qualitative information often provide firms with flexibility in choosing which information to disclose. These conditions provide the opportunity for strategic reporting. Firms may use platitudes and boilerplate language, limiting the informational content of their disclosures.

By strategic reporting, firms seek to avoid the risk of potential reputational costs. Tax aggressive firms have to be prepared to be targeted by the news (Chen et al. 2019). The concept of 'public shaming' can be used as a method to curb tax planning. Many managers fear reputational costs of tax planning, however, the literature on (ex post) reputational costs provides mixed results. As media coverage can usually not be controlled or biased by firms, it is an open question whether firms can take preventive actions to mitigate potential negative reactions to the revelation of tax planning activities. Recent studies show that corporate social responsibly (CSR) activities can provide an 'insurance-like protection' in case of adverse events. Nevertheless, the relationship between CSR and tax planning is complex and it is unclear whether customers' reactions to tax planning is magnified or alleviated by CSR performance.

1.2 Contribution and Main Findings

This thesis includes three essays. Table 1.1 provides an overview of the co-authors and current publication status of these studies.

Table 1.1: Essay Overview

Chapter	Titel	Co-authors	Current publication status
2	Do Better Tax Agency Services Improve Taxpayer Compliance?	Prof. Dr. Kay Blaufus Prof. Dr. Frank Hechtner	Working Paper
3	Public Disclosure of Tax Strategies and Firm's Actual Tax Policy	Prof. Dr. Kay Blaufus Jakob Reineke Dr. Ilko Trenn	Accepted for: Journal of Accounting, Auditing & Finance
4	The Effect of Corporate Social Responsibility on Reputational Costs of News about Tax Planning	–	Working Paper

Notes: The authors' contributions in the essays in chapter 2 and chapter 3 are approximately equally distributed.

The first article *Do Better Tax Agency Services Improve Taxpayer Compliance?* is presented in chapter 2. The research addresses the circumstance that many tax administrations worldwide try to effectively manage the taxpayers' service expectations and improve their service provisions (OECD 2007, 2013, 2020). The benefits or limitations of this service paradigm, however, are ambiguous. The article contributes to the literature by examining how the perceived service interaction quality affects individual taxpayers' tax compliance in Germany. It provides insights into behavioral responses to citizen-state interactions and is therefore informative for public administrations. The study provides support that service interaction quality between taxpayers and tax agency is positively associated with tax compliance. This effect is mainly driven by self-preparing taxpayers with low tax knowledge and high tax morale. Moreover, the positive association is influenced by tax offices with high perceived coercive power.

The second article *Public Disclosure of Tax Strategies and Firm's Actual Tax Policy*, presented in chapter 3, takes a closer look at another regulatory approach to curb tax avoidance – increasing corporate tax transparency. Whereas many studies focus on quantitative disclosures (e.g., Hoopes et al. 2018), little is known about the informativeness of qualitative tax disclosures and how firms adapt to these new regulations. The article fills this gap by investigating the content of U.K. tax strategy disclosures using a textual analysis. The U.K. released a regulation (Finance Act 2016, Schedule 19) requir-

ing certain firms as of September 15, 2016 to publically disclose information about their tax strategies. In contrast to other regulations, only qualitative information is required. This feature allows firms flexibility in what information they want to publish. If firms strategically report, they can avoid the steering purpose of the law and avoid 'public shaming' in case they fear reputational costs. The study converts the qualitative information of the firms' tax strategies into a numerical value and links it to the firm's actual tax policy. This methodology is applied to a sample consisting of FTSE 100 and FTSE 250 listed firms. The results support that the content of qualitative tax disclosures is managed as long as the likelihood of the misstatements being detected is sufficiently low.

In contrast to qualitative information, quantitative tax disclosures are more difficult for firms to bias. If tax aggressive firms fear reputational costs, they have to find another way to protect their reputation. Recent studies show that CSR activities can provide an 'insurance-like protection' in case of adverse events. However, it is not yet clear whether CSR actions can protect a firm against customer related reputational costs in the context of tax planning. Previous studies have shown that CSR and tax avoidance is rather viewed as inconsistent with one another (e.g., Inger and Vansant 2019). The third article *The Effect of Corporate Social Responsibility on Reputational Costs of News about Tax Planning*, presented in chapter 4, addresses the open question in literature whether tax planning is related to reputational costs. Furthermore, it contributes to the literature by investigating whether the risk of reputational damage varies with CSR performance. Although many firms fear reputational costs, the study cannot provide support that firms' sales (advertising expenses) are negatively (positively) related to news about firms' tax planning or that this effect is magnified by CSR performance.

Chapter 2

Do Better Tax Agency Services Improve Taxpayer Compliance?*

Abstract

In many countries, tax agencies are following Organisation for Economic Co-operation and Development (OECD) recommendations and are improving their services to increase voluntary tax compliance. By linking a survey of 10,000 individual German taxpayers with administrative data on the performance of tax offices, we find that better service interaction quality (SIQ) is indeed positively associated with tax compliance. However, there is great heterogeneity in this effect. First, improving SIQ mainly affects the tax compliance of self-preparers who have low tax knowledge and high tax morale, suggesting that tax agency services reduce the unintentional mistakes of these taxpayers. Second, our results show a positive association of SIQ with tax compliance when the perceived coercive power of the tax office is high, suggesting that deterrence and service provision are complementary instruments for increasing tax compliance. Furthermore, an increase in perceived SIQ is associated with a reduction in the probability of appeal. Thus, better services also help to significantly reduce the tax controversy costs for tax administrations and taxpayers.

* The following chapter is a co-authored paper with Prof. Dr. Kay Blaufus (Leibniz University Hannover) and Prof. Dr. Frank Hechtner (Friedrich-Alexander University Erlangen-Nürnberg).

2.1 Introduction

This paper examines the effect of the service interaction quality (SIQ) provided by the tax administration on taxpayers' compliance behavior. As an important trend in tax administration policies, the traditional "enforcement" paradigm of tax administration is considered to be incomplete and is thus complemented by a "service" paradigm. The U.S. Internal Revenue Service, for example, describes its mission as follows: "provide America's taxpayers top quality service by helping them understand and meet their tax responsibilities and enforce the law with integrity and fairness to all" (IRS 2020). According to the 2019 tax administration report of the Organisation for Economic Co-operation and Development (OECD), the "backbone of current efforts by tax administrations to manage compliance" supports positive compliance attitudes through education and taxpayer services (OECD 2019a, 22, 32). Similarly, in Germany, the country of our investigation, the Conference of State Finance Ministers stated that the central goal is a modern tax administration that sees itself as a service provider for citizens.[1]

In line with this reasoning, tax administrations worldwide aim to improve their service provision to become more "customer friendly" (OECD 2007, 2013, 2020). However, whether improved tax agency services increase taxpayer compliance is an open question. Prior studies in public administration research find a positive association between e-government services and trust in the government (Welch et al. 2005; Kim and Lee 2012; Im et al. 2014). However, whether this translates into higher compliance among citizens is empirically unclear. For example, May and Wood (2003) do not find an effect of building inspectors' enforcement style ("helpful and supportive" vs. "rigid and picky") on homeowners' compliance whereas the survey studies of Kirchler et al. (2006) and Gangl et al. (2013) as well as the lab experiments of Alm et al. (2010), McKee et al. (2018), and Vossler and McKee (2017) suggest a positive effect of agency services on compliance. Furthermore, it is unclear why service quality leads to greater compliance. Is it because it reduces intentional non-compliance or unintentional mistakes?

Our study fills this research gap by studying why and how individual taxpayers (either employed or self-employed) respond to differences in the perceived quality of tax agency services regarding their compliance and controversy behavior. Moreover, we

[1] Press release of the Annual Conference 2018 of the German Finance Ministers and Finance Ministers of the Federal States, 25.05.2018, FMK 5/18.

study how tax agencies can improve their SIQ. To this end, we use a large survey including more than 10,000 German citizens and match the survey data with administrative data on the performance and workload of tax offices and the satisfaction of tax officers. Our study, thus, responds to a recent call for more research on citizen-state interactions in public administration research (Jakobsen et al. 2019) by informing tax administrations and taxpayers about the potential benefits but also about the limitations of agencies' service paradigm.

2.2 Hypothesis Development

The standard economic model of tax compliance following Allingham and Sandmo (1972) assumes that taxpayers' compliance decisions depend only on their risk aversion, the tax rate, the detection probability and the fine in the case of detected noncompliance. However, in recent decades, research has significantly extended this model. The potential compliance effects of service quality can be derived from the following four approaches.

First, building on social psychology, Feld and Frey (2007) argue that tax compliance is based on a psychological tax contract between taxpayers and the government that establishes a fair, reciprocal exchange. This approach requires that taxpayers and the tax authority treat each other like partners. The psychological contract approach highlights the importance of interactional fairness, which is one specific aspect of the perceived SIQ of the tax administration. If taxpayers are treated unfairly and without respect by tax authorities, their intrinsic motivation to comply with the contract will diminish. In contrast, friendly, respectful, and trustful treatment should increase compliance due to the norms of reciprocity, which is seen "as one of the universal 'principal components' of moral codes". Thus, repaying for benefits received is regarded as a moral duty (Gouldner 1960).

Second, the *slippery slope framework* (Kirchler 2007) assumes that tax compliance depends on the power of the tax administration (ability to detect and punish noncompliance) as well as trust in the tax administration. Gangl et al. (2015) extend the slippery slope framework by differentiating between coercive/legitimate power and implicit/reason-based trust. The (extended) slippery slope framework assumes that an increase in the service quality of the tax administration increases trust in the administration and, thus, voluntary tax compliance. While implicit trust can be increased by friendly and

respectful treatment, providing competent advice to taxpayers increases reason-based trust (Gangl et al. 2015).

Third, the approach of *responsive regulation* (Braithwaite 2003) differentiates between different motivations of taxpayers and recommends adapting tax administrations' compliance strategy with regard to different types of motivational postures. Individuals who treat paying taxes as a "game" against the tax authorities should be countered with a deterrent approach, and better service provision should not affect these taxpayers. By contrast, for taxpayers with a positive attitude towards paying taxes, i.e., with high intrinsic motivation for tax honesty, better service quality should lead to higher tax compliance because those taxpayers' noncompliance is probably due to unintentional mistakes (misunderstanding of tax rules). There are estimates that 30 percent of incorrect tax returns in the U.S. are due to taxpayers' lack of tax knowledge (Christian 1994, Erard 1997). If tax administrations provide better information services, these taxpayers could learn about the correct treatment of certain revenues and expenditures, which could prevent them from making unintentional mistakes.

Fourth, according to fairness heuristic theory (Tyler and Lind 1992; Lind 2001), individuals use their judgments of interactional fairness as a heuristic guide to decide whether authorities can be trusted or will abuse their power. Thus, individuals who use the fairness heuristic will use the perceived service quality of the tax administration, particularly respectful and friendly treatment, as a proxy for the trustworthiness of the tax administration. Accordingly, for these individuals, an increase in perceived service quality increases trust in the tax administration and, thus, the willingness to cooperate by voluntarily filing a correct tax return.

In summary, all four approaches predict a positive effect of service quality on tax compliance, which leads to our first hypothesis:

H1: An increase in the SIQ of a tax office increases tax compliance.

However, the four approaches differ in their predictions regarding the potential heterogeneity in this effect. While the psychological contract approach and the slippery slope framework do not imply a heterogeneous effect of SIQ, the two other approaches do. Fairness heuristic theory assumes that individuals use the perceived SIQ of a tax office as a simplified heuristic to decide whether the tax office can be trusted and, thus, whether they should voluntarily cooperate with the tax office by being compliant. The

implication is that perceived service quality is mostly relevant to individuals who must rely on this heuristic because they are not able to judge the correctness of the administrations' tax assessment either based on their own tax knowledge or with the help of tax experts. In this case, taxpayers are assumed to use simple *available* indicators (friendly and respectful treatment by tax officers) as a substitute to assess the correctness of the tax administrations' assessment. In countries such as Germany, where the survey was conducted, taxpayers do not even have to calculate their taxes themselves. As opposed to self-reporting systems (such as in the U.S. and U.K.), German taxes are calculated by the fiscal authorities – the taxpayer only declares his/her taxable earnings and deductions and does not calculate his/her taxable income or tax payment. Thus, for taxpayers without the ability to check the correctness of their tax assessment, this assessment is actually a credence good (Darby and Karni 1973), and fairness heuristic theory predicts that these taxpayers will rely on perceived SIQ as a heuristic guide for their compliance behavior.

Therefore, from fairness heuristic theory, we derive that the positive effect of SIQ on compliance depends on the ability of the taxpayer to check the correctness of his/her tax assessment. This prediction is also in line with the assumption that some taxpayers make unintentional errors that can be reduced by the provision of administration information services (responsive regulation approach). Additionally, in this case, service quality should be most effective for individuals who do not have the ability to check the correctness of their tax assessment (i.e., taxpayers who self-prepare their tax returns and have only modest tax knowledge). Thus, we hypothesize as follows:

> ***H2:*** The positive effect of service quality depends on the ability of the taxpayer to check the correctness of his/her tax assessment.

Whereas fairness heuristic theory predicts that taxpayers with low tax literacy (i.e., a low ability to check the correctness of their tax assessment) rely on fairness heuristics and thus increase their voluntary compliance in response to improved SIQ, this positive effect depends on their tax morale according to the responsive regulation approach. Taxpayers with low tax literacy may make unintentional compliance errors; however, depending on their tax morale, they may additionally engage in aggressive tax avoidance. We expect that taxpayers with high tax morale do not engage in aggressive tax avoidance; thus, the noncompliance of high tax morale individuals mainly results from

unintentional mistakes. By contrast, we expect that the noncompliance of low tax morale individuals results from unintentional mistakes *and* aggressive tax avoidance. Thus, the effect of providing tax administration information services to taxpayers who are prone to unintentional mistakes (i.e., who have a low ability to check the correctness of their tax assessment) should depend on tax morale: high tax morale individuals should respond with full compliance by reducing unintentional mistakes, whereas low tax morale taxpayers might reduce unintentional mistakes but still engage in aggressive tax avoidance. Accordingly, we hypothesize the following:

H3: For taxpayers with a low ability to check the correctness of their tax assessment, the positive effect of service quality depends on their tax morale.

Finally, the positive effect of SIQ may depend on the tax office's perceived coercive power. On the one hand, one might expect that SIQ is more effective if coercive power is high because coercive power protects honest taxpayers from being exploited by potential tax evaders. If taxpayers conditionally act cooperatively (Frey and Torgler 2007), i.e., they comply as long as they expect all other taxpayers will also comply, sufficient coercive power might be a necessary precondition for a positive effect of improved service quality. Moreover, the incentive to use the information provided by higher levels of SIQ might be increased if coercive power is high. On the other hand, a very high level of coercive power might outweigh the effects of better services because the compliance of subjects is enforced and their intrinsic motivation to comply is crowded out (Feld and Frey 2007). From this perspective, one would expect that an increase in service quality is more effective under lower coercive power than under higher coercive power. Thus, theory is unclear about the direction of the effect, which makes it an empirical question. We hypothesize the following:

H4: The positive effect of service quality depends on the tax office's coercive power.

2.3 Sample Selection, Variable Measurement, Descriptive Statistics, and Estimation Strategy

2.3.1 Sample Selection

We use data from an online survey conducted by the Ministry of Finance of North Rhine-Westphalia from April 4, 2016, to December 31, 2016. North Rhine-Westphalia is the most populous federal state in Germany, having 8.48 million taxpayers (more than

20 percent of the German taxpayer population). In total, 29,632 taxpayers completed the survey.

The aim of the survey was to measure taxpayers' satisfaction with the tax agency's services. Taxpayers were asked to rate several aspects of services provided by the local tax office itself, the personal tax officer and the tax agency in general. The questionnaire included an item regarding audit adjustments made by the tax agency, demographic information (age, gender, income, income source, education of the taxpayer), and information about the use of external help, the e-filing mode, the type of filing (joint or single), the appeal decision, tax morale, the self-assessed tax knowledge level and an evaluation of the tax authorities' service quality.

As our survey data include a unique identifier for each local tax office, we are able to extend the primary data set by information on the percentage of granted appeals for each local tax office. Thus, we are able to control how consciously each of the 208 different tax offices works. In addition, we add information from a staff survey of all tax offices in North Rhine-Westphalia that contains information on employee satisfaction (average values at the tax office level), which we will use as an instrument (see section 2.3). This extended data set enables us to analyze the effects of SIQ on taxpayers' compliance behavior.

After merging the different data sets, we remove taxpayers with missing information on either tax audit adjustments or SIQ. In our baseline analyses, we additionally exclude taxpayers who have filed an appeal to ensure that the measurement of SIQ refers to the interaction before the tax assessment. After removing observations with missing data on the control variables and instrumental variables, we end up with a total sample of 10,443 taxpayers (see Table 2.1).

Table 2.1: Sample Selection

Filter	Observations
Initial sample size	29,632
./. missing values for audit adjustments	(5,697)
./. not all four items referring to service interaction quality were answered	(5,969)
./. taxpayers who have filed an appeal or give missing values for this	(4,335)
Sample size for calculating service interaction quality	13,631
./. missing values for the control variables	(2,700)
./. missing values for the instrumental variables	(488)
Final sample for IV regression	10,443

2.3.2 Variable Measurement and Descriptive Statistics

Table 2.6 in Appendix A displays the variable measurement. In line with prior research, we use tax audit adjustments to measure tax compliance (Mills 1996; e.g., Chan and Lan Mo 2000; Kleven et al. 2011). Our dependent variable COMPLIANCE equals one (zero) if the answer to the question "I had a deviation from my tax return" is no (yes).[2] In Germany, where tax assessments are made not by the taxpayer but by the revenue agency, a deviation from the taxpayer's tax return implies that the revenue agency has corrected the filed tax return. Thus, noncompliance encompasses both intentional noncompliance and unintentional mistakes.[3] Unintentional mistakes can occur due to errors in legal interpretation or due to simple calculation errors, since taxpayers often have to calculate the sum of their expenses. Moreover, the German tax law offers several tax planning opportunities for employees and the self-employed, concerning, for example, the amount of deductible commuting expenses, expenses for home offices, professionally induced travel expenses, double household deductions, and expenses for work equipment (Federal Statistical Office 2015, 16; Blaufus et al. 2017). Table 2.2 shows that 59 percent of survey participants comply with the tax rules (i.e., their tax returns were not corrected by the revenue agency).

We follow prior research and capture SIQ as the perceived quality (Parasuraman et al. 1988; Bansal et al. 2005; Parasuraman et al. 2005; Reimann et al. 2008)[4] from the perspective of taxpayers. To capture the SIQ of tax offices, we use the following four questions, which were answered on a five-point Likert scale (from 1: "I completely disa-

---

[2] As we do not have data on the number of audit adjustments, we are unable to make a statement about the association between SIQ and the extent of noncompliance.

[3] Deviations from the tax return also occur in the case of corrected "overcompliance". However, we assume that they represent very few cases and that they are randomly distributed over the tax offices; thus, the effect of SIQ is estimated without distortion. According to internal data of the financial administration, the ratio of overcompliance is approximately 0.5 percent. Another concern might be that not all cases of noncompliance are detected by the revenue agency. However, this should not bias our estimate of SIQ. Until 2016, no risk management system had been implemented by the German revenue agency, so that, in general, all tax returns had to be audited. Nevertheless, to consider the possibility that not all cases are examined equally strictly, but that the audit intensity depends on the complexity of the case, we control for case complexity using the variables age, income, and income type.

[4] One could argue that survey-based measurement of quality, although widely used, is problematic. It could be that all taxpayers are treated identically, but individuals differ in how they rate their services because of different personality types. Note that we address these concerns by including many taxpayer characteristics such as income, education, and age. In addition, to test whether reported service ratings vary with differences in actual treatment, we use data on actual tax offices' processing times and relate them to taxpayers' perceived ratings of processing times. We find that taxpayers' (perceived) adequacy of tax return processing time(s) (average scores at the tax office level) is significantly correlated with actual tax office processing time ($r=-0.55$ ($p<0.01$) (Pearson); $\rho=-0.56$ ($p<0.01$) (Spearman)). This confirms our assumption that differences in perceived SIQ are related to differences in actual treatment.

gree" to 5: "I fully agree"): (1) The employees are friendly and courteous; (2) the employees are helpful and support me; (3) the employees at my tax office are professionally competent; and (4) the employees deal conscientiously with my questions. We limit our analysis to the quality of the service interaction with the employees of the tax offices. Other aspects of service quality, such as the processing time of the tax return, cannot be used because taxpayers in the survey only evaluate the adequacy of the processing time of this year's tax return, which is known only when the tax assessment is received and, therefore, cannot influence the previous compliance decision. We further assume that the comprehensibility of written queries or other letters from the tax offices is largely standardized by means of sentence modules and that therefore, it should not vary much between tax offices. By contrast, the four questions we have selected ensure that there has been interaction with employees; thus, a transfer of support and sympathy towards the tax office is possible. Since the four questions are highly correlated, they cannot be used simultaneously in regression analysis. We use principal component analysis (PCA) to summarize the questions by one factor measuring SIQ. All questions load onto one component [rotated component loadings: (1) 0.9121, (2) 0.9404, (3) 0.8877, and (4) 0.9312], and the generated component scores represent our SIQ variable. The Cronbach's α amounts to 0.94.

Regarding H2, we measure the ability of the taxpayer to check the correctness of his/her tax assessment (ABILITY) by combining a question regarding the own evaluation of tax knowledge ("How would you rate your tax knowledge", 1: absolutely no knowledge, 5: expert) and a question concerning the preparation mode. ABILITY is a binary variable that equals one if the answer to the tax knowledge question is ≥ 4 (good/expert) or the taxpayer uses professional tax advice (a tax advisor or tax assistance association). In the sample, 50 percent have the ability to check the correctness of their tax assessment (Table 2.2).

Regarding H3, we split the subgroup of taxpayers with low ability (ABILITY=0) into taxpayers with high and low tax morale using a binary variable, HIGH_MORALE, that equals one if the taxpayer fully agrees with the following statement and zero otherwise: "I think one should be honest about the tax return" (1: completely disagree, 5: fully agree). In the full sample, we control for tax morale using the nondichotomized answer

to the question. The average value of MORALE is 4.73, indicating that most subjects strictly reject tax evasion.[5]

We measure coercive power (COERCIVE) by the answer to the question "My tax office is petty when it comes to checking my tax return(s)" (1: completely disagree, 5: fully agree). We assume that tax offices that are perceived as auditing more pedantic or petty, are more deterrent to taxpayers and thus exert more coercive power. This question was included in a survey similar to ours but conducted in 2012. We calculate the average values for each tax office and include them as our measure of coercive power. We use lagged data to ensure that the evaluation of pettiness is not influenced by the audit behavior of the tax office in 2016. In addition, the lagged data ensure that we measure the perception of deterrence before the compliance decision is made. On average, the evaluation of a tax office's coercive power is 2.64 on a 5-point Likert scale. With respect to H4, we differentiate between tax offices with high and low coercive power using the variable HIGH_COERCIVE, which equals one if the coercive power of the tax office is above the upper quartile.

As control variables, we include age, gender, income class, the type of filing (joint vs. single) INVESTMENT INCOME (a binary variable that equals one if the taxpayer earns income from capital assets or real estate), SELF-EMPLOYED (a binary variable indicating whether the taxpayer has income from self-employment), ELECTRONIC (a binary variable that equals one if the tax return is submitted electronically), EDUCATION (a binary variable that equals one if the taxpayer has at least a university entrance qualification), and CONSCIOUS (measuring how consciously each of the 208 different tax offices works). Through the latter variable, we control that the tax offices potentially also make mistakes.[6]

[5] 76 percent of taxpayers fully agree that people should be honest about their taxes. This proportion corresponds to the results of the World Value Survey (Wave 7). A total of 74.5 percent of respondents in Germany state that cheating on taxes if you have a chance is never justifiable (Haerpfer et al. 2020). However, to test the robustness of our results, we also vary the definition of tax morale. Within the robustness checks, we define a taxpayer as having high morale if he/she fully agrees that it is good that the tax authorities are buying "tax CDs" to combat tax evasion (tax CDs are electronic data sets that contain whistleblower information, e.g., bank accounts of potential evaders).

[6] In the sample selection, we already eliminate observations that involved filing an appeal or have missing values for this variable. This removes observations with obvious errors from the tax offices.

Table 2.2: Descriptive Statistics

Variables	N	mean	sd	min	p50	max
Dependent variable						
COMPLIANCE	10,443	0.59	0.49	0.00	1.00	1.00
Independent variables						
SIQ	10,443	0.05	0.98	-2.91	0.13	1.14
SIQ_COMP	9,454	0.00	0.19	-0.78	0.01	0.38
ABILITY	10,443	0.50	0.50	0.00	0.00	1.00
MORALE	10,443	4.73	0.54	1.00	5.00	5.00
COERCIVE	10,443	2.64	0.19	2.18	2.64	3.24
Instruments						
APPRECIATION	10,443	4.31	0.79	1.00	4.00	5.00
JOB SATISFACTION	10,443	3.58	0.16	3.15	3.59	4.06
Control variables						
CONSCIOUS	10,443	0.54	0.04	0.22	0.54	0.65
UNDER 36	10,443	0.13	0.33	0.00	0.00	1.00
36 TO 45	10,443	0.16	0.36	0.00	0.00	1.00
46 TO 55	10,443	0.25	0.43	0.00	0.00	1.00
56 TO 65	10,443	0.26	0.44	0.00	0.00	1.00
MALE	10,443	0.73	0.45	0.00	1.00	1.00
30,000 TO 50,000	10,443	0.37	0.48	0.00	0.00	1.00
50,000 TO 70,000	10,443	0.22	0.41	0.00	0.00	1.00
OVER 70,000	10,443	0.16	0.37	0.00	0.00	1.00
EDUCATION	10,443	0.74	0.44	0.00	1.00	1.00
JOINTLY	10,443	0.67	0.47	0.00	1.00	1.00
ELECTRONIC	10,443	0.79	0.40	0.00	1.00	1.00
INVESTMENT INC	10,443	0.24	0.43	0.00	0.00	1.00
SELF-EMPLOYED	10,443	0.23	0.42	0.00	0.00	1.00
Sample splitting						
HIGH_MORALE	10,443	0.76	0.43	0.00	1.00	1.00
HIGH_COERCIVE	10,443	0.25	0.43	0.00	0.00	1.00

Notes: All variables are defined in Table 2.6.

2.3.3 Estimation Strategy

To test H1, we estimate the following probit model:

$$\text{COMPLIANCE} = \beta_0 + \beta_1 \text{SIQ} + \beta_2 \text{ABILITY} + \beta_3 \text{MORALE} + \beta_4 \text{COERCIVE} + \beta_5 \text{CONTROLS} + \varepsilon. \quad (2.1)$$

We use robust standard errors to account for heteroscedasticity and cluster standard errors at the tax office level.

To test H2-H4, we estimate equation (2.1) separately for subsamples for which ABILITY=0 (ABILITY=1), HIGH_MORALE=0 (HIGH_MORALE=1) (for the sub-group of taxpayers with ABILITY=0), and HIGH_COERCIVE=0 (HIGH_COERCIVE=1).[7]

In the survey, SIQ was evaluated after receipt of the tax assessment notice. Consequently, the evaluation could have been distorted by the result of the tax assessment, which would lead to endogeneity. We address potential endogeneity using (i) a bivariate

[7] Using a subsample analysis, we avoid endogenous interaction terms as well as any problems that occur when interacting in logit models (Ai and Norton 2003).

probit model with IVs and (ii) a binary logit model that uses the average values of SIQ for each tax office based only on the evaluations of the subset of taxpayers submitting tax returns that have not been corrected (COMPLIANCE=1).[8]

We use the answer to the following question as an instrument: "I think it is a good thing that my personal tax burden [average tax rate] is mentioned in my income tax assessment notes" (1: completely disagree, 5: fully agree). This variable, APPRECIATION, measures the individual's appreciation of tax information provision such that we expect that the variable is positively correlated with SIQ (instrument relevance), and indeed, the correlation amounts to r=0.24 (p<0.01) (Pearson) or ρ=0.28 (p<0.01) (Spearman). In addition, we assume that appreciation has no direct effect (other than via SIQ) on the dependent variable, COMPLIANCE (i.e., APPRECIATION is uncorrelated with the error term of our outcome regression given our controls for taxpayer characteristics such as the taxpayer's know-how in our regression analysis, instrument exogeneity). Our second instrument, JOB SATISFACTION, measures the average satisfaction of tax officers with their work activities. We use a dataset provided to us by the revenue agency that includes information on controlling indicators (see section 2.5). The measure is based on the question "I am satisfied with my job (the content of the activities, the type of work tasks and the workload)" (1: completely disagree, 5: fully agree). The satisfaction of tax officers is most likely correlated with the service they provide. In our sample, the correlation amounts to r=0.03 (p<0.01) (Pearson) or ρ=0.03 (p<0.01) (Spearman). Although the correlation is small, it is highly significant. The first-stage F-statistic of our two instruments exceeds 10 (Stock and Watson 2015, 490). By definition, tax officers' satisfaction with their work should not affect taxpayers' compliance behavior other than via their service provision quality given our controls for taxpayer and tax office characteristics (in particular, we control for the way the tax office employees work by CONSCIOUS). Moreover, in unreported regressions, we include APPRECIATION and JOB SATISFACTION in our outcome equation and consistently find that they do not significantly affect COMPLIANCE. In addition, we do not reject the null hypothesis of the overidentifying restrictions test that all instruments are jointly exogenous (Stock and Watson 2015, 493 f.).

