
Corporate Misconduct and the Impact of Market Forces, Regulatory Change, and Auditor-Provided Services



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Am Fachbereich Rechts- und Wirtschaftswissenschaften
der Technischen Universität Darmstadt

angenommene

Dissertation

vorgelegt von

Christian Friedrich, M. Sc.

geboren am 16.10.1991 in Lindenfels

zur Erlangung des akademischen Grades
Doctor rerum politicarum (Dr. rer. pol.)

Erstgutachter: Prof. Dr. Reiner Quick
Zweitgutachter: Prof. Dr. Dirk Schiereck

Darmstadt 2020 (D17)

Friedrich, Christian: *Corporate Misconduct and the Impact of Market Forces, Regulatory Change, and Auditor-Provided Services*

Darmstadt, Technische Universität Darmstadt.

Jahr der Veröffentlichung der Dissertation auf TUprints: 2021

Tag der mündlichen Prüfung: 27.01.2021

Veröffentlicht unter CC BY-NC-SA 4.0 International

<https://creativecommons.org/licenses/>

Acknowledgement

Although my vita seems straightforward on paper, my journey to become an audit researcher has taken many unsuspected turns. I learned that this is common among audit researchers, making it even more fun to be part of this community. It goes without saying that such a journey would not be possible without mentors and the support of colleagues, family, and friends.

Coming from a non-academic family, I first have to mention the utmost support of my parents. They motivated me with interest rather than any pressure from the very beginning of my achievements as a student, although they had as little of a clue of what was going on when I entered the academic world as I had in the beginning.

Having developed a general, at this time still vague, fascination for knowledge and science at school age, I was dreaming of a doctoral degree when I progressed as a student. I had the idea that it must be hard to take on the challenge of a PhD. However, my supervisor and mentor, Reiner Quick, who introduced me to accounting, auditing, and the international academic community, did everything to make it as comfortable as it gets. With his strong understanding of the importance of finding the right balance of comfort and commitment and his unwavering integrity of being a scholar serving the public, he provides his doctoral students a secure and free environment to conduct their studies and develop their own research. When I finished my bachelor's degree, he already reached out to show me a potential path forward. This gave me plenty of opportunities to prepare and redevelop my interests and skills, time after time. Although never putting too much pressure on my progress, he pushed me to become involved in the academic community. He introduced me to conferences, international contacts, and the national and international publication game. My dream of earning a doctoral degree became a job, but being a young auditing scholar became a lot of fun. Professor Quick's support and his strong international network finally allowed me to end my journey of becoming a PhD with a semester at the University of Florida, visiting Robert Knechel at the Fisher School of Accounting.

The completion of my PhD would also not have been possible without the support of my co-referee, Dirk Schiereck. From being a great professor and moderator when I was still a student to giving great feedback in PhD seminars, he has always been very

likeable, thoughtful, and helpful. I also want to thank my committee members, professors Carolin Bock, Peter Buxmann, and Christoph Glock, for investing their time and effort for me to finish the process. Although not directly involved, Anette von Ahsen has always been a great professor and colleague at the same time.

Finally, the best thing about receiving a PhD from the Fachgebiet Rechnungswesen, Controlling, and Wirtschaftsprüfung was certainly my great colleagues. First and foremost, I am sure all would agree that Erika is vital for the smooth functioning of our team, for being hydrated, for an occasional drink, for the best lunches, for great cake, for a good talk any time, and, obviously, for lots of fun. Still, working with Christian, Nico, Sanjar, and Kevin, along with Dani, Dominik, Flo, Julia, and Viola, has made the job that goes along with being a PhD a pleasure. I also enjoyed working with our well-picked student assistants and want to thank Luca for becoming my backup when I am on parental leave. Lastly, I want to thank Vic to become my first close international colleague, co-author, and now a friend.

Although my stay in Florida is only some moth ago, Robert became a second mentor to me. He pushed me to think harder, showed me the flaws in my thinking and how to improve them, and always had an open ear for ideas. My gratitude also goes out to Jenny Tucker. In her accounting PhD seminar, she probably taught me as much about the financial accounting literature in one semester as I would have ever learned without taking part. Moreover, I want to thank a great group of PhD students that accepted my fellow international visitor Fynn and me as a part of them. Thanks for a great time, Brianna, Cicone, Diana, Jeffrey, Jenna, Justin, Krupa, Mark, and Q (special shout-out to Q for helping us cope with Covid psychologically and finally catch a flight to go home). Finally, I was fortunate to meet my office buddy, room-mate, and sparring partner, Fynn. I hope our paths cross many more times along the way in our (hopefully upcoming) academic careers.

I will end my acknowledgments with those to whom they belong the most, my family and friends. My development as a person was greatly supported by my parents Marion and Hartmut, from my first day, and my sister and friend Larissa. I owe you everything. Since I met my wife, Lisa, my family-in-law has become a second home. Thank you to all the Wehrmeisters for supporting me as a family member. My earliest companions,

my friends from my hometown Beerfurth just have the most reason and unconditional kindness I can imagine. They will always make me remember what is most important in life. Being a student of the Martin-Luther-Schule in Rimbach, I met the most inspiring friends that have all gone very different ways but never ceased to inspire each other. Finally, my Wirtschaftsingenieur buddies are the smartest people I have ever met. Special thanks are devoted to my dear friend and competitive villain Max Bauer, who never ceases to be a partner to push each other for the highest achievements in all of life and has not shied away from reading this dissertation as an outsider to provide valuable feedback.

Most importantly, the greatest gratitude of all has to go to my family, my beloved wife, Lisa, and most beautiful daughter, Charlotte Lilly. After all, of all people I have ever met in life, only Charlotte has, in only three months, managed to make me finish a dissertation. The support Lisa gave me along the way is simply invaluable. She backed me up for countless hours of extra work, many deadlines and promises I did not meet, and listened to all the crazy academic nobody could understand because it made no sense. I love you for all you do for me and all you are. Of course, this dissertation is for you.

Abstract

My dissertation addresses the fundamental economic question of whether regulation that restricts free markets can reduce corporate misconduct. I investigate different aspects of this issue in four related papers. They use different theories that all relate to basic economic decision-making theory. If the expected benefits of rule-abiding behavior or the expected costs of misconduct increase, decision-makers tend to develop a greater preference for rule-abiding behavior over misconduct and vice versa. Conversely, if the expected benefits of misconduct or the expected costs of rule-abiding behavior increase, decision-makers prefer misconduct. I study three forces, for which I argue that they may influence the balance of these four groups of benefits and costs. The two ends of a spectrum of regulation, regulatory changes and free market forces, can each shape the expectations of the decision-maker. As a third force, the auditor as a monitor in markets with information asymmetries can curb misconduct emerging from the abuse of these asymmetries. Because both regulation and free markets also shape auditor behavior, auditor-provided services are an indirect, third channel, through which the first two forces also affect behavior.

In my first paper, I conduct an experiment on perceptions of managers regarding misconduct punishment. I ask 71 board members of German listed companies about their expectations of detection likelihood and punishment severity upon detection. My findings suggest that detection perceptions are larger than estimates of factual detection likelihood from prior literature. However, managers underestimate both individual and organizational punishment severity. In an experiment with only 28 of the board members, I further find that teaching them about factual losses in market value for fraudulent firms decreases their tendency to commit misconduct. However, this is not the case when I also teach them about factual individual consequences for the perpetrator. In conclusion, I find that existing market forces are surprisingly strong, but biased perceptions make them ineffective.

The second paper is an in-depth qualitative study of how non-financial firms install anti-money laundering systems. Based on interviews in eight multinational organizations, I develop a Grounded Theory of the implementation process. A key takeaway is that the implementation benefits from clear regulatory structure and guidance. While I find that

similar guidance can emerge among peers in the free market, I find some clear indications that regulation is important for effective systems designed to curb misconduct in this case.

My third paper contains another experiment. It studies reactions of 128 private investors to a company that suspects misconduct internally when internal investigation and public detection are certain. The company can either obtain a credibility signal from using forensic services from a Big 4 to investigate or investigate with internal audit. The issue and investigation become public either by self-disclosure of the company or a press article. I find that, when the company engages the forensic services, self-disclosure improves investor perceptions over press disclosure. With the internal audit investigation, this is not the case. In the context of this dissertation, the unregulated market for forensic accounting can thus provide an improvement for investors and arguably the public (through greater transparency) under the strong assumptions of certain investigation and detection.

The fourth and final paper contains an archival analysis of two non-audit service scandals of KPMG. I aim to study whether, through reputation effects, the negative signals from these scandals spill over to the audit practice of KPMG. If this were the case, it would suggest that some market forces are strong enough to deter misconduct even across the different domains in which a company is active. In a short-term event study, I find abnormal negative returns for the average KPMG audit client. However, in longer-term analysis, I find that KPMG does not significantly lose clients or gain fewer clients in the first year after the scandals. Similarly, audit fees do not change. However, for the second scandal, there is some evidence that audit quality suffers. In conclusion, I infer a small reputation spillover from my results that has no substantial economic consequences for KPMG.

Zusammenfassung

In meiner Dissertation beschäftige ich mich mit der fundamentalen ökonomischen Fragestellung, ob Regulierung, welche freie Märkte einschränkt, Wirtschaftskriminalität verhindern kann. Ich beleuchte verschiedene Aspekte dieser Fragestellung in vier miteinander verbundenen Artikeln. Diese nutzen verschiedene Theorien, die alle auf die grundlegende ökonomische Entscheidungstheorie zurückgehen. Sofern der erwartete Nutzen von regelkonformem Verhalten oder die erwarteten Kosten von Wirtschaftskriminalität steigen, neigen Entscheidungsträger dazu, eine stärkere Präferenz für regelkonformes Verhalten zu entwickeln. Im Gegensatz dazu bevorzugen Entscheidungsträger Wirtschaftskriminalität, wenn der erwartete Nutzen von Wirtschaftskriminalität oder die erwarteten Kosten von regelkonformem Verhalten steigen. Ich untersuche drei Einflussfaktoren, für welche ich Argumente vorbringe, dass sie das Gleichgewicht dieser vier Arten von Nutzen und Kosten beeinflussen können. Als zwei Enden des Spektrums möglicher Regulierung können Regulierungsänderungen und freie Marktkräfte jeweils die Erwartungen des Entscheidungsträgers ändern. Als dritter Einflussfaktor kann der Wirtschaftsprüfer als Überwachungsinstitution in Märkten mit Informationsasymmetrien Wirtschaftskriminalität verhindern, die solche Informationsasymmetrien ausnutzt. Da sowohl Regulierungsänderungen als auch freie Marktkräfte das Verhalten des Wirtschaftsprüfers prägen, sind Leistungen des Wirtschaftsprüfers ein indirekter, dritter Weg, durch den die ersten beiden Einflussfaktoren Verhalten beeinflussen.

Mein erster Artikel enthält ein Experiment über die Wahrnehmungen von Vorständen bezüglich Wirtschaftskriminalität und deren Bestrafung. Ich befrage 71 Vorstände börsennotierter deutscher Unternehmen zu ihren Erwartungen bezüglich der Wahrscheinlichkeit, dass Wirtschaftskriminalität entdeckt wird, und der Höhe der Strafe bei Entdeckung. Meine Ergebnisse deuten darauf hin, dass die Vorstände aus bestehender Forschung entnommene tatsächliche Entdeckungswahrscheinlichkeiten überschätzen. Allerdings unterschätzen die Vorstände die Höhe von Strafen auf individueller und Unternehmensebene. Im anschließenden Experiment mit nur noch 28 Vorständen finde ich heraus, dass die Neigung der Vorstände zu Wirtschaftskriminalität dadurch gesenkt werden kann, dass sie über die tatsächlich drohenden Strafen auf Unternehmensebene aufgeklärt werden. Dieser Effekt verschwindet, wenn die

Vorstände zusätzlich über die tatsächlich drohenden individuellen Strafen aufgeklärt werden. Zusammenfassend finde ich überraschend wirkungsvolle Marktkräfte, die allerdings ihre volle Wirkung aufgrund verzerrter Wahrnehmungen in der Praxis nicht entfalten.

Der zweite Artikel enthält eine tiefgründige qualitative Untersuchung der Einführung von Geldwäschepräventionssystemen in Industrieunternehmen. Ausgehend von Interviews mit acht internationalen Konzernen entwickle ich eine „Grounded Theory“ des Einführungsprozesses. Eine wesentliche Erkenntnis ist, dass die Implementierung von klaren regulatorischen Strukturen und Anwendungshinweisen profitiert. Obwohl ich herausfinde, dass ähnliche Strukturen auch innerhalb von Branchennetzwerken ohne Regulierung entstehen können, gibt es klare Hinweise darauf, dass Regulierung wichtig für die Entstehung effektiver Systeme zur Verhinderung von Wirtschaftskriminalität ist.

Mein dritter Artikel basiert erneut auf einem Experiment. Es untersucht die Reaktionen von 128 Privatinvestoren auf ein Unternehmen mit einem Betrugsverdacht unter der Bedingung, dass eine interne Untersuchung sowie eine öffentliche Berichterstattung darüber sicher sind. Das Unternehmen kann sich für ein Glaubwürdigkeitssignal durch die Beauftragung von Forensic Services einer Big 4 für die interne Untersuchung entscheiden oder für eine interne Untersuchung durch die Interne Revision. Der Verdacht und die Untersuchung werden entweder durch das Unternehmen selbst oder von der Presse öffentlich berichtet. Ich finde heraus, dass sich die Wahrnehmung der Investoren von Unternehmen, die Forensic Services beauftragen, verbessert, wenn diese die Veröffentlichung selbst vornehmen und nicht die Presse zuvorkommen lassen. Dies ist nicht der Fall, wenn die Interne Revision die Untersuchung durchführt. Im Rahmen dieser Dissertation kann dies als Indiz dafür verstanden werden, dass der unregulierte Markt für Forensic Services unter den engen Annahmen einer sicheren Untersuchung und Berichterstattung einen positiven Beitrag zur Investorenwahrnehmung und möglicherweise, durch erhöhte Transparenz, zum öffentlichen Wohl beitragen kann.

Der vierte und letzte Artikel berichtet eine Archivstudie von zwei Skandalen im Rahmen von Nichtprüfungsleistungen von KPMG. Ich untersuche ob sich die negativen Signale dieser Skandale aufgrund von Reputationseffekten auf die nicht direkt betroffene

Abschlussprüfungspraxis von KPMG auswirken. Sofern dies der Fall wäre, würde dies darauf hindeuten, dass Marktkräfte in nur einem Bereich der Tätigkeit von Organisationen in solchen Fällen stark genug sind, um Wirtschaftskriminalität in allen Tätigkeitsbereichen dieser Organisationen zu verhindern. Eine kurzfristige Eventstudie ergibt negative abnormale Renditen für den durchschnittlichen Prüfungsmandanten von KPMG. Allerdings zeigt die Analyse längerfristiger Effekte, dass KPMG im jeweiligen Jahr nach den Skandalen keine signifikanten Mandantenverluste oder geringere Mandantengewinne hinnehmen muss. Genauso wenig ändern sich die Prüfungshonorare. Jedoch gibt es Anzeichen darauf, dass die Prüfungsqualität nach dem zweiten Skandal sinkt. Insgesamt schlussfolgere ich aus den Ergebnissen, dass es einen kleinen Reputationseffekt über Tätigkeitsbereiche hinweg gibt, der keine substantiellen ökonomischen Konsequenzen für KPMG hat.