[8] Alternatively, we could use only the evaluations of the subgroup of taxpayers submitting tax returns that have been corrected (COMPLIANCE=0). However, the disadvantage of doing so is that the evaluation of SIQ could also depend on the amount of deviation from the submitted tax return, which we do not observe in the data.

Our second approach to coping with the potential endogeneity of SIQ uses the variation in the average perceived SIQ between tax offices based on the evaluations of only those taxpayers whose tax assessments do not deviate from their tax return (COMPLIANCE =1). Thus, we estimate equation (2.1) using SIQ_COMP (instead of SIQ) as an independent variable. The advantage of this "only compliant" approach is that we definitely exclude endogeneity; however, the disadvantage is that we lose much variation in our independent variable. Moreover, while the model with IVs captures both the between and the within variation in SIQ, the logit regression ("only compliant" approach) investigates only whether the variation in SIQ *between* tax offices has an effect on tax compliance.

2.4 Results

Table 2.3 displays the results. Models (1) to (7) present the results of the IV approach, and models (8) to (14) present the results of the logit estimation using the average perceived SIQ between tax offices based on the evaluations of only those taxpayers whose tax assessments do not deviate from their tax return ("only compliant" approach).

Regarding the hypothesized positive effect on tax compliance (H1), models (1) and (8) show a significant effect of service interaction quality in the full sample. Moving from the lower quartile to the upper quartile of SIQ_COMP, we find that the average predicted probability of being compliant increases by 1.5 to 1.6 percentage points. This implies that by increasing service interaction quality from the lower quartile to the upper quartile, the absolute number of incorrect tax returns could be reduced by approximately 6.7 percent to 6.8 percent.[9] These are conservative estimates as we decided to use the lower quartile and the upper quartile of SIQ_COMP (p25=-0.11; p75=0.13) to calculate average marginal effects. The variation in SIQ_COMP is much smaller than the variation in SIQ because we calculate average values per tax office (see Table 2.2). If we use instead the lower and upper quartile of SIQ (p25= -0.63; p75=1.14)[10], we obtain an increase of approximately 11 percentage points. However, to provide conservative estimates, we subsequently use only the increase in SIQ_COMP to calculate the

[9] In 2016, 28.1 million taxpayers with unlimited tax liability filed their income tax return. Out of these tax returns, 6.4 million tax returns had to be corrected by the revenue agency. Multiplying the number of incorrect tax returns by the increase in compliance probability, we obtain approximately 430,000 fewer incorrect tax returns. This figure corresponds to a 6.7 percent reduction in incorrect tax returns.

[10] The upper quartile of SIQ (p75=1.14) equals the maximum value because 25.02 percent of the taxpayers rate all four service interaction quality questions with 5: "I fully agree".

average predicted probabilities. Regarding the expected heterogeneity in this effect (H2-H4), the results of the subsample analyses are displayed in models (2)-(7) and (9)-(14).

In line with H2, we find a positive (nonsignificant) association of SIQ for taxpayers with a low (high) ability to check the correctness of the tax office's tax assessment. For taxpayers with low ability, the average predicted probability of being compliant increases in both models by 1.8 percentage points if the tax administration is able to increase its SIQ from the lower quartile to the upper quartile. This finding is in line with fairness heuristic theory and the assumption that low-ability taxpayers make unintentional errors that can be reduced by the provision of administration information services. To shed further light on which of these two approaches best explains the data, we now turn to the test of H3.

According to fairness heuristic theory, taxpayers with a low ability to check the correctness of their tax assessment should increase their tax compliance in response to improved SIQ regardless of whether they have high or low tax morale. By contrast, if better SIQ mainly reduces unintentional compliance errors, we expect that the effect of SIQ differs between high and low tax morale subjects, as the noncompliance of taxpayers with low ability and high tax morale is mainly due to unintentional mistakes, not aggressive tax avoidance (H3). Our results support H3, suggesting that improving SIQ mainly helps reduce unintentional compliance errors.[11] We find a positive (nonsignificant) association of SIQ with compliance for taxpayers with low ability and high morale (low morale).[12] An increase from the lower quartile to the upper quartile of SIQ increases the average predicted probability of being compliant by 1.9 to 2.1 percentage points for taxpayers with low ability and high morale.

With respect to the difference in the perceived coercive power of the tax office (H4), our findings show a positive (nonsignificant) association of SIQ with compliance when perceived coercive power is high (low). This result indicates that increasing deterrence (by raising coercive power) and simultaneously increasing service provision are two approaches that are not mutually exclusive; rather, they complement each other. In-

[11] In unreported tests, we also examine whether there is an indirect effect of SIQ_COMP on COMPLIANCE mediated by MORALE. However, we do not find a significant effect of SIQ_COMP on MORALE and thus no indirect effect. This supports our interpretation that service provision mainly prevents taxpayers from making unintentional errors.

[12] We also perform the subsample analysis of HIGH_MORALE for taxpayers for whom ABILITY=1. However, we find only nonsignificant associations of SIQ with tax compliance. The p-values amount to 0.93 (ABILITY=1 & HIGH_MORALE=0) and 0.62 (ABILITY=1 & HIGH_MORALE=1) for the instrumental variable approach and 0.52 (ABILITY=1 & HIGH_MORALE=0) and 0.41 (ABILITY=1 & HIGH_MORALE=1) for the "only compliant" approach.

creasing SIQ provides more information and at the same time increasing deterrence increases the incentive to use this information. An increase from the lower quartile to the upper quartile of SIQ increases the average predicted probability of being compliant by 1.8 to 2.1 percentage points for tax offices with high perceived coercive power.

Table 2.3: Regression Results (Dependent Variable: Compliance)

MODEL	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
TYPE	IV	IV	IV	IV	IV	IV	IV	Logit	Logit	Logit	Logit	Logit	Logit	Logit
				ABILITY = 0							ABILITY = 0			
SAMPLE	Full Sample	ABILI. =0	ABILI. =1	HIGH_ MOR. =0	HIGH_ MOR. =1	HIGH_ COERC. =0	HIGH_ COERC. =1	Full Sample	ABILI. =0	ABILI. =1	HIGH_ MOR. =0	HIGH_ MOR. =1	HIGH_ COERC. =0	HIGH_ COERC. =1
SIQ	0.14**	0.23***	0.04	-0.02	0.31***	0.11	0.19*							
	(0.07)	(0.09)	(0.10)	(0.23)	(0.09)	(0.09)	(0.11)							
SIQ_COMP								0.27*	0.30**	0.24	0.13	0.36**	0.15	0.37*
								(0.14)	(0.15)	(0.21)	(0.30)	(0.16)	(0.18)	(0.20)
ABILITY	0.27***					0.30***	0.18***	0.38***					0.43***	0.23**
	(0.03)					(0.04)	(0.05)	(0.05)					(0.05)	(0.10)
MORALE	-0.00	-0.05	0.04			0.00	-0.01	0.07*	0.03	0.12**			0.06	0.13
	(0.03)	(0.04)	(0.05)			(0.03)	(0.05)	(0.04)	(0.06)	(0.06)			(0.04)	(0.11)
COERCIVE	0.11	0.25***	-0.03	0.20	0.28***			0.22*	0.40**	0.02	0.43	0.38**		
	(0.07)	(0.09)	(0.10)	(0.15)	(0.10)			(0.13)	(0.16)	(0.19)	(0.28)	(0.18)		
CONSCIOUS	-0.81**	-1.20***	-0.30	-3.60***	-0.60	-0.81*	-0.93	-0.95	-2.01**	0.33	-5.01**	-1.18	-0.84	-1.24
	(0.38)	(0.43)	(0.57)	(1.10)	(0.48)	(0.47)	(0.75)	(0.88)	(0.95)	(1.22)	(2.11)	(1.05)	(1.11)	(1.60)
UNDER 36	-0.17***	-0.28***	-0.05	-0.53***	-0.19*	-0.19***	-0.13	-0.43***	-0.70***	-0.13	-0.87***	-0.65***	-0.39***	-0.54***
	(0.06)	(0.09)	(0.08)	(0.16)	(0.11)	(0.07)	(0.09)	(0.08)	(0.12)	(0.12)	(0.23)	(0.14)	(0.10)	(0.14)
36 TO 45	-0.20***	-0.24***	-0.15**	-0.39***	-0.19**	-0.20***	-0.21*	-0.42***	-0.61***	-0.22*	-0.63***	-0.60***	-0.40***	-0.51**
	(0.05)	(0.07)	(0.07)	(0.13)	(0.09)	(0.06)	(0.12)	(0.08)	(0.10)	(0.12)	(0.20)	(0.12)	(0.08)	(0.20)
46 TO 55	-0.10**	-0.10	-0.10	-0.21	-0.06	-0.12**	-0.06	-0.29***	-0.34***	-0.24**	-0.36*	-0.33***	-0.30***	-0.28*
	(0.05)	(0.07)	(0.07)	(0.13)	(0.08)	(0.06)	(0.10)	(0.08)	(0.10)	(0.11)	(0.19)	(0.11)	(0.09)	(0.15)
56 TO 65	0.03	-0.01	0.08	-0.02	-0.03	0.02	0.06	-0.03	-0.12	0.07	-0.05	-0.15	-0.06	0.05
	(0.04)	(0.06)	(0.06)	(0.11)	(0.06)	(0.05)	(0.09)	(0.07)	(0.09)	(0.11)	(0.16)	(0.11)	(0.08)	(0.16)
MALE	-0.11***	-0.09**	-0.12***	-0.09	-0.11**	-0.12***	-0.09	-0.16***	-0.07	-0.25***	-0.14	-0.05	-0.16***	-0.14
	(0.03)	(0.05)	(0.04)	(0.09)	(0.05)	(0.03)	(0.06)	(0.04)	(0.07)	(0.07)	(0.14)	(0.09)	(0.05)	(0.10)
30K TO 50K	-0.02	-0.01	-0.05	-0.10	0.03	-0.03	-0.02	-0.02	-0.03	-0.02	-0.17	0.02	-0.03	0.01
	(0.03)	(0.04)	(0.05)	(0.10)	(0.04)	(0.03)	(0.07)	(0.05)	(0.07)	(0.09)	(0.16)	(0.08)	(0.05)	(0.13)
50K TO 70K	-0.11***	-0.12**	-0.10*	-0.14	-0.12**	-0.11**	-0.09	-0.13**	-0.18**	-0.09	-0.23	-0.17*	-0.13*	-0.13
	(0.04)	(0.05)	(0.06)	(0.11)	(0.06)	(0.04)	(0.08)	(0.07)	(0.09)	(0.11)	(0.18)	(0.09)	(0.08)	(0.14)
OVER 70K	-0.14***	-0.15**	-0.15**	-0.11	-0.14**	-0.15***	-0.11	-0.21***	-0.22**	-0.22*	-0.17	-0.22*	-0.24***	-0.14
	(0.04)	(0.06)	(0.07)	(0.14)	(0.07)	(0.05)	(0.10)	(0.08)	(0.11)	(0.12)	(0.22)	(0.12)	(0.09)	(0.18)
EDUCATION	-0.08**	-0.05	-0.11*	0.02	-0.06	-0.09**	-0.05	-0.09	-0.05	-0.13	0.04	-0.09	-0.12	-0.02
	(0.04)	(0.05)	(0.06)	(0.10)	(0.06)	(0.04)	(0.09)	(0.07)	(0.08)	(0.11)	(0.16)	(0.09)	(0.07)	(0.15)
JOINTLY	-0.21***	-0.27***	-0.14***	-0.31***	-0.25***	-0.20***	-0.24***	-0.33***	-0.45***	-0.18**	-0.53***	-0.43***	-0.30***	-0.43***
	(0.03)	(0.04)	(0.04)	(0.08)	(0.04)	(0.03)	(0.06)	(0.05)	(0.07)	(0.07)	(0.13)	(0.07)	(0.06)	(0.09)
ELECTRONIC	-0.27***	-0.27***	-0.26***	-0.37***	-0.24***	-0.29***	-0.21***	-0.43***	-0.41***	-0.42***	-0.59***	-0.35***	-0.46***	-0.33***
	(0.04)	(0.04)	(0.05)	(0.09)	(0.05)	(0.04)	(0.06)	(0.06)	(0.07)	(0.09)	(0.14)	(0.08)	(0.07)	(0.11)
INV. INC	-0.07**	-0.10**	-0.04	-0.13	-0.09*	-0.08**	-0.06	-0.07	-0.16**	0.01	-0.14	-0.17*	-0.07	-0.09
	(0.03)	(0.04)	(0.04)	(0.09)	(0.05)	(0.03)	(0.05)	(0.04)	(0.07)	(0.07)	(0.16)	(0.09)	(0.05)	(0.08)
SELF-EMPL.	0.07**	-0.01	0.11**	0.17	-0.06	0.06	0.12*	0.24***	0.15	0.27***	0.42**	0.06	0.20**	0.35**
	(0.04)	(0.05)	(0.05)	(0.10)	(0.06)	(0.04)	(0.07)	(0.09)	(0.12)	(0.10)	(0.20)	(0.14)	(0.10)	(0.16)
Constant	0.88***	0.97***	0.95**	2.34***	0.27	1.18***	1.21***	0.77	1.19	0.64	3.07**	0.89	1.40**	1.24
	(0.31)	(0.37)	(0.47)	(0.76)	(0.38)	(0.34)	(0.43)	(0.63)	(0.75)	(0.90)	(1.49)	(0.73)	(0.65)	(0.94)
# Tax Off.	208	208	208	187	207	156	52	132	132	132	131	132	99	33
Obs	10,443	5,240	5,203	1,306	3,934	7,812	2,631	9,454	5,005	4,449	1,250	3,755	7,070	2,384
Pseudo R^2	-	-	-	-	-	-	-	0.0235	0.0297	0.0140	0.0473	0.0261	0.0245	0.0231
Area under ROC curve	0.6271	0.6387	0.6031	0.6571	0.6370	0.6286	0.6252	0.6037	0.6163	0.5788	0.6471	0.6086	0.6055	0.6044

Table 2.3: Regression Results (Dependent Variable: Compliance) (continued)

Notes: The table shows the results of the probit regression (IV approach) and the logit estimation ("only compliant" approach) using the dependent variable, COMPLIANCE. All variables are defined in Table 2.6. The variables APPRECIATION and JOB SATISFACTION are used as instruments for the instrumented variable *SIQ* in models (1)-(7). Robust clustered standard errors are in parentheses. *, **, *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively (based on two-tailed tests). The number of observations in models (8)-(14) differs from that in models (1)-(7), as the variable SIQ_COMP is calculated only if the tax office has at least 20 observations of the SIQ of compliant taxpayers to obtain a valid mean value. Furthermore, taxpayers with missing values for the IV APPRECIATION are included in models (8)-(14). We split the subsamples with ABILITY=0 (ABILITY=1), HIGH_MORALE=0 (HIGH_MORALE=1) (for the subgroup with ABILITY=0), and HIGH_COERCIVE=0 (HIGH_COERCIVE =1).

2.5 Additional Analyses and Robustness Tests

In this section, we present additional analyses concerning (i) the effect of SIQ on the decision to file an appeal, (ii) the determinants of the tax office's service performance, and (iii) robustness checks. Table 2.7 in Appendix A displays an overview of the distribution of all the variables used in this section.

2.5.1 Appeal Decision

Thus far, we have investigated the influence of SIQ on tax compliance. In addition to the effect on compliance, SIQ is likely to impact appeal decisions. We adjust our sample and include all taxpayers whose tax return was corrected by the revenue agency and who have filed an appeal or not (6,543 taxpayers). The regression results are reported in Table 2.4. The results demonstrate that the SIQ of a tax office is negatively associated with the probability of filing an appeal. The average predicted probability of an appeal can be decreased by 2.5 to 3.8 percentage points (27.4 percentage points if we use the lower and upper quartile of SIQ instead of SIQ_COMP in the IV regression).[13] This decrease corresponds to a 4.8 percent to 7.3 percent reduction in the total number of appeals, which highlights the relevance of SIQ for tax controversy costs.[14] In the case of Germany, we estimate that improving SIQ from the lower quartile to the upper quartile would result in a reduction in taxpayer compliance costs and revenue agency costs of at least 30 million euros.[15]

[13] We find no systematic differences for taxpayers with low and high ability or for tax offices with low or high coercive power. In all these subsamples, the effect of SIQ on appeal decisions is significant.

[14] Given that 6.4 million tax returns had to be corrected by the revenue agency in 2016 and 3.3 million appeals were filed, we calculate the reduction in total appeals to be 160,000 and 243,000, which corresponds to a 4.8 percent to 7.3 percent reduction in the total number of appeals.

[15] We estimate taxpayers' compliance costs of filing an appeal using the lower-bound estimation of Blaufus et al. (2019a). To estimate the agency's administration costs, we assume that the average processing time for an appeal amounts to at least 0.5 hours. We then multiply the time burden by the average hourly wage of tax officers and neglect any other costs, e.g., the costs of materials and premises. In summary, our estimates present conservative lower-bound predictions.

Table 2.4: Regression Results (Dependent Variable: Appeal)

MODEL	(1)	(2)
TYPE	IV	Logit
SAMPLE	Full Sample	Full Sample
SIQ	-0.51***	
	(0.08)	
SIQ_COMP		-0.47***
		(0.16)
ABILITY	0.27***	0.64***
	(0.05)	(0.07)
MORALE	0.02	-0.06
	(0.04)	(0.06)
COERCIVE	0.13	0.16
	(0.12)	(0.21)
CONSCIOUS	0.08	0.09
	(0.61)	(1.17)
UNDER 36	-0.15**	0.09
	(0.08)	(0.12)
36 TO 45	-0.20***	-0.08
	(0.07)	(0.10)
46 TO 55	-0.15**	0.05
	(0.07)	(0.10)
56 TO 65	-0.04	0.13
	(0.06)	(0.10)
MALE	-0.05	-0.16***
	(0.04)	(0.06)
30K TO 50K	0.01	0.08
	(0.05)	(0.08)
50K TO 70K	0.05	0.13
	(0.06)	(0.10)
OVER 70K	0.18***	0.31***
	(0.06)	(0.09)
EDUCATION	0.12**	0.15
	(0.05)	(0.10)
JOINTLY	-0.05	-0.13*
	(0.04)	(0.07)
ELECTRONIC	-0.10**	-0.23***
	(0.05)	(0.07)
INV. INC	0.16***	0.23***
	(0.04)	(0.07)
SELF-EMPL.	0.17***	0.22**
	(0.05)	(0.10)
Constant	-1.12**	-1.12
	(0.44)	(0.81)
# Tax Off.	208	132
Obs	6,543	5,795
Pseudo R^2	-	0.0320
Area under ROC curve	0.7284	0.6199

Notes: The table shows the results of the probit regression (IV approach) and the logit estimation ("only compliant" approach) using the dependent variable, APPEAL. All variables are defined in Table 2.6. The variables APPRECIATION and JOB SATISFACTION are used as instruments for the instrumented variable SIQ in models (1)-(7). Robust clustered standard errors are in parentheses. *, **, *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively (based on two-tailed tests). The number of observations in models (8)-(14) differs from that in models (1)-(7), as the variable SIQ_COMP is calculated only if the tax office has at least 20 observations of the SIQ of compliant taxpayers to obtain a valid mean value. Furthermore, taxpayers with missing values for the IV APPRECIATION are included in models (8)-(14). We split the subsamples with ABILITY=0 (ABILITY=1), HIGH_MORALE=0 (HIGH_MORALE=1) (for the subgroup of ABILITY=0), and HIGH_COERCIVE=0 (HIGH_COERCIVE =1).

2.5.2 Determinants of Service Performance

Our previous analyses demonstrate that improving SIQ has the potential to increase tax compliance and reduce tax controversy costs. Thus, it is interesting to study how the revenue agency can improve its SIQ given the usually severe budget constraints that limit the opportunity to increase the number of revenue agents. As we assume that the number of revenue agents is fixed at least in the short run, we aim to explain the differences in SIQ between tax offices controlling for the number of employees (full-time equivalents) at each tax office. To benchmark the service performance of each tax office, we apply data envelopment analysis (DEA). DEA provides a linear programming model to assign a scalar efficiency value to each "decision-making unit" (DMU) (Charnes et al. 1978).[16] The idea of DEA is to maximize the ratio of weighted outputs to weighted inputs. We use the number of employees as the input variable and the answers to the four SIQ questions as output variables. We calculate the average evaluation per tax office based only on the evaluations of the subset of taxpayers submitting tax returns that have not been corrected (COMPLIANCE=1) and end up with four output variables in our DEA (FRIENDLY_COMP, HELPFUL_COMP, COMPETENT_COMP, QUESTIONS_COMP). We identify ten service-efficient tax offices; the efficiency score ranges between 0.77 and 1 (see Table 2.7 in Appendix A).

To identify the determinants of SIQ, we conduct an ordinary least squares (OLS) regression at the tax office level with the DEA efficiency measure (θ_{SIQ}) as the dependent variable. Regarding potential determinants, we exploit a unique dataset that is provided to us by the revenue agency. These data include information on controlling indicators such as the number of employees, the number of cases filed and closed, the job vacancy rate for 2015 and 2016 and information on average employee satisfaction collected by the revenue agency in 2015 for all tax offices in North Rhine-Westphalia. The data on employee satisfaction cover a large number of categories. We perform PCA to reduce the number of variables and create index measures, obtaining three components.[17] The first component captures any satisfaction with the work itself and career opportunities;

[16] A relative efficiency value is calculated by estimating a best practice frontier (Bogetoft and Otto 2011, 81). DEA evaluates the performance of DMUs by the deviation from this frontier line (envelope). In contrast, a classical regression analysis uses the distance to the fitted line; thus, it estimates efficiency relative to average performance (Cooper et al. 2007, 4; Demerjian et al. 2012). Another unique advantage of DEA, which we exploit in our study, is the possibility of considering multiple outputs.

[17] We limit the number of components to those that have eigenvalues greater than one. The components explain approximately 70 percent of the total variance. Table 2.8 in Appendix A displays the varimax rotated component loadings using Kaiser normalization.

it also includes satisfaction with colleagues and management behavior. The second component reflects satisfaction with the political framework. Finally, the third component measures the satisfaction of tax office employees with their external conditions (payment, working hours, IT equipment, office space, working environment, aids, technical equipment, noise, temperature, etc.). We expect that the service performance of tax offices is driven by the satisfaction of employees (e.g., Snipes et al. 2005). As an additional explanatory variable, we include the job vacancy rate and the workload (the number of incoming cases and incomplete cases from the previous year (in 1,000) divided by the number of employees). We expect that the higher the vacancy rate and the workload, the more overworked and stressed the employees are, which might reduce their focus on providing high-quality service. Finally, we include population density as a regressor to control for the potential differences between urban and rural areas. The regression results are reported in Table 2.5.

We find evidence for significant positive effects of employees' satisfaction with work and satisfaction with external conditions. Thus, by increasing the satisfaction of revenue agents, the revenue agency could indirectly improve tax compliance and reduce tax controversy costs.

Table 2.5: Regression Results (Dependent Variable: Θ_SIQ)

MODEL	(1)
SATISFACTION WORK	0.01***
	(0.00)
SATISFACTION POL. FRAMEWORK	-0.00
	(0.00)
SATISFACTION EXT. CONDITIONS	0.01*
	(0.00)
POPULATION DENSITY	-0.00
	(0.00)
JOB VACANCY RATE	0.07
	(0.08)
WORK EFFORT	0.01
	(0.01)
ASSESSMENT SELF-EMPLOYMENT	0.00
	(0.01)
Constant	0.85***
	(0.08)
Observations	132
Adjusted R^2	0.0770

Table 2.5: Regression Results (Dependent Variable: Θ_SIQ) (continued)

Notes: The table shows the results of the OLS regression using Θ_{SIQ} as the dependent variable. Θ_{SIQ} is the measure of service performance calculated by DEA. The number of employees is used as the input variable, and the variables FRIENDLY_COMP, HELPFUL_COMP, COMPETENT_COMP and QUESTIONS_COMP are used as output variables. The NUMBER OF EMPLOYEES refers to full-time equivalents in 2016. FRIENDLY_COMP measures the average rating of the tax office based on the question "The employees are friendly and courteous" (1: "I completely disagree" to 5: "I fully agree") based on the evaluations of taxpayers whose tax assessments do not deviate from their tax return. HELPFUL_COMP measures the average rating of the tax office based on the question "The employees are helpful and support me" (1: "I completely disagree" to 5: "I fully agree") based on the evaluations of taxpayers whose tax assessments do not deviate from their tax return. COMPETENT_COMP measures the average rating of the tax office based on the question "The employees in my tax office are professionally competent" (1: "I completely disagree" to 5: "I fully agree") based on the evaluations of taxpayers whose tax assessments do not deviate from their tax return. QUESTIONS_COMP measures the average rating of the tax office based on the question "The employees conscientiously deal with my questions" (1: "I completely disagree" to 5: "I fully agree") based on the evaluations of taxpayers whose tax assessments do not deviate from their tax return. SATISFACTION WORK measures the satisfaction of tax office employees with their work and career opportunities (rotated component calculated by PCA). SATISFACTION POL. FRAMEWORK measures the satisfaction of tax office employees with the conditions of the political framework (rotated component calculated by PCA). SATISFACTION EXT. CONDITIONS measures the satisfaction of tax office employees with their external conditions (payment, working hours, IT equipment, office space, working environment, aids, technical equipment, noise, temperature, etc.) (rotated component calculated by PCA). POPULATION DENSITY is the number of inhabitants (in 1,000) divided by the size of the tax office district (km^2). The JOB VACANCY RATE is the vacancy rate of the tax office in 2016. WORK EFFORT is the number of incoming cases and incomplete cases from the previous year (in 1,000) divided by the NUMBER OF EMPLOYEES. ASSESSMENT SELF-EMPLOYMENT is a dummy variable indicating a tax office for self-employed taxpayers. The data for SATISFACTION WORK, SATISFACTION POL. FRAMEWORK, SATISFACTION EXT. CONDITIONS, POPULATION DENSITY, and the JOB VACANCY RATE are available only in aggregated form for the tax office of self-employment assessment and the tax office for the assessment of others. The DEA is based on the 132 tax offices that have at least 20 observations of the SIQ of compliant taxpayers to obtain a valid mean value. Robust standard errors are in parentheses. *, **, *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively (based on two-tailed tests).