Table of Contents

Acknowledgement	III
Abstract	VI
Zusammenfassung	VIII
Table of Contents	XI
List of Tables	XIII
List of Figures	XIV
List of Abbreviations	XV
Chapter 1 Introduction	1
1.1 Motivation and Research Question	1
1.2 Theoretical Foundations	4
1.3 Thesis Structure and Synopsis.....	12
Chapter 2 The Deterrence Effect of Consequences of Fraud to Individuals and to Companies – Perception of Executive Board Members [Abstract only]	22
Chapter 3 An Analysis of Anti-Money Laundering in the German Non-Financial Sector [Abstract only]	23
Chapter 4 Private Investigations and Self-Disclosure of Suspected Fraud: Experimental Evidence	24
4.1 Introduction.....	25
4.2 Theoretical Background and Hypothesis Development	29
4.2.1 Engaging External Specialists.....	29
4.2.2 Self-Disclosure and Hypothesis Development	32
4.3 Method	34
4.3.1 Participants	34
4.3.2 Experimental Procedure	35
4.3.3 Treatment Variables	36
4.3.4 Dependent Variable and Additional Variables	37
4.4 Results	38
4.4.1 Manipulation and Comprehension Checks.....	38
4.4.2 Test of Hypothesis	39
4.4.3 Additional Analyses.....	41

4.5	Conclusion	45
4.6	Appendix A: Main body of the research instrument (screenshots)	48
Chapter 5 Do Non-Audit Service Failures Impair Auditor Reputation? An Analysis of KPMG Advisory Service Scandals in Germany		57
5.1	Introduction.....	58
5.2	Background and Hypothesis Development.....	62
5.2.1	Theory on Incentives for Audit Quality and its Relation to Non-Audit Services.....	62
5.2.2	Prior Research.....	64
5.2.3	German Setting	67
5.2.4	Timeline of Events.....	69
5.3	Short-Term Event Study	74
5.3.1	Models	75
5.3.2	Sample Selection.....	77
5.3.3	Results	79
5.3.4	Robustness checks.....	81
5.4	Longer-Term Effects on KPMG's Audit Practice	86
5.4.1	Models	87
5.4.2	Sample Selection.....	90
5.4.3	Results	91
5.4.3.1	Descriptive Statistics	91
5.4.3.2	Multiple Regression Analyses.....	96
5.4.4	Additional Analyses.....	100
5.5	Conclusion.....	103
5.6	Appendix B: Variable definitions.....	106
Chapter 6 Overall Conclusion		108
6.1	Contributions.....	108
6.2	Limitations and Directions for Future Research	109
References.....		112

List of Tables

Table 1: Investigated forces and related findings of my four research papers.....	13
Table 2: Results of ANOVA analyzing the difference in willingness to invest	40
Table 3: Mediation analyses	43
Table 4: Sample selection	78
Table 5: Market reaction to NAS failure events for German KPMG clients	80
Table 6: Robustness checks of market reaction to NAS failure events	82
Table 7: Descriptive statistics for additional analyses	92
Table 8: Logistic regression of auditor change	96
Table 9: OLS-regression of the natural logarithm of Audit Fees	97
Table 10: OLS-regression of accruals-based earnings management.....	99
Table 11: Robustness checks	101
Table 12: Variable definitions for Chapter 5.....	106

List of Figures

Figure 1: Theoretical background of the dissertation and research papers.	5
Figure 2: Predicted and observed effects of disclosure and investigation on willingness to invest.	34
Figure 3: Timeline of events.	72

List of Abbreviations

AICPA	American institute of certified public accountants
AktG	Aktiengesetz (German Stock Corporation Act)
AML	Anti-money laundering
ANOVA	Analysis of variance
ANCOVA	Analysis of covariance
AReG	Abschlussprüfungsreformgesetz (Auditing Reformation Act)
BGB	Bürgerliches Gesetzbuch (German Civil Code)
CDAX	Composite DAX
CSR	Corporate social responsibility
DAX	Deutscher Aktienindex (German stock index)
DGAP	Deutsche Gesellschaft für Ad-hoc-Publizität (German Society for Ad-hoc Publicity)
DPA	Deferred prosecution agreement
DV	Dependent variable
EBIT	Earnings before interest and taxes
e.g.	Exempli gratia (for example)
et al.	Et alii (and others)
EU	European Union
GBP	Great Britain Pounds
HGB	Handelsgesetzbuch (German Commercial Code)
HTML	Hypertext markup language
ISA	International standards of auditing
ISIN	International Securities Identification Number
i.e.	Id est (that is)
NAS	Non-audit services

Non-FS	Non-financial sector
NPA	Non-prosecution agreement
OLS	Ordinary least squares
PCAOB	Public company accounting oversight boards
pp.	Pages
RQ	Research question
SEC	Securities exchange commission
SIC	Standard industrial classification
Std. deviation	Standard deviation
UK	United kingdom
US	United states
VIF	Variance inflation factor
vs.	Versus

Chapter 1 Introduction

1.1 Motivation and Research Question

A fundamental problem in economics is why and how regulation is introduced into free markets to achieve the highest possible welfare (Posner, 1974). Regulation may be necessary when information asymmetries hinder efficient resource allocation in free markets and instead provide incentives for opportunistic, fraudulent behavior (Akerlof, 1970; Darby and Karni, 1973). Such theory has even modeled the regulator as an economic actor motivated by self-interest. Under information asymmetry, the optimal policy is obscure, and how the regulator chooses regulation and whether these choices are efficient becomes an equally difficult question (Downs, 1957; Posner, 1974).

This problem is central to accounting and auditing research, a field that borrows from economic theory, is characterized by information asymmetries, and is heavily regulated (Cooper and Robson, 2006; Knechel, 2016; Watts and Zimmerman, 1978). From its very beginning, accounting research has acknowledged the difficulty of answering whether regulation or market forces arrive at better solutions to reduce information asymmetry.¹ One argument in favor of regulation in this literature is that firms may fraudulently misrepresent accounting information in free markets. This issue of corporate misconduct has seen ongoing debates and uptake in regulatory activity (Coffee, 2019; Flesher and Flesher, 1986; Shibano, 1990; Zimbelman, 1997). However, the literature is critical about regulation, and early literature has found little empirical support for its effectiveness (Ball, 2009; Benston, 1969; Knechel, 2016). Recent auditing literature expresses deep concern that additional regulation disregarding the economic forces in the regulated market may have unintended negative consequences (Coffee, 2019; Knechel, 2016). In an attempt to contribute to this, theoretically and empirically, critical research stream, my dissertation covers multiple facets of the following research question:

¹ Throughout this dissertation, I use the term *regulation* to refer to regulatory change that originates from public authorities. The opposing term *market forces* describes all behavior, whether it follows explicit or implicit rules or no rules at all, that results from the free interaction of private actors. Market forces can create private rules, such as professional standards. I do not consider them as regulations as long as a public authority does not govern them. My definition of misconduct considers the violation of all explicit rules (see below). As such, it covers the breach of regulation and private rules that emerged from market forces.

RQ: Does regulation, or do free market forces, effectively reduce corporate misconduct?

To answer this question, I first need to define corporate misconduct. For the purpose of this dissertation, it is any behavior by an organization or individuals who act as a member of an organization that violates explicit rules, e.g., laws and privately agreed-upon rules (similar but broader in Singleton and Singleton, 2010). This definition is anything but obvious, and society may arrive at different definitions across space and time, for instance, because societal norms or technologies change (Cooper et al., 2013; Morales et al., 2014). Consequently, explicit rules changes, along with possible antecedents, behavioral schemes, incentives, investigation and detection, and consequences of their violation. Moreover, a reduction of misconduct is not always desirable. It comes at a cost and may have unintended consequences (Knechel, 2016). Therefore, not all prevention of misconduct will improve economic welfare. An effective regulatory regime may, hence, still experience single instances of misconduct (Camfferman and Wielhouwer, 2019).

If misconducts pile up, markets arguably break down, and action is required to repair their credibility and efficiency. History has shown repeated cycles of market breakdowns that led to regulation that, in turn, could not prevent the subsequent market breakdown, and so on (Jamal, 2008). In accounting and auditing, one of the most famous examples in recent history is the Enron fraud. It resulted in the demise of one of the then biggest companies in the United States (US) and its auditor Arthur Andersen. Suddenly, the Big 5 audit firms were only the Big 4. In response, US regulators enacted the Sarbanes-Oxley Act, increasing accounting and auditing regulation and regulatory scrutiny substantially (Knechel, 2016). Similar market breakdowns happened worldwide, for instance, the burst of the *Neuer Markt* in Germany with the *ComROAD* fraud as its infamous climax (Weber et al., 2008). The 2008 global financial crisis followed in close succession.

Since then, the regulator in the European Union (EU) has also increased accounting and auditing regulation drastically with the 2014 EU audit regulation (Regulation (EU) No. 537/2014). Importantly, in this dissertation's context, the EU audit regulation combines two different approaches: It contains strict regulations on conduct, such as auditor rotation or the prohibition of certain auditor-provided services, but it also

introduces an expanded auditor report in an attempt to reduce information asymmetry. While the latter is technically a regulation, it does not restrict supply and demand in the market for auditor-provided services. Instead, it relies on a combination of unrestricted market forces and less information asymmetry to promote proper behavior.

The latter mechanism is present in many other recent regulatory initiatives to mitigate corporate misconduct beyond egregious accounting fraud. They have in common that they privatize the detection, investigation, and mitigation of misconduct. From an agency perspective, they impose additional responsibilities on market participants that may increase monitoring and, hence, reduce information asymmetries between private organizations, potential perpetrators, and ultimately the public (Andrade, 2015). Two examples are the EU money laundering directive (Directive (EU) 2018/1673), which has been reformed five times in the last 20 years, and the recently enacted EU whistleblowing directive (Directive (EU) 2019/1937), which majorly encourages private detection, investigation, and mitigation of corporate misbehavior.

Despite all these changes, existing regulation and market forces have not effectively prevented significant scandals. The major fraud and bankruptcy case of *Wirecard* is only the latest and most prominent. As typical in such situations, the media and political actors have quickly reacted and proposed regulatory reform. The profession has been equally quick to share its views, and processes of change have begun, for instance, with the recent draft for a new law on improving the integrity of financial markets (Finanzmarktintegritätsstärkungsgesetz). However, academics caution these actors not to overreact, as has arguably been observed with earlier scandals, but to base their decisions on empirical evidence. To date, such empirical evidence is incomplete (Sellhorn, 2020).

Given this introductory discussion, it becomes clear that there is a continuum of regulatory intervention, from rigid regulation to subtle changes in market conditions that allow market forces to unfold. In addition, residual risk is unavoidable, and it is an empirical economic question which scandal might not be worth the cost of prevention (Camfferman and Wielhouwer, 2019). Under these fundamental economic considerations and given recent calls for evidence-based regulation, it is empirically unclear in various contexts where regulation is effective and where existing economic

forces, on average, already effectively reduce misconduct. This fundamental question continues to draw substantial research attention in auditing, accounting, and related disciplines (Camfferman and Wielhouwer, 2019; Leuz, 2018; Sellhorn, 2020). Single researchers can only provide pieces to the puzzle answering this bigger question, and those pieces may change along with the underlying issue across time and space (Cooper et al., 2013; Morales et al., 2014). In the next section, I outline my dissertation's overall theoretical foundations and the detailed theory I use in my four papers. Subsequently, I describe the pieces those four papers aim to add to answering my overarching research question and reveal what market forces, regulation, and the auditor have to offer to reduce misconduct.

1.2 Theoretical Foundations

Throughout this dissertation, I use several theories that all originate from basic microeconomic considerations. They cover different aspects of economic decision-making in the context of corporate misconduct. Ultimately, they aim to explain what determines a decision-maker's choice between rule-abiding behavior and misconduct.² Figure 1 summarizes the basic underlying theory and depicts how the single theories I use in the four papers align in this context. Below, I first explain the four basic building blocks of my theory, the benefits and costs of regulation-abiding behavior and misconduct, and how they form the fundamental decision-making problem I study. Then, I explain in more detail how each theory relates to one or more of these building blocks. Each box other than the four building blocks in the lower part of Figure 1

² As introduced above, I base my definition of corporate misconduct on explicit rules of proper conduct. As such, rule-abiding behavior differs slightly from ethical behavior, although both change across time and space (Cooper et al., 2013; Singleton and Singleton, 2010). Ethical behavior is broader and harder to define. A simple definition is that ethical behavior protects the integrity of the actor. (Lewis, 1985). Therefore, one can imagine rule-abiding behavior that is unethical if it threatens the integrity of the actor but does not violate any rule. An example is the abuse of tax benefits within legal limits but to an extent that the majority of the public views as egregious. This distinction is important to interpret the forces I study as potential promoters of rule-abiding conduct. By construction, regulators and auditors will always work with the boundaries of codified regulation because they have no mandate to restrict conduct within these boundaries. Market forces, however, can gravitate to any boundaries that will be economically beneficial for market participants. However, to the extent that certain behavior that society views as ethical has no utility in those markets, but a violation of this behavior has a positive utility, market forces will not curb such behavior. In conclusion, my definition of misconduct is the broadest definition that the forces I study fully address. The important question of how to align rule-abiding and ethical behavior to reduce behavior that is rule-abiding, but unethical is beyond the scope of my dissertation.

represents one theory.³ Note that the first and second papers (Chapter 2 and Chapter 3) use one theory each while I discuss multiple theories in the third and fourth papers (Chapter 4 and Chapter 5). The relations of the theories with the building blocks are the solid lines. The relations can be positive, negative, or ambiguous. Combining all theories leads to a situation in which there are competing theoretical arguments in some instances. As will be seen below, this provides tension to my empirical analysis. Next to depicting my theory, the dotted arrows in Figure 1 show how the papers and the employed theories align with the three forces I study in the context of my overarching research question (top part of Figure 1). I explain the relationship of the three forces with my research question and the four papers in the next section, which discusses the structure of the thesis and provides a contextualized synopsis of all four papers.

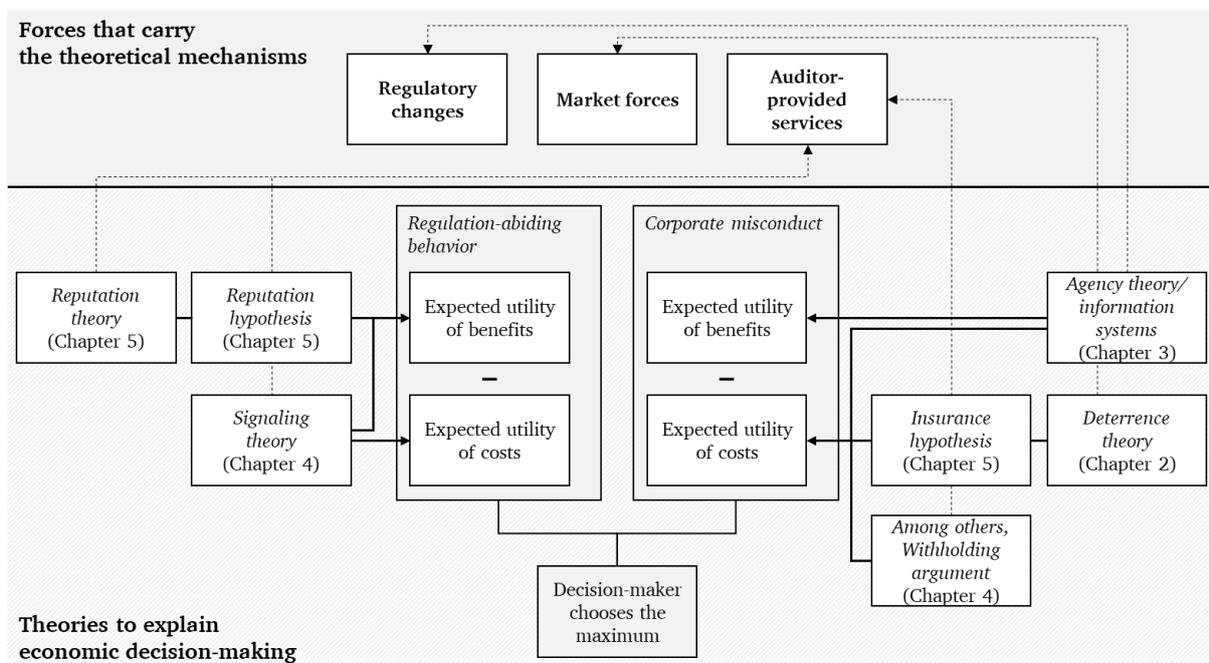


Figure 1: Theoretical background of the dissertation and research papers.

Economically, the choice between regulation-abiding behavior and misconduct is straightforward. Both options have a set of expected benefits and a set of expected costs. All benefits and costs are subject to some degree of uncertainty, and the decision-maker receives some positive (for benefits) or negative (for costs) utility from all possible outcomes. The decision-maker forms expectations about the likelihood of all possible outcomes and the individual utility if the outcomes occur. Each expected utility of

³ The arrangement of the boxes has no meaning.

benefits or costs, respectively, is the sum of the utility of all possible outcomes multiplied with their respective likelihood. Then, the total utility of each choice is the utility of all possible benefits minus the utility of all possible costs. A rational decision-maker will then choose the behavior with the higher total utility.

Once the decision-maker chooses and enacts a behavior, events will unfold, and some of the uncertain possibilities will manifest. An individual will not make many major accounting decisions, such as committing large-scale accounting fraud. Therefore, she will not receive the expected utility, on average, from many repeat decisions. Instead, her decision is a gamble. She might be better or worse off than the expected utility, based on the realized benefits and costs she receives from the universe of possible benefits and costs. This is different from frequent repeat decision-making, where the actors will eventually approach the expected utility. In such a situation, beliefs about the repeat events may be accurate and, even if not, the repeat behavior allows for rapid belief updating. This is not the case in the misconduct context. Therefore, beliefs may be biased due to a lack of first- and second-hand experience.

However, assuming that all decision-makers behave rationally and that the outcome of their decision is observable, a researcher can estimate whether, empirically, decision-makers behave consistently with the basic theory in Figure 1. Recent research estimates that between 8.1% and 43.5% of accounting fraud perpetrators experience a net benefit, although the fraud is discovered (Amiram et al., 2020). Based on earlier estimates that the detection likelihood is 31% (Dyck et al., 2017), the authors conclude that more than half of the fraudsters (combining detected and undetected fraud) benefit from their actions. Therefore, given that uncertainty is involved, and most decision-makers experience net benefits, the simplified theorization of these decisions in Figure 1 seems appropriate.

Each of the theories I use in my four research papers aligns with at least one of the four basic building blocks (benefits and costs of one of the two conducts) of this fundamental theory. They describe how a rational decision could change based on changes to these building blocks. I purposefully present the theories separately from the empirical results I summarize in the next section because they are valuable for understanding my dissertation, the overarching research question, and the overall theoretical foundation

beyond the way I employ them in the single papers. Below, I describe chronologically and in detail how the theories I use in each paper align with my overall theory. The solid arrows in Figure 1 depict these relationships.

The first paper in Chapter 2 uses deterrence theory based on Becker (1968) and Cornish and Clarke (1987). This theory uses the utility function of corporate misconduct and assumes that decision-makers will consider any form of punishment for corporate misconduct as costs in this utility function. Deterrence theory further details these costs into three components, punishment severity, punishment certainty, and punishment swiftness. In this form, multiplying these three components gives the loss of utility caused by punishment. Severity is the utility loss due if the perpetrator is punished (e.g., a monetary fine or the present value of career prospects lost due to the stigma of misconduct). Certainty is the likelihood that the misconduct will be detected and punished. Swiftness is the time discount of the negative utility dependent on the time that passes until detection and punishment of the misconduct. On a theoretical level, deterrence theory comprises all sorts of punishment. If we assume that the utility of other costs for the misconduct option (e.g., the effort to complete the behavior) are relatively minor, it covers the majority of the utility of costs for the misconduct option.