2.5.3 Robustness Checks

We next subject our analysis to a set of robustness tests. First, to address any potential unobserved heterogeneity between tax offices, we conduct a robustness check by adding tax office fixed effects to the IV models. Doing so eliminates all variables that do not vary within one tax office. The results remain qualitatively unchanged. Second, we address the concern that taxpayers are clustered in tax offices. In our baseline regression, we take this fact into account by using robust standard errors clustered at the tax office level. We control our logit regression results using multilevel models with random intercepts for the tax offices. The results remain qualitatively unchanged. Third, we vary the definitions of our sample splits. We capture taxpayers with low ability using a stricter definition, i.e., a taxpayer who assesses his/her tax knowledge not as good or expert, using no external help (i.e., a tax advisor or tax assistance association) and having no university degree. The results remain robust. However, moving from the lower quartile to the upper quartile of SIQ_COMP, we find that the increase in the average predicted probability of being compliant rises by 1.8 to 2.5 percentage points. For taxpayers with low ability and high tax morale, the average predicted probability of being compliant increases by 2.0 to 2.7 percentage points after adjusting the sample split definition. Furthermore, we modify the definition of ABILITY. We now define a taxpayer as having

low ability if he/she assesses his/her tax knowledge not as good or expert. In a second modification, low ability captures only taxpayers using no external help.[18] All reported results remain qualitatively unchanged except for the IV regression in which the effect for taxpayers with no external help is only significant at the 13.2 percent level. An increase from the lower to the upper quartile of SIQ increases the average predicted probability of being compliant by 1.6 to 1.8 (1.5 to 1.6) for taxpayers who assess their tax knowledge not as good or expert (taxpayers using no external help). The corresponding increase in the average predicted probability for taxpayers who assess their tax knowledge not as good or expert (taxpayers using no external help) and high tax morale amounts to 1.9 to 2.2 (1.5 to 1.6). We also change the definition of tax morale and now define a taxpayer as having high morale if he/she fully agrees that it is good that the tax authorities are buying tax CDs to combat tax evasion (tax CDs are electronic data sets that contain whistleblower information, e.g., bank accounts of potential evaders). The results remain qualitatively unchanged. Finally, we adjust our definition of high coercive power. In particular, we perform a split of high coercive power at the median and the 66.6 percent quartile instead of at the upper quartile. The results remain robust, except that in the logit regression, if splitting at the median, the effect of SIQ_COMP is significant only at the 14.5 percent level. As we define coercive power less strictly, the average predicted probability of being compliant rises by only 1.3 to 1.6 percentage points if moving from the lower quartile to the upper quartile of SIQ_COMP using tax offices with coercive power above the median. Fourth, we test whether the effect of SIQ is driven only by a particular component of service. Instead of using our aggregated variable SIQ or SIQ_COMP, we separately use the four items that constitute our previous variable. Taking the IV approach (models (1) to (7)), we find the same results as those in our baseline regressions for each of the four items. However, the results of the logit regressions (models (8) to (14)) suggest that the answers to the items "The employees conscientiously deal with my questions" and "The employees are friendly and courteous" are particularly important for a positive effect of perceived SIQ between tax offices. Similarly, we use the four SIQ questions separately to explain the likelihood of an appeal. We find significant effects for all four SIQ questions. According to this, the result is not driven by one type of service in particular. Moreover, in a further analysis, we replace our SIQ variable with the question "The explanations for the deviations in

[18] We use the respective other characteristic as control variable.

the tax assessment notice were understandable". In line with our interpretation that the provision of information services is the main driver of observed taxpayer responses, we find evidence that the comprehensibility of the explanations for the deviations is negatively associated with the probability of appeal.

2.6 Conclusion

Prior research considers service quality, as perceived by customers, to be one of the most important nonfinancial performance measures in service organizations (Snipes et al. 2005). Tax administrations worldwide are now trying to improve the tax compliance of taxpayers through better service provision. Our paper informs taxpayers and tax administrations of the effectiveness of the service paradigm in preventing tax noncompliance.

We provide evidence that this approach can indeed be effective. We find that by increasing SIQ from the lower quartile to the upper quartile, the absolute number of incorrect tax returns could be reduced by approximately 6.7 percent. The positive effect of SIQ is particularly due to the subgroup of taxpayers with low tax literacy and high tax morale. Our interpretation is that these taxpayers make unintentional mistakes that can be avoided by providing tax agency information services to them. Moreover, we find a negative association of SIQ with the likelihood of making an appeal. By increasing SIQ from the lower quartile to the upper quartile, the average predicted probability of an appeal could be decreased by 2.5 to 3.8 percentage points, which corresponds to a 4.8 percent to 7.3 percent reduction in the total number of appeals. Thus, improving tax agency services not only improves tax compliance but also may help to reduce tax controversy costs. Finally, our results shed light on a potential channel through which tax agencies might improve their SIQ given a specific number of tax officers. In particular, our results suggest that tax officers' satisfaction with their work is positively associated with the SIQ perceived by taxpayers. Thus, by improving tax officers' work satisfaction, tax agencies might indirectly also improve tax compliance and reduce tax controversy costs.

Regarding the limitations of our study, we would like to highlight that we measure the effect of SIQ on the probability of being fully tax compliant. It would be interesting if future research could use data that make it possible to also measure the effect on the extent of noncompliance. Moreover, our analysis is restricted to cross-sectional differences in perceived SIQ. Thus, future research should extend our research using longitu-

dinal data. In addition, while we show that SIQ improves tax compliance, future research should examine which drivers of the positive effect of service provision on tax compliance (e.g., telephone and electronic availability, the quality of website information, or electronic filing support) are the most important. Furthermore, we cannot make a statement about the potential net benefit of better services as we do not have information about the costs imposed by the service paradigm. It would be interesting to have access to this information for further research. Finally, our study concerns individual tax compliance. As decisions in firms are made by individuals, our results should also be descriptive for firms, especially small firms (OECD 2019b). However, in large firms, the professionalism of the firms' tax control framework may shape the relation to the tax authority, and other aspects of agency services beyond information services may be more important (e.g., offering cooperative tax compliance programs).

2.7 Appendix A: Additional Tables

Table 2.6: Variable Measurement

Variable	Description
Dependent variable:	
COMPLIANCE	A binary variable that equals one (zero) if the answer to the question "I had a deviation from my tax return" is no (yes).
Independent variables:	
SIQ	Service interaction quality. The rotated component score is calculated by a PCA that includes the following questions: (1) "The employees are friendly and courteous"; (2) "The employees are helpful and support me"; (3) "The employees in my tax office are professionally competent"; and (4) "The employees conscientiously deal with my questions" (1: "I completely disagree" to 5: "I fully agree"). We apply varimax rotation using Kaiser normalization and limit the number of components to those that have eigenvalues greater than one.
SIQ_COMP	The average rotated component score per tax office calculated by a PCA that includes the following questions: (1) "The employees are friendly and courteous"; (2) "The employees are helpful and support me"; (3) "The employees in my tax office are professionally competent"; and (4) "The employees conscientiously deal with my questions" (1: "I completely disagree" to 5: "I fully agree"). The PCA is performed only for those who did not have an adjustment, i.e., who are compliant.
ABILITY	A dummy variable indicating whether the taxpayer has the ability to check the correctness of his/her tax assessment, that is, whether the taxpayer assesses his/her tax knowledge as *good* or *expert* or uses external help (i.e., a tax advisor or tax assistance association).
MORALE	Tax morale. A variable indicating tax morale, i.e., whether the taxpayer thinks that people should be honest about their taxes (1: "I completely disagree" to 5: "I fully agree").
COERCIVE	Coercive power. The average perceived deterrence per tax office as of 2012 measured by the question "My tax office is petty when it comes to checking my tax return(s)" (1: "I completely disagree" to 5: "I fully agree").
Instruments	
APPRECIATION	Measures the individual's appreciation of tax information provision. It is the answer to the question whether the taxpayer agrees with the following statement: "I think it is a good thing that my personal tax burden [average tax rate] is mentioned in my income tax assessment notes" (1: completely disagree, 5: fully agree).
JOB SATISFACTION	Measures the average satisfaction of tax office employees with their work activities. The averages are calculated based on the question "I am satisfied with my job (the content of the activities, the type of work tasks and the workload)" (1: completely disagree, 5: fully agree).
Control variables	
CONSCIOUS	Measures how consciously or thoroughly the tax office works. It is defined as one minus the percentage of granted appeals. Since the tax assessment office and the tax appeal office can grant appeals, we use the average of both offices.
UNDER 36	A dummy variable indicating whether the taxpayer is younger than 36 years old.
36 TO 45	A dummy variable indicating whether the taxpayer is between 36 and 45 years old.
46 TO 55	A dummy variable indicating whether the taxpayer is between 46 and 55 years old.
56 TO 65	A dummy variable indicating whether the taxpayer is between 56 and 65 years old.
MALE	A dummy variable indicating whether the taxpayer is male.
30,000 TO 50,000	A dummy variable indicating whether the taxpayer has a gross annual income of more than 30,000 euro to 50,000 euro (45,000 euro to 75,000 euro for joint filers).
50,000 TO 70,000	A dummy variable indicating whether the taxpayer has a gross annual income of more than 50,000 euro to 70,000 euro (75,000 euro to 105,000 euro for joint filers).
OVER 70,000	A dummy variable indicating whether the taxpayer has a gross annual income over 70,000 euro (105,000 euro for joint filers).
EDUCATION	A dummy variable indicating whether the taxpayer has at least a university entrance qualification *(Abitur)*.
JOINTLY	A dummy variable indicating whether the taxpayer files jointly.
ELECTRONIC	A dummy variable indicating whether the taxpayer submitted the tax return with the ELSTER program, ELSTER forms or software (it is assumed that every taxpayer who hires external help submits electronically).
INVESTMENT INC	A dummy variable indicating whether the taxpayer has income from capital assets or real estate.
SELF-EMPLOYED	A dummy variable indicating whether the taxpayer has income from self-employment.
Sample splitting	
HIGH_MORALE	A dummy variable indicating whether the taxpayer fully agrees that people should be honest about their taxes.
HIGH_COERCIVE	A dummy variable indicating that the coercive power of the tax office is above the upper quartile (the upper quartile is calculated based on 208 tax office observations).

Table 2.6: Variable Measurement (continued)

Notes: We have data from 208 tax offices (tax offices for the assessment of self-employed taxpayers and tax offices for the assessment of others). However, the variables COERCIVE and JOB SATISFACTION are available only in aggregated form for the tax office of self-employment assessment and the tax office for the assessment of others. The variable SIQ_COMP is calculated only if the tax office has at least 20 observations of the SIQ of compliant taxpayers. The number of tax offices decreases from 208 to 132.

Table 2.7: Descriptive Statistics (Additional Analyses)

Variables	N	mean	sd	min	p50	max
APPEAL	6,543	0.35	0.48	0.00	0.00	1.00
NUMBER OF EMPLOYEES	132	26.72	11.19	8.84	24.46	61.20
FRIENDLY_COMP	132	4.06	0.19	3.36	4.05	4.45
HELPFUL_COMP	132	3.94	0.20	3.19	3.96	4.36
COMPETENT_COMP	132	4.01	0.18	3.34	4.02	4.40
QUESTIONS_COMP	132	3.89	0.21	3.08	3.91	4.27
Θ_{SIQ}	132	0.94	0.04	0.77	0.94	1.00
SATISFACTION WORK	132	-0.02	1.01	-3.21	-0.11	2.39
SATISFACTION POL. FRAMEWORK	132	0.01	1.01	-2.39	-0.01	1.99
SATISFACTION EXT. CONDITIONS	132	-0.06	1.05	-2.59	0.03	3.24
POPULATION DENSITY	132	1.54	1.98	0.10	0.75	12.75
JOB VACANCY RATE	132	0.93	0.05	0.81	0.93	1.06
WORK EFFORT	132	2.74	0.91	1.02	2.82	4.96
ASSESSMENT SELF-EMPLOYMENT	132	0.23	0.42	0.00	0.00	1.00

Notes: APPEAL is a dummy variable indicating whether the taxpayer filed an appeal. NUMBER OF EMPLOYEES refers to full-time equivalents in 2016. FRIENDLY_COMP measures the average rating of the tax office based on the question "The employees are friendly and courteous" (1: "I completely disagree" to 5: "I fully agree") based on the evaluations of taxpayers whose tax assessments do not deviate from their tax return. HELPFUL_COMP measures the average rating of the tax office based on the question "The employees are helpful and support me" (1: "I completely disagree" to 5: "I fully agree") based on the evaluations of taxpayers whose tax assessments do not deviate from their tax return. COMPETENT _COMP measures the average rating of the tax office based on the question: "The employees in my tax office are professionally competent" (1: "I completely disagree" to 5: "I fully agree") based on the evaluations of taxpayers whose tax assessments do not deviate from their tax return. QUESTIONS_COMP measures the average rating of the tax office based on the question "The employees conscientiously deal with my questions" (1: "I completely disagree" to 5: "I fully agree") based on the evaluations of taxpayers whose tax assessments do not deviate from their tax return. Θ_{SIQ} is the measure of service performance calculated by the DEA. The number of employees is used as the input variable, and the variables FRIENDLY_COMP, HELPFUL_COMP, COMPETENT_COMP and QUESTIONS_COMP are used as output variables. SATISFACTION WORK measures the satisfaction of tax office employees with their work and career opportunities (rotated component calculated by PCA). SATISFACTION POL. FRAMEWORK measures the satisfaction of tax office employees with the conditions of the political framework (rotated component calculated by PCA). SATISFACTION EXT. CONDITIONS measures the satisfaction of tax office employees with their external conditions (payment, working hours, IT equipment, office space, working environment, aids, technical equipment, noise, temperature, etc.) (rotated component calculated by PCA). POPULATION DENSITY is the number of inhabitants (in 1,000) divided by the size of the tax office district (km^2). The JOB VACANCY RATE is the vacancy rate of the tax office in 2016. WORK EFFORT is the number of incoming cases and incomplete cases from the previous year (in 1,000) divided by the NUMBER OF EMPLOYEES. ASSESSMENT SELF-EMPLOYMENT is a dummy variable indicating a tax office for self-employed taxpayers. The number of observations of APPEAL represents all taxpayers who have an adjustment and filed an appeal or not. The other variables are measured at the tax office level.

Table 2.8: Component Loadings

	Scale item	RC1	RC2	RC3
(1)	I am satisfied with the way in which the political leaders are promoting the reputation and positive image of the tax authorities among the public.		0.8630	
(2)	I feel valued in my work by the political leaders.		0.7981	
(3)	Politics arouses certain expectations among citizens via the media (e.g., advisory services, processing time). I can live up to these expectations in my daily work.		0.7692	
(4)	I feel burdened by the development of tax laws.		-0.7321	
(5)	I consider our weekly working hours to be acceptable, particularly in view of the working hours in the private sector and the overall situation in the labor market.			0.5766
(6)	I am satisfied with my work (the content of the activity, the nature of the work tasks and the workload).	0.6776		
(7)	The work situation in my job promotes my motivation and my commitment.	0.7363		
(8)	I am satisfied with the cooperation with colleagues in my immediate environment.	0.6654		
(9)	The IT equipment of my workplace (devices and applications) usefully supports me in fulfilling my tasks.			0.7666
(10)	I am satisfied with the conditions under which I work (office space, the working environment, aids, technical equipment, noise, temperature, etc.).			0.8028
(11)	I am satisfied with the organization, the working procedures and the cooperation of the subdepartments within our department.	0.8195		
(12)	I am satisfied with the information flow (volume, timeliness and quality) and the communication culture within our department.	0.8484		
(13)	I am satisfied with the leadership behavior of my immediate superior.	0.7774		
(14)	I am satisfied with the managers in my department in terms of cooperation, competence and the reliability of their management decisions.	0.8895		
(15)	I am satisfied with the implementation (initiation and execution) of changes and with the way I am involved.	0.8893		
(16)	I feel that I belong to my office and overall feel good here.	0.7833		
(17)	I was satisfied with my employee interview.	0.7148		
(18)	I am satisfied with how the results of the last staff survey were handled in my department.	0.7432		
	Eigenvalues (unrotated)	9.07	2.20	1.40
	Cronbach's α (scale reliability coefficient)	0.95	0.88	0.68

Notes: Only component loadings greater than 0.5 are displayed. The Cronbach's α is calculated with standardized items. All items were measured using a five-point Likert scale (from 1: "I completely disagree" to 5: "I fully agree"). The data measure the average employee satisfaction per tax office. However, the data on employee satisfaction are available only in aggregated form for the tax office of self-employment assessment and the tax office for the assessment of others. Thus, the PCA is based on 104 observations. We deleted items with cross-loadings <0.2.

Chapter 3

Public Disclosure of Tax Strategies and Firm's Actual Tax Policy

For copyright reasons this chapter is not available in this published version. This paper was published as Blaufus, K., Jarzembski, J. K., Reineke, J., & Trenn, I. 2023. Public Disclosure of Tax Strategies and Firm's Actual Tax Policy. *Journal of Accounting, Auditing & Finance*. https://doi.org/10.1177/0148558X231200913

Chapter 4

The Effect of Corporate Social Responsibility on Reputational Costs of News about Tax Planning

Abstract

A fundamental question in tax research is to understand the costs of corporate tax planning. While many managers fear reputational damage from tax planning, research shows mixed results. This study aims to investigate whether corporate social responsibility (CSR) provides insurance-like protection for a firm's reputation or, on the contrary, magnifies the customers' reactions to corporate tax planning. Using tax news data from Lee et al. (2021), I analyze (1) whether the revelation of tax planning behavior affects firms' sales or advertising expenses and (2) whether this relationship is moderated by CSR performance. I hypothesize that reputational damage increases with CSR performance. The regression analysis, however, does not find support that firms' sales or advertising expenses are related to news about firms' tax planning regardless of the CSR performance. The results, based on recent tax planning data, contribute to the literature by providing insights into how CSR affects the relationship between tax news and sales and advertising expenses, thus addressing the customer point of view.

4.1 Introduction

A firm's reputation is considered one of the most valued of assets of a company (Barnett and Pollock 2012, 4; Rindova and Martins 2012, 24 ff.).[1] As a consequence, reputational risks are a key corporate concern and are considered in the corporate tax risk management (PwC 2004; Neuman et al. 2020). A positive reputation is an important intangible assets as it provides the firm with benefits like "premium prices for products, lower costs of capital and labor, improved loyalty from employees, greater latitude in decision making, and a cushion of goodwill when crises hit" (Fombrun 1996, 57). However, a firm's reputation is not a static good (Rhee and Valdez 2009). While it takes a long time to build a strong reputation, it can easy be damaged by negative events, making building and maintaining a positive reputation an important corporate task (Rhee and Valdez 2009).

A potential threat to a firm's reputation is its approach to taxes. The public is increasingly focused on whether firms pay their 'fair share' of taxes (EY 2017). If a firm engages in tax planning, it risks being labeled a "poor corporate citizen" (Hanlon and Slemrod 2009, 127).[2] A recent survey conducted by EY revealed that 89% of the largest firms are "somewhat or significantly concerned about media coverage of taxes" (EY 2014, 6). As a result, many companies focus on how to manage their public tax profile and communicate tax-related information (EY 2014, 7). Austin and Wilson (2017) find that firms with valuable brands have higher effective tax rates (ETRs), i.e., engage in less tax avoidance. Likewise, Dyreng et al. (2016) find a decrease in tax avoidance for noncompliant firms following the public pressure to disclose subsidiary locations. In line with these findings, Graham et al. (2014) report that 69% of the surveyed corporate tax executives assess reputation concerns as "important" or "very important" reason to not engage in tax planning. These results suggest that managers act as if they (ex ante) believe that avoiding taxes causes reputational damage (Austin and Wilson 2017). However, are these concerns justified? Empirical evidence on (ex post) reputational costs of activities that aim to minimize the tax burden is mixed (e.g., Hanlon and Slemrod 2009; Gallemore et al. 2014; Hardeck and Hertl 2014; Antonetti and Anesa 2017; Lee et al. 2021). Furthermore, factors such as the increasing tax competition

[1] There is no generally accepted definition of reputation. Gallemore et al. (2014), for example, define reputation as "a general perception of the firm by all interested stakeholders" (Gallemore et al. 2014, 1104).

[2] For the purpose of this study, the term *tax planning* encompasses illegal tax evasion as well as legal tax avoidance activities or the involvement in tax havens.

(Dyreng et al. 2017; Thomsen and Watrin 2018), the relevance of social media (Wilde and Wilson 2018), or changing market perceptions (EY 2017) call into question the generalizability of previous findings.

Although empirical evidence on reputational costs of tax planning is mixed, real-life examples, like Starbucks or Burger King, illustrate that the media regularly reports about negative consumer reactions and boycotts (Escobales and McVeigh 2012; Puzzanghera 2014). But do all firms face the same risk of consumer backlash? Is there a tool that enables firms to hedge against reputational risks? Or contrary, do some factors magnify the response to tax planning? To explain the extreme consumer reaction in the case of Starbucks, Campbell and Helleloid (2016) argue:

> *"While the tax avoidance practices Starbucks used were common among multinational companies, Starbucks had been very public in its commitment to being socially responsible and a good citizen of the communities in which it operated. […] Thus, its critics found it easy to point out that not paying its fair share of taxes was inconsistent with the image Starbucks was portraying to consumers". (Campbell and Helleloid 2016, 38)*

The question arises whether a firm's corporate social responsibility (CSR) performance affects the level of reputational costs of tax planning.

The role and importance of CSR has expanded in recent years. Public interest increased steadily and firms were urged to incorporate socially responsible business practices into their corporate strategies. The sustainability reporting rate for the world's 250 largest companies by revenue increased from 35% in 1999 to 96% in 2020 (KPMG 2020). Moreover, the expenses for CSR activities are rising. At the beginning of 2020, U.S. assets under professional management using sustainable, responsible and impact (SRI)-strategies amounted to USD 17,1 trillion which corresponds to an 25-fold increase since 1995 (Social Investment Forum 2020). At the same time, research shows that CSR performance is related to lower analysts forecast errors, increasing analyst coverage, lower cost of equity capital (Dhaliwal et al. 2011; Dhaliwal et al. 2012; Dhaliwal et al. 2014), less earnings smoothing and earnings loss avoidance (Chih et al. 2008), lower likelihood of earnings management (Hong and Andersen 2011; Kim et al. 2012), lower information asymmetry (Cho et al. 2013), or lower idiosyncratic risk (Lee and Faff 2009). But can CSR activities also protect the firm's reputation?

Reasons can be given both for why CSR performance could exacerbate and for why it could moderate reputational costs of tax planning. On the one hand, socially responsible business practices can act as a "risk management tool" (Shiu and Yang 2017, 456) by providing an 'insurance-like protection' (Godfrey 2005).[3] The philanthropic activities build moral capital which protects the relationship between stakeholder and firm against a decrease in economic value in case of adverse events (Godfrey 2005; Godfrey et al. 2009). Based on the concept of reciprocity, stakeholders are good to the company if the company is good to the stakeholders in doing social good (Lins et al. 2017). Moreover, CSR performance could influence the stakeholders' perceptions of the firm's intention to engage in tax planning (Godfrey 2005; Godfrey et al. 2009). On the other hand, there are arguments why CSR activities exacerbate reputational costs of tax planning. Inger and Vansant (2019) find support that engaging simultaneously in tax avoidance and CSR is interpreted as inconsistent with one another. Socially responsible business practices raise the expectation that the company will act (truthfully) in accordance with ethical standards and norms and do social good including acting in the spirit of the law. As a consequence, news about tax planning might be a bigger surprise, and thus magnify reputational costs (Bartov et al. 2021). Furthermore, engaging in tax planning while performing in high CSR can be interpreted as an ingratiating attempt to earn favor (i.e., hypocrisy) which creates negative moral capital (Godfrey 2005).

Thus, it remains an open question (1) whether tax planning leads to reputational costs and (2) whether CSR performance affects the link between tax planning and reputational costs. Based on prior research results, I hypothesize that news about tax planning is negatively related to reputational costs and that this response is exacerbated by CSR performance. Although there are also arguments for an opposite effect of CSR performance, I predict that the *negative surprise effect* dominates the reactions because, unlike many other adverse events, tax planning is usually a deliberate act that cannot be treated as a "one-off event beyond management control" (Bartov et al. 2021, 82).

In contrast to other studies that measure reputational effects by attitudes or intentions, the following research aims to capture reputational costs through real economic consequences (i.e., financial reputational effects). For this purpose, firms' sales and advertising expenses are analyzed. The goal is to investigate whether tax planning causes customers to boycott firms and stop purchasing their products, leading to increased adver-

[3] Empirical support is provided by several studies (e.g., Minor and Morgan 2011; Christensen 2016; Lins et al. 2017; Shiu and Yang 2017; Wans 2020; Bartov et al. 2021).

tising expenses. Thus, this study focuses on reputational costs exerted by costumers. Following a survey of PwC, consumers and clients are the "most influential stakeholder group" (PwC 2013, 7). To identify firms that engage in tax planning, news data on tax planning activities from S&P 500 firms between 2008 and 2017 from Lee et al. (2021) are used. News data has the advantage of being unpredictable, easy to understand and reaching a broad audience. Lee et al. (2021) include the terms "*tax avoidance*", "*tax evasion*", and "*tax haven*" in their news research (see Lee et al. 2021, Appendix C). Consequently, this study covers both legal and illegal tax planning activities. In the following, I simply refer to the data as tax news or 'tax avoidance' news. The CSR performance is captured by the environment, social, and governance (ESG) score of Thomsen Reuters. Financial data is collected from Compustat.

Due to the staggered design of the tax planning revelations, I use a generalized difference-in-differences regression design to analyze my hypotheses (see also Lee et al. 2021). Although many managers fear reputational costs of tax planning, the results of this study do not show that news about tax planning behavior is related to a decrease in sales or an increase in advertising expenses. Lee et al. (2021), on the other hand, uses the same tax news data and find support that employee perceptions of managers and firms are negatively affected by tax news. The question arises whether tax news merely affects perceptions but does not result in behavioral change, or whether employees react differently than customers. The results of this study are consistent with the findings of Gallemore et al. (2014). The authors find no changes in sales, sales growth, or advertising expenses for firms whose tax shelter was revealed between 1995 and 2005 compared to the control firms. Using more recent data, my findings indicate that the effects have not changed. However, it should be noted that the present study covers a broader spectrum of tax planning (i.e., tax avoidance, tax havens, and tax evasions) than the research of Gallemore et al. (2014). Because my research data includes large U.S. firms, the results may also be affected by the fact that the public already widely believes that these firms are not paying their 'fair share' of taxes (Hoopes et al. 2018). Correspondingly, the findings of Hoopes et al. (2018) indicate that consumer sentiment in the context of tax disclosure in Australia is more fragile for smaller firms. Similarily, Brooks et al. (2016) find that the negative reactions to negative news about corporate tax payments are more pronounced for smaller firms. Furthermore, this study cannot provide support that the relationship between tax news and reputational costs is moderated by

the level of CSR performance. Thus, neither 'the insurance-like protection theory' nor the 'negative surprise theory' is supported in the context of costumers' reactions to tax planning.[4] My findings are validated by a number of robustness tests. An explanations could be that stakeholders, just like firms, may decouple CSR commitment and tax policy, as firms regularly fail to incorporate CSR commitments into all areas of the business (Mayberry and Watson 2021). In addition, there is the question of the extent to which customers are aware of the level of CSR performance of firms and whether they are able to assess it. Nonetheless, I carefully acknowledge that the non-results of this study could also be caused by a low power of the empirical tests. Although this study is not able to empirically identify reputational costs of tax planning, reputational effects may exist.

This study contributes to the research on reputational costs of tax planning in the following ways: First, as pointed out by Austin and Wilson (2017), there is a gap in literature as research suggests that managers act as if they believe tax avoidance causes reputational damage, however, empirical evidence on (ex post) reputational costs is ambiguous. The following study adds to this stream of literature by analyzing reputational costs of tax planning using recent data that fit to the era of social media, thus following the call of Wilde and Wilson (2018) for future research using new data. By using the same variables to capture reputational costs as Gallemore et al. (2014) – whose last tax shelter revelation is from 2005 – this study provides insight into whether the results have changed over time.[5] Although there have been recent studies on consumer responses to tax avoidance, these studies differ from the present one in terms of research design (Hardeck and Hertl 2014; Antonetti and Anesa 2017; Hardeck et al. 2019) or the country studied (Hoopes et al. 2018). Second, this study contributes to understanding the role of CSR by analyzing whether costumers' responses to tax planning are sensitive to the level of CSR a firm engages in. Several prior studies focus on how CSR is related to tax avoidance (e.g., Hoi et al. 2013; Davis et al. 2016; Mayberry and Watson 2021; Davis et al. 2022). This study, on the other hand, focuses on the implications of managers' decisions about the level of tax planning and CSR they choose to undertake in the context of consumer reactions. The role of CSR performance in case of tax planning revelations is still unclear. While existing research examines the effects of CSR on interplay between

[4] I find weak support that the human rights score is negatively related to the link between tax news and advertising expenses. However, this effect cannot be supported by the other models.

[5] However, as mentioned before, Gallemore et al. (2014) analyze tax shelters, while this study covers a broader spectrum of tax planning (i.e., tax avoidance, tax evasion, and tax haven).

tax avoidance and stock price reactions or firm value[6] (Choy et al. 2017; Huang et al. 2017; Inger and Vansant 2019; Rudyanto and Pirzada 2020; Inger and Stekelberg 2022), the present study adds to this stream of literature by focusing on reputational costs of tax planning exerted by customers (i.e., decrease in sales revenues or increasing advertising expenses). Attitudes towards tax planning might differ among each group of stakeholders as they face different costs and benefits (Austin and Wilson 2017).

From a practical perspective, my results are informative for managers and their decision making process about the level of CSR performance they want to engage in.[7] This study provides useful information about the effectiveness of CSR performance as a tool to manage reputational damage of tax planning. Furthermore, this study is of interest for managers as it provides insights on whether they should indeed be concerned about negative media coverage with respect to their approach to taxes (i.e., the effects of 'public shaming' by the media).

The remainder of this study is organized as follows: Section 4.2 reviews the relevant literature and derives the hypotheses. Section 4.3 outlines the research design. Section 4.4 presents the main results, and section 4.5 discusses the robustness checks. Finally, section 4.6 concludes.

4.2 Literature Review and Hypotheses Development

4.2.1 Reputational Costs of Tax Planning

In recent years, reputational risks have become a key concern as policymakers, activist groups, and the media focus more on tax avoidance behavior and whether firms pay their 'fair share' of taxes (EY 2017). The global 2014 tax risk and controversy survey from EY revealed that 89% of the largest firms are "somewhat or significantly concerned about media coverage of taxes" (EY 2014, 6) compared to *only* 60% in 2011 (EY 2014, 6). Moreover, 65% of the firms reported that they "developed a more structured approach to managing their public tax profile" (EY 2014, 7) and 42% stated they have "changed the way they communicate tax-related information to external stakeholders such as the investment community" (EY 2014, 7). Likewise, Graham et al. (2014)

[6] Other studies on reputational costs of tax avoidance analyze whether the effect varies heterogeneously, but do not consider CSR performance as a moderator (e.g., Hanlon and Slemrod 2009; Gallemore et al. 2014; Blaufus et al. 2019b). Antonetti and Anesa (2017) and Hardeck et al. (2019) examine the role of perceived ethically or CSR perception as mediators for the link between reputational costs and tax avoidance.

[7] The potential costs of CSR are outside the scope of this study, but should be taken into account in terms of practical implications (Christensen 2016).

report that 69% of the surveyed corporate tax executives classify reputation concerns as "important" or "very important" reason to not engage in tax planning. In addition, approximately 58% rate the "risk of adverse media attention" (Graham et al. 2014, 1001) as "important" or "very important" reason. In line with these findings, Austin and Wilson (2017) find that firms with valuable brands have higher ETRs suggesting that managers act as if they believe that tax avoidance can lead to reputational costs. Dyreng et al. (2016) show that public pressure to disclose subsidiary locations led noncompliant firms to engage in less tax avoidance. In contrast, Chen et al. (2019) find no support that firms engage in less tax avoidance after tax news coverage. However, the authors point out that the managers may have already incorporated the costs of media coverage in the decision about the level of tax avoidance.

While the advantages of tax planning are obvious (i.e., reduced tax payments and higher after-tax earnings), negative consumer reactions and reputational damages are potential costs. Accordingly, reputational risks are also a component of corporate tax risks and considered in the tax risk management (PwC 2004; Neuman et al. 2020). The following arguments support the existence of reputational effects: First, prior research shows that tax avoidance is perceived as unethical and morally wrong (Antonetti and Anesa 2017; Christian Aid 2017; DeZoort et al. 2018) and even can be considered socially irresponsible (Avi-Yonah 2014; Dowling 2014; Hardeck et al. 2019). For example, 43% of the British consider boycotting products of firms that do not pay their 'fair share' of taxes (Christian Aid 2017). Research shows that costumers take tax planning behavior into account when making purchasing decisions (Hardeck and Hertl 2014; Antonetti and Anesa 2017; Christian Aid 2017; Hardeck et al. 2019). Second, aggressive tax behavior could create the impression that the firm has a tendency to dishonesty (Bosch and Eckard Jr 1991). The reliability of other financial statements could also be doubted (Hanlon and Slemrod 2009). As a result, distrust increases and costumers may switch (Bosch and Eckard Jr 1991). In addition to these arguments, it is not clear whether costumers derive any benefit from corporate tax avoidance (e.g., lower prices). Literature on who bears the corporate tax incidence provides mixed results (e.g., Harberger 1962; Gordon 1967; Baker et al. 2020; Jacob et al. 2022). Despite these arguments, it has to be kept in mind that, unlike tax evasion, tax avoidance and tax havens are usually legal tax planning strategies that do not violate the letter of the law even though they may be frowned upon. The aspect of legality may influence the stakehold-

ers' reactions (Gallemore et al. 2014; Blaufus et al. 2019b). Moreover, views on tax planning may vary between countries which differ with regard to public perceptions, religion, legislation, etc. (Kountouris and Remoundou 2013; Hardeck et al. 2019). However, prior studies show that even shareholder-oriented countries such as the United States (Bottenberg et al. 2017) have negative attitudes towards tax planning (e.g., Pew Research Center 2015; Antonetti and Anesa 2017; Lee et al. 2021).

Although prior research suggests that managers are concerned about reputational costs of tax planning, it is unclear whether these concerns are justified (Austin and Wilson 2017). The empirical evidence on whether tax planning actually leads to (ex post) reputational costs is mixed. Hanlon and Slemrod (2009) are one of the first authors to address this research question. The authors analyze stock price reactions after the revelation of tax shelter news and find evidence for a price decline. This reaction is more negative for firms in the retail sector and less negative for firms with relatively high cash ETRs. These findings indicate that the price decline may be caused by a consumer backlash and that the market reacts positively to the information that firms which have high ETR are more tax aggressive than assumed. Gallemore et al. (2014) expand the research of Hanlon and Slemrod (2009) by increasing the dataset of tax shelter revelations. They find that although tax shelter revelations are followed by negative abnormal returns, this effect is only temporally and reverses within 30 days. In addition, Gallemore et al. (2014) find no evidence for reputational effects of tax shelter revelations on corporate managers (i.e., chief executive officer (CEO) or chief financial officer (CFO) turnover), auditor turnover, customer behavior (i.e., sales, sales growth, advertising expense, advertising expense growth, media reputation), or with the tax authority (i.e., change in ETR). This finding does not change even in subsamples of high-reputation firms or firms for which the tax shelter revelation should be a surprise (measured by the likelihood of being in a tax shelter). However, the authors are cautious in interpreting their results. They point out that there may be a reputational effect, but their empirical analysis is not able to identify it.

Like Gallemore et al. (2014), many researchers address the reputational costs exerted by *customers*. Hoopes et al. (2018) investigate the consumer sentiment after the public disclosure of tax return information in Australia using survey data from YouGov and TurkPrime. The authors find a small reduction in consumer sentiment for Australian-owned private companies but not for large foreign-owned companies and Australian

public companies. The findings suggest that global brands of large public companies are more resistant. Using an experimental setting, Hardeck and Hertl (2014) find that aggressive tax strategies negatively influence a firm's reputation as well as the consumers' purchase intentions. Responsible tax strategies, on the other hand, have positive effects. However, while consumers have lower willingness to pay for aggressive tax strategies, they are not willing to pay a price premium for responsible tax strategies. Furthermore, the authors provide some evidence that the tax morale and attitude towards tax avoidance are moderating variables. In line with these results, Antonetti and Anesa (2017) find that consumers react negatively to aggressive corporate tax strategies. The authors show that this association is mediated by consumers' perceived corporate ethicality. Likewise, Hardeck et al. (2019) find evidence that the link between the tax strategy and the consumers' attitudes towards the firm and the willingness to pay is mediated by the CSR perceptions. Further experimental support for the negative effect of tax planning on consumer purchase behavior is provided by Asay et al. (2018). In contrast to these results, Baudot et al. (2020) find no link between reputation and firms' tax behavior.

Nonetheless, reputational costs can also be exerted by other parties. Lee et al. (2021) find that employee perceptions of managers and firms – measured by Glassdoor.com ratings – are negatively affected by tax avoidance news. The negative reaction is larger in consumer-facing industries and smaller for well-performing firms. Lee et al. (2021) show that not only the firm itself but also the managers are exposed to reputational costs. Besides negative employee perceptions, managers can face negative board responses. While Gallemore et al. (2014) find no evidence that managers bear reputational costs (i.e., CEO-, CFO-, auditor turnover), Chyz and Gaertner (2018) show that forced CEO turnovers are more likely for firms with low tax rates (for periods after the Sarbanes-Oxley Act) or high tax rates relative to its peers. In contrast to prior studies, the authors are able to differentiate between forced and unforced turnovers. The analysis demonstrates that CEOs bear (individual) reputational costs when they avoid a lot of taxes but also when they pay too much tax. In line with the latter, Lanis et al. (2019) find that engaging in corporate tax avoidance positively affects board and CEO reputation (measured by outside directorships held).

Like Gallemore et al. (2014) and Hanlon and Slemrod (2009), other studies measure reputational costs by analyzing the firm value and abnormal returns. These studies relate to the question of whether the market anticipates reputational consequences and reacts.

However, it should be noted that negative returns can also be influenced by determinants other than reputational costs (Hanlon and Slemrod 2009; Gallemore et al. 2014). Huang et al. (2017), Cloyd et al. (2003), and Desai and Hines (2002) investigate the market reactions to corporate inversions, a tax strategy that is legal but considered aggressive. Both Huang et al. (2017) and Desai and Hines (2002) find (on average) positive market reactions to inversion announcements. Nevertheless, Desai and Hines (2002) show that not all stock price reactions were positive. In contrast, Cloyd et al. (2003) find that the average return based on 20 single-company inversion announcements is negative but does not significantly differ from zero. However, also these authors find different reactions for the separate inversion announcements (positive significant, negative significant or insignificant). In contrast to these studies, Desai and Dharmapala (2009) use book-tax differences to capture tax avoidance and analyze the effect on firm value. The authors find a positive, yet not significant, effect. The effect, however, is significant positive for high levels of institutional ownership. Thus, Desai and Dharmapala (2009) conclude that agency problems diminish the shareholders' benefits of tax avoidance. Several more studies analyze the relationship between tax rate measures and firm value and provide mixed results (Wahab and Holland 2012; Inger and Vansant 2019; Rudyanto and Pirzada 2020; Inger and Stekelberg 2022). Another method of capturing tax avoidance is through news data. Brooks et al. (2016) find no clear evidence for a relationship between tax rates and stock returns for firms of the Financial Times Stock Exchange (FTSE) All-Share Index. However, the authors find (small) negative returns for firms with bad news coverage regarding their tax payments. This finding is in line with Lee et al. (2021), who find that employee perceptions are affected by tax avoidance news but not by the cash ETRs, suggesting that the public is more likely to learn about tax avoidance behavior through the media than through information from financial statements. Choy et al. (2017) and Dyreng et al. (2016) both use the release of the ActionAid report in 2011, thus protest periods, to examine the effect of tax avoidance. The report increased public scrutiny with respect to the firms' subsidiary locations. Choy et al. (2017) provide evidence that the market reacted negatively to the ActionAid report. The stock prices of nonfinancial companies of the FTSE 100 recorded an abnormal decline of 0.9%. The negative response was larger for well-governed firms and firms with a high presence in tax havens. The authors also examined whether reputation, among others, was a channel through which the event lowered the

firm value. They find that firms with high overall social responsibility or high reputation within the general community reacted more negatively to the release of the report suggesting that the ActionAid report caused reputational damage. Dyreng et al. (2016) complement the research of Choy et al. (2017) and find that FTSE 100 firms that did not initially disclose their full subsidiary list had lower returns around the release date than other firms in the market, so that the negative market responses were especially concentrated around them. Similarly, Dhaliwal et al. (2022) find evidence that firms that engage in more tax avoidance are valued less during the protest period of 2011. Using tax news data, Blaufus et al. (2019b) demonstrate that market responses to tax planning behavior depend of the type of tax planning. Using listed German firms, the authors find negative abnormal returns when news about illegal tax strategies (i.e., tax evasion) is released; however, in case of news about legal tax strategies (i.e., tax avoidance), they find positive market reactions when the tax risk is low. Moreover, the authors find no evidence that reputational risks impact the stock price responses to tax avoidance news negatively. If tax avoidance news leads to reputational costs, one would expect that firms with high reputational risks respond more negatively.

This review of literature illustrates that results on the reputational costs of tax planning are mixed. In addition to that, results could be no longer suitable. For example, the last tax shelter revelation of Gallemore et al. (2014) is from 2005, thus, over 15 years ago. Tax competition has increased over the last years and has led to decreasing ETRs (Dyreng et al. 2017; Thomsen and Watrin 2018). Moreover, social media and smartphones revolutionized the speed, processing, and addressing of information, making it easier for media to disseminate information quickly to the desired audience (Wilde and Wilson 2018). PwC points out that due to today's rapid information delivery "Great damage can be done before a company has a chance to explain their position." (PwC 2013, 8). In addition to that, corporate taxation became more and more a public interest over the last years (EY 2017). Initiatives such as the Base Erosion and Profit Shifting (BEPS) project of the Organisation for Economic Co-operation and Development (OECD) or recent reporting and disclosure jurisdictions may have sensitized the population regarding their awareness and perception of tax planning practices and tax fairness. A poll conducted by Christian Aid revealed that 89% of British adults said that "Tax avoidance by large companies is morally wrong, even if it is legal" (Christian Aid 2017). Already at that time, Hanlon and Slemrod (2009) indicate the (potentially) lim-

ited generalizability of their results as market's perception of tax avoidance change in response to changing regulations. Following Wilde and Wilson (2018) and Austin and Wilson (2017) on future research topics, the present study aims to analyze whether managers' concerns about reputational costs of tax planning are justified using recent data that fit to the era of social media and rapid information processing by costumers. Based on the previously discussed arguments for negative costumer reactions, I state the following hypothesis:

H1: Firm sales (advertising expenses) are negatively (positively) related to news about firms' tax planning activities.

4.2.2 The Role of Corporate Social Responsibility

In the last decade, there has not only been much development in the topic of tax avoidance and tax planning but also the role of CSR has evolved.[8] Public interest in CSR increased steadily and forced companies to incorporate socially responsible business practices into their corporate strategies. As a consequence, investments in SRI-strategies (Social Investment Forum 2020) as well as the sustainability reporting rate (KPMG 2020) have increased drastically. The concept of CSR deals with the question of what responsibilities a firm has towards society. Views on this vary. Under the *agency thesis* investments in CSR activities are considered a waste of resources, because the only responsibility of a corporation is to increase profit and thus maximizing shareholder value (Friedman 1970; Bartov et al. 2021).[9] On the other hand, the *stakeholder theory* posits that the purpose of a company is to consider all important corporate constituencies (customers, employees, suppliers, government, etc.) and create wealth for all of them (Freeman 1984, 24 ff.; Clarkson 1995). This includes not prioritizing the shareholders at the cost of the others (Clarkson 1995).[10] Consistent with this view, CSR is often considered a firm's social obligation. According to Baudot et al. (2020), the ac-

[8] The literature provides several definitions of CSR. For example, Mackey et al. (2007) define CSR as "voluntary firm actions designed to improve social or environmental conditions" (Mackey et al. 2007, 818). Another description is provided by Carroll (1991), who posits four kinds of responsibilities for firms to be socially responsible: economic ("*Be profitable*"), legal ("*Obey the law*"), ethical ("*Be ethical*"), and philanthropic ("*Be a good corporate citizen*") (Carroll 1991, 42).

[9] CSR investments are seen a result of agency problems as these expenses are not devoted to increase the wealth of the shareholders but rather create benefits for the managers at the expense of the shareholders (Krüger 2015; Bartov et al. 2021). In line with this theory, Di Giuli and Kostovetsky (2014) find that the market react negatively (i.e., negative future stock returns) to increases in CSR ratings.

[10] To balance the competing demands of the different stakeholders, Jensen (2001) proposes the *enlightened* stakeholder theory and *enlightened* value maximization based on the same ideas of the stakeholder theory but adds the maximization of the long-term market value as objective function.

countability of a firm to society arises from the argument that its existence is based on the consent of society. As a result of this view, tax planning can be considered as inconsistent with CSR (Inger and Vansant 2019). Proponents refer to taxes as "the most fundamental way in which private and corporate citizens engage with broader society" (Christensen and Murphy 2004, 37) and consider it the social obligation of a firm to pay its 'fair share' of taxes and not engage in tax planning activities for the sole purpose of tax minimizing (Preuss 2012; Avi-Yonah 2014). In line with this view, the Global Reporting Initiative (GRI) recently added the corporate approach to tax to their sustainability reporting framework, thus considering tax payments as an important contribution to the society (Global Reporting Initiative 2019). Opponents, however, argue that tax payments are not the best means to act socially responsible because (1) tax payments are said to negatively affect corporate investment, entrepreneurship, innovation and job creation (e.g., Djankov et al. 2010; Davis et al. 2016) and (2) the private sector is said to more efficiently utilize resources than the government (e.g., McGee 2010).[11] Empirical results on the association between CSR and taxes, however, are mixed (Huseynov and Klamm 2012; Lanis and Richardson 2012; Hoi et al. 2013; Landry et al. 2013; Lanis and Richardson 2015; Muller and Kolk 2015; Watson 2015; Davis et al. 2016; Garcia 2016; Zeng 2016; Huang et al. 2017; Lanis and Richardson 2018; Mayberry and Watson 2021; Benlemlih et al. 2023).

While these studies analyze the association between taxes and CSR, the present study examines the implications of managers' decisions about the level of tax planning and CSR they choose to undertake in the context of consumer reactions. Based on the concept "Doing Well by Doing Good" (Bartov et al. 2021, 85), accounting research shows that CSR performance is related to lower analysts forecast errors, increasing analyst coverage, lower cost of equity capital (Dhaliwal et al. 2011; Dhaliwal et al. 2012; Dhaliwal et al. 2014), less earnings smoothing and earnings loss avoidance (Chih et al. 2008), lower likelihood of earnings management (Hong and Andersen 2011; Kim et al. 2012), lower information asymmetry (Cho et al. 2013), and lower idiosyncratic risk (Lee and Faff 2009). In line with these results, CSR can be considered as "consistent with maximizing shareholder wealth as well as achieving broader societal goals" (*good governance view*) (Ferrell et al. 2016, 586). But what role does CSR play in the reputational costs of tax planning? Can CSR act as a tool to mitigate the risk and thus maxim-

[11] See Preuss (2010) and Sikka (2010) for qualitative support that socially responsible firms also engage in tax planning activities.

izing the corporate value? Or contrary, are the reputational costs higher when firms engage in both tax planning and high CSR? Based on prior literature, this study derives the following two theories about the role of CSR in the context of reputational costs of tax planning.

CSR as Risk Management

CSR performance may serve as "risk management tool" (Shiu and Yang 2017, 456) in case of adverse events. Prior literature often refers to an 'insurance-like protection' or 'insurance policy' (e.g., Godfrey 2005; Godfrey et al. 2009; Christensen 2016; Lins et al. 2017; Shiu and Yang 2017; Bartov et al. 2021), because CSR performance provides value to the firm by protecting the relationship between stakeholder and firm against a decrease in economic value (Godfrey et al. 2009). Empirical evidence can be found for a variety of different negative events. Lins et al. (2017) show that the stock returns of high-CSR firms were four to seven percentage points higher during the 2008-2009 financial crisis compared to low-CSR firms. Shiu and Yang (2017) analyze whether CSR performance can protect the stock and bond prices against a variety of adverse events (e.g., fraud, lawsuits, labor rights, pollution, etc.). Their investigation shows that continuous and long-term engagement in CSR indeed provides an insurance-like protection and diminishes the loss in shareholder value; however, effects quickly disappear if the negative events occur repeatedly. Likewise, Christensen (2016) finds evidence that CSR reporting helps protecting the reputation, i.e., the stock price, in case of high-profile misconduct. If the firm issues a CSR report, the market reaction is less negative. Additionally, Minor and Morgan (2011) find that CSR performance can protect stock prices in case of product recalls. Two recent studies address this research question in the context of investors' reactions to financial restatements. Wans (2020) finds that investors' reactions to financial restatements are less negative if the company has a strong CSR performance. Moreover, more irresponsible CSR practices increase the likelihood of litigation measured by class action lawsuits. In contrast, Bartov et al. (2021) illustrate that the relationship is more complex. The authors find that the effect of CSR performance depends on the cause of the restatement announcement (inadvertent vs. fraudulent). The authors show that CSR performance magnifies the reaction for fraudulent restatement announcements, but on the other side, alleviates the reaction for inadvertent restatement announcements.

The theory behind the 'insurance-like protection' is that stakeholders assess, evaluate, and impute the firm's philanthropic activities and attribute moral capital to the firm, which in turn provides the insurance-like protection (Godfrey 2005). Godfrey (2005) argues that the moral capital influence the perception of the corporate's *mens rea*. When deciding about the punishment, stakeholders not only consider the act itself but also the mind or intention of the offender (Godfrey et al. 2009). In this context, CSR performance can, for example, persuade stakeholders that companies avoid taxes to do social well, i.e., use the tax savings for CSR activities (Rudyanto and Pirzada 2020). However, to provide 'insurance-like protection', the community must believe that the philanthropic activities are an honest manifestation of the company's intentions (Godfrey 2005). Based on the concept of reciprocity, stakeholders are good to the company if the company is good to the stakeholders by doing social good (Lins et al. 2017). In line with these arguments, one would expect that CSR performance moderates reputational costs of tax planning. Inger and Stekelberg (2022) provide evidence that socially responsible behavior enhances reputational capital in the context of tax avoidance. The authors find that in the electricity industry investors value other forms of tax avoidance more positively if the firm reduces its tax burden also in a socially responsible way. In line with this idea, Rudyanto and Pirzada (2020) find some evidence that sustainability reporting mitigates the negative effect of tax avoidance on firm value for non-environmentally sensitive firms listed on the Indonesian Stock Exchange.

CSR as Negative Surprise

The second theory is based on the idea of the stakeholder theory and assumes it is considered a firm's social obligation to pay its proper amount of taxes and not engage in tax planning activities for the sole purpose of tax minimizing (Preuss 2012; Avi-Yonah 2014; Huang et al. 2017). Correspondingly, tax avoidance is perceived as socially irresponsible behavior (Avi-Yonah 2014; Dowling 2014; Hardeck et al. 2019). Inger and Vansant (2019) find support that equity market participants view tax avoidance and CSR as inconsistent with one another when firms engage in both activities simultaneously. The authors find positive first-order effects of CSR performance and tax avoidance on the firm value; however, the interaction term is negative and significant. Thus, increasing the engagement in one of these activities lowers the value of the other, i.e., CSR and tax avoidance activities are counterproductive when simultaneously engaged. Likewise, Huang et al. (2017) find that equity investors react positively to inversion

announcements, but the reaction is less positive for firms with high CSR performance. In line with the stakeholder theory, the authors provide support that high-CSR firms that violate their social obligation to pay taxes suffer higher reputational costs. In addition, Choy et al. (2017) provide evidence that the market reacts more negatively to the release of the ActionAid report on the tax haven locations of the FTSE 100 in 2011 for firms with high overall social responsibility, which is used as a proxy for reputation.[12] The theory behind this *negative surprise effect* is that a high CSR performance creates expectations among stakeholders. They expect high-CSR firms to act (truthfully) in accordance with ethical standards and norms and do social good including acting in the spirit of the law. As the market expectations are higher, tax avoidance news should be a bigger surprise (Bartov et al. 2021). As a result, a high CSR performance would enhance costumers' reactions to tax avoidance. In line with this argument, Godfrey (2005) posits that philanthropic activities, which are seen as an ingratiating attempt to earn favor, create negative moral capital. If the philanthropic activities are not consistent with the firm's other moral behavior, an insurance-like protection is unlikely. A high CSR performance combined with tax planning activities can be seen as such an ingratiating attempt, i.e., hypocrisy. The loss of trust can cause stakeholders to no longer grant the company a "reputational premium" (Bartov et al. 2021, 86). Additionally, Shiu and Yang (2017) find that the insurance-like effects provided by CSR performance disappear if negative events occur repeatedly. Recurring violations, thus, could trigger a "credibility cliff" (Shiu and Yang 2017, 458). A firm's decision to avoid taxes is deliberate and conscious – similar to the fraudulent restatement announcement analyzed by Bartov et al. (2021)[13] – and thus, cannot be considered a "one-off event beyond management control" (Bartov et al. 2021, 82). Hence, I expect that this theory dominates the effect of CSR performance on the link between tax planning and reputational costs. I state the following hypothesis:

> ***H2:*** The negative (positive) relation between firm sales (advertising expenses) and news about firms' tax planning activities is magnified by CSR performance.

[12] In one of the specifications of the baseline regression, the authors also use the overall social responsibility as a covariate. The effect is insignificant.

[13] It should be emphasized that tax minimization strategies are not necessarily illegal.

4.3 Research Design

4.3.1 Model Specification

A common approach to identify the causal effect of an event is to apply the "difference-in-differences" (DiD) estimator.[14] Opposed to the traditional 2×2 DiD framework with two time periods and two groups, firms in the following research setting are exposed to the treatment (i.e., tax news) at different points in time (staggered treatment adoption). Staggered treatment events occur in different economic settings (e.g., policy changes) and have been examined by a number of researchers (e.g., Bertrand and Mullainathan 2003; Beck et al. 2010; Giroud and Mueller 2010; Giroud 2013; Fauver et al. 2017; Jiang et al. 2019; Lee et al. 2021; Mayberry and Watson 2021; Wang et al. 2021; Jacob et al. 2022).[15] A standard procedure in literature in case of a staggered treatment design is to estimate a generalized DiD specification: the two-way fixed effects (TWFE) model (Baker et al. 2022). The following equation is estimated:

$$\text{REP. COSTS}_{i,t} = \propto_i + \lambda_t + \beta_1 \text{ TAX\_NEWS}_{i,t} + \sum \beta_k \text{CONTROLS}_{i,t} + \varepsilon_{i,t} \qquad (4.1)$$

The staggered treatment variable TAX_NEWS takes the value 1 in the year in which a firm's tax planning activities are mentioned in the media for the first time and in all subsequent years and 0 otherwise (Lee et al. 2021). Hence, the variable is coded as the product of the post and treatment variable in a traditional 2×2 DiD design (Lee et al. 2021). The (traditional) post and treatment variable are subsumed by firm and year fixed effects, i.e., α_i and λ_t (Baker et al. 2022). To test H2, TAX_NEWS is interacted with the firm's CSR performance.

Goodman-Bacon (2021) shows that the TWFE DiD estimator is "a weighted average of all possible two-group/two-period DD estimators in the data" (Goodman-Bacon 2021, 254), where the weights of the separate 2×2 DiD estimators depend on group sizes and the treatment variances (Goodman-Bacon 2021; Baker et al. 2022).[16] In contrast to the two-group/two-period DiD framwork, the TWFE estimator not only uses never-treated firms as an effective control group, but also firms that are already treated serve as comparison for late-treated firms (Baker et al. 2022). Outcome changes of the early-

[14] Under the parallel-trends assumption, this estimator compares the change in average outcome of the treated units before and after the treatment with the change in average outcome of the (untreated) control group at the same periods. The control group serves as counterfactual outcome of the treated units, since the outcome of the treated units, if they had not been treated, is not observable (Barrios 2021; Baker et al. 2022).

[15] An overview of the use in finance and accounting is provided by Baker et al. 2022.

[16] An illustration is provided in Appendix A.

treated firms may reflect treatment effects over time that are subtracted from change in outcome from the late-treated firms (Baker et al. 2022). As a result, the TWFE DiD estimator requires treatment effect heterogeneity (e.g., Barrios 2021; Goodman-Bacon 2021; Baker et al. 2022). Thereby, treatment effects can vary over time or across groups.[17] Time-varying treatment effects cause the TWFE estimator to differ from the sample-average average treatment effect on the treated (ATT) even if the parallel-trends assumption holds (Goodman-Bacon 2021; Baker et al. 2022).[18] On the other hand, treatment effects that vary across treatment groups (but are static over time), cause the TWFE DiD estimator to deviate from the sample ATT because the weights differ from sample shares. Therefore, the TWFE estimator is not necessarily "wrong", but aggregates the overall treatment effect differently (Baker et al. 2022).[19] However, the risk of a bias is less problematic as the sample consists of a large proportion of never treated firms (Baker et al. 2022).

I estimate the static TWFE estimator (equation (4.1)) under the assumption that the treatment effect does not vary across treatment groups.[20] To minimize the risk of a treatment effect variation across time, I restrict my pre- and post-treatment periods to three years (plus the period of initial tax news) (Fauver et al. 2017; Jiang et al. 2019).[21] This also reduces the risk of confounding events (Fauver et al. 2017). In the second step, I estimate a dynamic TWFE estimation, so called "event study" specification, which replaces the treatment variable (TAX_NEWS) by a set of relative-time indicators (Baker et al. 2022). It allows the treatment effect to vary for each year.[22] The following regression equation is estimated:

---

17 Units that are initially treated at the same time period are categorized to groups/cohorts.

18 The bias is arbitrary with respect to sign and magnitude (Baker et al. 2022).

19 For more detailed econometric explanations see Goodman-Bacon (2021) and Baker et al. (2022).

20 As explained above, the estimator would not be "wrong" but weighted differently (Baker et al. 2022).

21 Prior studies find that market reactions to tax avoidance revelations reverse quiet quickly (i.e., fully or partially reversed within a month) (Gallemore et al. 2014; Brooks et al. 2016). Nonetheless, I expect that customers react differently and remember negative events longer. I assume that restoring a firm's reputation is a long-term process. Survey results from the global public relation firm Weber Shandwick (in partnership with KCR Research) indicate that the reputation recovery process takes approximately three and a half years (Weber Shandwick/KRC Research 2006). However, I additionally estimate a dynamic TWFE model which allows the effect to vary over time. Lee et al. (2021) do not restrict the treatment period for treated firms in the main analysis, thus the authors assume that the tax news result in lower employee perceptions in all subsequent periods.