Beyond that, deterrence theory adds another theoretical layer that is conceptually transferrable and important to all building blocks of my overall theory. The theoretical discussion in Chapter 2 clarifies the importance of distinguishing expected utility from factual costs and benefits. As the theory captures changes in expected utility, changes in factual costs or benefits will only affect the choice if they also change expectations about the respective utilities. This notion is of particular importance, given the above discussion that beliefs about misconduct may be biased due to a lack of misconduct experience. They may be very far from reality and may not update if actual costs and benefits change. This generally reduces the effectiveness of any factual intervention targeted at the net utility of the misconduct option. Conversely, interventions tailored at changing perceptions are more likely to be effective. For the rule-abiding option, this is less of a concern to the extent that first- and second-hand experience with this option is somewhat frequent.

From a theoretical standpoint, Chapter 3 differs significantly from the other three chapters. In the qualitative study, I use Grounded Theory (Glaser and Strauss, 1967) to develop my own theoretical framework studying an area that is generally under-theorized. I discuss this in more detail in the synopsis in the next section. However, this self-developed theoretical framework is based on traditional agency theory (Eisenhardt, 1989; Jensen and Meckling, 1976). Within it, I focus on information systems as a monitoring tool. In the context of the paper, specifically, the information system's only goal is to detect misconduct. If this information system is effective, it allows the principal to anticipate adverse selection and detect hidden action.⁴ Theoretically, this can drastically reduce the likelihood of a decision-maker to receive benefits from misconduct. In the adverse selection case, a perpetrator who intends to commit misconduct will lose her opportunity to receive these benefits if the information system detects her a priori. For instance, if a burglar cannot sell off her stolen goods because any potential buyer can detect their fraudulent origin, she loses the expected utility of the crime. The likelihood of successfully selling the stolen goods becomes zero, while it was positive before and the expected utility given such sales was probably large. In the hidden action case, the information system can reduce the expected utility and increase the expected costs of the perpetrator. Consider that the information system detects a tax evader. Upon detection, authorities may confiscate her funds from the tax evasion, removing the benefits from the misconduct. At the same time, she may receive a fine or another punishment, increasing the misconduct costs.

Next to its direct effects on expected utilities, the information system closely relates to deterrence theory. The information system's ability to detect hidden action increases the punishment certainty element used in deterrence theory. This shows how the theories in my dissertation interact. It emphasizes that joint consideration of the theory

⁴ Adverse selection occurs before a contract under information asymmetry is agreed upon. It arises if there are agents of a good type and a bad type and the principal is only willing to contract with good agents. However, she cannot observe the type of agents a priori. In the case of adverse selection, she would unknowingly choose a bad contractor. Such a situation could arise if a company plans to work with a new supplier and does not know her type. If the company uses an additional information system and requires the potential contractor to reveal information through the system, it can avoid adverse selection (Akerlof, 1970). Hidden action occurs when a contract already exists. The principal cannot observe all the agent's actions, and the agent may choose to act in her own interest that may misalign with the principal's interest. An information system that provides additional information on the agent's actions can reduce the likelihood of such hidden action being undetected (Eisenhardt, 1989).

mosaic helps to understand the entire decision process and the potential interplay between certain decision-making aspects. This provides opportunities to detect mutually reinforcing or substitutive effects, which may allow for more precise predictions of empirical phenomena. In this case, theory suggests a mutually reinforcing effect because the elements in deterrence theory are multiplicative. Hence, an increase in the two elements will be mutually reinforcing instead of additive or even substitutive.

Chapter 4 uses several theories that help develop a hypothesis how investors react to the behavior of an organization that faces a suspicion of misconduct within the organization. This situation comprises two distinct links with my overarching theory. First, the suspected misconduct has already happened⁵ and has been detected. Therefore, the chosen behavior of the organization will influence the costs of misconduct. Second, the organization's behavior not only covers the reaction to the suspected misconduct, which happened in the past but also a commitment to future behavior. More specifically, it signals the choice of whether and how the organization plans to change the alleged misconduct to rule-abiding conduct or not. Hence, the theory also encompasses the net expected utility of future rule-abiding behavior.

The first link to the expected utility of costs of the misconduct is through a direct link to deterrence theory from Chapter 2. I use the withholding argument, the persuasion knowledge model, correspondent inference and attribution theory, and the disconfirmation of expectations theory to discuss how post-detection behavior can influence the market-assessed present value of punishment costs. These theories cover how information disclosure and investigation behavior of the alleged perpetrator organization affect the punishment costs. The withholding argument suggests that they increase when the press reveals information that the organization withholds because this increases the likelihood that more undisclosed and unknown negative information exists (Dolnik et al., 2003). The persuasion knowledge model covers proactive disclosure of the organization and suggests that it can backfire when investors perceive persuasive intent. If they do, they will doubt the credibility of the self-disclosed

⁵ Note that there is uncertainty that allegations are true. However, economically, this is only a discount factor for the consequences of the allegations because I consider the expected utility of the consequences. As mentioned, it is the product of a consequences' likelihood (in this case, the likelihood that the allegation is true) and utility.

information, which increases the likelihood that self-disclosed information understates the severity of the true underlying issue (Friestad and Wright, 1994). Conversely, correspondent inference and attribution theory suggest that proactive self-disclosure reduces the attribution of the misconduct to the organization's character (Weiner, 1986, 1985). Consequently, expected costs decrease because the misconduct seems less egregious if it is uncharacteristic of the organization. Finally, the disconfirmation of expectations theory relies on belief updating. It suggests that at least some investors believe that organizations will not self-disclose negative information. If these beliefs update following a disconfirmation of this expectation due to self-disclosure, expectations become more positive, reducing expected costs (Eagly et al., 1978).

Turning to the second link, I use signaling theory to explain how signals following misconduct allegations can change the net expected utility of future rule-abiding behavior. Signaling theory can be understood from the agency theory framework (Morris, 1987). Specifically, it considers an adverse selection situation in which the agent (in this case, the perpetrator) provides a signal with certain credibility to partly overcome information asymmetry that induces uncertainty of the agent's type. Alleged misconduct could increase the expectation that the perpetrator is a bad type, aggravating the adverse selection problem under information asymmetry. Therefore, signaling theory suggests that the perpetrator organization could alleviate adverse selection concerns by providing a high-credibility signal that it is of good type despite the allegations. In turn, the uncertainty of whether the organization is of the bad or good type becomes more favorable, increasing the net expected utility of future rule-abiding behavior. Possible mechanisms are lower risk premiums or capital costs for investments in rule-conform investment projects (decreasing expected utility of costs) or more potential customers who believe that the organization will fulfill contracts in the future (increasing expected utility of benefits).

Finally, Chapter 5 uses two classic theories from the auditing literature, which traditionally explain why auditors provide high-audit quality. Within the overarching theoretical framework of this dissertation, the arguments naturally extend beyond the auditing context. Briefly, the insurance hypothesis suggests that auditors provide high-audit quality (or, in the context of the overarching theory, avoid misconduct) because they try to avoid costly litigation (Dye, 1993; Lennox, 1999). In a broad sense, this is

just the adaptation of deterrence theory to auditing as a specific domain. This includes the domain-specific identification of existing punishment certainty, severity, and swiftness in auditing. However, it only considers regulatory punishment. Besides, a particularity of this domain-specific adaptation is that academics and practitioners have plenty of experience with auditor litigation, making beliefs and belief updating less biased than in the context of Chapter 2. The second theory, the reputation hypothesis, is concurrent to the litigation hypothesis and similarly theorizes auditor misconduct costs (Davis and Simon, 1992; DeAngelo, 1981; Skinner and Srinivasan, 2012). It differs from the litigation hypothesis in that it covers all non-regulatory punishment when studying corporate misconduct through the lens of deterrence theory. Again, the important domain-specific particularity is that auditing researchers and the auditing profession have firm beliefs of the expected utility of costs from this non-regulatory reputation loss. Importantly, there is consensus in the auditing literature that these two competing theories are (at least partly) substitutes. In contrast, my application of deterrence theory is silent on whether prosecution and reputational penalties are substitutes or complements.

In addition to these theories resembling deterrence theory, the reputation hypothesis adds the opposite of deterrence: incentives to gain and maintain a high reputation, which will increase the expected utility of benefits of rule-abiding behavior. This consideration is part of neither deterrence theory nor the insurance hypothesis. However, in the auditing domain, there are firm theoretical arguments from economic decision-making theory that conceptualize expected benefits for the auditor from rule-abiding behavior, such as increasing the customer base and the pricing of services (Klein and Leffler, 1981). To leverage those arguments beyond the auditing domain, I add theory on organizational reputation (Lange et al., 2011). It suggests that perceptions of organizations and their brand names can influence the benefits and costs of any behavior of an organization subject to information asymmetry. In other words, reputation effects from one branch of an organization spill over to all branches of the organization. This can conceptually be linked to signaling theory, as a brand name, regardless of its source, is one possible signal to overcome information asymmetry (Akerlof, 1970).

To sum up, my dissertation covers several theoretical levers to affect whether a rational decision-maker chooses misconduct or not. Deterrence theory and the related insurance hypothesis suggest ways to increase expected corporate misconduct costs before it happens. Similarly, according to agency theory, organizations can design information systems that, a priori, decrease the expected benefits of corporate misconduct and increase deterrence. These mechanisms make corporate misconduct a less favorable choice. Several theories on the disclosure behavior of an alleged perpetrator indicate that ex-post, disclosure increases costs for alleged perpetrators that do not signal a change to future rule-abiding behavior. Simultaneously, self-disclosure and accompanying signals of future rule-abiding behavior can improve the net expected utility of rule-abiding behavior. Together, the increasing costs of misconduct and benefits of rule-abiding behavior may shift the decision-maker's choice. Finally, the necessity to build and maintain a reputation leads to a promotion of building a history of rule-abiding behaviors, which will improve expected benefits from such conduct. I next describe my empirical findings regarding and testing certain links in this overall theoretical framework.

1.3 Thesis Structure and Synopsis

As described in my motivation, my overarching research question considers how regulatory change or the improved efficiency of market forces can curb corporate misconduct. In this context, these two primary forces directly influence decision-making behavior. Besides, I consider the auditor as a prime institution in the highly regulated domain of accounting. As a monitor, the auditor contributes to curbing corporate misconduct and is herself subject to regulation and market forces. Hence, through the auditor as a third force, the two primary forces indirectly influence decision-maker behavior. Below, I describe how each of my papers aligns with these three forces (dotted lines in Figure 1) and contributes to my overarching research question. Note that the six research questions I list below are not necessarily the original research questions or hypotheses from the papers. Rather, they summarize the conceptual idea behind each paper as it relates to the overarching research question of my dissertation. Table 1 gives an overview of the general force and more specific concept each paper studies, how I operationalize this force as an independent variable, whether and under which

conditions I find evidence that the force is effective, and for which theories I find evidence.

Table 1: Investigated forces and related findings of my four research papers

Chapter (paper)	Investigated force (concept)	Operationalization of independent variable	Effectiveness of force (condition)	Supported theory
<i>Chapter 2 (first paper)</i>	Market forces (Punishment severity)	Decreases in stock prices Managerial job loss	Effective (if biased perceptions are corrected)	Deterrence theory
<i>Chapter 3 (second paper)</i>	Regulation (obligation to implement anti- money laundering systems)	Whether interviewed organizations follow the regulation	More effective than existing, weaker implementations based on market forces	None (Agency theory only framework; not tested)
<i>Chapter 4 (third paper)</i>	Market forces (Credibility signal of forensic services)	Investigation with forensic services vs. internal audit	Effective (only if combined with transparent disclosure)	Signaling theory
<i>Chapter 5 (fourth paper)</i>	Market forces (Audit clients loss of investor trust; Lost success with audit services)	Press coverage of non-audit service failure	Effective only in the short run	Reputation theory and hypothesis (limited support)

My first paper (Chapter 2) links deterrence theory to observable reactions of capital markets (organizational level) and labor markets (individual level) in the case of detected fraud. Hence, it directly studies market forces. To the extent that markets price information correctly, the negative utility of detected fraud will immediately correct stock prices or individual career prospects downward. There is convincing empirical evidence that shows this phenomenon (Karpoff et al., 2008a, 2008b). Since detection is certain in this context and efficient markets discount all prices to the current day, the market reaction precisely measures all punishments on an organizational level. This includes potential regulatory punishment. However, earlier research has shown that

market punishment many times larger than regulatory punishment (Karpoff et al., 2008b; Karpoff and Lott, 1993). Therefore, the paper studies a context in which there is empirical evidence that market forces are much stronger than regulatory forces. In the grand scheme of this dissertation, this raises the question of whether existing market forces are sufficient to curb misconduct. If they are, the question remains why we can still observe misconduct empirically. In light of the discussion at the beginning of this dissertation, there are two possible answers. First, there is information asymmetry, which hinders the market reaction to take its full effect. More precisely, because fraud is a rare event and organizations may have little to no experience with it, they might have incorrect expectations about the market behavior.⁶ Second, the remaining misconduct may just be the residual risk that remains because some organizations irrationally choose misconduct, although they could know about its lower expected utility compared with rule-abiding behavior.

Because my overarching research is concerned with the relative effectiveness of regulations versus market forces, I compare factual market forces, perceived market forces, and factual regulatory forces. In light of the above discussion, I ask the following research questions in my first paper:

RQ 1: Do decision-makers' beliefs about market punishment differ from factual market punishment for corporate fraud?

RQ 2: Does factual market punishment deter corporate fraud better than factual regulatory punishment?

I conduct an experiment with 28 executive board members. For another 43 board members, I collect additional data on their beliefs about market punishment, resulting in a sample of 71 board members for my hypothesis tests related to RQ 1. In the experiment, I give participants the decision to overstate revenue fraudulently to achieve the consensus analyst forecast. I manipulate the organizational and individual consequences for the executives if they commit fraud. The manipulations only change punishment severity. On the organizational level, the high punishment condition

⁶ In this case, information asymmetry is inverse of what it typically is in accounting (i.e., the organization knows more than the capital market). Here, the market knows more because of more frequent experiences with fraud. This situation arises because the market continually monitors many organizations and therefore observes all instances of the rare event of fraud.

contains the information that, if detected, the stock price will plummet by 300% of the fraudulent act's total value. In the low punishment condition, the rate is only 35%. On the individual level, the high punishment condition combines a 90% chance to lose the managerial job and a 30% chance to get a criminal sentence. The low punishment condition only contains a 25% chance to lose the job. I take the numbers from earlier empirical results in Karpoff et al. (2008a, 2008b). The high condition resembles all empirically observable consequences. The low condition only resembles regulatory consequences, such as fines, on the organizational level, and a turnover rate in the absence of misconduct on the individual level. My additional questions for the larger sample cover beliefs about detection likelihood and organizational and individual punishment severity, all before manipulation.

Concerning RQ 1, I compare my sample means from the additional questions with estimates of factual detection likelihood, market value loss of companies, and job losses from earlier literature. I find that executive board members overestimate the detection likelihood but underestimate both types of punishment severity. Answering RQ 2, the experimental results suggest that changing manager beliefs about organizational punishment from factual regulatory punishments to factual market punishments significantly reduces their tendency to commit fraud. This effect is only present when beliefs about individual punishment are low. As the first set of results shows, this is the case in practice. Interestingly, I also find that, although explicitly held constant in the experiment, beliefs about detection likelihood seem to increase when both punishment severities are high. However, this effect is insignificant on conventional levels.

In conclusion, the first paper shows that, for rare events with little personal experience like accounting fraud, beliefs are far from reality. Therefore, already existing market forces seem strong enough to curb at least some misconduct. This suggests that information asymmetries, rather than a lack of regulation, drive unnecessary high rates of existing misconduct. From the perspective of my overarching research question, I uncover a setting in which regulatory change, at best, would be inefficient. At the same time, efforts to raise awareness of already existing market forces may have substantial effects on the relative expected net utility of the decision-maker.

Using the specific example of anti-money laundering (AML) systems in the non-financial sector, my second paper (Chapter 3) studies the introduction of a new information system targeted at curbing misconduct, specifically, money laundering. However, since money laundering is a secondary crime (i.e., there is always a primary crime creating the funds that need laundering), it indirectly reduces the net benefit of various misdeeds. In the context of my overarching research question, I study a regulatory change that can directly and indirectly, by introducing new market dynamics, change the information asymmetry in an agency relationship. As my primary concern is with the balance of regulation versus market forces, the qualitative analysis focuses on the mechanisms that determine whether the regulation is effective. I consider both actions by the regulator and market dynamics introduced through the regulation.

As this research endeavor is explorative, I use a qualitative interview-based approach to address the following research question:

RQ 3: What dynamics does the introduction of AML systems bring to the non-financial sector and the mitigation of money-laundering related misconduct?

I use Grounded Theory (Glaser and Strauss, 1967) to analyze the transcripts from semi-structured interviews with 13 managers responsible for AML in eight multinational companies. The result is a framework that describes a generic process of how the interviewed organizations and their peers develop the newly required information system, the AML system. Along with other results with specific importance for the narrower topic of money laundering, I observe regulatory and market dynamics that may generalize from the introduction of AML in the non-financial sector to other regulatory changes considering corporate misconduct. Therefore, they are of specific interest to the overall theme of this dissertation. Such dynamics promote the implementation of AML, which is in itself rule-abiding behavior.

Concerning regulatory change, the interpretation of the interviews suggests that the mere regulatory change is unlikely to change behavior or even beliefs about the net utility of particular behavior unless there is some incentive or pressure to abide by the changed regulation. Such incentives or pressures can take different forms. The regulation end of the spectrum is strict enforcement. Firms that were most likely to be prosecuted due to regional differences cited this as a motivation to create better AML

systems. A step closer to the middle of the spectrum, regulatory guidance can also serve as an incentive to promote rule-abiding behavior. Specifically, it lowers the costs of rule-abiding behavior because it relieves firms from having to find their own way to implement AML systems. Besides, it reduces the uncertainty of whether the regulator will accept efforts to comply with AML-regulation as sufficient. By the same mechanism, standardization among peers, as a market force, promotes proper conduct. Finally, market forces can multiply regulatory pressure, or create their own pressure, once large players want to ensure compliance along the supply chain. Consequently, they will require all their contractors to implement, in this case, AML systems, which rapidly increases the reach of such regulation.

In addition to these detailed regulatory and market forces that drive proper AML implementation, my second paper considers the overall role of regulation according to its agency lens. The analysis suggests that regulation can be critical beyond market dynamics to change the design of AML as an information system that could curb misconduct. Specifically, the process I develop with Grounded Theory suggests that some companies follow the general idea of AML to avoid dirty money incentivized by market forces only. The generic process contrasts them to another group of companies that care to fulfill regulatory compliance of AML-regulation. The comparison indicates that the AML system as an information system appears more effective for the companies that follow the regulatory forces. Hence, in the second paper, I find the opposite of my first paper with regard to my overarching research question. The first paper uncovered a situation with market forces dominating regulation, whereas it is the other way around in the second paper. Ironically and similar to my first paper, biased beliefs about the misconduct and its consequences possibly explain the observations in the second paper. Intuitively, the non-financial sector has comparably low risks regarding money laundering. Therefore, many companies have incomplete beliefs about it. A consideration of the regulation helps to update these beliefs and develop a structured approach to detect and mitigate money laundering even if an organization has no first-hand experience. A mere market-incentivized approach will likely be incomplete.

Chapter 4 (my third paper) investigates how choices by an organization that has allegedly committed fraud can alter the expected costs of the alleged misconduct and the future prospects of rule-abiding behavior. I purposefully study a context in which

the fraud suspicion becomes public. This choice has two purposes. First, I aim to explore whether, in this situation, the economically optimal strategy of the firm aligns with the strategy that is arguably most favorable for the public. Second, I am interested in whether auditor-provided services can serve as a credibility signal that facilitates this possibly optimal strategy. Again, I take the perspective of capital markets by measuring the success of the firm's strategy based on investor perceptions.

In the context of my overarching research question, I study auditor-provided services that grow in practice as a third potential contributor to the relative balance of rule-abiding behavior and corporate misconduct. Specifically, I study forensic services, a consulting service for companies that suspect misconduct, and plan to address it internally. One motivation is the observation that there is a growing demand for these services and that they are not regulated. Given this rare context of a free market in accounting, the question arises whether regulation is necessary to achieve optimal welfare outcomes from this situation or whether the markets that created the supply and demand of these services have already increased welfare.⁷ A second motivation is to investigate whether firms and the public can concurrently benefit from these services. There is natural tension to this question because the public can only benefit from these services when they are transparent. Otherwise, information asymmetry does not allow pricing these services, hinders their potential deterrence effect, and allows opportunistic hidden action within the forensic service. However, transparency about the service also always reveals the suspected misconduct. As already discussed in the first paper, this leads to negative reactions that punish the firm. Therefore, I need to establish that non-transparent firms detected by a third party (in my case, the press) suffer more severe consequences than self-disclosing firms. Also, I have to assume a severe threat of detection. In conclusion, I address the following research question:

RQ 4: Does investor willingness to invest differ when an alleged fraudster engages forensic services, depending on whether the suspicion is self-disclosed or press-disclosed?

⁷ Technically, these two alternatives are not mutually exclusive. Even if the free market increases welfare, regulation may be necessary to achieve optimal market outcomes. However, this is beyond the scope of the paper. Instead, the paper investigates whether the market in its unregulated form is beneficial for public welfare.

I conduct an experiment with 128 private investors. Participants have to make an initial investment decision on a specific company. Then, I inform them that there is a fraud suspicion concerning this company. This information contains two manipulations. First, I manipulate whether the company engages forensic services or whether the internal audit department provided an investigation. Holding everything else constant, I hence just manipulate the investigation's credibility signal, not its actual conduct. Second, I manipulate whether the firm self-discloses this investigation and the underlying suspicion or whether a newspaper reveals it. This follows multiple theoretical arguments described above, which suggest that unforthcoming behavior increases costs from misbehavior. In this case, I hypothesize that it also depletes any potential benefits from the credibility signal. Participants make the same investment decision with this additional information. I use their change in decision to measure the consequences to the allegedly fraudulent firm.

My findings suggest that, when forensic services are engaged, investors are more willing to invest when the firm self-discloses. However, I find no differences between the two disclosure conditions with the internal audit investigation. Willingness to invest in this case is below forensic services and self-disclosure but above forensic services and press disclosure. Concerning the paper's positioning in my dissertation, these results suggest that auditor-provided services can provide a positive signal. This signal requires transparency. Hence, the best outcome that market forces for this unregulated service seem to create for the market participants align with arguably the best outcome for the public (high-quality service and full transparency). However, recall that this result is in a context that requires a severe threat of press disclosure. The first paper suggests that manager's beliefs support such a threat. Estimates from earlier literature indicate that, factually, such a threat is missing. Moreover, as I do not study what decision a firm makes but only the consequences of the different options, I cannot judge whether firms' decisions align with the optimal option. As shown in the first two papers, biased beliefs can disturb such optimal decision-making. Hence, some regulation may be necessary to effectively create a situation that resembles the mechanisms revealed in the experiment.

My fourth and final paper (Chapter 5) again considers the auditor as a critical institutional force curbing misconduct. My main goal is to contribute to the long-standing discussion in auditing research whether market-based reputation incentives

substitute regulation-based litigation incentives for auditors in improving the auditors' efforts against misconduct. However, it does not focus on auditing as the auditor's traditional main service but on non-audit services (NAS). As such, it studies to what extent incentives from one domain (auditing, where I assume the auditor faces reputation incentives) translate to another domain of the same organization (in this case, NAS). This provides additional insight into the overarching question, in which situations regulation is necessary. In this regard, I depart from the original continuum that looks at market forces as an alternative to regulation in only one domain. Now, I also include a force from outside the narrow domain. Hence, I broaden the perspective on markets and accompanying institutions and how spillovers may or may not be effective forces to consider when trying to improve efficiency in a market.

Consequently, changes in one narrow domain may not be necessary if there are changes in related domains that spill over. Specifically, in the auditing context, if reputation effects from the NAS failure spill over to auditing, the NAS provider's audit practice will suffer negative consequences. In this case, the auditor's single reputation incentive from the audit market would promote high service quality in all auditor-provided services. This, in turn, could strengthen the institution of auditing and contribute to curbing misconduct. I use an empirical-archival study to answer the following research questions that emerge from this discussion:

RQ 5: Do non-audit service failures have short-term negative effects on the reputation of the accounting firm's audit services?

RQ 6: Do non-audit service failures have longer-term negative effects on the economic success of the accounting firm's audit services?

I identify two scandals of KPMG's consulting branch in 2013 and 2017, respectively, and first use an event study to analyze short-term consequences. The scandals have no direct relation to KPMG's audit branch, and the litigation hypothesis is irrelevant in the concrete context due to the absence of material litigation in Germany. To examine spillovers, I estimate short-term market reactions to KPMG's listed audit clients on the two event days. Consistent with a reputation spillover in the context of RQ 5, I find negative reactions. Addressing RQ 6, I examine longer-term consequences to KPMG's audit branch. Specifically, I estimate changes to their ability to gain and retain clients,

charge audit fees, and their audit quality after the NAS scandals. Overall, I do not find evidence for longer-term effects, except for some indications of lower audit quality for the second event only. In conclusion, the last paper suggests that spillovers, even though they may play a role, are unlikely to be strong enough to substitute domain-specific regulation, market incentives, and institutions such as auditors that promote proper conduct and demote misconduct.

The following four chapters report the four research papers that I have just summarized. Except for the third research paper, which I wrote as a sole author, I was the primary co-author of a research team. Therefore, papers one, two, and four use first-person plural (“we”, “us”, “our”). Note that the first and second papers have already been published. I decided against including the body of these two papers in this dissertation to ensure consistency of their original language and formatting.⁸ Instead, readers can find these two papers with the bibliographic reference given on the title page of Chapter 2 and Chapter 3. I submitted the third paper to *Behavioral Research in Accounting*, where it is currently under review. To ensure consistency, besides careful formatting changes, including the numbering of sections and tables and reference formatting, Chapter 4 of this dissertation is identical with the submitted version. The fourth paper is a working paper that we still prepare for submission. After these four chapters containing the research papers, the sixth and final chapter draws overall conclusions by summarizing contributions, limitations, and avenues for future research.

⁸ The publication language of the second paper is German. However, the original publication contains an English title and abstract. I use those on the title page of Chapter 2. The bibliographic reference uses the German title because it appears in the journal’s table of contents. Together with the goal to avoid any, even smaller editorial, changes in the body of the published work, this language difference drives my decision not to include the body of these papers in my dissertation.

Chapter 2 The Deterrence Effect of Consequences of Fraud to Individuals and to Companies – Perception of Executive Board Members [Abstract only]

Published in

Friedrich, C.; Quick, R. (2019) “Der Abschreckungseffekt von Konsequenzen von Bilanzbetrug für Individuen und Unternehmen – Wahrnehmung von Vorständen” In: *Betriebswirtschaftliche Forschung und Praxis*, 71, pp. 21-54.

Abstract

Fraud is an egregious type of corporate crime which entails serious consequences once it is detected. Still, it is common. This raises the question why these serious consequences do not effectively deter fraud. The present study hypothesizes that this question can be answered by biased perceptions of these consequences. A case based experiment with executive board members of listed German companies indicates that the probability of detection is overestimated while the severity of sanctions is underestimated. Therefore, it is proposed to increase the saliency of existing consequences. It is suggested that an experimental manipulation of this saliency could decrease fraud propensity. Thereby, the study contributes to current discussions about increasing legal penalties for corporate crime.

Chapter 3 An Analysis of Anti-Money Laundering in the German Non-Financial Sector [Abstract only]

Published in

Friedrich, C.; Quick, R. (2019) “An Analysis of Anti-Money Laundering in the German Non-Financial Sector” In: *Journal of Management and Governance*, 23(4), pp. 1099-1137.

Abstract

European anti-money laundering (AML) law obliges both financial and non-financial private companies to contribute to combatting money laundering. Since the financial sector has implemented largely effective AML in the meantime, money launderers are increasingly moving their activities to the non-financial sector (non-FS). This study examines how AML obligations are implemented in the German non-FS. We intend to systemize different implementations of these obligations to provide a basis for improving future AML implementations, guidance, and research in the non-FS. The German setting is especially suited for this research, because its AML law is stricter than European AML law with respect to non-FS obligations, and because Germany has a large and interconnected economy. We use Grounded Theory to collect and analyze rich data from semi-structured interviews with 13 managers from eight multinational companies. Our result is a theory which systemizes the identified AML implementations. This helps explain how these various AML approaches emerge, and identifies ways in which the identified non-compliance and unaddressed AML risks can be mitigated. Another key observation is that guidance by the regulatory authorities is lacking. For both practice and regulators, the findings imply that the non-FS should follow a systematic AML process and that the regulatory authorities should support this approach through additional guidance. We also close some gaps in the literature, which has largely neglected the non-FS and rarely collected original data of actual AML implementation. The developed theory contributes to a better understanding of how AML effectiveness can be assessed and enhanced.

Chapter 4 Private Investigations and Self-Disclosure of Suspected Fraud: Experimental Evidence

A revised version of this chapter is submitted as

Friedrich, C. (2020) "Private Investigations and Self-Disclosure of Suspected Fraud: Experimental Evidence on Forensic Accounting Services" *Working Paper, Second Round at Behavioral Research in Accounting*.

Abstract

When a company internally suspects fraud, it faces difficult choices. For instance, it can choose to investigate internally or engage an external specialist. Additionally, it can choose to disclose the investigation or take the risk that the investigation is leaked to the public. I analyze whether the choice to engage an external specialist, rather than investigate internally, changes investors' willingness to invest in the company. I argue that, when engaging external specialists, disclosure choices matter. Conceptualizing the engagement of external specialists as an external credibility signal, I hypothesize that, when engaging external specialists, self-disclosure increases investors' willingness to invest compared to when the press reveals the investigation. My results from a 2x2 between-subjects experiment with 128 private investors support my hypothesis. This suggests that aligning a signal of credible investigation efforts with forthcoming disclosure could be beneficial. Hence, companies conducting genuine investigations could benefit from resisting any temptation of non-disclosure.

4.1 Introduction

The present research seeks to explore investor reactions to companies that suspect a fraud when the suspicion becomes public.⁹ When fraud is suspected, a typical reaction is to investigate the issue internally. Companies commonly engage external specialists (e.g., forensic service specialists from Big 4 audit firms), rather than relying on internal resources only, to help conduct the investigation.¹⁰ Given the prevalence of such investigations, I ask the following research question: Does investor willingness to invest differ when a company, facing a potential fraud, engages an external specialist? In this case, does willingness to invest differ based upon whether a company self-discloses the suspicion or whether it is revealed by a third party (e.g., the press)?

The use of investigations involving external specialists in high-profile corporate scandals has become more prevalent in the last decade (ENSSCPA, 2017; IBISWorld, 2017; Ott, 2017) and has a multiple-decade history in the US (Levick, 2017). Questions surrounding these investigations are gaining importance with the ongoing growth of the forensic service lines of audit firms, a major provider of private investigation services (e.g. Hegazy et al., 2017; Williams, 2014). Moreover, practitioners and researchers have suggested mandating periodic private investigations since the accounting scandals in the wake of the century (Asare et al., 2015; Christensen et al., 2005; DiPiazza et al., 2006). Others expect regulations or standards on private investigations (Huber, 2012; Smieliauskas, 2006). Private investigations also gain academic attention, for instance with the recent foundation of the Journal of Forensic Accounting Research by the American Accounting Association. At the same time, there is a dearth of high-quality forensic accounting research (Brody et al., 2012; Trompeter et al., 2013). Little is known about private investigations in general, and forensic services as separate services

⁹ A fraud suspicion is a rare, but highly relevant event with a longstanding tradition in accounting research (Amiram et al., 2018; Hogan et al., 2008; Trompeter et al., 2013). My study focuses on situations in which the suspicion becomes public because historical scandals have shown that many irregularities, which have been concealed, eventually reveal themselves. About half of SEC financial fraud investigations are voluntarily disclosed to the public, including suspicions with a comparably low likelihood of a material fraud (Blackburne et al., 2020; Solomon and Soltes, 2019).

¹⁰ My study focuses on the nature of the company's investigation, specifically, on whether it engages external specialists because most frauds are eventually investigated internally. For instance, a PWC (2018) survey with 500 German companies reveals that 85% of the companies internally investigate instances of fraud. In 54% of these investigations, they engage external specialists. Recent global survey evidence of PWC (2020) shows that more than half of the companies experiencing a fraudulent incident conduct an internal investigation.

from the financial statements audit specifically (Williams, 2014). Therefore, it is worthwhile to address the fundamental question whether, and under which conditions, voluntary, private investigations with external specialists improve investor perceptions.

Conceptually, the engagement of external specialists could serve as a credibility signal that a company genuinely conducts a high-quality investigation of the suspected issue. However, depending on the context in which the company sends this potential credibility signal, it might backfire. Investors may perceive the company to abuse the assurance signal of hiring an external specialist as a persuasive tactic instead of a genuine effort to assure a high-quality investigation. One contextual factor is how investors learn about the suspicion. When the company self-discloses the suspicion, the company is forthcoming and sends strong signals of integrity, because investors do not expect forthcoming bad news disclosure (disconfirmation of expectations, Eagly et al., 1978). If the engagement of external specialists is disclosed after such forthcoming disclosure behavior, investors likely perceive it as a genuine signal that the company is taking the issue seriously and takes all necessary steps to resolve it. Hence, there is an alignment of genuine behaviors. In contrast, when a third party reveals the suspicion, expectations of investors change. Investors tend to become suspicious that the company might withhold private information (Dolnik et al., 2003; Gottschalk, 2017). From the involvement of external specialists, they might infer a higher likelihood that material bad news has been uncovered and withheld. Moreover, it might induce the perception that the issue may be more severe. Hence, investors are unlikely to perceive external specialists as a high credibility signal, when it is misaligned with the low credibility signal of revelation by a third party.

These arguments lead to the following hypothesis: When a company engages external specialists, investors are relatively more willing to invest in it when the company self-discloses the suspicion and investigation than when they are revealed by a third party; but, for an investigation with internal resources only, the source of disclosure does not matter. To test this hypothesis, I conduct a 2 x 2 between-subjects experiment with 128 private investors. I manipulate whether the company engages external specialists (specifically, Big 4 forensic services) to help their internal audit department investigate the suspicion versus relying on the internal audit department only. I also manipulate whether the suspicion and investigation are self-disclosed or revealed by the press.