22 Sun and Abraham (2021) show that the dynamic TWFE DiD estimates are biased when the treatment effects are heterogeneous across groups. Even under treatment effect homogeneity, the estimates can be contaminated by the causal effects of other periods which can be avoided by excluding periods with zero treatment effect (Sun and Abraham 2021; Baker et al. 2022). Therefore, I still assume treatment effect homogeneity across groups.

$$\text{REP. COSTS}_{i,t} = \alpha_i + \lambda_t + \sum_{l=-K}^{-2} \beta_l D_{it}^{l} + \sum_{l=0}^{L} \mu_l D_{it}^{l} + \sum \beta_k \text{CONTROLS}_{i,t} + \varepsilon_{i,t}, \quad (4.2)$$

where α_i and λ_t are firm and year fixed effects and D_{it}^{l} are relative-time indicators (e.g., D_{it}^{-2} equals 1 two years prior to the tax news and 0 otherwise) (Sun and Abraham 2021).[23] To test H2, the tax news indicator variables are interacted with the CSR performance. The research design is validated through a robustness test (section 4.5). A variety of alternative estimates or remedies have been developed in the literature due to the additional model assumptions (e.g., Cengiz et al. 2019; Callaway and Sant'Anna 2021; Goodman-Bacon 2021; Sun and Abraham 2021).

4.3.2 Operationalization of Variables

Reputational Costs

I follow Gallemore et al. (2014) and use sales, sales growth, advertising expenses, and growth in advertising expenses as my measure of reputational costs. Therefore, I define reputational costs in the sense of real economic consequences. I investigate whether firms experience a loss in sales after tax planning acitivities become public (i.e., consumer backlash). A poll in U.K. found that 43% of the British adults considering boycotting and 25% actually currently boycotting products of firms that not paying their 'fair share' of taxes (Christian Aid 2017). By using firm sales as dependent variable, I aim to investigate whether tax news trigger behavioral changes in customers, rather than merely influencing their perceptions. The variable SALES measures the firms' sales (Compustat SALE) scaled by lagged total assets (Compustat AT). ΔSALES measures the firms' sales minus the sales from the previous year scaled by lagged total assets. It is expected that sales and sales growth are negatively related to news about firms' tax planning ($\beta_1 < 0$). Furthermore, I examine the effects on the firms' advertising expenses. The variable AD_EXPENSE measures the firms' advertising expenses (Compustat XAD) scaled by lagged total assets. ΔAD_EXPENSE measures the firms' advertising expenses minus the advertising expenses from the previous year scaled by lagged total assets. In line with Gallemore et al. (2014), missing values of advertising expenses are set to zero (see also Dyreng et al. 2008; Fee et al. 2009; Lev et al. 2010; Servaes and Tamayo 2013; Dyreng et al. 2017; Bird et al. 2018; Benlemlih et al. 2023).

[23] To avoid multicollinearity, I exclude the relative-time indicator D_{it}^{-1}, i.e., the indicator for the period before the initial tax news (Baker et al. 2022).

Firms are expected to increase their advertising spending when their tax planning activities become public in order to offset negative media coverage and prevent consumer boycotts. Thus, I expect that advertising expenses and growth in advertising expenses are positively related to news about firms' tax planning ($\beta_1 > 0$).

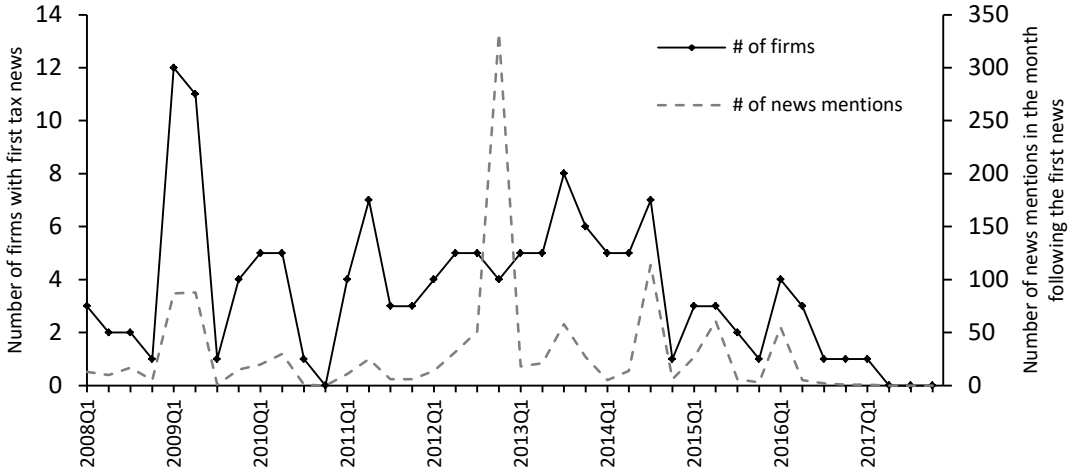
Tax News

Data on corporate tax planning activities are taken from Lee et al. (2021). The authors hand-collected tax avoidance news from S&P 500 firms between 2008 and 2017. The data is available on quarterly basis (Lee et al. 2021, Appendix B). However, I aggregate the data on annual basis due to variable availabilities. The news research of Lee et al. (2021) covers the terms "*tax avoidance*", "*tax evasion*", and "*tax haven*" (see Lee et al. 2021, Appendix C).[24] As a consequence, not only illegal but also legal tax planning activities are included. However, in the poll by Christian Aid, 89% of the respondents indicated that tax avoidance *even if it is legal* is morally wrong (Christian Aid 2017). In total, Lee et al. (2021) identify tax avoidance news for 143 of the 495 firms.[25] Figure 4.1 shows the number of firms with tax news coverage over time. In addition to the information whether tax avoidance news are revealed, Lee et al. (2021) collect data on media coverage intensity by identifying the number of media articles in the month that follows the initial news.

[24] Further features of the news research of Lee et al. (2021) are the following: (1) the authors focus on income tax avoidance, (2) the news research includes all global media sources in LexisNexis, (3) the authors use the first mention of the firms' activities, and (4) the article does not have to focus on a specific firm and its tax avoidance activity.

[25] Five firms are eliminated due to missing data on their dependent variable (employee perceptions).

Figure 4.1: Tax Avoidance News from Lee et al. (2021)



Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
No. of firms with first tax news	8	28	11	17	18	24	18	9	9	1

Notes: This figure represents the tax news data from Lee et al. (2021) graphically (see Lee et al. 2021, Appendix B). Lee et al. (2021) collect the data on quarterly basis. The solid line displays the number of firms with initial tax news; the dashed line represents the number of news mentions in the month following the first news. The noticeably high number of news mentions in fourth quarter of 2012 is driven by Starbucks Corp for which Lee et al. (2021) found 316 articles. The second highest number of articles is found for Abbvie Inc in 2014 Q3 with 62 articles.

Using this tax news data, I construct an identical TAX_NEWS variable (Lee et al. 2021) which takes the value 1 for the year in which the first news mention of the firm's tax planning is identified as well as in all subsequent years; 0 otherwise. Thus, it is assumed that the reputation is also influenced in the following periods of the tax news. The actual years in which the firm engaged in tax planning is not known.

Like Hanlon and Slemrod (2009), Gallemore et al. (2014), or Lee et al. (2021), I decide to use news data to capture avoiders. Since the revelation of tax planning in the news is usually unpredictable, firms most likely cannot influence the timing (Lee et al. 2021). It is likely that costumers learn about tax planning through news. Information is spread within seconds and reaches a wide audience. Moreover, the information in the media is often prepared in such a way that it is understandable for non-experts (Lee et al. 2021). In contrast, the acquisition and processing of the information from financial statements is often associated with higher costs for costumers. Additionally, it is unclear whether costumers are familiar with tax information from financial statements, whether

they check it regularly, or whether they understand it correctly.[26] In line with these arguments, Lee et al. (2021) find that employee perceptions are affected by tax avoidance news but the authors cannot find an effect for the ETR. One drawback of using news data, however, is there is the risk to measure the reaction to bad media rather than the reaction to corporate tax planning.

CSR Performance

The CSR performance of the firms is measured by the ESG score of Thomson Reuters Eikon (a replacement of the ASSET4 rating) (Thomson Reuters 2017).[27] The database collects information from several publicly available sources (e.g., annual reports, CSR reports, company websites, etc.) in order to calculate a relative performance measure of the firm's social responsible behavior (Refinitiv 2021). The ESG score is a percentile rank score that ranges between 0 and 100, with a higher value reflecting better performance (Refinitiv 2021).[28] Data are available on an annual basis. The ESG score is composed of three pillars (10 categories): environmental (resource use, emission, innovation), social (workforce, human rights, community, product responsibility), and governance (management, shareholders, CSR strategy) (Refinitiv 2021).[29] The database publishes a separate score for each of the three pillars and 10 categories.

The variable ESG_SCORE measures the Thomson Reuters ESG score in the year before the tax news. [30] It is assumed that costumers use the most recent ESG performance before the news as a reference for their response.[31] As the variable TAX_NEWS is always zero for never-treated firms, the interaction term (TAX_NEWS × ESG_SCORE) is by calculation always zero for never-treated firms regardless of their ESG score. Using a time-constant measure, the main effect of ESG_SCORE is subsumed by the firm

[26] Prior studies provide support that taxpayer even struggle with their own tax rates (Bartolome 1995; Rupert and Fischer 1995; Slemrod 2010).

[27] Data was drawn in September 2021.

[28] A detailed description of the scoring methodology is provided by Refinitiv (2021). Besides the overall ESG score, Thomson Reuters Eikon also calculates an ESG combined (ESGC) score which discounts the ESG score for controversies (including tax fraud controversies) (Refinitiv 2021). The present study uses the ESG score and not the ESGC score to eliminate associations with the dependent variable.

[29] A description for each of the ten categories is provided in Appendix B.

[30] Using a time-constant ESG score variable has the advantage that the recommendation of Baker et al. (2022, Fn. 2) to estimate a variant of the TWFE without time-varying covariates can be followed.

[31] A similar variable measurement is used by Jiang et al. (2019). The authors investigate the effect of Big N acquisitions on audit quality using a staggered DiD estimation. The authors analyze whether the effect varies by firm size by splitting the treatment sample based on whether the value of total assets in the year before the Big N acquisition is above or below the treatment sample median.

fixed effects.[32] It is hypothesized that the coefficient of the interaction term TAX_NEWS × ESG_SCORE is negative for regressions with SALES or ΔSALES as dependent variable and positive for regression with AD_EXPENSE or ΔAD_EXPENSE as dependent variable.

Controls

The choice of firm-specific control variables is inspired by Gallemore et al. (2014). This includes: SIZE, PPE, ΔPPE, LEVERAGE, INTANG, R&D, NOL_DUMMY, ΔNOL, SPECIAL_ITEMS, FOREIGN_INC, FOREIGN_INC_DUMMY.[33] In model specification with SALES or ΔSALES as a dependent variable, AD_EXPENSE are also a control variable. In line with Gallemore et al. (2014), I set missing values for long-term debt (DLTT), intangibles (INTAN), research and development expense (XRD), tax loss carryforwards (TLCF) and special items (SPI) to zero. I also set missing values for foreign pretax income (PIFO) to zero (Watson 2015).[34]

A description of the variable definitions is provided, for example, in Table 4.1.

4.3.3 Sample Selection

Following Lee et al. (2021), the analysis is based on S&P 500 firms for years 2008 to 2017.[35] The data is extracted from the following resources: Refinitiv/Thomson Reuters and Compustat. I retrieve the list of S&P 500 firms from Thomson Reuters Datastream.[36] All financial variables are extracted from Compustat. ESG data is retrieved from Thomsen Reuters Eikon. I merge the data via the International Securities Identification Number (ISIN). I eliminate the five firms that were not included in the analysis of Lee et al. (2021). For these five companies, I have no information on the tax avoidance news. Moreover, I lose 12 firms due to matching errors or data availability, and further 20 firms due to missing values in PPE. I end up with a sample of 463 firms (135 treatment firms), and 3,995 firm-year observations.

[32] Since the ESG score of the never-treated firms is not accounted for in the regression, it is not relevant which value is assigned to them. For completeness, never-treated firms are assigned to their average ESG score between 2008 and 2017.
[33] In contrast to Gallemore et al. (2014), I do not use extraordinary items as control variable.
[34] The abbreviations refer to Compustat variables.
[35] I use the 2013 S&P 500 firms.
[36] Historical Index Constituents as of December 2013.

4.4 Empirical Results

4.4.1 Descriptive Statistics

Descriptive statistics are provided in Table 4.1. The dependent variable SALES has a mean (median) of 0.87 (0.68). The mean (median) of ΔSALES amounts to 0.04 (0.02). The values are in a comparable range to the statistics of Gallemore et al. (2014) who uses the same dependent variables. The mean and median values of AD_EXPENSE and ΔAD_EXPENSE are close to zero. As firms do not have to disclose advertising expenses if they are deemed immaterial (Fee et al. 2009; Servaes and Tamayo 2013; Benlemlih et al. 2023), the Compustat variable XAD has a large amount of missing values that I set to zero consistent with the procedure of Gallemore et al. (2014) (see also Dyreng et al. 2008; Fee et al. 2009; Lev et al. 2010; Servaes and Tamayo 2013; Dyreng et al. 2017; Bird et al. 2018; Benlemlih et al. 2023). The mean of TAX_NEWS amounts to 0.13, i.e., 507 of the firm-year observations are treated firm-year observations. In total, the sample includes 135 treatment firms. These treatment firms vary with respect to their CSR performance. The sample includes treatment firms with both high and low ESG score. The quartile interval (mean) of ESG_SCORE for treated firms is equal to 30.31 (56.35). The mean of SIZE amounts to 9.68, which shows that this study includes mostly large firms. In Table 4.1 (Panel B), statistics are separated for treated firms (pre-treatment and post-treatment) and never-treated firms.[37] The mean values of SALES and ΔSALES in the post-treatment periods (that is the period of initial tax news and three periods after) are smaller than in the pre-treatment periods. Although this could indicate a reduction in sales after news about firms' tax planning is revealed, this bivariate result may also be driven by increases in firm size as the dependent variable is calculated as a quotient. With respect to the firms' advertising expenses, Table 4.1 (Panel B) records no major differences between pre- and post-treatment periods. One firm in particular attracts attention when it comes to the tax news data of Lee et al. (2021). Starbucks Corp records by far the highest media coverage. The authors count a total of 316 articles in the first month after the initial news. Moreover, the media has reported negative consumer reactions and boycotts, so that sales are expected to decline. Figure 4.2 displays a time series plot of the dependent variable SALES and AD_EXPENSES (Panel A) as well as their unscaled values (Panel B) for Starbucks Corps. However, no changes can

[37] I recognize that the never-treated firms differ from the treated firms with respect to some variables. However, I find no indication of a violation of the parallel-trends assumption.

be observed after the first news about the tax planning was revealed in 2012. Appendix C, Table 4.9 presents correlations for the regression variables.

Table 4.1: Descriptive Statistics

PANEL A: Descriptive Statistics - Full Sample

Variables	N	Mean	Standard Deviation	25%	Median	75%
SALES	3,995	0.87	0.74	0.38	0.68	1.09
ΔSALES	3,995	0.04	0.14	-0.01	0.02	0.08
AD_EXPENSE	3,995	0.01	0.03	0.00	0.00	0.01
ΔAD_EXPENSE	3,995	0.00	0.00	0.00	0.00	0.00
TAX_NEWS	3,995	0.13	0.33	0.00	0.00	0.00
ESG_SCORE (treatment firms)	826	56.35	20.08	40.99	59.85	71.30
SIZE	3,995	9.68	1.26	8.76	9.56	10.48
PPE	3,995	0.28	0.27	0.07	0.17	0.44
ΔPPE	3,995	0.01	0.05	-0.00	0.00	0.02
LEVERAGE	3,995	0.25	0.18	0.12	0.22	0.34
INTANG	3,995	0.25	0.26	0.03	0.18	0.40
R&D	3,995	0.02	0.04	0.00	0.00	0.02
NOL_DUMMY	3,995	0.55	0.50	0.00	1.00	1.00
ΔNOL	3,995	0.01	0.04	0.00	0.00	0.00
SPECIAL_ITEMS	3,995	-0.01	0.03	-0.01	-0.00	0.00
FOREIGN_INC	3,995	0.04	0.05	0.00	0.01	0.05
FOREIGN_INC_DUMMY	3,995	0.70	0.46	0.00	1.00	1.00

PANEL B: Descriptive Statistics - Separated for Treated Firms (Pre-Treatment and Post-Treatment) and Never-Treated Firms

	Pre-treatment [-3;-1]		Post-treatment (incl. initial tax news period) [0;+3]		Never-treated	
Variables	N	Mean	N	Mean	N	Mean
SALES	319	0.89	507	0.77	3,169	0.88
ΔSALES	319	0.06	507	0.04	3,169	0.04
AD_EXPENSE	319	0.02	507	0.02	3,169	0.01
ΔAD_EXPENSE	319	0.00	507	0.00	3,169	0.00
TAX_NEWS	319	0.00	507	1.00	3,169	0.00
ESG_SCORE	319	56.58	507	56.21	3,169	49.32
SIZE	319	10.14	507	10.45	3,169	9.51
PPE	319	0.25	507	0.24	3,169	0.29
ΔPPE	319	0.02	507	0.01	3,169	0.01
LEVERAGE	319	0.22	507	0.24	3,169	0.25
INTANG	319	0.25	507	0.28	3,169	0.24
R&D	319	0.04	507	0.04	3,169	0.02
NOL_DUMMY	319	0.50	507	0.56	3,169	0.56
ΔNOL	319	0.01	507	0.00	3,169	0.01
SPECIAL_ITEMS	319	-0.01	507	-0.01	3,169	-0.01
FOREIGN_INC	319	0.05	507	0.06	3,169	0.03
FOREIGN_INC_DUMMY	319	0.79	507	0.82	3,169	0.67

Table 4.1: Descriptive Statistics (continued)

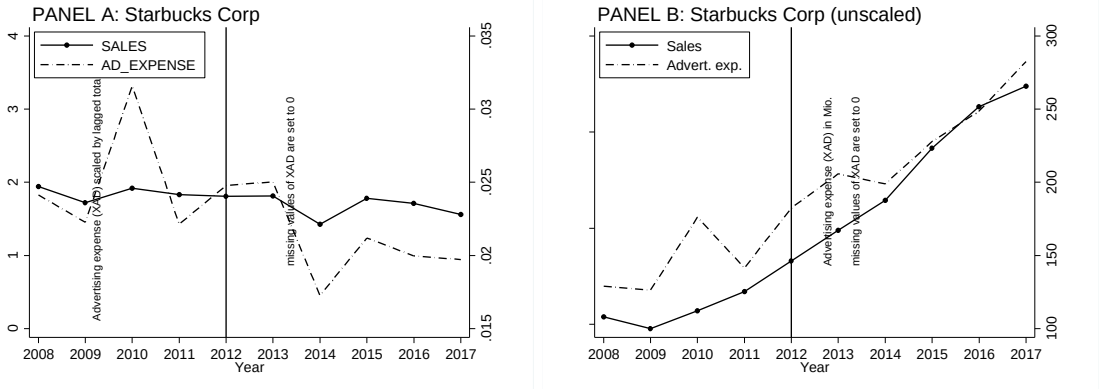
_____

Notes:

Variable Definitions:
SALES = Sales/turnover (net) (SALE) scaled by lagged total assets;
ΔSALES = Sales/turnover (net) (SALE) minus sales/turnover (net) (SALE) from the previous year scaled by lagged total assets;
AD_EXPENSE = Advertising expense (XAD) scaled by lagged total assets, missing values of XAD are set to 0;
ΔAD_EXPENSE = Advertising expense (XAD) minus advertising expense (XAD) from the previous year scaled by lagged total assets, missing values of XAD are set to 0;
TAX_NEWS = 1, for the year in which Lee et al. (2021) initially identified tax avoidance news for the firm as well as in all subsequent years; 0 otherwise;
ESG_SCORE = Environmental, social and governance performance score from Thomsen Reuters/Eikon in the year before the tax news; never-treated firms are assigned their average ESG performance between 2008 and 2017, if the CSR performance is missing in the year before the treatment, the CSR performance in the year of the treatment is used;
SIZE = Natural logarithm of total assets (AT) ($ in millions);
PPE = Total property, plant and equipment (net) (PPENT) scaled by lagged total assets;
ΔPPE = Total property, plant and equipment (net) (PPENT) minus total property, plant and equipment (net) (PPENT) from the previous year scaled by lagged total assets;
LEVERAGE = Total long-term debt (DLTT) scaled by lagged total assets, missing values of DLTT are set to 0;
INTANG = Total intangible assets (INTAN) scaled by lagged total assets, missing values of INTAN are set to 0;
R&D = Research and development expense (XRD) scaled by lagged total assets, missing values of XRD are set to 0;
NOL_DUMMY = 1, if tax loss carry forward (TLCF) is positive, and 0 otherwise, missing values of TLCF are set to 0;
ΔNOL = Tax loss carry forward (TLCF) minus tax loss carry forward (TLCF) from the previous year scaled by lagged total assets, missing values of TLCF are set to 0;
SPECIAL_ITEMS = Special items (SPI) scaled by lagged total assets, missing values of SPI are set to 0;
FOREIGN_INC = Absolute foreign pretax income (PIFO) scaled by lagged total assets, missing values of PIFO are set to 0; and
FOREIGN_INC_DUMMY = 1, if foreign pretax income (PIFO) is nonzero, and 0 otherwise, missing values of PIFO are set to 0.

All continuous variables (except of ESG_SCORE) are winsorized by year at the 1st and 99th percentiles.

Figure 4.2: Time Series Plot of Sales and Advertising Expenses for Starbucks Corp



_____

Notes: The vertical line visualizes the year of first media coverage, i.e., the year in which the TAX_NEWS first takes the value 1.

4.4.2 Regression Results

Regression results are reported in Table 4.2 (H1) and Table 4.3 (H2).[38] Based on the recommendations of Baker et al. (2022), estimations are performed with and without covariates. Standard errors are clustered by firm (cluster-robust standard errors). Hypothesis 1 predicts that firm sales (advertising expenses) are negatively (positively) related to news about firms' tax planning. Accordingly, the coefficient of TAX_NEWS is expected to be negative (positive) for regressions with SALES and ΔSALES (AD_EXPENSE and ΔAD_EXPENSE) as dependent variable. Table 4.2 shows that the regression coefficients of TAX_NEWS are not statistically significant in any of the models.[39] The dynamic model replaces TAX_NEWS by relative-time indicators in order to allow the effect to vary by time. The insignificant post-treatment indicators support the results of the static model.[40] To validate that these findings are not affected by the scaling of the dependent variables, I repeat the regressions from Table 4.2 using the natural logarithm of sales and advertising expenses[41] instead of scaling the values to the lagged total assets. Furthermore, I use the unscaled growth rates. However, the coefficients of interest remain insignificant. For the dependent variable SALES, the coefficient of TAX_NEWS even becomes positive (but insignificant). To conclude, this empirical analysis cannot find support that firm sales (advertising expenses) are negatively (positively) related to news about firms' tax planning.[42] Similar results can be found in literature. Gallemore et al. (2014) find no differential changes in their dependent variables sales, sales growth, advertising expenses, or growth in advertising expenses for firms whose tax shelter was revealed compared to the controls firms. The authors analyze tax shelter revelations between 1995 and 2005, thus, data earlier to mine. Similarly, Hoopes et al. (2018) find no change in consumer sentiment for large, public firms with strong global brands after disclosing private tax information, even if they disclose having paid no taxes. One explanation could be that the public already widely believed that these firms were not paying their 'fair share' of taxes (Hoopes et al. 2018). Lastly, Chen

[38] The estimations are based on the reghdfe command of Stata. Although reghdfe also returns an estimate for a constant, I will refrain from reporting this value in the following tables.

[39] With respect to the control variables, no directional predictions were made, so no interpretation of the effects is provided.

[40] The relative-time indicators are interpreted as the difference of the difference in the dependent variable between treated and untreated observations in the k-th year (k= -3, -2, 0, 1, 2, 3) before/after the tax news and difference in the dependent variable between these observations one period before the treatment (i.e., the excluded period) (Baker et al. 2022).

[41] Due to the zero values, I used the natural logarithm of the advertising expenses plus one.

[42] The results must also be interpreted taking into account that the adjusted within R^2's of the models are rather low. All variance inflation factors (VIFs) are below the threshold of 10 (Wooldridge 2016, 86).

et al. (2019) find no evidence that firms decrease their level of tax avoidance following media coverage, suggesting that managers do not believe that media coverage increase existing costs of tax avoidance.[43]

Nevertheless, I carefully acknowledge that my non-results may be affected by a low power of the empirical tests. Even though, the empirical analysis cannot provide support for a decline in sales or an increase in advertising expenses after firms' tax planning has been revealed by the media, reputational effects may exist. When interpreting the results of the present study, it is also important to take into consideration that reputational effects are measured by actual behavior change among customers. According to the poll by Christian Aid, 43% of the respondents state that they considering boycotting products or services of firms that do not pay their 'fair share' of taxes, however, only 25% actually do so currently (Christian Aid 2017). This finding suggests that while customers often have attitudes about a given topic, not all of them translate those attitudes into action. The boycott of products or services is often associated with costs for consumers (e.g., changing behavior patterns or the renunciation of products/services). It is possible that tax news lead to negative customer perceptions, however these negative attitudes are not always sufficient to drive a change in behavior. This might also explain why Lee et al. (2021), using the same tax news data, find an effect on employee perceptions on managers and firms. Further research on whether tax news only affect perceptions but do not result in actions, or whether employees react differently than customers, is recommended. Furthermore, it must also be taken into account that this analysis examines U.S. firms, thus, a rather shareholder-oriented system (Bottenberg et al. 2017). The results may not be generalizable to stakeholder-oriented countries (e.g., Germany) (Bottenberg et al. 2017).

To investigate whether reputational effects only occur in specific industries (e.g., retail sector as it is a consumer-facing industry), the analysis is repeated for the following industries: manufacturing (Standard Industrial Classification (SIC) Code 2000-3999), transportation, communications, electric, gas and sanitary service (SIC Code 4000-4999), retail trade (SIC Code 5200-5999), finance, insurance and real estate (SIC Code 6000-6799), and services (SIC Code 7000-8999).[44] The only weak support that could be found in line with H1 is an increase of advertising expenses growth in the service indus-

[43] However, the costs of media coverage may have already incorporated in the decision about the level of tax avoidance (Chen et al. 2019).

[44] Due to the number of observations, I do not regress the other industries.

try.[45] In addition, I take the market concentration in consideration. Customers of companies in less concentrated industries can be expected to find it easier to switch to competitors if they want to boycott a company because of its tax planning.[46] Market concentration is measured by the sales Herfindahl-Hirschman index (e.g., Kubick et al. 2015). Thus, an industry-specific measure is generated. I identify the manufacturing, transportation, communications, electric, gas and sanitary service, and finance, insurance and real estate industry sectors as less-concentrated industries (i.e., below average). However, I find no clear support that the reputational effect of tax news varies between high-concentrated and low-concentrated industries.

To validate my research design, I additionally address the parallel-trends assumption. Even though this assumption is not directly testable, researchers often use the coefficients of the leads to test for pre-trends (Baker et al. 2022). Sun and Abraham (2021) show that the relative-time period estimates are contaminated by the causal effect of the other periods, which is why pre-period coefficients should only be used as 'pre-trend test' under strong assumptions (Sun and Abraham 2021; Baker et al. 2022). The present analysis assumes that treatment effects do not vary across groups and only pre-treatment periods are excluded from the model. As I find no clear pattern of significant pre-treatment coefficients (especially in the models with control variables), I carefully interpret this finding as an indication that the models do not violate the parallel-trends assumption.[47]

[45] Contrary to H1, firms in the transportation, communications, electric, gas and sanitary service industry even show weak support for increase in sales and sales growth. As this industry has one of the lowest tax rates (Dyreng et al. 2008), one possible explanation could be that tax planning is more accepted. On the other hand, firms of this industry report weak support that advertising expenses are positively related to tax news which would be in line with H1. Moreover, contrary to H1, there is some weak evidence that advertising expenses are negatively related to tax news in the manufacturing sector.

[46] Similarly, Rhee and Haunschild (2006) find support that market reactions to product recalls are less negative if firms have fewer substitutes with an equivalent reputation.

[47] In the static TWFE estimation, the parallel-trends assumption is investigated by using a pseudo-event (Fauver et al. 2017; Jiang et al. 2019). It is simulated that the firms are covered by the news two years prior to the actual tax news timing. The event window is set to one period before and one period after the pseudo-event as well as the period of the pseudo-event. The 'true' treatment periods are excluded for this test. Significant outcome differences before and after the pseudo-event would indicate a violation of the parallel-trends assumption. The coefficient of the pseudo tax news event is, however, insignificant.

Table 4.2: Effect of Tax News Coverage on Sales and Advertising Expenses

PANEL A: Static TWFE DiD Estimation

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	SALES	SALES	ΔSALES	ΔSALES	AD_E.	AD_E.	ΔAD_E.	ΔAD_E.
Variables								
TAX_NEWS	-0.0078	-0.0054	-0.0012	-0.0049	-0.0008	-0.0007	0.0000	-0.0000
	(0.0181)	(0.0151)	(0.0090)	(0.0088)	(0.0008)	(0.0008)	(0.0003)	(0.0003)
AD_EXPENSE		6.4424***		2.0216***				
		(1.3802)		(0.6011)				
SIZE		-0.1744***		-0.0043		-0.0031***		-0.0008*
		(0.0450)		(0.0144)		(0.0011)		(0.0004)
PPE		0.3272		-0.0445		0.0122***		-0.0037**
		(0.2258)		(0.0685)		(0.0042)		(0.0015)
ΔPPE		0.5731**		0.4521***		-0.0039		0.0075***
		(0.2535)		(0.0960)		(0.0045)		(0.0022)
LEVERAGE		0.1279		0.0232		0.0016		0.0004
		(0.0847)		(0.0271)		(0.0029)		(0.0012)
INTANG		0.1759***		0.1369***		0.0049*		0.0035***
		(0.0596)		(0.0280)		(0.0029)		(0.0011)
R&D		2.2711***		0.5921**		0.0375		-0.0032
		(0.5627)		(0.2606)		(0.0304)		(0.0108)
NOL_DUMMY		-0.0107		-0.0057		-0.0004		-0.0003
		(0.0253)		(0.0105)		(0.0006)		(0.0003)
ΔNOL		-0.1606*		-0.0824		0.0044		0.0041
		(0.0922)		(0.0792)		(0.0038)		(0.0031)
SPECIAL_ITEMS		0.1871		0.0995		-0.0089		-0.0039
		(0.1617)		(0.1003)		(0.0149)		(0.0046)
FOREIGN_INC		1.7331***		0.9315***		0.0071		0.0092*
		(0.2451)		(0.1590)		(0.0102)		(0.0053)
FOREIGN_INC_DUMMY		0.0169		-0.0147		0.0012		0.0002
		(0.0361)		(0.0134)		(0.0019)		(0.0006)
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	3,995	3,995	3,995	3,995	3,995	3,995	3,995	3,995
Adjusted within R^2	-0.00	0.23	-0.00	0.11	0.00	0.03	-0.00	0.02

PANEL B: Dynamic TWFE DiD Estimation

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	SALES	SALES	ΔSALES	ΔSALES	AD_E.	AD_E.	ΔAD_E.	ΔAD_E.
Variables								
D^{-3}	-0.0052	-0.0051	0.0083	0.0116	0.0011	0.0010	0.0007	0.0007
	(0.0190)	(0.0187)	(0.0169)	(0.0167)	(0.0010)	(0.0010)	(0.0007)	(0.0007)
D^{-2}	0.0237*	0.0113	0.0138	0.0089	0.0001	-0.0000	0.0001	0.0001
	(0.0142)	(0.0128)	(0.0160)	(0.0157)	(0.0007)	(0.0007)	(0.0005)	(0.0005)
D^{0}	0.0114	-0.0074	0.0170	0.0075	0.0006	0.0005	0.0006	0.0005
	(0.0164)	(0.0121)	(0.0143)	(0.0132)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
D^{1}	0.0028	-0.0075	0.0021	-0.0065	-0.0003	-0.0005	-0.0000	-0.0001
	(0.0172)	(0.0146)	(0.0142)	(0.0141)	(0.0007)	(0.0007)	(0.0004)	(0.0004)
D^{2}	-0.0156	0.0038	-0.0043	0.0001	-0.0013	-0.0011	-0.0000	-0.0000
	(0.0211)	(0.0179)	(0.0144)	(0.0140)	(0.0009)	(0.0009)	(0.0003)	(0.0004)
D^{3}	-0.0120	0.0008	0.0026	0.0013	-0.0011	-0.0008	0.0005	0.0005
	(0.0252)	(0.0210)	(0.0137)	(0.0128)	(0.0010)	(0.0010)	(0.0004)	(0.0004)
Control variables	✗	✓	✗	✓	✗	✓	✗	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	3,995	3,995	3,995	3,995	3,995	3,995	3,995	3,995
Adjusted within R^2	-0.00	0.23	-0.00	0.11	0.00	0.03	-0.00	0.02

Table 4.2: Effect of Tax News Coverage on Sales and Advertising Expenses (continued)

Notes: *, **, and *** denote the statistical significance at the 10, 5, and 1 percent levels, respectively (two-tailed tests). Standard errors are clustered by firm.

Variable Definitions:
SALES = Sales/turnover (net) (SALE) scaled by lagged total assets;
ΔSALES = Sales/turnover (net) (SALE) minus sales/turnover (net) (SALE) from the previous year scaled by lagged total assets;
AD_EXPENSE = Advertising expense (XAD) scaled by lagged total assets, missing values of XAD are set to 0;
ΔAD_EXPENSE = Advertising expense (XAD) minus advertising expense (XAD) from the previous year scaled by lagged total assets, missing values of XAD are set to 0;
TAX_NEWS = 1, for the year in which Lee et al. (2021) initially identified tax avoidance news for the firm as well as in all subsequent years; 0 otherwise;
D^k = 1, in the k-th year (k= -3, -2, 0, 1, 2, 3) before/after the tax news; 0 otherwise;
SIZE = Natural logarithm of total assets (AT) ($ in millions);
PPE = Total property, plant and equipment (net) (PPENT) scaled by lagged total assets;
ΔPPE = Total property, plant and equipment (net) (PPENT) minus total property, plant and equipment (net) (PPENT) from the previous year scaled by lagged total assets;
LEVERAGE = Total long-term debt (DLTT) scaled by lagged total assets, missing values of DLTT are set to 0;
INTANG = Total intangible assets (INTAN) scaled by lagged total assets, missing values of INTAN are set to 0;
R&D = Research and development expense (XRD) scaled by lagged total assets, missing values of XRD are set to 0;
NOL_DUMMY = 1, if tax loss carry forward (TLCF) is positive, and 0 otherwise, missing values of TLCF are set to 0;
ΔNOL = Tax loss carry forward (TLCF) minus tax loss carry forward (TLCF) from the previous year scaled by lagged total assets, missing values of TLCF are set to 0;
SPECIAL_ITEMS = Special items (SPI) scaled by lagged total assets, missing values of SPI are set to 0;
FOREIGN_INC = Foreign pretax income (PIFO) scaled by lagged total assets, missing values of PIFO are set to 0; and
FOREIGN_INC_DUMMY = 1, if foreign pretax income (PIFO) is nonzero, and 0 otherwise, missing values of PIFO are set to 0.

All continuous variables (except of ESG_SCORE) are winsorized by year at the 1st and 99th percentiles. Panel B uses the same control variables as Panel A.

Even though I cannot find an on-average effect of tax news on firm sales or advertising expenses, reputational effects may occur for firms with higher CSR performances. Consistent with the findings of Inger and Vansant (2019), Huang et al. (2017), and Choy et al. (2017), hypothesis 2 predicts that the negative (positive) relation of tax news on firm sales (on advertising expenses) is magnified for firms with higher CSR performance. In order to analyze this hypothesis, TAX_NEWS is interacted with ESG_SCORE. Table 4.3 shows that the regression coefficients of the interaction terms are statistically insignificant at any of the three significance levels (1, 5, or 10 percent level). The results of the dynamic model are consistent to the ones of the static model: the interactions terms of the post-treatment indicators are statistically insignificant. This finding indicates that the effect of tax news does not depend on CSR performance. H2 cannot be supported.[48] A possible explanation for the non-results could be that customers are not aware of or cannot properly assess the level of CSR performance of the firms. Nowadays, nearly all large firms publish a CSR report and emphasize their sustainability, which leads to the inflationary use of the 'sustainability image'. The ques-

[48] The findings are also validated for the unscaled dependent variables. Furthermore, no clear support in line with H2 can be found for the different industries. Although significant coefficients occur in isolated cases, effects cannot be supported in all models or dependent variables.

tion is to what extent consumers can interpret or distinguish this information. Thus, further research is recommended. Moreover, it could be that customers decouple CSR commitment and tax strategy as firms regularly fail to incorporate their CSR commitment into all areas of the business (Mayberry and Watson 2021). However, I would like to emphasize once again that the failure to detect significant effects could be attributed to a low power of the statistical tests at well as further limitations of this econometric analysis (see section 4.6 for discussion). Although prior research demonstrates a variety of effects of CSR, I cannot provide support for the 'risk management' theory or the 'negative surprise' theory. Therefore, my findings are not in line with the results of Inger and Stekelberg (2022), whose analysis provided support that CSR can enhance reputational capital in the context of tax avoidance, nor with the results of Inger and Vansant (2019), who find that tax avoidance and CSR are perceived as inconsistent with one another and therefore counterproductive. It may suggest that investors and customers differ in their perceptions.

Table 4.3: Variation in the Effect of Tax News Coverage on Sales and Advertising Expenses Depending on CSR Performance

PANEL A: Static TWFE DiD Estimation

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	SALES	SALES	ΔSALES	ΔSALES	AD_E.	AD_E.	ΔAD_E.	ΔAD_E.
Variables								
TAX_NEWS	0.0093	0.0003	0.0092	-0.0109	0.0052	0.0055	0.0009	0.0008
	(0.0512)	(0.0438)	(0.0213)	(0.0219)	(0.0046)	(0.0045)	(0.0013)	(0.0012)
TAX_NEWS× ESG_SCORE	-0.0003	-0.0001	-0.0002	0.0001	-0.0001	-0.0001	-0.0000	-0.0000
	(0.0008)	(0.0007)	(0.0004)	(0.0004)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
AD_EXPENSE		6.4368***		2.0275***				
		(1.3860)		(0.6043)				
SIZE		-0.1745***		-0.0041		-0.0033***		-0.0009**
		(0.0451)		(0.0144)		(0.0011)		(0.0004)
PPE		0.3268		-0.0442		0.0118***		-0.0038**
		(0.2260)		(0.0685)		(0.0041)		(0.0015)
ΔPPE		0.5737**		0.4514***		-0.0031		0.0076***
		(0.2539)		(0.0960)		(0.0045)		(0.0022)
LEVERAGE		0.1279		0.0232		0.0016		0.0004
		(0.0847)		(0.0271)		(0.0029)		(0.0012)
INTANG		0.1759***		0.1369***		0.0048*		0.0035***
		(0.0596)		(0.0280)		(0.0029)		(0.0011)
R&D		2.2717***		0.5914**		0.0380		-0.0031
		(0.5629)		(0.2605)		(0.0305)		(0.0108)
NOL_DUMMY		-0.0107		-0.0057		-0.0004		-0.0004
		(0.0253)		(0.0105)		(0.0006)		(0.0003)
ΔNOL		-0.1607*		-0.0823		0.0043		0.0041
		(0.0923)		(0.0792)		(0.0039)		(0.0031)
SPECIAL_ITEMS		0.1876		0.0990		-0.0083		-0.0038
		(0.1617)		(0.1004)		(0.0150)		(0.0046)
FOREIGN_INC		1.7327***		0.9318***		0.0067		0.0091*
		(0.2451)		(0.1590)		(0.0102)		(0.0053)
FOREIGN_INC_DUMMY		0.0168		-0.0146		0.0012		0.0002
		(0.0361)		(0.0134)		(0.0019)		(0.0006)
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	3,995	3,995	3,995	3,995	3,995	3,995	3,995	3,995
Adjusted within R^2	-0.00	0.23	-0.00	0.11	0.01	0.04	-0.00	0.02

Table 4.3: Variation in the Effect of Tax News Coverage on Sales and Advertising Expenses Depending on CSR Performance (continued)

PANEL B: Dynamic TWFE DiD Estimation

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	SALES	SALES	ΔSALES	ΔSALES	AD_E.	AD_E.	ΔAD_E.	ΔAD_E.
Variables								
D^{-3}	-0.0172	-0.0319	-0.0318	-0.0272	-0.0033	-0.0043	-0.0010	-0.0016
	(0.0703)	(0.0712)	(0.0488)	(0.0502)	(0.0042)	(0.0041)	(0.0043)	(0.0042)
D^{-3}× ESG_SCORE	0.0002	0.0005	0.0007	0.0007	0.0001	0.0001	0.0000	0.0000
	(0.0011)	(0.0011)	(0.0008)	(0.0008)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
D^{-2}	0.0060	-0.0022	-0.0140	-0.0112	-0.0041	-0.0046	-0.0017	-0.0019
	(0.0358)	(0.0400)	(0.0460)	(0.0475)	(0.0028)	(0.0028)	(0.0020)	(0.0021)
D^{-2}× ESG_SCORE	0.0003	0.0002	0.0005	0.0004	0.0001*	0.0001*	0.0000	0.0000
	(0.0007)	(0.0006)	(0.0007)	(0.0007)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
D^{0}	0.0330	-0.0172	0.0308	0.0100	0.0028*	0.0023	0.0005	0.0003
	(0.0438)	(0.0358)	(0.0365)	(0.0336)	(0.0016)	(0.0016)	(0.0018)	(0.0018)
D^{0}× ESG_SCORE	-0.0004	0.0002	-0.0002	-0.0000	-0.0000	-0.0000	0.0000	0.0000
	(0.0007)	(0.0006)	(0.0007)	(0.0006)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
D^{1}	0.0103	-0.0437	-0.0100	-0.0441	0.0028	0.0022	-0.0004	-0.0008
	(0.0443)	(0.0367)	(0.0360)	(0.0365)	(0.0035)	(0.0034)	(0.0011)	(0.0012)
D^{1}× ESG_SCORE	-0.0001	0.0006	0.0002	0.0007	-0.0001	-0.0000	0.0000	0.0000
	(0.0007)	(0.0006)	(0.0007)	(0.0006)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
D^{2}	0.0082	0.0319	-0.0157	-0.0201	0.0031	0.0032	0.0001	-0.0001
	(0.0541)	(0.0388)	(0.0388)	(0.0347)	(0.0046)	(0.0045)	(0.0009)	(0.0010)
D^{2}× ESG_SCORE	-0.0004	-0.0005	0.0002	0.0004	-0.0001	-0.0001	-0.0000	0.0000
	(0.0009)	(0.0007)	(0.0006)	(0.0006)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
D^{3}	-0.0334	0.0000	-0.0286	-0.0383	0.0029	0.0037	-0.0001	-0.0002
	(0.0663)	(0.0498)	(0.0374)	(0.0323)	(0.0042)	(0.0040)	(0.0009)	(0.0010)
D^{3}× ESG_SCORE	0.0004	0.0000	0.0006	0.0007	-0.0001	-0.0001	0.0000	0.0000
	(0.0010)	(0.0008)	(0.0006)	(0.0006)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
Control variables	✗	✓	✗	✓	✗	✓	✗	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	3,995	3,995	3,995	3,995	3,995	3,995	3,995	3,995
Adjusted within R^2	-0.00	0.23	-0.00	0.11	0.01	0.04	-0.00	0.02

Notes: *, **, and *** denote the statistical significance at the 10, 5, and 1 percent levels, respectively (two-tailed tests). Standard errors are clustered by firm. Panel B uses the same control variables as Panel A.

Variable Definitions:
ESG_SCORE = Environmental, social and governance performance score from Thomsen Reuters/Eikon in the year before the tax news; never-treated firms are assigned their average ESG performance between 2008 and 2017.
The remaining definitions are found in Table 4.2.

Philanthropic CSR Performance:

Inger and Vansant (2019) provide support that the effect of CSR performance on the link between tax avoidance and firm value is sensitive to the type of CSR the firm engages in. The authors differentiate between philanthropic and non-philanthropic CSR and find that philanthropic CSR activities significantly affects the relation between tax avoidance and firm value, but non-philanthropic CSR does not. Similarly, costumers may differentiate between the types of CSR when considering their reactions to tax planning. Non-philanthropic CSR includes investments that are rather focused on future profitability (e.g., improving product quality or workplace conditions) (Inger and

Vansant 2019). These CSR activities might be perceived as less inconsistent with tax planning, which is also targeted on increasing after-tax profits. Philanthropic CSR, on the other hand, is focused in human welfare (Inger and Vansant 2019). Tax planning is assumed to be rather contradictive to these types of CSR. I follow Inger and Vansant (2019) and define the ESG score categories *Community* and *Human Rights* as philanthropic CSR.[49] I investigate whether H2 can be supported for philanthropic CSR. Regression results are displayed in Table 4.4. To increase the level of information, the two ESG categories are not aggregated to one measure of philanthropic CSR but rather regressed separately (i.e., two separate regressions). Moreover, for reasons of simplicity, only the interactions of the treatment period (D^0) and the period thereafter (D^1) are reported but all relative-time indicators are included in the regression. The regression analysis provides some support that the effect of tax news on advertising expense differs by firms' human right scores. In contrast to H2, advertising expenditures of firms with high human rights performance appear to be negatively related to tax news. However, such an effect cannot be found for SALES or ΔSALES. As no clear pattern is shown, I conclude that, unlike Inger and Vansant (2019), I find no clear evidence that customers are more surprised and react more negatively when firms engage in tax planning and high level of philanthropic CSR at the same time.

Table 4.4: Variation in the Effect of Tax News Coverage on Sales and Advertising Expenses Depending on Philanthropic CSR Performance

PANEL A: Static TWFE DiD Estimation

Model	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Dep. Var.	**SALES**	**SALES**	**ΔSALES**	**ΔSALES**	**AD_E.**	**AD_E.**	**ΔAD_E.**	**ΔAD_E.**
Variables								
TAX_NEWS× HUMANRIGHTS	-0.0004	-0.0004	-0.0001	-0.0001	-0.0000**	-0.0000**	-0.0000	-0.0000
	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
TAX_NEWS× COMMUNITY	0.0001	0.0004	-0.0004	-0.0002	-0.0001	-0.0001	-0.0000	-0.0000
	(0.0008)	(0.0007)	(0.0003)	(0.0004)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
Control variables	✗	✓	✗	✓	✗	✓	✗	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓

[49] Refinitiv (2021) defines *Community* as "company's commitment to being a good citizen, protecting public health and respecting business ethics" (Refinitiv 2021, 22) and *Human Rights* as "a company's effectiveness in terms of respecting fundamental human rights conventions" (Refinitiv 2021, 22). In line with Inger and Vansant (2019), governance categories are defined as rather non-philanthropic.

Table 4.4: Variation in the Effect of Tax News Coverage on Sales and Advertising Expenses Depending on Philanthropic CSR Performance (continued)

PANEL B: Dynamic TWFE DiD Estimation

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	SALES	SALES	ΔSALES	ΔSALES	AD_E.	AD_E.	ΔAD_E.	ΔAD_E.
Variables								
D^0× HUMANRIGHTS	-0.0003	-0.0001	-0.0005	-0.0004	-0.0000**	-0.0000**	-0.0000	-0.0000
	(0.0004)	(0.0003)	(0.0004)	(0.0004)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
D^1× HUMANRIGHTS	-0.0003	-0.0001	-0.0003	-0.0001	-0.0000*	-0.0000	-0.0000	-0.0000
	(0.0005)	(0.0004)	(0.0004)	(0.0004)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
D^0× COMMUNITY	-0.0008	-0.0006	-0.0006	-0.0007	-0.0000	-0.0000	0.0000	0.0000
	(0.0009)	(0.0007)	(0.0007)	(0.0007)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
D^1× COMMUNITY	-0.0006	0.0005	-0.0005	0.0002	-0.0000	-0.0000	0.0000	0.0000
	(0.0008)	(0.0007)	(0.0006)	(0.0007)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
Control variables	✗	✓	✗	✓	✗	✓	✗	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: *, **, and *** denote the statistical significance at the 10, 5, and 1 percent levels, respectively (two-tailed tests). Standard errors are clustered by firm. Each ESG category is regressed separately. However, for the sake of brevity, the coefficients are listed below each other. The table displays only the interactions of the treatment period (D^0) and the period thereafter (D^1), but just as in Table 4.3, all relative–time indicators are included in the regression. The same control variables are used as in Table 4.2.

Variable Definitions:
HUMANRIGHTS = Refinitiv ESG human rights score from Thomsen Reuters/Eikon in the year before the tax news; never-treated firms are assigned their average score between 2008 and 2017;
COMMUNITY = Refinitiv ESG community score from Thomsen Reuters/Eikon in the year before the tax news; never-treated firms are assigned their average score between 2008 and 2017.
The remaining definitions are found in Table 4.2.

4.5 Robustness Checks

I conduct a battery of robustness tests to corroborate the findings from Table 4.2 to Table 4.4. First, the estimation approach is validated. As discussed in section 4.3, the TWFE estimator requires specific model assumptions. As a result, numerous alternative estimators and remedies have been developed that aim to compare treated units to "clean controls" (Baker et al. 2022, 383). Callaway and Sant'Anna (2021) derive estimators for group-time-specific treatment effects which allow for treatment effect heterogeneity and dynamic effects. Sun and Abraham (2021), on the other hand, focus on the "event study" (i.e., dynamic) specification and derive an interacted-weighted (IW) estimator. In the following robustness test, I apply the so-called stacked regression estimator from Cengiz et al. (2019). It can be applied to both static and dynamic specifications and I can integrate interactions. In the first step, 2 × 2 datasets are created for each cohort using "clean" control units (i.e., never-treated units). These 'clean' datasets are then stacked together and identified with dataset-specific identifiers. Lastly, I apply the

TWFE DiD estimation on the expanded dataset using dataset-specific fixed effects (unit and time) (Baker et al. 2022).[50] Regression results are summarized in Table 4.5. However, the findings remain qualitatively unchanged. The analysis does not provide support that firm sales (advertising expenses) are negatively (positively) related to news about firms' tax planning. No support is found that this effect is sensitive to the CSR performance.

Table 4.5: Stacked Regression Estimator

PANEL A: Robustness Check for H1

Model / **Dep. Var.**	**(1)** **SALES**	**(2)** **SALES**	**(3)** **ΔSALES**	**(4)** **ΔSALES**	**(5)** **AD_E.**	**(6)** **AD_E.**	**(7)** **ΔAD_E.**	**(8)** **ΔAD_E.**
Static estimation:								
TAX_NEWS	-0.0079	-0.0043	-0.0013	-0.0048	-0.0008	-0.0007	0.0000	-0.0001
	(0.0167)	(0.0140)	(0.0086)	(0.0083)	(0.0008)	(0.0008)	(0.0003)	(0.0003)
Dynamic estimation:								
D^{-3}	-0.0053	-0.0050	0.0081	0.0111	0.0010	0.0011	0.0007	0.0008
	(0.0191)	(0.0188)	(0.0174)	(0.0171)	(0.0010)	(0.0010)	(0.0007)	(0.0007)
D^{-2}	0.0233	0.0106	0.0134	0.0085	0.0001	-0.0000	0.0001	0.0001
	(0.0146)	(0.0132)	(0.0166)	(0.0162)	(0.0007)	(0.0007)	(0.0005)	(0.0005)
D^{0}	0.0090	-0.0074	0.0144	0.0061	0.0005	0.0004	0.0005	0.0005
	(0.0164)	(0.0124)	(0.0149)	(0.0137)	(0.0005)	(0.0005)	(0.0004)	(0.0004)
D^{1}	0.0056	-0.0044	0.0039	-0.0044	-0.0004	-0.0005	0.0001	-0.0001
	(0.0170)	(0.0146)	(0.0145)	(0.0144)	(0.0007)	(0.0007)	(0.0004)	(0.0004)
D^{2}	-0.0148	0.0038	-0.0032	0.0004	-0.0012	-0.0010	-0.0000	-0.0000
	(0.0203)	(0.0171)	(0.0145)	(0.0140)	(0.0009)	(0.0009)	(0.0003)	(0.0004)
D^{3}	-0.0116	-0.0006	0.0038	0.0010	-0.0011	-0.0008	0.0005	0.0004
	(0.0241)	(0.0197)	(0.0135)	(0.0124)	(0.0010)	(0.0010)	(0.0004)	(0.0004)
Control variables	✗	✓	✗	✓	✗	✓	✗	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	29,319	29,319	29,319	29,319	29,319	29,319	29,319	29,319

[50] To avoid a loss of power, always-treated units are left in the sample in the main analysis (see also Lee et al. 2021). In the following robustness test, always-treated units are dropped.

Table 4.5: Stacked Regression Estimator (continued)

PANEL B: Robustness Check for H2

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	SALES	SALES	ΔSALES	ΔSALES	AD_E.	AD_E.	ΔAD_E.	ΔAD_E.
Static estimation:								
TAX_NEWS × ESG_SCORE	-0.0003	-0.0001	-0.0002	0.0001	-0.0001	-0.0001	-0.0000	-0.0000
	(0.0008)	(0.0007)	(0.0004)	(0.0004)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
Dynamic estimation:								
D^{-3} × ESG_SCORE	0.0002	0.0004	0.0007	0.0006	0.0001	0.0001	0.0000	0.0000
	(0.0011)	(0.0011)	(0.0008)	(0.0008)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
D^{-2} × ESG_SCORE	0.0003	0.0002	0.0005	0.0003	0.0001*	0.0001*	0.0000	0.0000
	(0.0007)	(0.0006)	(0.0007)	(0.0007)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
D^{0} × ESG_SCORE	-0.0002	0.0001	-0.0001	-0.0001	-0.0000	-0.0000	-0.0000	0.0000
	(0.0007)	(0.0005)	(0.0007)	(0.0006)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
D^{1} × ESG_SCORE	-0.0002	0.0007	0.0002	0.0007	-0.0001	-0.0000	0.0000	0.0000
	(0.0007)	(0.0006)	(0.0006)	(0.0006)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
D^{2} × ESG_SCORE	-0.0004	-0.0004	0.0001	0.0003	-0.0001	-0.0001	-0.0000	-0.0000
	(0.0009)	(0.0007)	(0.0006)	(0.0006)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
D^{3} × ESG_SCORE	0.0002	-0.0000	0.0004	0.0006	-0.0001	-0.0001	0.0000	0.0000
	(0.0011)	(0.0008)	(0.0006)	(0.0005)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
Control variables	✗	✓	✗	✓	✗	✓	✗	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	29,319	29,319	29,319	29,319	29,319	29,319	29,319	29,319

Notes: *, **, and *** denote the statistical significance at the 10, 5, and 1 percent levels, respectively (two-tailed tests). Standard errors are clustered by firm. Always-treated firms are dropped for this regression. Panel B only displays the interaction terms. However, the main effect of TAX_NEWS or the relative-time indicators are also included in the regression. The same control variables are used as in Table 4.2.

Variable Definitions:
ESG_SCORE = Environmental, social and governance performance score from Thomsen Reuters/Eikon in the year before the tax news; never-treated firms are assigned their average ESG performance between 2008 and 2017.
The remaining definitions are found in Table 4.2.

Second, I vary my measurement of CSR performance. In the main analysis, the ESG performance in the period before treatment is used. Shiu and Yang (2017) find that while short-term CSR commitment does not provide a significant insurance-like protection, long-term CSR activities do. Based on the formula of Shiu and Yang (2017), I construct a long-term ESG score:

$$\text{ESG\_LONG} = \tfrac{1}{2}\,\text{ESG}_{i,t-1} + \tfrac{1}{4}\,\text{ESG}_{i,t-2} + \tfrac{1}{8}\,\text{ESG}_{i,t-3,} \tag{4.3}$$

where t denote the period of the initial tax news and ESG represents the Thomson Reuters Eikon ESG score in the corresponding period. Accordingly, this long-term measure takes into account the ESG performance in the three periods preceding the tax news. Table 4.6 summarizes the regression results. The results, however, remain qualitatively robust.