Consistent with my hypothesis, I find a significant interaction between the type of investigation and the type of disclosure. Specifically, when forensic services are engaged, investors' willingness to invest is highest with self-disclosure and lowest when the press reveals the suspicion and investigation (and in-between those two extremes in the control conditions of internal investigations with internal audit only). This suggests that alignment of forthcoming disclosure with a credible investigation is beneficial. Consistent with my theoretical arguments, additional analyses reveal that the engagement of external specialists improves investor perceptions of the quality of the investigation. Moreover, investors judge management to be more credible when the company self-discloses the suspicion. Finally, the combination of an external investigation with self-disclosure decreases perceptions of management withholding information. Deviating from the theoretically predicted pattern, perceived severity of the suspicion, and perceived persuasive intent are highest in the internal audit and self-disclosure condition. These additional results provide some initial evidence for the theory underlying my hypothesis.

My research contributes to prior experimental accounting research, which analyzes situations, in which certain disclosures might improve or impair ("backfire") how stakeholders perceive the actor, depending on whether stakeholders perceive the disclosure to be consistent with other characteristics of the disclosure.¹¹ Prior studies examine different settings, different theoretical constructs, and rely on subtly different theories. Specifically, Barton and Mercer (2005) study analyst perceptions of managers who explain bad firm performance with external factors. They find that explanations improve analyst perceptions when they are plausible, but backfire when they are implausible. Grant et al. (2018) study investor willingness to invest, dependent on how and through which medium managers disclose positive firm performance. Their results indicate that bragging (versus modest communication) improves willingness to invest when it is aligned with expectations that managers brag (during a conference call), but it backfires when it is misaligned with expectations (disclosure through Twitter).

¹¹ A recent archival study also documents such an interaction. Bartov et al. (2020) analyze negative stock market reactions to restatements. They find that, for companies with high corporate social responsibility (CSR) performance, reactions are less negative when restatements are based on errors, but more negative when restatements are based on fraud (i.e., inconsistent with ethical signal of high CSR).

Finally, Grenier et al. (2012) study negligence assessments of auditors accused of an audit failure related to undetected fraud. They find that remedial tactics by the accused auditor (claiming compliance with professional standards) reduce negligence assessments when auditor credibility is high, but backfire when credibility is low.

This literature analyzes how disclosure choices shape stakeholder perceptions dependent on their alignment with perceptions formed from characteristics of the disclosing party. I extend it to the analysis of how the alignment of two choices shapes shareholder perceptions. Additionally, I introduce an independent third party (the Big 4 forensic service) as a potential source of credibility that contributes to such an alignment. Moreover, I show the applicability of concepts from the literature on crisis communication and stealing thunder in the accounting context. This adds another theoretical perspective which helps to explain how disclosure shapes investor perceptions of companies in crisis situations.

My findings also contribute to research on private investigations in general, and forensic accounting (as a non-audit service) in particular, by providing preliminary evidence of how forensic services by audit firms affect investor perception of companies under investigation. Some authors are concerned that private investigations (as a substitute for, or complement to, state prosecution) impose the interests of client companies and entail negative consequences for society (Button et al., 2015; Williams, 2005). Given the increasing prevalence of such investigations, it is important to empirically investigate such concerns. These concerns seem to be consistent with investors' willingness to invest in a company that is caught with a suspicion of fraud and forensic services. However, they do not seem to be a general problem of private investigations. Specifically, investors seem to perceive forensic services positively when the company self-discloses the suspicion and investigation.

Finally, my research may be of practical relevance by showing the effects of private investigations and concurrent disclosure choices in the course of suspected accounting fraud. This may provide companies new evidence to underpin their disclosure decisions when facing corporate scandals. It can also help external specialists to adjust their services to disclosure needs and risks of their clients, and potentially remediate suspicions more transparently. Finally, the study may inform recent regulatory

discussions about regulations for private investigations on corporate fraud. As my results indicate, an appropriate disclosure regime is critical for the success of such a potential regulation.

I next develop my hypothesis based on a discussion of potential positive and negative effects of engaging forensic services and how they relate to disclosure choices. Section 4.3 describes my experimental design and Section 4.4 the results. A conclusion follows in Section 4.5.

4.2 Theoretical Background and Hypothesis Development

4.2.1 Engaging External Specialists

Given the suspicion of accounting fraud, there are two main reasons why the company investigating the suspicion would engage external specialists.¹² First, the company might not feel competent enough to conduct an investigation with internal resources only. Second, despite having the necessary skills, the company might feel that it cannot credibly signal to internal or external stakeholders that an internal investigation was carried out with sufficient competence and independence to fully investigate the issue. Following traditional theory from auditing, these factors determine the perceived (audit) quality of the investigation (DeAngelo, 1981).

Hence, from an investor perspective, the independent external specialist provides assurance that the investigation is of a certain quality. Conceptually, such a situation is close to a financial statements audit. The auditor provides both the necessary competence to judge whether financial statements are free from material misstatements, and embodies public trust that the audit opinion is stated truthfully and is complete. Therefore, forensic accounting services provided by Big 4 audit firms likely provide a strong setting to study whether stakeholders perceive external specialists to fulfill this role. I argue that the assurance of a credible investigation of a certain quality obtained by engaging forensic services provides value to a forensic service client,

¹² To date, fraud audits and related disclosure remain voluntary and largely unregulated. In July 2019, the AICPA issued the Statement on Standards for Forensic Services No. 1. However, this standard is short and has little detail (Schwartz, 2020). ISA 240.A34 refers to forensic experts in the context of the auditor's responsibility with respect to fraud but does not provide any specificities about forensic experts beyond general guidance on expert usage. The German federal government plans on regulating private investigations carried out in the context of corporate misbehavior (Anger, 2018; KPMG Law, 2019).

irrespective of whether the client would have been competent enough to investigate the issue without external help.

From an economic decision-making perspective, the choice to engage forensic services is likely the positive result of a cost-benefit analysis (in the context of voluntary audits, Kausar et al., 2016; Lennox and Pittman, 2011).¹³ The public engagement of forensic services signals this result to capital markets. Forensic accountants from Big 4 audit firms may be especially suited to provide this signal because reputational concerns incentivize forensic accountants to provide high service quality and to resist client pressures to support opportunistic client behavior (Davis and Simon, 1992; DeAngelo, 1981). Consistent with this economic argument, forensic services potentially restore investors' confidence in a company's financial information after a scandal (Rezaee and Crumbley, 2007).

From a legal perspective, forensic services can alleviate punishment through prosecutors. In the US and the UK, companies can settle charges by entering into non- or deferred prosecution agreements (NPA/DPA), which generally include a monetary fine and the requirement of remedial actions, such as installing effective compliance programs or engaging an external monitor (Arlen, 2012; Serious Fraud Office, 2017). Similarly, discussions on regulatory changes in Germany also include a rule under which private investigations should be accepted as a means to reduce or avoid prosecution and punishment (KPMG Law, 2019). Again, the credible signal that the investigation is of high quality may be necessary to convince prosecutors and/or the public of the effectiveness of such actions, and hence, successfully reduce expected damages.

However, there are several counterarguments to such potential positive effects of external specialists that rely on features of the setting that are distinct from financial statements auditing. In situations of suspicions of fraud, companies fear reputational

¹³ Prior literature has documented benefits of forensic services. They likely detect (and deter) fraud to some extent (Boritz et al., 2015). Jenkins et al. (2018) provide evidence that the benefits of including forensic experts in financial statements audits exceed the cost for forensic experts. Arlen (2012) argues that private investigations are economically more efficient than public investigations. Similarly, forensic accountants argue that public authorities do not have enough resources to pursue all major corporate crimes, are not interested in misbehavior that does not lead to criminal prosecution, and do not go beyond prosecution to mitigate systematic failure to avoid future scandals (Heißner, 2014, pp. 73–79).

damage from disclosing information about potential corporate scandals.¹⁴ Hence, investors may expect managers to withhold private information. They might interpret the engagement of external specialists as a signal that managers have private information of a higher likelihood of a severe fraud, which would require such assistance. From an economic perspective, investors would consequently perceive the fundamental value of the company to be lower, due to a higher risk of a severe irregularity.

Next to this fundamental value argument, there are arguments contrasting above claims that external specialists could signal a high-quality investigation. In fact, a backfiring effect might occur. That is to say, investors perceive management to strategically use the presumable assurance signal of engaging external specialists to pretend transparency and integrity without the intention to fully investigate and remediate the issue (Claeys, 2017; Lee, 2016). This is theoretically grounded in the Persuasion Knowledge Model, which indicates that investors, who perceive a persuasive intent in a signal, become suspicious about this signal (Friestad and Wright, 1994). In this case, investors would perceive external specialists to be dependent on management, a concern expressed in the private policing literature which argues that external specialists are limited by the scope, and dependent on the goals, of the engagement dictated by the client (Gottschalk, 2017; Williams, 2014).

Overall, it is hence theoretically unclear whether external specialists could successfully provide (or would risk harming) assurance to investors that investigations are genuine and of a certain quality. However, the context in which private investigations take place might reinforce one or the other group of arguments from above, making directional theoretical predictions possible. Next, I introduce self-disclosure (versus third party revelations) as a potential contextual mechanism and develop a directional hypothesis on the effect of disclosure when external specialists are engaged.

¹⁴ Amiram et al. (2018) provide an overview of research estimating the reputational damage of detected corporate misconduct. Overall, these studies indicate that there are substantial reputational penalties for perpetrator firms if the misconduct directly affects contract partners (vs. third parties). At the same time, they discuss evidence that investments in good corporate governance can rebuild reputational capital. They do not report any research on the role of private investigations in this context.

4.2.2 Self-Disclosure and Hypothesis Development

Different theories from psychology suggest that self-disclosure can improve perceptions of those who self-disclose (Weiner et al., 1991). Following correspondent inference theory, self-disclosure of bad news by a manager reduces the attribution of the manager's character to the bad news (Jones and Davis, 1965). Attribution theory suggests that self-disclosure weakens the causal link of the two by making the bad news event seem more external to the control of the manager, more time-variant, and less volitional (Weiner, 1986, 1985). Research on criminal trials shows that the self-disclosure of negative evidence by the defendant, instead of the revelation of this evidence by a witness ("stealing thunder"), significantly reduces the perceived guilt of the defendant through an increase in credibility (Williams et al., 1993). On a company level, the literature on crisis communication supports these findings (e.g. Claeys et al., 2016; Easley et al., 1995). Fennis and Stroebe (2014) report two similar theoretical explanations for differential effects of self-disclosure and third-party revelations. First, the disconfirmation of expectations theory suggests that, if a company disconfirms investors' expectations to withhold bad news, it increases its perceived integrity (Eagly et al., 1978).

The second theory they discuss, the withholding argument, indicates how the revelation of bad news by a third party could influence perceptions about a company. It suggests that investors likely perceive bad news which is revealed by third parties (e.g., the press) as information a company actively concealed. Therefore, they judge companies as less trustworthy (Dolnik et al., 2003). Hence, it is more likely that they infer the withholding of private information based on other signals of the company. More specifically, as argued above, they might infer from engagements of external specialists that the company has private information on which the decision to engage the specialists was made, even if the company does not have such information. This could include perceptions that the suspicion might be more severe compared to suspicions investigated without external specialists. Moreover, perceptions that firms engage external specialists in order to pass responsibility, or to get support concealing the scandal, most likely occur when investors perceive the company to be less trustworthy.

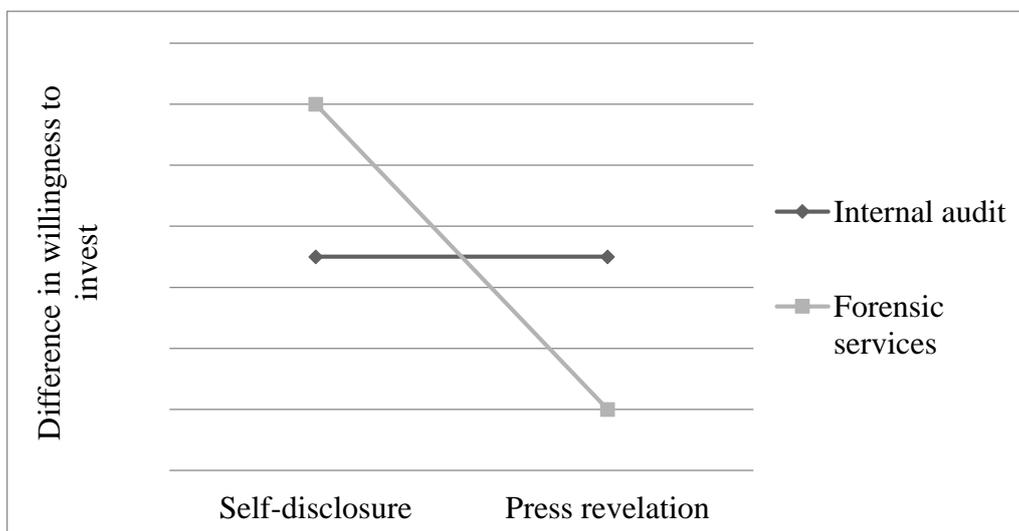
Taking the above-discussed arguments together, when the company adds self-disclosure to the engagement of external specialists, there are aligned credibility signals. Hence, I

expect investors to be convinced that the company genuinely investigates the suspicion and is willing to take appropriate remediation measures. Accordingly, I expect an attenuation of potential decreases in willingness to invest due to the suspicion. Contrastingly, a potential credibility signal through external specialists would be misaligned with negative signals from third party revelations, when the suspicion is revealed by the press. Therefore, I expect that, when a company engages external specialists, getting caught by the press exacerbates decreases in willingness to invest due to the suspicion. I do not expect such a disclosure effect in my control condition of internal investigations with internal resources only. To sum it up, I expect the following disordinal interaction:

Hypothesis: When a company engages forensic services because it suspects fraud, self-disclosure (versus press revelation) of the suspicion improves investor willingness to invest; but, for an investigation with internal resources only, the source of disclosure does not matter.

The pattern of expected results, which is predicted by my hypothesis, is shown in Figure 2, Panel A.

Panel A Hypotheses: Predicted results pattern for willingness to invest



Panel B Results: Means of difference of willingness to invest after manipulation minus willingness to invest before manipulation

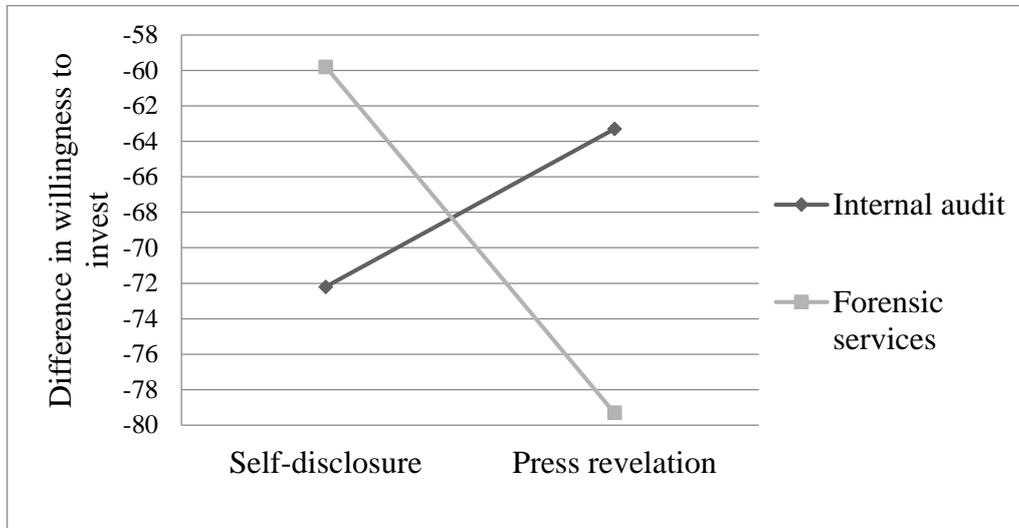


Figure 2: Predicted and observed effects of disclosure and investigation on willingness to invest.

Note that my hypothesis is based on potentially multifaceted differences in how investors *perceive* the disclosure strategies of a company investigating a suspected fraud with the help of external specialists. As explained in the theoretical discussion above, this could include differences in how investors *perceive* the severity of the issue, the private knowledge of the company (and whether it is withheld), and the quality of the investigation. Importantly, I expect these *perceptions* to differ regardless of *factual* differences along these dimensions. In the next section, I describe how my experimental design holds *factual* information on these dimensions constant across conditions.

4.3 Method

I employ a 2 x 2 between-subjects experimental design, manipulating (1) disclosure of the scandal and (2) the type of investigation. The appendix in section 4.6 contains the main body of my instrument. In the following subsections, I describe participants, the experimental procedure, my treatments, and the dependent and additional variables.

4.3.1 Participants

I recruited private investors as participants through the online participant pool Prolific.¹⁵ The platform is similar to MTurk but is primarily built to recruit participants

¹⁵ <https://www.prolific.co/>

for academic research. Prior comparisons of MTurk, Prolific, and other platforms and participant pools have shown that the quality of data generated through Prolific participants is at least equal to other alternatives (Palan and Schitter, 2018; Peer et al., 2017). Prolific allows to screen participants for a variety of criteria, which participants fill into their profile, and I restricted my sample to participants that have invested in stock. 137 participants completed the experiment.¹⁶ Participants consist of 37 (27.0%) females and 99 (72.3%) males (one did not disclose gender) with a mean age of 39.9 years (median: 37, standard deviation: 13.3). They have moderate self-assessed expertise in accounting (mean: 41.6 on a scale from 0 to 100; median: 38, standard deviation: 26.3) and with regard to capital markets (mean: 39.2 on a scale from 0 to 100; median: 36, standard deviation: 25.9). About half of the participants (71, 51.8%) were familiar with forensic services.¹⁷ None of the sample characteristics differ significantly across cells ($p > .49$). I paid participants 1.25 Great Britain Pounds (GBP) for completing the experiment, resulting in an average payment of 7.10 GBP/hour.

4.3.2 Experimental Procedure

Participants received a link to an online questionnaire administered through *soscisurvey*.¹⁸ I assigned each participant randomly to one of the experimental cells. Participants assume the role of a private investor considering an investment in the chemical industry. In all conditions, they read a brief description of the fictitious chemical company *Spechem* including information about its operations, its business segments, and key financial indicators. After they view the company information, I ask participants to make an initial assessment of their willingness to invest in the company. In the next step, participants view their respective treatments. After receiving the treatment, I ask participants for their final willingness to invest. The experiment concludes with a debriefing questionnaire containing manipulation and comprehension checks, variables for additional analyses, and questions on their experience. Participants

¹⁶ Originally, I have 139 observations. However, one participant accessed the experiment twice and I hence remove both observations from this participant.

¹⁷ Responses to my main dependent variable do not differ between those participants who were familiar with forensic services and those who were not ($p = .19$).

¹⁸ <https://www.soscisurvey.de/admin/index.php>. *Soscisurvey* is a German platform for conducting survey research which allows researchers to build HTML-based questionnaires and store the data on German servers.

could not change responses after each step of the experimental procedure they completed.

4.3.3 Treatment Variables

I manipulate *disclosure* as follows: In the *self-disclosure* condition, the experimental materials contain an ad-hoc disclosure¹⁹ of *Spechem* revealing that there is a suspicion of a material accounting fraud in a subsidiary which *Spechem* is investigating. In the *press revelation* condition, the experimental materials contain a short press release of “the renowned business magazine ‘Shareholder Weekly’”, which reveals the suspicion, followed by a later ad-hoc disclosure of *Spechem*, which informs about the investigation.

Besides the source of the initial disclosure, both conditions contain the same information and wording. Hence, across both conditions, I hold the timeline of the fraud suspicion, the severity of the potential fraud, and the private information of management constant, on a *factual* level. I provide precise information that the report is two weeks old, give identical and concrete numbers for the potential scope of the fraud, and explicate that there will have to be further analyses in the coming weeks. While I argue that there is, thus, no *factual* information in the manipulations that would allow inferring differential levels of private knowledge of management or severity of the suspicion, it is possible that participants’ *perceptions* of private knowledge or fraud severity differ; for instance because it is more likely that they suspect management to withhold information in the press revelation (versus self-disclosure) condition. In fact, my hypothesis predicts such differences in *perception*, despite no information on *factual* differences, due to the investigation and disclosure choices.

For my *investigation* manipulation, the ad-hoc disclosure informs participants in the *forensic services (internal audit)* condition, that the forensic service specialists of a Big 4 accounting firm form an investigation team in cooperation with the internal audit department (the internal audit department alone forms an investigation team). Because I did not screen participants for whether they were familiar with forensic services or internal auditing, participants can also access (via clicking on a button) a short

¹⁹ Ad-hoc disclosures are the German form of disclosures of inside information pursuant to Art. 17 Par. 1 of the European market abuse regulation, REGULATION (EU) No 596/2014.

explanation of forensic services (internal audit departments) and a typical forensic service engagement (internal audit investigation).²⁰

Similar to the disclosure manipulation, note that the experimental design choices aim to hold important *factual* dimensions constant across conditions, for which my hypothesis predicts differences on a *perceptual* level. The wording describing the goal, scope, and process of the investigation is identical in both conditions. Moreover, I describe a team of investigators of the same size and experience in both conditions to make sure there is no *factual* difference of scope or quality because the forensic experts work in cooperation with the internal audit department in the forensic service condition. Hence, any differences in participants' *perception* of whether the investigation will be of sufficient quality to resolve the suspicion could only be driven by the (assurance) signal of the forensic services engagement.

4.3.4 Dependent Variable and Additional Variables

The dependent variable for my main tests is the difference in willingness to invest after manipulation minus willingness to invest before manipulation. Willingness to invest is the sum of the answers to the following two items on a 101-point scale ranging from 0 (very unattractive; very unlikely) to 100 (very attractive; very likely):

Attractiveness: “How attractive is an investment in Spechem stock as part of your diversified portfolio?”

Willingness to buy: “How likely are you to invest in Spechem stock as part of your diversified portfolio?”

Both items are from Elliott et al. (2018). As in their paper, I ask both questions before manipulation (informing participants that they make an initial assessment) and after manipulation (final assessment). I use the difference of the two assessments as my main

²⁰ Boritz et al. (2015) suggest that fraud experts spend 52% and 37% of their time with fraud- and litigation-related activities, respectively. My explanation of forensic services focuses on fraud-related activities and does not mention litigation-related activities, because only the former is relevant in my experimental setting. 105 of 137 participants (76.6%) clicked the information button at least once. The number of clicks per participant does not significantly differ across cells ($p = .46$) and responses to my main dependent variable do not differ between those participants who clicked on the information button at least once and those who did not click on the button ($p = .23$).

dependent variable to address differences in individual baseline propensity to invest in the company which is unrelated to my experimental case.

To gain additional insight into the theoretical mechanisms underlying my hypothesis, I collect a number of additional items in the debriefing questionnaire. I measure all items on a 101-point scale ranging from 0 (e.g., very unlikely, completely disagree) to 100 (e.g., very likely, completely agree). Firstly, I use three items to explore arguments that the assurance signal of my forensic services manipulation affects perceived investigation quality (as a combination of perceived competence and perceived independence of the investigation team). I ask participants to rate the independence, competence, and trustworthiness of the investigation team and combine these three items to form an investigation quality score.²¹ Secondly, as I argue that self-disclosure could improve the credibility of the board, I ask participants to rate the competence and trustworthiness of the board and follow prior literature to combine these two items to form a credibility score (Koonce and Lipe, 2010). Thirdly, I combine three items to form a perceived severity score. I explicitly ask for a severity judgment and add questions on the perceived likelihood that the investigation leads to future losses and the perceived extent to which it affects future earnings. Finally, I ask participants how much they perceive the board to withhold information and to try to influence investor opinions through the investigation. I use each of these items as a separate measure of withholding of information and persuasive intent, respectively. The debriefing questionnaire also includes open-ended questions for participants to describe the most important factors underlying their investment decisions.

4.4 Results

4.4.1 Manipulation and Comprehension Checks

I use two questions to check whether participants understood their manipulation. For the *disclosure* manipulation, I ask participants who revealed the potential fraud in Spechem's subsidiary to the public. For the *investigation* manipulation, I ask participants who conduct the internal investigation considering the potential fraud in Spechem's subsidiary. In addition, I use three comprehension checks to check whether participants

²¹ I added trustworthiness, because prior literature combines competence and trustworthiness to a credibility score which typically has high validity, and because I use this scale to assess the credibility of the board (Koonce and Lipe, 2010).

attentively read and understood the experimental material. They had to recall case information on whether Spechem's revenue grew from 2018 to 2019, whether the chemical industry was described as a growing market, and whether the suspected fraud led to an overstatement of 2018 and 2019 revenue. No participant failed more than one of these three checks, and 116 participants (84.7%) answered all three correctly. 114 participants (83.2 %) and 122 participants (89.1 %) correctly answered the manipulation check for the investigation manipulation and the disclosure manipulation, respectively.²² Four participants failed both manipulation checks. I hence conclude that participants, on average, paid sufficient attention and understood their manipulation, and remove these four participants from my sample. Moreover, I remove two participants who took an unusual amount of time (more than four standard deviations away from the mean time), two participants who provided completely nonsensical reasoning for their investment decision,²³ and the only participant who failed one manipulation check, one additional comprehension checks, and failed to provide any reference to the fraud in the open-ended responses.²⁴

4.4.2 Test of Hypothesis

Panel A of Table 2 shows the number of participants in each experimental cell as well as means and standard deviations of the difference in their willingness to invest in the company after manipulation minus before manipulation.²⁵

²² The frequency of failures of manipulation checks did not differ between cells (before exclusion of the nine participants: $p > .54$; after exclusion of the nine participants: $p > .38$).

²³ One participant simply replied "unsure" and one participant claimed that there was no online information available about the company (which was obvious, because I inform participants that the company is fictitious).

²⁴ My results remain qualitatively unchanged when I retain these nine participants. When I additionally remove all participants who fail any of the two manipulation checks (remaining sample $N = 102$) or any of the five checks (remaining sample $N = 90$) my results remain qualitatively unchanged, but some tests become insignificant at conventional levels ($p < .14$ for the sample of 102, $p < .21$ for the sample of 90).

²⁵ Cronbach's alpha for the two questions which form the willingness to invest score is .87 before manipulation and .90 after manipulation. This confirms that they measure the same underlying construct. When I use either of the questions instead of the combined measure, my results remain qualitatively unchanged when I use the *Willingness to buy* question and are qualitatively similar, but insignificant ($p < .31$) for the *Attractiveness* question. Moreover, when I run an ANCOVA with willingness to invest after manipulation as the dependent variable and willingness to invest before manipulation as a covariate, results are qualitatively unchanged, but some tests become insignificant at conventional levels ($p < .14$).

Table 2: Results of ANOVA analyzing the difference in willingness to invest

Panel A Descriptive statistics of willingness to invest for each experimental cell				
Mean of difference of willingness to invest after minus before manipulation (Std. deviation)				
	Internal audit		Forensic services	
	<i>N</i> = 32		<i>N</i> = 28	
Self-disclosure	-72.2 (46.65)		-59.8 (41.53)	
	<i>N</i> = 40		<i>N</i> = 28	
Press revelation	-63.3 (44.64)		-79.3 (42.22)	
Panel B Results of ANOVA with difference of willingness to invest as dependent variable				
Variable	Degrees of freedom	Type III sum of squares	F-statistic	p-value
Intercept	1	160402	82.922*	<.001
<i>Investigation</i>	1	4196	2.169	.143
<i>Disclosure</i>	1	1406	0.727	.396
<i>Investigation x Disclosure</i>	1	6314	3.264*	.073
Error	124	239863		
Panel C Follow-up simple effects tests				
Comparison			T-statistic	p-value
<i>Forensic services and self-disclosure</i> versus <i>Forensic services and press revelation</i>			1.74*	.044
<i>Internal audit and self-disclosure</i> versus <i>Internal audit and press revelation</i>			-0.82	.416

Notes: Panel A presents the number of observations (*N*), and the mean and standard deviation (in parentheses) of the difference of willingness to invest for each experimental cell. Columns denote the investigation manipulation and rows denote the disclosure manipulation. The difference of willingness to invest is willingness to invest after manipulation minus willingness to invest before manipulation. Willingness to invest is the sum of the responses to the following two questions (on a scale from 0 = very unattractive; very unlikely to 100 = very attractive; very likely): “How attractive is an investment in Spechem stock as part of your diversified portfolio?”; “How likely are you to invest in Spechem stock as part of your diversified portfolio?”

Panel B reports results of a full-factorial ANOVA with the difference of willingness to invest as the dependent variable, and the investigation and disclosure manipulation as factors.

Panel C reports results of t-tests of difference of willingness to invest between the two cells mentioned in the comparison. The first comparison reflects my hypothesis. Therefore, the p-value is one-tailed. For the second comparison, I make no prediction as to the direction of the effect. Therefore, the p-value is two-tailed

* denotes significance at the 10%-level or better.

The pattern of means is largely consistent with my hypothesis, as the difference in willingness to invest is highest (i.e., least negative) for *Self-disclosure and forensic services* (-59.8) and lowest for *Press revelation and forensic services* (-79.3), and it is

higher for *forensic services* (vs. *internal audit*) in the *self-disclosure* conditions (-59.8 vs. -72.2), but lower in the *press revelation* conditions (-79.3 vs. -63.3).

Panel B of Table 2 presents the results of a full-factorial ANOVA with difference in willingness to invest as the dependent variable. The interaction effect of *Investigation* and *Disclosure* is significant at the 10%-level, and both main effects are insignificant. Panel B of Figure 2 depicts the mean difference in willingness to invest for each cell in order to get a better understanding of the pattern of the significant interaction effect. My hypothesis predicts a higher willingness to invest for self-disclosure versus press revelation in the forensic services condition and makes no prediction for the internal audit condition. The pattern is as expected. When the company engages forensic services, willingness to invest is higher when the firm self-discloses than when the firm is caught by the press. A follow-up simple effects test for this positive difference between self-disclosure and press revelation in the forensic services condition (Table 2, Panel C) is significant at the 5%-level. In contrast, when the company relies on internal audit only, willingness to invest is lower when it self-discloses than when it is caught by the press. However, a follow-up simple effects test shows that this negative difference is insignificant.²⁶ Overall, these results support my hypothesis.

4.4.3 Additional Analyses

As discussed above, I collected additional variables in the debriefing questionnaire to further explore the theoretical arguments underlying my hypothesis. In this section, I separately consider whether each of these variables mediates the significant interaction effect reported above (Baron and Kenny, 1986). I first run a series of ANOVAs with the additional variables as a dependent variable (replacing willingness to invest).²⁷ Panel A, B, and C of Table 3 present descriptive statistics for three of the additional variables, and Panel D of Table 3 shows the ANOVA results. For brevity, I suppress the intercepts and error terms as well as degrees of freedom and Type III sums of squares.

²⁶ In the forensic service condition, I use a one-tailed test due to the directional hypothesis. In the internal audit condition, I use a two-tailed test because I have no directional hypothesis. I do not report other possible simple effects tests, because they are not subject to my hypothesis. Untabulated results reveal no significant results for any other potential comparison of two cells ($p > .13$, two-tailed).

²⁷ Cronbach's alpha of these three scales is .86 for investigation quality, .87 for board credibility, and .70 for severity.

The results of the ANOVAs are largely consistent with the theory underlying my hypothesis. The first column of Table 3, Panel D shows that perceptions of investigation quality significantly differ between *Investigation* conditions. This is in line with arguments that the engagement of external specialists signals a higher quality investigation. Moreover, the insignificant interaction effect indicates that this positive signal is unconditional on an alignment of credible signals, i.e. it improves perceptions regardless of the source of disclosure. Therefore, although it generally supports theoretical arguments discussed herein, it cannot be a mediator for the interaction effect described in the main analysis.

Next, the second column of Table 3, Panel D reveals that participants perceive the board to be more credible when the suspicion is self-disclosed. Again, this is unconditional of an interaction with the type of investigation. Hence, despite being consistent with the theory discussed above, this phenomenon cannot mediate the main result.

The third, fourth, and fifth columns of Table 3, Panel D do reveal potential mediators, as investors perceive the severity of the suspicion, the withholding of information, and the persuasive intent of the board significantly differently based on an interactive effect of both of my manipulations. However, for severity and persuasive intent, there is a significant main effect of *Disclosure* in addition to the significant interaction. Hence, I next inspect the pattern of results to get a better understanding of whether these potential mediators mirror the pattern of results of my main analysis. Patterns of cell means of these variables shown in Panels A, B, and C of Table 3 reveal that only the results with respect to the withholding of information fit to my main results. Specifically, perceptions of withholding of information follow the exact same pattern as my main results (best perception for forensic services and self-disclosure, worst perception for forensic services and press revelation, and perceptions in-between for internal audit conditions). Perceptions of severity and of the persuasive intent show a different pattern; internal audit with self-disclosure leads to the worst perceptions, internal audit with press revelation leads to the best perceptions, and the forensic services conditions are in-between.²⁸ In conclusion, the withholding of

²⁸ I also conduct the same follow-up simple effects tests as in my main analysis (untabulated). For withholding of information, withholding perceptions are significantly lower for forensic and self-disclosure compared to forensic and press revelation. No other comparison is significant. This is the same pattern as in my main analysis. In contrast, severity perceptions are significantly higher for internal audit and self-disclosure compared to internal audit and press revelation, and no other comparison is significant. Perceptions of persuasive intent are significantly higher for internal audit and self-disclosure compared to any other cell, and no other comparison is significant.

information is the only additional variable with a significant interaction that resembles my main results. Therefore, this is the only potential mediator I investigate further.

To do so, I take the model from my main analysis and add withholding of information as a control to explore whether the interaction effect loses its significance. This results in the following model, which I now estimate using OLS regression:

$$\Delta WTI_i = b_1 + b_2 Inv_i + b_3 Disc_i + b_4 Inv_i \times Disc_i + b_5 withholding\ information_i + \varepsilon_i \quad (1)$$

where ΔWTI_i is the difference in willingness to invest after minus before manipulation, Inv_i is an indicator variable (1 for forensic services, 0 for internal investigations), $Disc_i$ is an indicator variable (1 for self-disclosure, 0 for press revelation), and *withholding information* is the withholding of information scale. Full mediation occurs when the coefficient of withholding of information, b_5 , is significant, and the coefficient of the interaction term, b_4 , becomes insignificant.