Table 4.6: Long-Term CSR Commitment

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.	SALES	SALES	ΔSALES	ΔSALES	AD_E.	AD_E.	ΔAD_E.	ΔAD_E.
Static estimation:								
TAX_NEWS × ESG_LONG	-0.0002	-0.0000	-0.0001	0.0002	-0.0001	-0.0001	-0.0000	-0.0000
	(0.0009)	(0.0008)	(0.0004)	(0.0004)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
Dynamic estimation:								
D^{-3} × ESG_LONG	0.0004	0.0007	0.0009	0.0009	0.0001	0.0001	0.0000	0.0000
	(0.0013)	(0.0012)	(0.0009)	(0.0009)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
D^{-2} × ESG_ LONG	0.0005	0.0003	0.0005	0.0003	0.0001	0.0001*	0.0000	0.0000
	(0.0007)	(0.0007)	(0.0008)	(0.0008)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
D^{0} × ESG_ LONG	0.0000	0.0006	-0.0000	0.0001	-0.0000	-0.0000	-0.0000	0.0000
	(0.0007)	(0.0006)	(0.0007)	(0.0007)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
D^{1} × ESG_ LONG	-0.0000	0.0008	0.0004	0.0009	-0.0001	-0.0001	0.0000	0.0000
	(0.0008)	(0.0007)	(0.0007)	(0.0007)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
D^{2} × ESG_ LONG	-0.0004	-0.0006	0.0002	0.0003	-0.0001	-0.0001	-0.0000	-0.0000
	(0.0011)	(0.0008)	(0.0007)	(0.0006)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
D^{3} × ESG_ LONG	0.0007	0.0002	0.0008	0.0009	-0.0001	-0.0001	0.0000	0.0000
	(0.0013)	(0.0010)	(0.0007)	(0.0006)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
Control variables	✗	✓	✗	✓	✗	✓	✗	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	3,979	3,979	3,979	3,979	3,979	3,979	3,979	3,979

Notes: *, **, and *** denote the statistical significance at the 10, 5, and 1 percent levels, respectively (two-tailed tests). Standard errors are clustered by firm. Only the interaction terms are displayed. However, the main effect of TAX_NEWS or the relative-time indicators are also included in the regression. The same control variables are used as in Table 4.2.

Variable Definitions:

ESG_LONG = $½ESG_{i,t-1} + ¼ESG_{i,t-2} + ⅛ESG_{i,t-3}$, where $ESG_{t-1,t-2,t-3}$ is the environmental, social and governance performance score from Thomsen Reuters/Eikon in the year, two years, three years before the tax news; never-treated firms are assigned their average ESG performance between 2008 and 2017; since the scores from the last three years must be available, the number of observations decreases compared to the main analysis.

The remaining definitions are found in Table 4.2.

Third, I validate my measures of reputational costs. A limitation of using firm sales as a measure of reputational costs is that tax avoiders may inflate their sales. Thus, a decline in sales could be caused by a reduction in sales-inflation instead of a costumer boycott (Gallemore et al. 2014, Fn. 23). I replace sales and sales growth by operating cash flow (Compustat: OANCF – Operating Activities Net Cash Flow) and its growth (Gallemore et al. 2014) scaled to lagged total assets to check the robustness of the results. A further drawback is the data availability of firms' advertising expenses. Thus, I replace advertising expenses by selling, general and administrative expenses (Compustat: XSGA: Selling, General and Administrative Expense).[51] Nevertheless, results should be interpreted with caution because a wide range of expenses are covered. The findings remain qualitatively robust. I continue to find no clear support for H1 or H2.[52]

[51] Missing values are replaced by zeros. The dependent variables are still scaled to lagged total assets.

[52] I find a positive, weak significant effect for D^{1} × ESG_SCORE for the operativing cash flow growth, contrary to H2. However, this effect cannot be found in the other models.

Fourth, I analyze whether my results change with the intensity of public attention. Table 4.7 illustrates that the treatment intensity varies widely across the 143 firms for which Lee et al. (2021) find tax news. The likelihood that customers are unaware of tax planning behavior is greater for firms with low media coverage especially due to the fact that the news search mechanism does not require the firm to be the focus of the article. In an untabulated analysis, I repeat the analysis for the 398 firms who had at least three news mentions in the first month. However, still no clear support of H1 or H2 can be found.[53]

Table 4.7: Treatment Intensity

# news mentions in the month following the first news	1	2	3	4	5	6	7	8	9	10	11	12	15
N	43	22	13	9	10	7	4	4	4	4	2	1	2

16	18	19	21	22	23	25	26	34	36	62	316	
3	1	2	2	2	2	1	1	1	1	1	1	

Notes: Data is based on Lee et al. (2021) (Appendix B). The table displays how treatment intensity varies across the 143 firms for which Lee et al. (2021) find tax news.

Fifth, I vary my amount of pre-treatment and post-treatment periods (i.e., event window). The main analysis includes three periods before and after the treatment (as well as the period of initial tax news itself). In an untabulated analysis, the treatment periods are not restricted at all (see Lee et al. 2021) to increase the power of the tests. The findings, however, remain robust. Sixth, I redefine the control sample of never-treated firms. Each treated firm is matched with a never-treated firm from the same industry (Fama-French 17 industry classification[54]) that is closest in firm size (measured by total assets) in the year before the tax news (Gallemore et al. 2014). I use a one-to-one matching without replacement.[55] The new sample includes 134 treatment firms and 134 never-treated firms.[56] I repeat the regression analysis. I continue to find no evidence that firm sales (advertising expenses) are negatively (positively) related to news about firms' tax

[53] I find a negative, signicant effect of $D^2 \times$ ESG_SCORE for the dependent variable SALES. However, this effect is not found in the prior periods or the other models. Thus, I conclude that I still cannot find clear support for H2.

[54] Unmapped SIC codes are assigned to the "other" category.

[55] Nevertheless, I still recognize differences between never-treated and treated firms with respect to some variables.

[56] One treatment firm cannot be matched.

planning. I also continue to find no support that this effect is sensitive to the firms' CSR performance. Seventh, I repeated my analysis with quarterly data. Brooks et al. (2016) and Gallemore et al. (2014) find support that reputational costs quickly reverse. As Compustat does not include quarterly data for advertising expenses, this validity test is only applied to the dependent variables of sales and sales growth. Independent variables that are only available yearly are assigned the same value for all for quarters of the year. To ensure that the same data is used, I still restrict the event window to three pre-treatment and post-treatment years as in the main analysis. The findings of Table 4.2 and Table 4.3 remain qualitatively unchanged. However, in the dynamic estimtion, I find weak evidence in support of H2 that the effect of tax news on sales differs by firms' community scores (i.e., philanthropic CSR). Further research regarding the time horizon of reputational effects is recommended.

4.6 Conclusion

One of the most fundamental questions in tax research is why some companies engage in tax avoidance and others do not, i.e., "under-sheltering puzzle" (Gallemore et al. 2014, 1103). To decide on the level of tax planning, managers must weigh the costs against the benefits. Although reputational effects are considered a potential cost of tax planning, it is still controversial whether avoiding taxes causes reputational damage (ex post). This study addresses this need for research and further provides insights on whether the reputational costs are sensitive to the level of CSR the firm engages in. Thereby, reputational costs are measured by a decrease in sales or increase in advertising expenses (i.e., financial reputational costs). Using a staggered difference-in-differences regression design, this study cannot find support that firm sales (advertising expenses) are negatively (positively) related to news about firms' tax planning. The results are consistent with the findings of Gallemore et al. (2014), who find no support that tax shelter revelations between 1995 and 2005 impose any reputational costs exerted by customers. Similarly, Hoopes et al. (2018) find no change in consumer sentiment for large public firms following the disclosure of private tax information in Australia. Moreover, this study cannot provide support that reputational costs of tax news are magnified for firms with higher CSR performance. Prior studies find that CSR activities can act as insurance-like protection in case of adverse events. Reputational goodwill is built and acts as protection. Other researches rather support the theory that CSR and tax avoidance are incompatible with one another. The present analysis does not provide

clear support for either of these theories in the context of customer reactions to tax news. One possible explanation could be that stakeholders, just like firms, may decouple CSR commitment and tax strategy as firms regularly fail to incorporate their CSR commitment into all areas of the business (Mayberry and Watson 2021). Moreover, customers might not be aware of firms' CSR performance nor are able to assess it. I carefully acknowledge that my findings could be affected by a low power of the empirical tests. Although the present analysis is unable to empirically identify reputational costs of tax planning, reputational effects may exist.

The results of this study must be interpreted in consideration of certain caveats and limitations. First, one difficulty in measuring reputational costs of tax planning is the risk of a selection bias. Firms that face high reputational costs may not engage in tax planning (Gallemore et al. 2014; Lee et al. 2021). Although Gallemore et al. (2014) find no evidence that reputation is associated with the likelihood of tax shelter participation, I cannot rule out that my non-results are affected by a selection bias. This relationship needs to be further explored. Second, this research is limited to the news search criteria of Lee et al. (2021). As the authors include the terms "tax avoidance", "tax evasion", and "tax haven" in their news search (see Lee et al. 2021, Appendix C), a broad spectrum of tax planning is covered. The present study uses information on whether tax planning was reveled but has no information whether the revelations refer to legal or illegal tax planning strategies. Prior studies show that reactions vary with the type of tax planning (Blaufus et al. 2019b). The non-results of this study could be driven by articles describing legal tax avoidance behavior which may do not cause reputational costs, even though the revelation of illegal tax evasion maybe does. In addition, there is the risk that tax news about the firm may have been public before the investigation period and may still be anticipated by customers. Another caveat of using news data is that it cannot be ruled out that customers may respond only to negative news, rather than news about tax planning in particular. Third, due to the difficulties of measuring the CSR performance, there are different metrics and approaches. Chatterji et al. (2016) analyze several raters and find little overlap between the assessments which is why the authors advise to be cautious when drawing conclusion from theses rating. Results of this study might be affected by the choice of ESG measure. Alternative CSR metrics would be the KLD database/MSCI ESG Rating (e.g., Hoi et al. 2013; Watson 2015; Davis et al. 2016; Huang et al. 2017; Lanis and Richardson 2018; Bartov et al. 2021) or the Sustainalytics

database (Bartov et al. 2021). Fourth, the Compustat variable XAD has a large number of missing values which I set to zero in line with Gallemore et al. (2014) (see also Dyreng et al. 2008; Fee et al. 2009; Lev et al. 2010; Servaes and Tamayo 2013; Dyreng et al. 2017; Bird et al. 2018; Benlemlih et al. 2023). Advertising expenses that are deemed to be immaterial do not have to be disclosed (Fee et al. 2009; Servaes and Tamayo 2013; Benlemlih et al. 2023). I recommend validating the results with a more refined variable of advertising expenses. One possibility would be to use information from the Kantar Media database (Ad$pender). Fifth, findings may not be generalized to smaller firms as the research sample consists of S&P 500 firms. Hoopes et al. (2018) find changes in consumer sentiment after firms are subject to tax disclosure only for smaller private firms but not for large public firms with strong brands. Similarily, Brooks et al. (2016) find more pronounced negative reactions to negative news about corporate tax payments for smaller firms. Sixth, reputational effects may reverse so quickly that that they cannot be detected even with quarterly data (Gallemore et al. 2014; Brooks et al. 2016).

This analysis provides useful insights on the effects of tax planning and the effectiveness of CSR activities. It helps mangers to decide on the level of tax planning and CSR they want to engage in. The study offers a number of avenues for further research and hopefully encourages other researchers to investigate this research field in more detail. First, it is important to understand whether costumers' reactions to tax news depend on the type of tax planning. Although shareholders are most likely interested in the legality and aggressiveness of the tax planning strategy, it is unclear whether the same applies for costumers. Costumers are often no tax-experts and may not be able to assess the tax aggressiveness on their own. It would be valuable to understand if the findings of Blaufus et al. (2019b) apply to all stakeholders (customers, employees, suppliers, government, etc.). Second, Lee et al. (2021), using the same tax news data, find support that tax news affects the employee perception of managers and firms. Further research should address the question of whether tax planning merely influences perceptions or attitudes but does not lead to actions, or whether employees respond differently than customers. Third, as noted above, future research on whether only firms that do not suffer reputational costs engage in tax planning is recommended. Fourth, more insights on whether firms can protect themselves from reputational costs or whether costs may vary

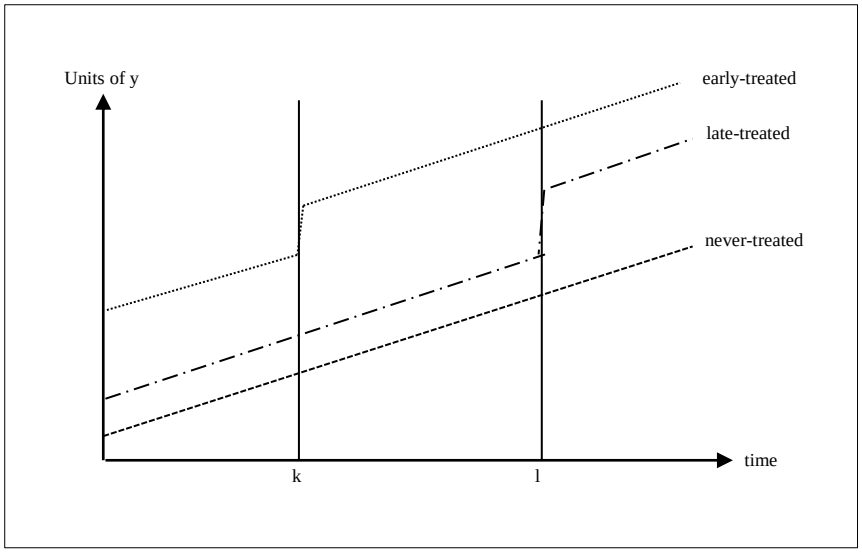
between firms is needed[57] (e.g.: Does Apple face the same likelihood of a consumer boycott due to aggressive tax planning as Starbucks, despite a potentially similar customer base and country of origin?, Do consumers overlook bad press of firms they feel close to?[58]). Fifth, my research sample focuses on large U.S. firms and may not be generalizable to other sample (e.g., smaller firms). The findings of Hoopes et al. (2018) suggest that the resilience of consumer sentiment may vary between large public firms and smaller private ones. Research on whether my findings also apply to smaller firms would yield valuable insights. Furthermore, it is important to understand whether the results differ with culture, religion, country, or government system (Kountouris and Remoundou 2013; Hardeck et al. 2019). The U.S. is a rather shareholder-orientated system (Bottenberg et al. 2017). It is recommended to examine whether the same results can be found for stakeholder-oriented systems. Finally, to better understand the power of social media, future research using tax news data should additionally include social media coverage.

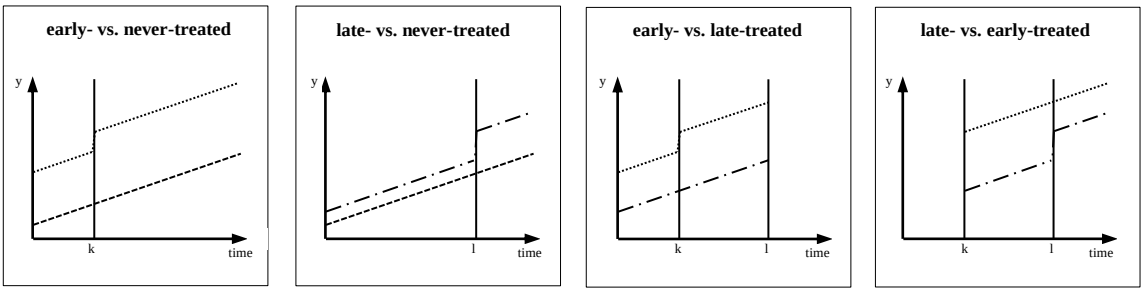
[57] See Gatzert (2015) for a discussion of further potential moderators of (financial) reputational effects (i.e., prior reputation, substitutability and generalism/specialism, crisis communication ability, control over social media).

[58] This context is also discussed by Antonetti and Anesa (2017).

4.7 Appendix A: Staggered TWFE DiD Estimation

Figure 4.3: Staggered TWFE DiD Estimation





Notes: This graphical example is taken from Goodman-Bacon (2021), including the notation. The depiction, however, may be adjusted. k denotes the treatment timing of the early-treated units, l denotes the treatment timing of the late-treated units.

In a staggered treatment model, units are treated at different periods of time. Thereby, the time period when units are first treated categorizes units into "groups" (or "cohorts") (Callaway and Sant'Anna 2021; Sun and Abraham 2021). The TWFE DiD estimator calculates a "weighted average of all possible two-group/two-period DD estimators in the data" (Goodman-Bacon 2021, 254). This figure illustrates the estimation by an example of three groups of units (early-treated, late-treated, never-treated). In total, four 2×2 DiD comparisons ((1) early vs. never-treated, (2) late vs. never-treated, (3) early vs. late-treated, (4) late vs. early-treated) are made by the TWFE estimator (Goodman-Bacon 2021). Not only never-treated units but also, for example, early-treated units are used as control group for late-treated units.

4.8 Appendix B: ESG Categories

Table 4.8: Definition of ESG Categories

ESG Category	Definition
Resource use	"The resource use score reflects a company's performance and capacity to reduce the use of materials, energy or water, and to find more eco-efficient solutions by improving supply chain management." (Refinitiv 2021, 22)
Emissions reduction	"The emission reduction score measures a company's commitment and effectiveness towards reducing environmental emissions in its production and operational processes." (Refinitiv 2021, 22)
Innovation	"The innovation score reflects a company's capacity to reduce the environmental costs and burdens for its customers, thereby creating new market opportunities through new environmental technologies and processes, or eco-designed products." (Refinitiv 2021, 22)
Workforce	"The workforce score measures a company's effectiveness in terms of providing job satisfaction, a healthy and safe workplace, maintaining diversity and equal opportunities and development opportunities for its workforce." (Refinitiv 2021, 22)
Human rights	"The human rights score measures a company's effectiveness in terms of respecting fundamental human rights conventions." (Refinitiv 2021, 22)
Community	"The community score measures the company's commitment to being a good citizen, protecting public health and respecting business ethics." (Refinitiv 2021, 22)
Product responsibility	"The product responsibility score reflects a company's capacity to produce quality goods and services, integrating the customer's health and safety, integrity and data privacy." (Refinitiv 2021, 22)
Management	"The management score measures a company's commitment and effectiveness towards following best practice corporate governance principles." (Refinitiv 2021, 22)
Shareholders	"The shareholders score measures a company's effectiveness towards equal treatment of shareholders and the use of anti-takeover devices." (Refinitiv 2021, 22)
CSR strategy	"The CSR strategy score reflects a company's practices to communicate that it integrates economic (financial), social and environmental dimensions into its day-to-day decision-making processes." (Refinitiv 2021, 22)

Notes: All definitions are taken from (Refinitiv 2021).

4.9 Appendix C: Correlation Matrix

Table 4.9: Correlation Matrix

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1)	SALES		**0.41**	**0.31**	**0.12**	**-0.03**	**0.10**	**-0.43**	**0.15**	**0.09**	-0.02
(2)	ΔSALES	**0.40**		**0.14**	**0.29**	-0.01	**-0.07**	**-0.17**	**0.04**	**0.33**	-0.02
(3)	AD_EXPENSE	**0.23**	**0.13**		**0.28**	**0.09**	**0.07**	**-0.16**	**-0.13**	**-0.07**	-0.03
(4)	ΔAD_EXPENSE	**0.11**	**0.29**	**0.52**		-0.01	**-0.04**	**-0.07**	-0.01	**0.13**	-0.03
(5)	TAX_NEWS	**-0.05**	-0.01	**0.07**	0.02		**0.12**	**0.21**	**-0.04**	-0.02	-0.00
(6)	ESG_SCORE	**0.06**	**-0.06**	0.00	**-0.11**	**0.12**		**0.37**	**0.12**	-0.02	0.01
(7)	SIZE	**-0.30**	**-0.11**	**-0.21**	**-0.11**	**0.24**	**0.36**		**-0.07**	0.00	-0.02
(8)	PPE	-0.01	-0.01	**-0.09**	**-0.05**	**-0.06**	**0.08**	0.01		**0.41**	**0.36**
(9)	ΔPPE	**0.05**	**0.26**	-0.01	**0.10**	-0.01	**-0.04**	0.01	**0.42**		**0.14**
(10)	LEVERAGE	**-0.06**	-0.00	0.03	0.01	-0.01	**-0.04**	**-0.06**	**0.27**	**0.18**	
(11)	INTANG	-0.01	**0.10**	**0.06**	**0.08**	**0.05**	**-0.08**	**-0.11**	**-0.35**	0.03	**0.31**
(12)	R&D	**-0.03**	**0.08**	0.00	**0.04**	**0.14**	0.02	**-0.22**	**-0.24**	-0.01	**-0.14**
(13)	NOL_DUMMY	**0.07**	**0.04**	**0.08**	**0.07**	0.01	**-0.07**	**-0.21**	**-0.17**	**-0.06**	**0.06**
(14)	ΔNOL	-0.00	0.03	0.02	**0.08**	-0.02	**-0.03**	**-0.04**	**0.06**	**0.09**	**0.10**
(15)	SPECIAL_ITEMS	-0.01	0.03	-0.02	-0.02	-0.00	-0.01	**0.07**	**0.05**	**0.04**	**-0.10**
(16)	FOREIGN_INC	**0.11**	**0.14**	**0.23**	**0.14**	**0.17**	**0.06**	**-0.22**	**-0.13**	0.02	**-0.07**
(17)	FOREIGN_INC_ DUMMY	**0.11**	0.01	**0.15**	**0.06**	**0.10**	0.01	**-0.25**	**-0.28**	**-0.13**	**-0.03**

		(11)	(12)	(13)	(14)	(15)	(16)	(17)
(1)	SALES	**0.18**	**0.19**	**0.16**	**0.04**	**-0.07**	**0.30**	**0.25**
(2)	ΔSALES	**0.12**	**0.07**	**0.07**	0.03	**0.08**	**0.12**	**0.04**
(3)	AD_EXPENSE	**0.12**	**0.06**	**0.08**	0.01	-0.02	**0.11**	**0.09**
(4)	ΔAD_EXPENSE	**0.05**	0.01	**0.06**	**0.03**	**0.05**	**0.04**	0.02
(5)	TAX_NEWS	**0.05**	**0.14**	0.01	**-0.04**	**-0.05**	**0.15**	**0.10**
(6)	ESG_SCORE	**-0.04**	**0.09**	**-0.07**	-0.02	**-0.04**	**0.04**	-0.00
(7)	SIZE	**-0.14**	**-0.22**	**-0.20**	**-0.03**	**0.10**	**-0.29**	**-0.26**
(8)	PPE	**-0.31**	**-0.17**	**-0.11**	**0.05**	0.01	**-0.09**	**-0.14**
(9)	ΔPPE	**-0.08**	**-0.06**	**-0.09**	**0.05**	**0.10**	**-0.06**	**-0.15**
(10)	LEVERAGE	**0.19**	**-0.09**	**0.04**	0.03	**-0.21**	**-0.08**	**-0.05**
(11)	INTANG		**0.28**	**0.27**	**0.05**	**-0.32**	**0.27**	**0.35**
(12)	R&D	**0.09**		**0.21**	-0.00	**-0.22**	**0.55**	**0.46**
(13)	NOL_DUMMY	**0.24**	**0.17**		**0.13**	**-0.16**	**0.28**	**0.34**
(14)	ΔNOL	**0.09**	0.01	**0.13**		**-0.08**	0.03	**0.06**
(15)	SPECIAL_ITEMS	**-0.15**	**-0.11**	**-0.07**	**-0.09**		**-0.21**	**-0.23**
(16)	FOREIGN_INC	**0.05**	**0.38**	**0.15**	**0.04**	**-0.07**		**0.80**
(17)	FOREIGN_INC_ DUMMY	**0.28**	**0.29**	**0.34**	**0.04**	**-0.11**	**0.48**	

Notes: The lower triangle report the Pearson correlation coefficients, the upper triangle the Spearman correlation coefficients. Variable definitions are found in Table 4.1. Correlations that are significant at the 5 percent level are marked bold.

Chapter 5

Concluding Remarks

A large amount of tax research aims to understand tax planning with all its related aspects. This thesis addresses this research need by investigating behavioral responses in relation to tax planning and tax compliance. First of all, it addresses questions regarding the effectiveness of efforts to promote tax compliance and provides insides into behavioral responses to citizen-state interactions. Second, it extends research on corporate behavioral responses to anti-avoidance rules that aim to increase tax transparency. Thirdly, this thesis provides insights on how customers respond to news about firms' tax planning and whether firms can influence the risk to be publically shamed by their CSR performance.

The first article shows that service interaction quality between tax administration and taxpayers is positively related to the compliance of taxpayers. The findings provide useful information for tax administrations by demonstrating that their efforts to be more 'customer friendly' can be worthwhile. However, in the absence of information on the costs of these service provisions, further research on the actual net benefits is needed. Moreover, it is necessary to determine the extent to which the tax compliance is affected. A further field of research would be to identify the drivers of the positive effect of these services and explore the research questions in different cultural and institutional settings.

The second article informs policymakers about limitations of regulations that aim to increase tax transparency. The fundamental idea of these regulations is that firms disclose truthful and valid information. The study shows that firms avoid the purpose of the U.K. tax strategy disclosure (Finance Act 2016, Schedule 19) by reporting strategically. However, external monitoring can limit this behavior. Future research should address the characteristics by which policymakers increase the effectiveness of tax trans-

parency rules. It is recommended to investigate whether strategic reporting depends on qualitative or qualitative information and how the monitoring of disclosed data can be optimized. To evaluate the consequences of strategic reporting, it is also necessary to examine how shareholders use the information disclosed.

The third article informs policymakers about the effectiveness of public shaming. Countries increasing rely on the perceptions of the public to curb tax planning (e.g., Okafor and Farrar 2021). The results of this study, however, do not support that firm sales (advertising expenses) are negatively (positively) related to news about firms' tax planning. Nonetheless, reactions might differ in other cultural settings. From a practical perspective, the last article offers managers insights into deciding how much CSR to engage in. The analysis does not find support that reputational costs are magnified by CSR performance. Future research on whether reactions differ depending on the legality of the tax planning activities is recommended (Blaufus et al. 2019b).

Bibliography

Ai, C., and E. C. Norton. 2003. Interaction terms in logit and probit models. *Economics Letters* 80 (1): 123–29.

Allingham, M. G., and A. Sandmo. 1972. Income tax evasion: A theoretical analysis. *Journal of Public Economics* 1 (3-4): 323–38.

Alm, J., T. Cherry, M. Jones, and M. McKee. 2010. Taxpayer information assistance services and tax compliance behavior. *Journal of Economic Psychology* 31 (4): 577–86.

Antonetti, P., and M. Anesa. 2017. Consumer reactions to corporate tax strategies: The role of political ideology. *Journal of Business Research* 74: 1–10.

Asay, H. S., J. L. Hoopes, J. R. Thornock, and J. H. Wilde. 2018. Consumer responses to corporate tax planning. Working Paper. Available at: https://ssrn.com/abstract=3224670.

Austin, C. R., and R. J. Wilson. 2017. An examination of reputational costs and tax avoidance: Evidence from firms with valuable consumer brands. *The Journal of the American Taxation Association* 39 (1): 67–93.

Avi-Yonah, R. S. 2014. Corporate taxation and corporate social responsibility. *New York University Journal of Law & Business* 11 (1): 1–30.

Baker, A. C., D. F. Larcker, and C. C. Y. Wang. 2022. How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics* 144 (2): 370–95.

Baker, S. R., S. T. Sun, and C. Yannelis. 2020. Corporate taxes and retail prices. NBER Working Paper No. 27058. Available at: https://ssrn.com/abstract=3586190.

Bansal, H. S., S. F. Taylor, and Y. St. James. 2005. "Migrating" to new service providers: Toward a unifying framework of consumers' switching behaviors. *Journal of the Academy of Marketing Science* 33 (1): 96–115.

Barnett, M. L., and T. G. Pollock. 2012. Charting the landscape of corporate reputation research. In *The Oxford handbook of corporate reputation*. Edited by M. L. Barnett and T. G. Pollock: Oxford University Press.

Barrios, J. M. 2021. Staggeringly problematic: A primer on staggered DiD for accounting researchers. Working paper. Available at: https://ssrn.com/abstract=3794859.

Bartolome, C. am de. 1995. Which tax rate do people use: Average or marginal? *Journal of Public Economics* 56 (1): 79–96.

Bartov, E., A. Marra, and F. Momente. 2021. Corporate social responsibility and the market reaction to negative events: Evidence from inadvertent and fraudulent restatement announcements. *The Accounting Review* 96 (2): 81–106.

Baudot, L., J. A. Johnson, A. Roberts, and R. W. Roberts. 2020. Is corporate tax aggressiveness a reputation threat? Corporate accountability, corporate social responsibility, and corporate tax behavior. *Journal of Business Ethics* 163 (2): 197–215.

Beck, T., R. Levine, and A. Levkov. 2010. Big bad banks? The winners and losers from bank deregulation in the United States. *The Journal of Finance* 65 (5): 1637–67.

Benlemlih, M., J. Jaballah, S. Schochet, and J. Peillex. 2023. Corporate social responsibility and corporate tax avoidance: The channel effect of consumer awareness. *Journal of Business Finance & Accounting* 50 (1-2): 31–60.

Bertrand, M., and S. Mullainathan. 2003. Enjoying the quiet life? Corporate governance and managerial preferences. *Journal of Political Economy* 111 (5): 1043–75.