Table 3: Mediation analyses

Panel A Descriptive statistics of severity for each experimental cell		
	Mean of <i>severity</i>	
	Internal audit	Forensic services
Self-disclosure	231.7	217.4
Press revelation	206.3	219.9
Panel B Descriptive statistics of withholding information for each experimental cell		
	Mean of <i>withholding information</i>	
	Internal audit	Forensic services
Self-disclosure	60.1	50.8
Press revelation	58.7	64.0
Panel C Descriptive statistics of persuasive intent for each experimental cell		
	Mean of <i>persuasive intent</i>	
	Internal audit	Forensic services
Self-disclosure	61.3	46.1
Press revelation	43.5	45.0

Panel D ANOVAs with additional variables as dependent variables

Variable	DV = <i>investigation quality</i>	DV = <i>board credibility</i>	DV = <i>severity</i>	DV = <i>withholding information</i>	DV = <i>persuasive intent</i>
	F-statistic (p-value)	F-statistic (p-value)	F-statistic (p-value)	F-statistic (p-value)	F-statistic (p-value)
<i>Investigation</i>	7.970* (.006)	0.699 (.405)	1.446 (.231)	0.952 (.331)	0.050 (.824)
<i>Disclosure</i>	1.214 (.273)	2.936* (.089)	5.431* (.021)	0.078 (.780)	7.076* (.009)
<i>Investigation x Disclosure</i>	0.001 (.973)	0.072 (.789)	2.888* (.092)	3.412* (.067)	2.754* (.100)

Panel E OLS regression with willingness to invest as dependent variable and potential mediators as control

Variable	DV = <i>difference in willingness to invest</i>	
	Mediator = <i>withholding information</i>	
	T-statistic	(p-value)
<i>Investigation</i>	-1.384	(.169)
<i>Disclosure</i>	-0.828	(.409)
<i>Investigation x Disclosure</i>	1.627	(.106)
<i>withholding information</i>	-0.947	(0.346)

Notes: Panel A presents the mean of *severity* for each experimental cell. *Severity* is the sum of the responses to the following three questions (on a scale from 0 = not severe at all; very unlikely; not at all affected to 100 = very severe; very likely; very affected): “How severe is the investigation of the suspicion?”; “How likely is Spechem to incur a future material loss due to the investigation?”; “To what extent are Spechem’s future earnings affected by the investigation?” Panel B presents the mean of *withholding information* for each experimental cell. *Withholding information* is the response to the following question (on a scale from 0 = very unlikely to 100 = very likely): “How likely does Spechem’s board withhold private information about the investigation?” Panel C presents the mean of *persuasive intent* for each experimental cell. *Persuasive intent* is the response to the following question (on a scale from 0 = completely disagree to 100 = completely agree: “To what extent do you agree with the following statement?: The aim of the investigation is to influence your opinion on Spechem.” In Panels A, B, and C, columns denote the investigation manipulation, and rows the disclosure manipulation. In Panel D, each column presents results from a full-factorial ANOVA with the variable mentioned in the column title as the dependent variable (DV), and the investigation and disclosure manipulation as factors. I suppress intercepts and errors for brevity. *Investigation quality* is the sum of the responses to the following three questions (on a scale from 0 = very low independence; very untrustworthy to 100 = very high independence; very competent; very trustworthy: “How do you assess the independence of the team that investigates the suspicion?”; “How competent is the team that investigates the suspicion?”; “How trustworthy is the team that investigates the suspicion?” *Board credibility* is the sum of the responses to the following two questions (on a scale from 0 = very incompetent; very untrustworthy to 100 = very competent; very trustworthy: “How competent is Spechem’s board?”; “How trustworthy is Spechem’s board?” Panel E presents results from OLS regressions of the following model:

$$\Delta WTI_i = b_1 + b_2 Inv_i + b_3 Disc_i + b_4 Inv_i \times Disc_i + b_5 withholding\ information_i + \varepsilon_i$$

where ΔWTI_i is the difference in willingness to invest after minus before manipulation, Inv_i is an indicator variable (1 for forensic services, 0 for internal audit), $Disc_i$ is an indicator variable (1 for self-disclosure, 0 for press revelation), and *withholding information* is as defined above. I suppress intercepts for brevity. * denotes significance at the 10%-level or better.

Panel E of Table 3 shows OLS regression results. Again, I suppress the intercepts for brevity. The results do not reveal mediation because the coefficient of *withholding information* is insignificant. While the interaction effect becomes (just) insignificant, this missing direct effect does not allow to conclude that there is mediation.

Finally, responses to my open-ended question underline that the suggested mechanisms may be at play for the average participant. For example, participants in the self-disclosure and forensic services condition underlined their investment decisions with the following statements: “They have shown honesty with shareholders and potential shareholders”, “Independent team investigating”, “The investigation will look into any misdemeanors”, and “The fraudulent behaviour is most likely an isolated incident”. Participants in the press revelation and forensic services condition, however, explained their investment decisions as follows: “The investigation makes it hard to believe either part”, “Culture that created the lying”, “Information not forthcoming enough from board”, and “Makes you curious as to whether there are other investigations that need to take place at Spechem”.

Taken together, my additional analyses reveal that investor perceptions and behavioral intentions are in line with my theoretical predictions, but I am unable to establish a clear causal path for my results based on mediation analysis.

4.5 Conclusion

Given the prevalence of private investigations of suspicions of fraud, I analyze whether engaging external specialists for such investigations changes investors' willingness to invest in the company under investigation, based on whether the company self-discloses the situation or whether it is revealed by the press. If the suspicion is self-disclosed, investors likely perceive the engagement of external specialists as a genuine signal that the company is undertaking a credible investigation (alignment of genuine behaviors). In contrast, when the press reveals the suspicion, investors may suspect that the company withholds private information, and the engagement of external specialists might signal a higher likelihood of a more severe issue. Importantly, these differences in perception may arise even if there is no reason to believe that the factual investigation quality is different in these situations. In an experiment with 128 private investors, holding factual information suspicion and investigation quality constant, I find evidence

for my hypothesized interaction. My results suggest that alignment of credible investigation efforts with forthcoming disclosure could be beneficial. Companies with a genuine interest in transparent investigation could be better off resisting any possible temptation of non-disclosure of the issue.

My paper is subject to several limitations. As in many experimental studies, my sample is rather small and only consists of a subgroup of share- and broader stakeholders of a company. Therefore, while I can explore causal relations and can contribute to our theoretical understanding of how investors perceive firm disclosure and mitigating behavior in potential fraud situations, I do not observe the real-life dynamics of similar suspicions. A strategy to identify a broad sample of publicly known private investigations in archival data sets could be a worthwhile avenue for future research addressing this limitation.

As discussed above, my experiment only covers situations, in which the suspicion eventually reveals itself. In reality, there might be situations in which decision-makers deem it a realistic option to conceal such an issue. While this does not jeopardize the internal validity of the experiment, the real decision-making process on whether to disclose dependent on the perceived probability of being able to conceal an issue is a complex task. This must be taken into account when interpreting the present findings with respect to decisions to disclose. This opens an important avenue for future research. Such research could investigate how decision-makers choose to disclose under real-life conditions, which strategies can be observed, and how self-disclosure could be promoted considering that it may be beneficial when compared to being caught. In addition, it would be desirable to explore to what extent, and under which conditions, companies under investigation are able to conceal an investigation.

Finally, in my additional analysis, I use several variables that are intended to capture constructs from my hypothesis development. Most results with respect to these variables are consistent with my theoretical arguments. However, I am unable to provide a clear causal link from my manipulations through these additional variables to investors' willingness to invest. Therefore, my additional analyses can only give a limited additional understanding of the mechanisms underlying the overall result of my

experiment. Future research is necessary to enhance our understanding of external forensic specialists, the firm's crisis communication strategies, and investor behavior.

Despite these limitations, the present research provides novel evidence on the potential value of external specialists for private investigations, conditions of such value creation, and potential mechanisms through which such value is created, an area which the literature has rarely investigated. This area warrants more research to get a more detailed understanding of how stakeholders perceive private investigations and the use of external specialists, how high actual competence and independence of forensic services are, whether regulation on fraud audits and forensic services could decrease incidents of corporate fraud, and what role forensic experts could and should play in the financial statements audit.

4.6 Appendix A: Main body of the research instrument (screenshots)

For brevity, I excluded the formal welcome, informed consent, technical instructions for Prolific compensation, and demographics; all of the excluded pages come before or after the main body presented below. Unless noted differently, each page in this Appendix equals one page which participants view in their browser. They were unable to navigate backwards.

Unless noted differently, all participants view all information.

Please assume that you own a significant portfolio of stock. You are doing a regular review of that portfolio, and you are considering investing in the special chemicals industry to maintain a diversified portfolio. On the following page, we will present you financial and non-financial information on the (fictitious) company Spechem AG, a listed chemical company. Please use this information to **make your investment decision**. There are no right or wrong answers and we are not interested in how others would decide in these scenarios.

Leave and delete my data

Next

Short description of Spechem AG:

Spechem AG is a leading specialty chemicals company. Despite its comparably high maturity, the chemical industry remains a growing market. To participate in that growth, Spechem AG concentrates on attractive growth markets, e.g. resource efficiency, and conducts market oriented research and development in close collaboration with its customers.

Since the chemical industry faces a radical transformation, the challenge for many chemical companies currently is to serve existing and developing areas of application with sustainable products and to design global supply changes sustainably, efficiently, securely, and fairly. This is necessary to remain competitive to producers with easy access to cheap resources and energy (e.g. Middle-East, USA) which expand their value chains.

Spechem AG operates internationally with three operative chemical segments and the segment Services.

- The segment *Health & Nutrition* produces customized primary and intermediate products as well as chemically active ingredients for applications in consumer goods.
- The segment *Efficiency Chemicals* offers process and functional chemicals for environmental-friendly and energy-efficient systems solutions for the automobile, paints, adhesives, and construction industry. Both segments predominantly operate in attractive markets with above average growth and offer customized, individual and innovation-driven solutions.
- The segment *Performance Materials* is characterized by resource- and energy-intensive processes and comprises activities in the area of technical polymers and artificial rubber. Through cost optimized technology platforms, efficient processes, and economies of scale, Spechem aims to secure its market position for the future.
- The segment *Services* offers services to internal and external customers at Spechem facilities and supports the operative business as well as the management holding with company-wide, standardized administrative services.

Important KPIs of Spechem AG:

KPI	2019	2018	change
Revenue (million €)	20,900	19,200	+ 8.9%
Adjusted EBITDA (million €)	3,612	3,360	+ 7.5%
Adjusted EBITDA-return (%)	17.28	17.50	- 0.22 points
Earnings per share (€)	3.84	3.06	+ 25.5%
Operating cashflow (million €)	3,101	2,073	+ 49.6%

Recall that you own a diversified portfolio of stocks and you are considering investing in the chemical industry. One of the firms you are considering is Spechem. It is June 30, 2020, and you are performing an initial assessment of Spechem.

How attractive is an investment in Spechem stock as part of your diversified portfolio?

Please select a value between 0 and 100 by using the slide bar below.

very unattractive ————— very attractive

How likely are you to invest in Spechem stock as part of your diversified portfolio?

Please select a value between 0 and 100 by using the slide bar below.

very unlikely ————— very likely

Next

Leave and delete my data

Only for participants in condition *self-disclosure and forensic services*

Before answering the questions below, please assume that the following public disclosure has just been issued by Spechem:

DGAP-AdHoc July 1, 2020: Spechem announces potential issue in 2018 and 2019 financial reports:

"Two weeks ago, Spechem received an internal report on the performance of its international subsidiaries in the Performance Materials segment. We decided to proactively inform our shareholders on a potentially critical issue which will have to be analyzed in further detail in the coming weeks. It relates to sales of the international subsidiaries of € 200 million (3.7% of segment sales) for fiscal year 2018 and € 250 million (5.2% of segment sales) for fiscal year 2019 which are potentially fraudulently inflated. Spechem's board apologizes for the potentially intentional misbehavior within the firm and takes responsibility for any material misbehavior, although we assure that the board has not been involved. The board hired the forensic service specialists of a Big 4 accounting firm who have formed a team of 5 experienced auditors in cooperation with our internal audit department that currently carefully investigates the suspicion. In case of a fraud, they will investigate whether systematic deficiencies led to this potential misbehavior or whether it was the fault of individuals. As soon as this investigation is completed, we will take the appropriate remedial actions."

[Click on this button to view some basic information on Forensic Services.](#)

[Click on this button to view the information on Spechem AG from the previous page.](#)

It is time to make your final investment assessment. Please note that your answer to each question should reflect your final investment assessment, not how you would adjust your preliminary assessment.

How attractive is an investment in Spechem stock as part of your diversified portfolio?

Please select a value between 0 and 100 by using the slide bar below.

very unattractive ————— very attractive

How likely are you to invest in Spechem stock as part of your diversified portfolio?

Please select a value between 0 and 100 by using the slide bar below.

very unlikely ————— very likely

[Next](#)

[Leave and delete my data](#)

Only for participants in condition *self-disclosure and internal audit*

Before answering the questions below, please assume that the following public disclosure has just been issued by Spechem:

DGAP-AdHoc July 1, 2020: Spechem announces potential issue in 2018 and 2019 financial reports:

"Two weeks ago, Spechem received an internal report on the performance of its international subsidiaries in the Performance Materials segment. We decided to proactively inform our shareholders on a potentially critical issue which will have to be analyzed in further detail in the coming weeks. It relates to sales of the international subsidiaries of € 200 million (3.7% of segment sales) for fiscal year 2018 and € 250 million (5.2% of segment sales) for fiscal year 2019 which are potentially fraudulently inflated. Spechem's board apologizes for the potentially intentional misbehavior within the firm and takes responsibility for any material misbehavior, although we assure that the board has not been involved. Our internal audit department has formed a team of 5 experienced auditors that currently carefully investigates the suspicion. In case of a fraud, it will investigate whether systematic deficiencies led to this potential misbehavior or whether it was the fault of individuals. As soon as this investigation is completed, we will take the appropriate remedial actions."

[Click on this button to view some basic information on Internal Audit.](#)

[Click on this button to view the information on Spechem AG from the previous page.](#)

It is time to make your final investment assessment. Please note that your answer to each question should reflect your final investment assessment, not how you would adjust your preliminary assessment.

How attractive is an investment in Spechem stock as part of your diversified portfolio?

Please select a value between 0 and 100 by using the slide bar below.

very unattractive ————— very attractive

How likely are you to invest in Spechem stock as part of your diversified portfolio?

Please select a value between 0 and 100 by using the slide bar below.

very unlikely ————— very likely

[Next](#)

[Leave and delete my data](#)

Only for participants in condition *press revelation and forensic services*

Before answering the questions below, please assume that the following press release has just been issued, followed by the subsequent public disclosure by Spechem:

The renowned business magazine "Shareholder Weekly" reports on its website as follows (July 1, 2020):

"Our investigative journalists received a secret internal report of chemical company Spechem from two weeks ago on the performance of its international subsidiaries in the Performance Materials segment. In the report, Spechem identified a potentially critical issue which will have to be analyzed in further detail in the coming weeks. It relates to sales of the international subsidiaries of € 200 million (3.7% of segment sales) for fiscal year 2018 and € 250 million (5.2% of segment sales) for fiscal year 2019 which are potentially fraudulently inflated."

Related DGAP-AdHoc July 1, 2020: Response to press statement published by "Shareholder Weekly":

"Spechem shareholders are referred to the statements made by "Shareholder Weekly". We received the cited report two weeks ago and it will have to be analyzed in further detail in the coming weeks. Spechem's board apologizes for the potentially intentional misbehavior within the firm and takes responsibility for any material misbehavior, although we assure that the board has not been involved. The board hired the forensic service specialists of a Big 4 accounting firm who have formed a team of 5 experienced auditors in cooperation with our internal audit department that currently carefully investigates the suspicion. In case of a fraud, they will investigate whether systematic deficiencies led to this potential misbehavior or whether it was the fault of individuals. As soon as this investigation is completed, we will take the appropriate remedial actions."

[Click on this button to view some basic information on Forensic Services.](#)

[Click on this button to view the information on Spechem AG from the previous page.](#)

It is time to make your final investment assessment. Please note that your answer to each question should reflect your final investment assessment, not how you would adjust your preliminary assessment.

How attractive is an investment in Spechem stock as part of your diversified portfolio?

Please select a value between 0 and 100 by using the slide bar below.

very unattractive ————— very attractive

How likely are you to invest in Spechem stock as part of your diversified portfolio?

Please select a value between 0 and 100 by using the slide bar below.

very unlikely ————— very likely

[Next](#)

[Leave and delete my data](#)

Only for participants in condition *press revelation and internal audit*

Before answering the questions below, please assume that the following press release has just been issued, followed by the subsequent public disclosure by Spechem:

The renowned business magazine "Shareholder Weekly" reports on its website as follows (July 1, 2020):

"Our investigative journalists received a secret internal report of chemical company Spechem from two weeks ago on the performance of its international subsidiaries in the Performance Materials segment. In the report, Spechem identified a potentially critical issue which will have to be analyzed in further detail in the coming weeks. It relates to sales of the international subsidiaries of € 200 million (3.7% of segment sales) for fiscal year 2018 and € 250 million (5.2% of segment sales) for fiscal year 2019 which are potentially fraudulently inflated."

Related DGAP-AdHoc July 1, 2020: Response to press statement published by "Shareholder Weekly":

"Spechem shareholders are referred to the statements made by "Shareholder Weekly". We received the cited report two weeks ago and it will have to be analyzed in further detail in the coming weeks. Spechem's board apologizes for the potentially intentional misbehavior within the firm and takes responsibility for any material misbehavior, although we assure that the board has not been involved. Our internal audit department has formed a team of 5 experienced auditors that currently carefully investigates the suspicion. In case of a fraud, it will investigate whether systematic deficiencies led to this potential misbehavior or whether it was the fault of individuals. As soon as this investigation is completed, we will take the appropriate remedial actions."