Bird, A., A. Edwards, and T. G. Ruchti. 2018. Taxes and peer effects. *The Accounting Review* 93 (5): 97–117.

Blaufus, K., F. Hechtner, and J. K. Jarzembski. 2019a. The income tax compliance costs of private households: Empirical evidence from Germany. *Public Finance Review* 47 (5): 925–66.

Blaufus, K., F. Hechtner, and A. Möhlmann. 2017. The effect of tax preparation expenses for employees: Evidence from Germany. *Contemporary Accounting Research* 34 (1): 525–54.

Blaufus, K., J. K. Jarzembski, J. Reineke, and I. Trenn. 2023. Public Disclosure of Tax Strategies and Firm's Actual Tax Policy. *Journal of Accounting, Auditing & Finance*. Available at: https://doi.org/10.1177/0148558X231200913.

Blaufus, K., A. Möhlmann, and A. N. Schwäbe. 2019b. Stock price reactions to news about corporate tax avoidance and evasion. *Journal of Economic Psychology* 72: 278–92.

Bogetoft, P., and L. Otto. 2011. *Benchmarking with DEA, SFA, and R.* New York, NY: Springer.

Bosch, J.-C., and E. W. Eckard Jr. 1991. The profitability of price fixing: evidence from stock market reaction to federal indictments. *The Review of Economics and Statistics* 73 (2): 309–17.

Bottenberg, K., A. Tuschke, and M. Flickinger. 2017. Corporate governance between shareholder and stakeholder orientation: Lessons from Germany. *Journal of Management Inquiry* 26 (2): 165–80.

Braithwaite, V. 2003. Dancing with tax authorities: Motivational postures and non-compliant actions. *Taxing Democracy*, 15–39.

Brooks, C., C. Godfrey, C. Hillenbrand, and K. Money. 2016. Do investors care about corporate taxes? *Journal of Corporate Finance* 38: 218–48.

Callaway, B., and P. H. C. Sant'Anna. 2021. Difference-in-differences with multiple time periods. *Journal of Econometrics* 225 (2): 200–30.

Campbell, K., and D. Helleloid. 2016. Starbucks: Social responsibility and tax avoidance. *Journal of Accounting Education* 37: 38–60.

Carroll, A. B. 1991. The pyramid of corporate social responsibility: Toward the moral management of organizational stakeholders. *Business Horizons* 34 (4): 39–48.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer. 2019. The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics* 134 (3): 1405–54.

Chan, K. H., and P. L. Lan Mo. 2000. Tax holidays and tax noncompliance: An empirical study of corporate tax audits in China's developing economy. *The Accounting Review* 75 (4): 469–84.

Charnes, A., W. W. Cooper, and E. Rhodes. 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2 (6): 429–44.

Chatterji, A. K., R. Durand, D. I. Levine, and S. Touboul. 2016. Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal* 37 (8): 1597–614.

Chen, S., K. Schuchard, and B. Stomberg. 2019. Media coverage of corporate taxes. *The Accounting Review* 94 (5): 83–116.

Chih, H.-L., C.-H. Shen, and F.-C. Kang. 2008. Corporate social responsibility, investor protection, and earnings management: Some international evidence. *Journal of Business Ethics* 79 (1): 179–98.

Cho, S. Y., C. Lee, and R. J. Pfeiffer Jr. 2013. Corporate social responsibility performance and information asymmetry. *Journal of Accounting and Public Policy* 32 (1): 71–83.

Choy, S. K., T.-K. Lai, and T. Ng. 2017. Do tax havens create firm value? *Journal of Corporate Finance* 42: 198–220.

Christensen, D. M. 2016. Corporate accountability reporting and high-profile misconduct. *The Accounting Review* 91 (2): 377–99.

Christensen, J., and R. Murphy. 2004. The social irresponsibility of corporate tax avoidance: Taking CSR to the bottom line. *Development* 47 (3): 37–44.

Christian, C. W. 1994. Voluntary compliance with the individual income tax: Results from the 1988 TCMP study. *The IRS Research Bulletin* 1500 (9-94): 35–42.

Christian Aid. 2017. Tax Avoidance Survey. Accessed May 15, 2021. Available at: https://mediacentre.christianaid.org.uk/89-of-british-adults-say-tax-avoidance-by-large-companies-is-morally-wrong-christian-aid-comres-poll/.

Chyz, J. A., and F. B. Gaertner. 2018. Can paying “too much” or “too little” tax contribute to forced CEO turnover? *The Accounting Review* 93 (1): 103–30.

Clarkson, M. B. E. 1995. A stakeholder framework for analyzing and evaluating corporatesocial performance. *Academy of Management Review* 20 (1): 92–117.

Cloyd, C. B., L. F. Mills, and C. D. Weaver. 2003. Firm valuation effects of the expatriation of US corporations to tax-haven countries. *Journal of the American Taxation Association* 25 (Supplement): 87–109.

Cooper, W. W., L. M. Seiford, and K. Tone. 2007. *Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software* 2. New York, NY: Springer.

Darby, M. R., and E. Karni. 1973. Free competition and the optimal amount of fraud. *The Journal of Law & Economics* 16 (1): 67–88.

Davis, A. B., R. D. Moore, and T. J. Rupert. 2022. Corporate social responsibility and tax management: The moderating effect of beliefs about corporate tax duty. *The Journal of the American Taxation Association* 44 (2): 35–53.

Davis, A. K., D. A. Guenther, L. K. Krull, and B. M. Williams. 2016. Do socially responsible firms pay more taxes? *The Accounting Review* 91 (1): 47–68.

Demerjian, P., B. Lev, and S. McVay. 2012. Quantifying managerial ability: A new measure and validity tests. *Management Science* 58 (7): 1229–48.

Desai, M. A., and D. Dharmapala. 2009. Corporate tax avoidance and firm value. *The Review of Economics and Statistics* 91 (3): 537–46.

Desai, M. A., and J. R. Hines. 2002. Expectations and expatriations: Tracing the causes and consequences of corporate inversions. *National Tax Journal* 55 (3): 409–40.

DeZoort, F. T., T. J. Pollard, and E. J. Schnee. 2018. A study of perceived ethicality of low corporate effective tax rates. *Accounting Horizons* 32 (1): 87–104.

Dhaliwal, D., O. Z. Li, A. Tsang, and Y. G. Yang. 2014. Corporate social responsibility disclosure and the cost of equity capital: The roles of stakeholder orientation and financial transparency. *Journal of Accounting and Public Policy* 33 (4): 328–55.

Dhaliwal, D. S., T. H. Goodman, P. J. Hoffman, and C. M. Schwab. 2022. The incidence, valuation and management of tax-related reputational costs: Evidence from a period of protest. *Journal of the American Taxation Association* 44 (1): 49–73.

Dhaliwal, D. S., O. Z. Li, A. Tsang, and Y. G. Yang. 2011. Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review* 86 (1): 59–100.

Dhaliwal, D. S., S. Radhakrishnan, A. Tsang, and Y. G. Yang. 2012. Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure. *The Accounting Review* 87 (3): 723–59.

Di Giuli, A., and L. Kostovetsky. 2014. Are red or blue companies more likely to go green? Politics and corporate social responsibility. *Journal of Financial Economics* 111 (1): 158–80.

Djankov, S., T. Ganser, C. McLiesh, R. Ramalho, and A. Shleifer. 2010. The effect of corporate taxes on investment and entrepreneurship. *American Economic Journal: Macroeconomics* 2 (3): 31–64.

Dowling, G. R. 2014. The curious case of corporate tax avoidance: Is it socially irresponsible? *Journal of Business Ethics* 124 (1): 173–84.

Dyreng, S. D., M. Hanlon, and E. L. Maydew. 2008. Long-run corporate tax avoidance. *The Accounting Review* 83 (1): 61–82.

Dyreng, S. D., M. Hanlon, E. L. Maydew, and J. R. Thornock. 2017. Changes in corporate effective tax rates over the past 25 years. *Journal of Financial Economics* 124 (3): 441–63.

Dyreng, S. D., J. L. Hoopes, and J. H. Wilde. 2016. Public pressure and corporate tax behavior. *Journal of Accounting Research* 54 (1): 147–86.

Erard, B. 1997. Self-selection with measurement errors A microeconometric analysis of the decision to seek tax assistance and its implications for tax compliance. *Journal of Econometrics* 81 (2): 319–56.

Escobales, R., and T. McVeigh. 2012. "Starbucks hit by UK uncut protests as tax row boils over." *The Guardian*, 2012. Accessed May 15, 2021. Available at: https://www.theguardian.com/business/2012/dec/08/starbucks-uk-stores-protests-tax.

EY. 2014. *Bridging the divide: Highlights from the 2014 tax risk and controversy survey*. Accessed May 14, 2021. Available at: https://www.eytax.jp/tax-library/other-report/pdf/Tax-Risk-and-Controversy-Survey-2014_E.pdf.

EY. 2017. *Dimming the glare: 2017 Tax risk and controversy survey series*. Accessed September 03, 2023. Available at: https://www.ey.com/en_be/tax/dimming-the-glare--trends-in-tax-controversy.

Fauver, L., M. Hung, X. Li, and A. G. Taboada. 2017. Board reforms and firm value: Worldwide evidence. *Journal of Financial Economics* 125 (1): 120–42.

Federal Statistical Office. 2015. *Finanzen und Steuern: Jährliche Einkommensteuerstatistik Sonderthema: Werbungskosten 2011*. Wiesbaden: Federal Statistical Office.

Fee, C. E., C. J. Hadlock, and J. R. Pierce. 2009. Investment, financing constraints, and internal capital markets: Evidence from the advertising expenditures of multinational firms. *The Review of Financial Studies* 22 (6): 2361–92.

Feld, L. P., and B. S. Frey. 2007. Tax compliance as the result of a psychological tax contract: The role of incentives and responsive regulation. *Law & Policy* 29 (1): 102–20.

Ferrell, A., H. Liang, and L. Renneboog. 2016. Socially responsible firms. *Journal of Financial Economics* 122 (3): 585–606.

Fombrun, C. J. 1996. *Reputation: Realizing value from the corporate image.* Boston, Mass. Harvard Business School Press.

Freeman, E. R. 1984. *Strategic management: A stakeholder approach.* Boston, Mass. [among others]: Pitman.

Frey, B. S., and B. Torgler. 2007. Tax morale and conditional cooperation. *Journal of Comparative Economics* 35 (1): 136–59.

Friedman, M. 1970. "A Friedman doctrine - The social responsibility of business is to increase its profits." *The New York Times Magazine*, 1970. Accessed May 15, 2021. Available at: https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html.

Gallemore, J., E. L. Maydew, and J. R. Thornock. 2014. The reputational costs of tax avoidance. *Contemporary Accounting Research* 31 (4): 1103–33.

Gangl, K., E. Hofmann, and E. Kirchler. 2015. Tax authorities' interaction with taxpayers: A conception of compliance in social dilemmas by power and trust. *New Ideas in Psychology* 37: 13–23.

Gangl, K., S. Muehlbacher, M. de Groot, S. Goslinga, E. Hofmann, C. Kogler, G. Antonides, and E. Kirchler. 2013. " How can I help you?" Perceived service orientation of tax authorities and tax compliance. *FinanzArchiv/Public Finance Analysis*, 487–510.

Garcia, J. 2016. The influence of corporate social responsibility on lobbying effectiveness: Evidence from effective tax rates. Working Paper. Available at: https://ssrn.com/abstract=2745506.

Gatzert, N. 2015. The impact of corporate reputation and reputation damaging events on financial performance: Empirical evidence from the literature. *European Management Journal* 33 (6): 485–99.

Giroud, X. 2013. Proximity and investment: Evidence from plant-level data. *The Quarterly Journal of Economics* 128 (2): 861–915.

Giroud, X., and H. M. Mueller. 2010. Does corporate governance matter in competitive industries? *Journal of Financial Economics* 95 (3): 312–31.

Global Reporting Initiative. 2019. *GRI 207: Tax 2019*. Accessed May 16, 2021. Available at: https://www.globalreporting.org/how-to-use-the-gri-standards/gri-standards-english-language/.

Godfrey, P. C. 2005. The relationship between corporate philanthropy and shareholder wealth: A risk management perspective. *Academy of Management Review* 30 (4): 777–98.

Godfrey, P. C., C. B. Merrill, and J. M. Hansen. 2009. The relationship between corporate social responsibility and shareholder value: An empirical test of the risk management hypothesis. *Strategic Management Journal* 30 (4): 425–45.

Goodman-Bacon, A. 2021. Difference-in-differences with variation in treatment timing. *Journal of Econometrics* 225 (2): 254–77.

Gordon, R. J. 1967. The incidence of the corporation income tax in US manufacturing, 1925-62. *The American Economic Review* 57 (4): 731–58.

Gouldner, A. W. 1960. The norm of reciprocity: A preliminary statement. *American Sociological Review*, 161–78.

Graham, J. R., M. Hanlon, T. Shevlin, and N. Shroff. 2014. Incentives for tax planning and avoidance: Evidence from the field. *The Accounting Review* 89 (3): 991–1023.

Haerpfer, C., R. Inglehart, A. Moreno, C. Welzel, K. Kizilova, J. Diez-Medrano, M. Lagos et al. 2020. *World Values Survey: Round Seven - Country-pooled datafile.* Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat. doi.org/10.14281/18241.1.

Hanlon, M., and J. Slemrod. 2009. What does tax aggressiveness signal? Evidence from stock price reactions to news about tax shelter involvement. *Journal of Public Economics* 93 (1-2): 126–41.

Harberger, A. C. 1962. The incidence of the corporation income tax. *Journal of Political Economy* 70 (3): 215–40.

Hardeck, I., J. W. Harden, and D. R. Upton. 2019. Consumer reactions to tax avoidance: Evidence from the United States and Germany. *Journal of Business Ethics* 170 (1): 75–96.

Hardeck, I., and R. Hertl. 2014. Consumer reactions to corporate tax strategies: Effects on corporate reputation and purchasing behavior. *Journal of Business Ethics* 123 (2): 309–26.

Hoi, C. K., Q. Wu, and H. Zhang. 2013. Is corporate social responsibility (CSR) associated with tax avoidance? Evidence from irresponsible CSR activities. *The Accounting Review* 88 (6): 2025–59.

Hong, Y., and M. L. Andersen. 2011. The relationship between corporate social responsibility and earnings management: An exploratory study. *Journal of Business Ethics* 104 (4): 461–71.

Hoopes, J. L., L. Robinson, and J. Slemrod. 2018. Public tax-return disclosure. *Journal of Accounting and Economics* 66 (1): 142–62.

Huang, H. H., Li Sun, and T. Yu. 2017. Are socially responsible firms less likely to expatriate? An examination of corporate inversions. *The Journal of the American Taxation Association* 39 (2): 43–62.

Huseynov, F., and B. K. Klamm. 2012. Tax avoidance, tax management and corporate social responsibility. *Journal of Corporate Finance* 18 (4): 804–27.

Im, T., W. Cho, G. Porumbescu, and J. Park. 2014. Internet, trust in government, and citizen compliance. *Journal of Public Administration Research and Theory* 24 (3): 741–63.

Inger, K., and J. Stekelberg. 2022. Valuation implications of socially responsible tax avoidance: Evidence from the electricity industry. *Journal of Accounting and Public Policy* 41 (4).

Inger, K. K., and B. Vansant. 2019. Market valuation consequences of avoiding taxes while also being socially responsible. *Journal of Management Accounting Research* 31 (2): 75–94.

IRS. 2020. The Agency, its mission and statutory authority. Available at: https://www.irs.gov/about-irs/the-agency-its-mission-and-statutory-authority.

Jacob, M., M. A. Müller, and T. Wulff. 2022. Do consumers pay the corporate tax? TRR 266 Accounting for Transparency Working Paper Series No. 15. Available at: https://ssrn.com/abstract=3468142.

Jakobsen, M., O. James, D. Moynihan, and T. Nabatchi. 2019. JPART virtual issue on citizen-state interactions in public administration research. *Journal of Public Administration Research and Theory* 29 (4): e8-e15.

Jensen, M. 2001. Value maximisation, stakeholder theory, and the corporate objective function. *European Financial Management* 7 (3): 297–317.

Jiang, J., I. Y. Wang, and K. P. Wang. 2019. Big N auditors and audit quality: New evidence from quasi-experiments. *The Accounting Review* 94 (1): 205–27.

Kim, S., and J. Lee. 2012. E-participation, transparency, and trust in local government. *Public Administration Review* 72 (6): 819–28.

Kim, Y., M. S. Park, and B. Wier. 2012. Is earnings quality associated with corporate social responsibility? *The Accounting Review* 87 (3): 761–96.

Kirchler, E. 2007. *The economic psychology of tax behaviour:* Cambridge University Press.

Kirchler, E., A. Niemirowski, and A. Wearing. 2006. Shared subjective views, intent to cooperate and tax compliance: Similarities between Australian taxpayers and tax officers. *Journal of Economic Psychology* 27 (4): 502–17.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez. 2011. Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica* 79 (3): 651–92.

Kountouris, Y., and K. Remoundou. 2013. Is there a cultural component in tax morale? Evidence from immigrants in Europe. *Journal of Economic Behavior & Organization* 96: 104–19.

KPMG. 2020. *The time has come: The KPMG survey of sustainability reporting 2020*. Accessed September 03, 2023. Available at: https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2020/11/the-time-has-come.pdf.

Krüger, P. 2015. Corporate goodness and shareholder wealth. *Journal of Financial Economics* 115 (2): 304–29.

Kubick, T. R., D. P. Lynch, M. A. Mayberry, and T. C. Omer. 2015. Product market power and tax avoidance: Market leaders, mimicking strategies, and stock returns. *The Accounting Review* 90 (2): 675–702.

Landry, S., M. Deslandes, and A. Fortin. 2013. Tax aggressiveness, corporate social responsibility, and ownership structure. *Journal of Accounting, Ethics & Public Policy* 14 (3): 611–45.

Lanis, R., and G. Richardson. 2012. Corporate social responsibility and tax aggressiveness: An empirical analysis. *Journal of Accounting and Public Policy* 31 (1): 86–108.

Lanis, R., and G. Richardson. 2015. Is corporate social responsibility performance associated with tax avoidance? *Journal of Business Ethics* 127 (2): 439–57.

Lanis, R., and G. Richardson. 2018. Outside directors, corporate social responsibility performance, and corporate tax aggressiveness: An empirical analysis. *Journal of Accounting, Auditing & Finance* 33 (2): 228–51.

Lanis, R., G. Richardson, C. Liu, and R. McClure. 2019. The impact of corporate tax avoidance on board of directors and CEO reputation. *Journal of Business Ethics* 160 (2): 463–98.

Lee, D. D., and R. W. Faff. 2009. Corporate sustainability performance and idiosyncratic risk: A global perspective. *Financial Review* 44 (2): 213–37.

Lee, Y., S. Ng, T. J. Shevlin, and A. Venkat. 2021. The Effects of Tax Avoidance News on Employee Perceptions of Managers and Firms: Evidence from Glassdoor. com Ratings. *The Accounting Review* 96 (3): 343–72.

Lev, B., C. Petrovits, and S. Radhakrishnan. 2010. Is doing good good for you? How corporate charitable contributions enhance revenue growth. *Strategic Management Journal* 31 (2): 182–200.

Lind, E. A. 2001. Fairness heuristic theory: Justice judgments as pivotal cognitions in organizational relations. *Advances in Organizational Justice* 56 (8).

Lins, K. V., H. Servaes, and A. Tamayo. 2017. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance* 72 (4): 1785–824.

Mackey, A., T. B. Mackey, and J. B. Barney. 2007. Corporate social responsibility and firm performance: Investor preferences and corporate strategies. *Academy of Management Review* 32 (3): 817–35.

May, P. J., and R. S. Wood. 2003. At the regulatory front lines: Inspectors' enforcement styles and regulatory compliance. *Journal of Public Administration Research and Theory* 13 (2): 117–39.

Mayberry, M. A., and L. Watson. 2021. Is Corporate Social Responsibility Related to Corporate Tax Avoidance? Evidence from a Natural Experiment. *The Journal of the American Taxation Association* 43 (1): 79–106.

McGee, R. W. 2010. Ethical issues in transfer pricing. *Manchester Journal of International Economic Law* 7 (2): 24–41.

McKee, M., C. A. Siladke, and C. A. Vossler. 2018. Behavioral dynamics of tax compliance when taxpayer assistance services are available. *International Tax and Public Finance* 25 (3): 722–56.

Mills, L. F. 1996. Corporate tax compliance and financial reporting. *National Tax Journal*, 421–35.

Minor, D., and J. Morgan. 2011. CSR as reputation insurance: Primum non nocere. *California Management Review* 53 (3): 40–59.

Muller, A., and A. Kolk. 2015. Responsible tax as corporate social responsibility: The case of multinational enterprises and effective tax in India. *Business & Society* 54 (4): 435–63.

Müller, R., C. Spengel, and H. Vay. 2020. On the determinants and effects of corporate tax transparency: Review of an emerging literature, ZEW - Centre for European Economic Research Discussion Paper No. 20-063. Available at: https://ssrn.com/abstract=3736747.

Neuman, S. S., T. C. Omer, and A. P. Schmidt. 2020. Assessing tax risk: Practitioner perspectives. *Contemporary Accounting Research* 37 (3): 1788–827.

OECD. 2007. *Improving taxpayer service delivery: Channel strategy development*. Available at: https://www.oecd.org/tax/administration/38528306.pdf.

OECD. 2013. *A practical guide to help revenue bodies better meet taxpayers' service expectations*. Available at: https://dx.doi.org/10.1787/9789264200821-en.

OECD. 2017. *Background brief: Inclusive framework on BEPS*. Accessed June 06, 2023. Available at: https://www.oecd.org/tax/beps/background-brief-inclusive-framework-on-beps.pdf.

OECD. 2019a. *Tax Administration 2019: Comparative information on OECD and other advanced and emerging economies.* Paris: OECD Publishing. Available at: https://doi.org/10.1787/74d162b6-en.

OECD. 2019b. *Tax Morale: What Drives People and Businesses to Pay Tax?* Paris: OECD Publishing. Available at: https://doi.org/10.1787/f3d8ea10-en.

OECD. 2020. *Tax Administration 3.0: The Digital Transformation of Tax Administration*. Available at: http://www.oecd.org/tax/forum-on-tax-administration/publications-and-products/.

Okafor, O. N., and J. Farrar. 2021. Punishing in the public interest: Exploratory Canadian evidence pertaining to convictions and incarcerations for tax offences. *Journal of Accounting and Public Policy*, 106848.

Parasuraman, A., V. A. Zeithaml, and L. L. Berry. 1988. SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing* 64 (1): 12–40.

Parasuraman, A., V. A. Zeithaml, and A. Malhotra. 2005. ES-QUAL: A multiple-item scale for assessing electronic service quality. *Journal of Service Research* 7 (3): 213–33.

Pew Research Center. 2015. Federal tax system seen in need of overhaul. Top complaints: Wealthy, corporations 'don't pay fair share'. Accessed May 15, 2021. Available at: https://www.pewresearch.org/politics/2015/03/19/federal-tax-system-seen-in-need-of-overhaul/.

Preuss, L. 2010. Tax avoidance and corporate social responsibility: You can't do both, or can you? *Corporate Governance* 10 (4): 365–74.

Preuss, L. 2012. Responsibility in paradise? The adoption of CSR tools by companies domiciled in tax havens. *Journal of Business Ethics* 110 (1): 1–14.

Puzzanghera, J. 2014. "Burger King, Warren Buffett under fire for Canadian inversion deal." *Los Angeles Times*, 2014. Accessed May 15, 2021. Available at: https://www.latimes.com/business/la-fi-burger-king-tim-hortons-boyctott-warren-buffett-20140826-story.html.

PwC. 2004. *Tax risk management*. Accessed May 15, 2021. Available at: https://www.pwc.co.za/en/assets/pdf/tax-risk-management-guide.pdf.

PwC. 2013. *Tax strategy and corporate reputation: A business issue*. Accessed June 24, 2021. Available at: https://www.pwc.com/gx/en/tax/publications/assets/pwc-tax-strategy-and-corporate-reputation-a-tax-issue-a-business-issue.pdf.

Refinitiv. 2021. Environmental, social and governance (ESG) scores from Refinitiv. Accessed September 17, 2021. Available at: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf.

Reimann, M., U. F. Lünemann, and R. B. Chase. 2008. Uncertainty avoidance as a moderator of the relationship between perceived service quality and customer satisfaction. *Journal of Service Research* 11 (1): 63–73.

Rhee, M., and P. R. Haunschild. 2006. The liability of good reputation: A study of product recalls in the US automobile industry. *Organization Science* 17 (1): 101–17.

Rhee, M., and M. E. Valdez. 2009. Contextual factors surrounding reputation damage with potential implications for reputation repair. *Academy of Management Review* 34 (1): 146–68.

Rindova, V. P., and L. L. Martins. 2012. Show me the money: A multidimensional perspective on reputation as an intangile asset. In *The Oxford handbook of corporate reputation*. Edited by M. L. Barnett and T. G. Pollock: Oxford University Press.

Rudyanto, A., and K. Pirzada. 2020. The role of sustainability reporting in shareholder perception of tax avoidance. *Social Responsibility Journal* 17 (5): 669–85.

Rupert, T. J., and C. M. Fischer. 1995. An empirical investigation of taxpayer awareness of marginal tax rates. *Journal of the American Taxation Association* 17: 36.

Servaes, H., and A. Tamayo. 2013. The impact of corporate social responsibility on firm value: The role of customer awareness. *Management Science* 59 (5): 1045–61.

Shiu, Y.-M., and S.-L. Yang. 2017. Does engagement in corporate social responsibility provide strategic insurance-like effects? *Strategic Management Journal* 38 (2): 455–70.

Sikka, P. 2010. Smoke and mirrors: Corporate social responsibility and tax avoidance. *Accounting Forum* 34 (3-4): 153–68.

Slemrod, J. 2010. Old George Orwell got it backward: some thoughts on behavioral tax economics. *FinanzArchiv/Public Finance Analysis* 66 (1): 15–33.

Snipes, R. L., S. L. Oswald, M. LaTour, and A. A. Armenakis. 2005. The effects of specific job satisfaction facets on customer perceptions of service quality: An employee-level analysis. *Journal of Business Research* 58 (10): 1330–39.

Social Investment Forum. 2020. *Report on US sustainable, responsible and impact investing trends 2020*. Accessed May 01, 2023. Available at: https://www.ussif.org/files/Trends%20Report%202020%20Executive%20Summary.pdf.

Stock, J. H., and M. W. Watson. 2015. *Introduction to econometrics:* Pearson.

Sun, L., and S. Abraham. 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics* 225 (2): 175–99.

Thomsen, M., and C. Watrin. 2018. Tax avoidance over time: A comparison of European and US firms. *Journal of International Accounting, Auditing and Taxation* 33: 40–63.

Thomson Reuters. 2017. Thomson Reuters ESG Scores. Accessed November 06, 2022. Available at: https://www.esade.edu/itemsweb/biblioteca/bbdd/inbbdd/archivos/Thomson_Reuters_ESG_Scores.pdf.

Tyler, T. R., and E. A. Lind. 1992. A relational model of authority in groups. In *Advances in experimental social psychology*. Vol. 25, 115–91: Elsevier.

Vossler, C. A., and M. McKee. 2017. Efficient tax reporting: The effects of taxpayer liability information services. *Economic Inquiry* 55 (2): 920–40.

Wahab, N. S. A., and K. Holland. 2012. Tax planning, corporate governance and equity value. *The British Accounting Review* 44 (2): 111–24.

Wang, Z., Q. E. Yin, and L. Yu. 2021. Real effects of share repurchases legalization on corporate behaviors. *Journal of Financial Economics* 140 (1): 197–219.

Wans, N. 2020. Corporate social responsibility and market-based consequences of adverse corporate events: Evidence from restatement announcements. *Journal of Accounting, Auditing & Finance* 35 (2): 231–62.

Watson, L. 2015. Corporate social responsibility, tax avoidance, and earnings performance. *The Journal of the American Taxation Association* 37 (2): 1–21.

Weber Shandwick/KRC Research. 2006. *Safeguarding reputation*.

Welch, E. W., C. C. Hinnant, and M. J. Moon. 2005. Linking citizen satisfaction with e-government and trust in government. *Journal of Public Administration Research and Theory* 15 (3): 371–91.

Wilde, J. H., and R. J. Wilson. 2018. Perspectives on corporate tax planning: Observations from the past decade. *The Journal of the American Taxation Association* 40 (2): 63–81.

Wooldridge, J. M. 2016. *Introductory econometrics: a modern approach.* 6th ed. Boston: Cengage Learning.

Zeng, T. 2016. Corporate social responsibility, tax aggressiveness, and firm market value. *Accounting Perspectives* 15 (1): 7–30.