[Click on this button to view some basic information on Internal Audit.](#)

[Click on this button to view the information on Spechem AG from the previous page.](#)

It is time to make your final investment assessment. Please note that your answer to each question should reflect your final investment assessment, not how you would adjust your preliminary assessment.

How attractive is an investment in Spechem stock as part of your diversified portfolio?

Please select a value between 0 and 100 by using the slide bar below.

very unattractive ————— very attractive

How likely are you to invest in Spechem stock as part of your diversified portfolio?

Please select a value between 0 and 100 by using the slide bar below.

very unlikely ————— very likely

[Next](#)

[Leave and delete my data](#)

Information that pops up if the first button is clicked. If the second button is clicked, the information from the previous page pops up again.

Only for participants in any of the *forensic service* conditions

[Click on this button to view some basic information on Forensic Services.](#)

Forensic Services provide help for investigating and preventing corporate crime. In this case, they are provided by a special division of an accounting firm. This division consists of an interdisciplinary team, e.g. CPAs, lawyers, criminologists, and IT specialists. A typical example for a Forensic Service is the investigation of the suspicion of an instance of corporate crime in a company, just as in the presented case. Forensic Services can also be engaged for fraud prevention, e.g. by administering a whistle blowing process or by offering anti-fraud trainings.

When Forensic Services are engaged in such a case, they typically entail the goal to investigate the issue in a systematic way with the help of the Forensic Services team. To do so, the team helps with an internal investigation which produces results that can be used in court. They also deliver decision relevant information to executives and the board. The team typically has considerable relevant experience with corporate crime, as opposed to internal employees who have no or limited such experience because corporate crime is a comparably rare event. Further particularities of Forensic Services are access to technological specialists, e.g. to analyze deleted data, and to an international network, e.g. when a foreign subsidiary is involved and foreign law may play a role.

[Hide](#)

Only for participants in any of the *internal audit* conditions

[Click on this button to view some basic information on Internal Audit.](#)

Internal audit is a division of a company which internally provides independent and objective auditing and consulting services to analyze and improve the effectivity of risk management, controls, and leadership and governance processes. These systematic and goal-oriented services can be both reactive and proactive and also include cases of corporate crime, especially its prevention. Internal auditors should therefore have sufficient knowledge about risks with respect to malicious acts. This knowledge, however, is not expected to reach the level of a corporate crime expert.

The investigation of a suspicion of corporate crime or an already occurred fraud are also part of the activities of internal auditors and become increasingly important. Systematic approaches in this area can comprise the detection, investigation and remediation of corporate crime. Internal auditors sometimes collaborate with law enforcement authorities to produce results that can be used in court and deliver decision relevant information to executives and the board. After fraud cases, they can also help with fraud prevention, e.g. by administering a whistle blowing process or by offering anti-fraud trainings. However, the main tasks and main areas of expertise of the internal audit function remain the comprehensive audit of diverse operational processes and structures and a critical evaluation thereof.

[Hide](#)

Please answer the following questions.

For each item please select a value between 0 and 100 by using the slide bar below.

How forthcoming is Spechem in providing information to investors?

very unforthcoming very forthcoming

To what extent do you agree with the following statement?: The specialists who investigate the suspicion have a reputation of conducting high-quality fraud investigations.

completely disagree completely agree

How competent is Spechem's board?

very incompetent very competent

How trustworthy is Spechem's board?

very untrustworthy very trustworthy

How likely does Spechem's board withhold private information about the investigation?

very unlikely very likely

How do you assess the independence of the team that investigates the suspicion?

very low independence very high independence

How competent is the team that investigates the suspicion?

very incompetent very competent

How trustworthy is the team that investigates the suspicion?

very untrustworthy very trustworthy

How severe is the investigation of the suspicion?

not severe at all very severe

How likely is Spechem, to incur a future material loss due to the investigation?

very unlikely very likely

To what extent are Spechem's future earning affected by the investigation?

not at all affected very affected

To what extent do you agree with the following statement?: The aim of the investigation is to influence your opinion on Spechem.

completely disagree completely agree

To what extent do you agree with the following statement?: The aim of the investigation is to carefully investigate the suspicion and take the appropriate remedial actions.

completely disagree completely agree

Please list below 1 to 3 key factors supporting your investment decision.

Next

Leave and delete my data

On this page, we do no longer ask for your opinion, but ask you about the scenario you have been presented on the previous page.

Spechem's revenue increased from 2018 to 2019.

Is this statement correct?

- True
 False

The chemicals industry remains a growing market.

Is this statement correct?

- True
 False

Spechem's revenue is potentially fraudulently overstated in 2018 and 2019.

Is this statement correct?

- True
 False

Who revealed the potential fraud in Spechem's subsidiary to the public?

- The company itself
 Investigative journalists from the renowned business magazine "Shareholder Weekly"

Who conducts the internal investigations considering the potential fraud in Spechem's subsidiary?

- Forensic service specialists of a Big 4 accounting company together with Spechem's internal audit department
 Spechem's internal audit department alone

Next

Leave and delete my data

Chapter 5 Do Non-Audit Service Failures Impair Auditor Reputation? An Analysis of KPMG Advisory Service Scandals in Germany

A revised version of this chapter is submitted as

Friedrich, C.; Quick, R. (2020) "Do non-audit service failures impair auditor reputation? An analysis of KPMG advisory service scandals in Germany" *Working Paper, under Review at Critical Perspectives on Accounting.*

Abstract

Non-audit services have become a major focus of the Big 4 accounting firms. Failures in these services may impair the accounting firm's reputation as an auditor. Along with short-term reputation effects, this may induce longer-term economic consequences and reactions by the accounting firm to counteract reputation losses. Accordingly, we analyze whether two events of observable non-audit service deficiencies of KPMG Germany impair KPMG's reputation as an auditor in the short run. Considering longer-term consequences, we investigate KPMG's ability to retain and gain clients, audit fees, and indicators of audit quality. With an event study, we find evidence for short-term negative market reactions to the average KPMG client. The follow-up analyses of a broad sample of German listed firms do not suggest that KPMG attracts less or loses more clients nor do we find a difference in their audit fees. However, we find indications of more aggressive earnings management by KPMG clients after the second event. Overall, we present evidence that, in a low-litigation setting, markets punish an audit firm for non-audit service failures in the short run, but longer-term economic consequences seem marginal.

5.1 Introduction

The Big 4 accounting firms have heavily increased their non-audit services revenue in recent years so that it now surpasses audit revenue (Rapoport, 2018; and Schmitt, 2019; for Germany: Streicher and Deveci, 2016). Since profit margins are higher for non-audit services (NAS) than for audits, the Big 4 are focusing on further increasing their non-audit revenue (Fröndhoff, 2018). Hence, in addition to possible threats of auditor-provided NAS to auditor independence, there are concerns that the auditing practice will suffer from a shift of focus of accounting firms from auditing to NAS. Such concerns are aggravated when turmoil due to a NAS failure demands additional attention from the accounting firm and its management (Baugh et al., 2018). These concerns may threaten an accounting firm's reputation, in general, and regarding their auditing service. Therefore, we ask the following research question: Do non-audit service failures have short-term negative effects on the reputation of the accounting firm's audit services? We also investigate potential longer-term consequences to their audit practice, including the ability to retain and gain clients, the level of audit fees, and indicators of audit quality.

To answer those research questions, we study the short- and longer-term consequences of two presumable NAS failures of KPMG in a low-litigation setting. In both cases, KPMG managed the sale of a (mostly) state-owned company in Western Germany, the racetrack and event location *Nürburgring* and the airport *Frankfurt-Hahn*. KPMG showed lax operations, giving a dubious candidate access to confidential data in the first case and failing to identify a shell company that finally became the preferred buyer candidate in the second case. As is typical and expected for the low-litigation setting in Germany, KPMG did not suffer any legal or regulatory consequences. Therefore, this setting lends itself to a clean investigation of reputation effects from NAS failures on the audit practice of an accounting firm.

Such an investigation provides a new perspective to auditing research regarding the reputation hypothesis, i.e. that high audit quality is motivated by auditor reputation concerns, especially in low-litigation settings (DeFond and Zhang, 2014; Weber et al., 2008). Conversely, in high-litigation settings, the insurance hypothesis that litigation risks motivate high audit quality is empirically hard to disentangle from the reputation

hypothesis. Following the latter, a major success factor for the Big 4 is the generation and preservation of a brand name with an impeccable reputation (Shapiro, 1983). Therefore, the Big 4 accounting firms, which present themselves as professional service firms with a broad service portfolio rather than pure audit firms, have to ensure that their entire range of services contribute to their reputation, or at least do not impair their brand name.

Baugh et al. (2018) have recently adopted this viewpoint. They investigate whether a sanctioned misconduct of KPMG's non-audit practice in the US impaired KPMG's reputation as an auditor, measured by audit fees and client gains and losses, compared to other Big 4 accounting firms. They do not find a negative effect on KPMG's reputation. Analyzing the same misconduct with a short-term event study, however, Brown et al. (2013) find significant negative capital market reactions for KPMG audit clients. In the high-litigation US setting, they largely attribute this result to the insurance hypothesis. Extending the reputation hypothesis even further, Donelson et al. (2019) study a context with a reputation impairment that is outside of the auditor's responsibility. They analyze the effect of non-accounting securities violations by audit clients on the auditor, although the auditor does not monitor the non-accounting conduct of the client. They use theory from the organizational reputation literature (e.g. Lange et al., 2011) that suggests different dimensions of reputation. Those include "being known for something", i.e. that any behavior of an organization influences the reputation of anything the organization does, and "general favorability", i.e. that any connotation of an organizations brand name, regardless of whether it stems from the organization's behavior, influences the reputation of anything the organization does. Consistent with the "general favorability" dimension, their results suggest that the reputation of the auditor (measured by bargaining power) decreases, but only for clients in the same industry and audited by the same audit office as the scandal client. As the setting of Baugh et al. (2018) resembles the narrower "being known for something" dimension, this inconsistency of results is surprising.

Therefore, we argue that revisiting the "being known for something" dimension of reputation in a low-litigation setting and the NAS context is worthwhile. It allows us to include direct short-term reactions that provide a conceptually different angle on

reputation compared to longer-term consequences. It addresses reputation perceptions of different stakeholders (shareholders vs. the contractor of the auditor, i.e., audit committees) and considers that they may revise their reputation perceptions over time (e.g. because the audit firm takes measures to counteract reputation losses or other signals to update reputation perceptions emerge). Besides, we can investigate both short-term and longer-term consequences in the same setting, potentially providing a more complete picture of these conceptually different reputation effects.

We follow Donelson et al. (2019) in using organizational reputation theory, but focus on the “being known for something” dimension, and first hypothesize that accused NAS failure events lead to negative abnormal stock returns of the audit clients of the accused audit firm. Our findings are consistent with this expectation for both events as we find an abnormal decline in one-day stock returns of about -0.5% for the average KPMG client at either event. However, robustness checks indicate that these results are sensitive to some design choices for the second event. To exclude that our findings are random or merely a result of heightened attention, we also analyze two pseudo-events as a falsification test. We choose these based on search engine activity taken from Google Trends as a measure of general attention (Da et al., 2011). Our results do not support this concern.

Following the basic economic underpinnings of the reputation hypothesis, when looking at longer-term consequences, this decline in reputation should impair the auditor’s ability to retain existing clients, gain new ones, and charge high audit fees (Klein and Leffler, 1981). However, the auditor may implement measures to counteract both the negative reputation effect and/or the expected loss of clients. To improve reputation, the auditor may improve audit effort to signal high audit quality, which may lead to less earnings management of audit clients and higher audit fees charged for the extra effort. Conversely, to attract clients despite reputation loss, the auditor may provide fee discounts or acquiesce to the client’s preferred accounting choices, which potentially increases earnings management. Finally, small changes in reputation may not affect auditor-changing behavior at all, because transaction costs of finding a new auditor are high, and optimal auditor-client pairings can have significant economic benefits above alternative pairings (Gerakos and Syverson, 2015). In conclusion, we hypothesize that, after the NAS failure events, KPMG clients are more likely to change to a new auditor

and KPMG is less likely to gain new clients. Given ambiguous theoretical arguments, we refrain from making directional hypotheses with respect to changes in audit fees and earnings management of KPMG clients.

Results from our regression analyses of large samples of German listed firms between 2012 and 2018 do not support our hypothesis that the NAS failures affected KPMG's ability to retain and gain audit clients. We also find no effect on audit fees of KPMG clients. However, we find some indication that accruals-based earnings management of KPMG clients became more aggressive after the second event, but not after the first event. Specifically, signed discretionary accruals, total accruals, and signed abnormal working capital accruals increase in the first full year after event 2. However, absolute discretionary accruals and abnormal working capital accruals do not change, suggesting that overall earnings management activity does not increase. These results are robust to a firm fixed effects model but become significantly weaker when excluding new KPMG clients and when changing the treatment period to the year of the scandal, in which part of the audit work had already been completed when the scandal period ended.

We contribute to the auditing literature concerned with the insurance and reputation hypotheses as incentives for audit quality. Our analyses extend the non-results in Baugh et al. (2018) with the use of a low-litigation setting that allows us to disentangle reputation from insurance effects in a short-term event study. Results indicate that non-audit service failures may have significant effects on the reputation of the accounting firm's audit practice in the short run. However, as in Baugh et al. (2018), we do not find results for longer-term economic consequences. The different findings in our short- and longer-term analyses suggest that perceptions may differ between different stakeholder groups (shareholders vs. audit committees) or may change over time. One potential reason that emerges from our analysis of audit quality is that KPMG was more lenient with aggressive client earnings management.

This is potentially concerning from the perspective of regulators and users of audited financial statements and points toward a potential audit quality threat in low-litigation jurisdictions. To the best of our knowledge, we provide the first non-US evidence of the overall effects of potential negative reputation signals from outside auditing on the audit firm and its audit clients. Therefore, we provide the first evidence in a low-litigation

environment. In a pure audit setting in low-litigation jurisdictions, research has shown that negative effects of failures or misconduct, which are not punished through litigation, are punished through short-term market reactions and longer-term economic consequences (Skinner and Srinivasan, 2012; Weber et al., 2008). Our results in the context of NAS deficiencies suggest that reputation compensates for a lack of litigation only for short-term reactions. The lack of longer-term consequences is potentially interesting to regulators in low-litigation jurisdictions, who are concerned with changing the litigation environment to enhance incentives for the provision of high quality (non-) audit services.

We also add to the literature on reputation in auditing by substantiating the use of organizational reputation theory introduced by Donelson et al. (2019). In addition to their findings that the “general favorability” dimension of reputation may play a role for audit firms, our findings suggest that the “being known for something” dimension of reputation also applies in the auditing setting.

The remainder of our paper is structured as follows. Section 5.2 discusses the theoretical background, prior empirical literature, our setting, and the timeline of events. In Section 5.3, we present the short-term event study. We present our analyses of longer-term consequences to KPMG’s audit practice in Section 5.4. The paper concludes with Section 5.5.

5.2 Background and Hypothesis Development

5.2.1 Theory on Incentives for Audit Quality and its Relation to Non-Audit Services

A fundamental theoretical question in research on audit quality is what drives the supply of high-quality audits. The literature has established two hypotheses, each on a different theoretical incentive for high audit quality: the insurance hypothesis and the reputation hypothesis (Skinner and Srinivasan, 2012). According to the insurance hypothesis, auditors have an incentive to provide high-quality audits because of the risk of litigation that may result in large payments to shareholders of the audit client in the event of an audit failure. This incentive is stronger for larger (presumably, financially stronger) audit firms because they can cover greater claims by plaintiffs (Dye, 1993; Lennox, 1999). Therefore, when doubts arise about the ability of the auditor to provide such insurance, this may have negative effects on the valuation of audit clients by investors (Brown et al., 2013; Dee et al., 2011).

The reputation hypothesis, on the other hand, proposes that an auditor can only gain new audit clients, retain existing ones, and charge satisfactorily high audit fees if the auditor generates and maintains a reputation of high trustworthiness and service quality. Therefore, the auditor has an incentive to avoid any kind of failure that impairs this reputation and harms the overall revenue-generation potential of the auditor (Davis and Simon, 1992; DeAngelo, 1981).

Traditionally, these theories focus on auditing services performed by the auditor. However, the reputation of audit firms is often considered a brand name reputation (Craswell et al., 1995; Ratzinger-Sakel and Schönberger, 2018), i.e. Big 4 brand names create reputation for all services of the audit firm. In this logic, a decrease of this brand name reputation harms the perceived value of all audit firm services, regardless of whether the failure occurred in the non-audit or audit practice of the audit firm (Selnes, 1993). This is consistent with theory from the literature on organizational reputation. Lange et al. (2011) argue that organizational reputation is a multidimensional construct and differentiate three dimensions, “being known for something”, “general favorability”, and “prominence”. We do not consider the latter dimension (i.e., stakeholders knowing an organization, regardless of their opinion about it) because stakeholders of NAS and auditing service know the Big 4 well. Besides, some authors consider prominence an antecedent of the other two dimensions (Agarwal et al., 2015). “Being known for something” refers to the reputation stemming from opinions about observed behavior or characteristics of the auditor, such as audit or NAS outcomes. “general favorability” broadly considers a good (bad) name regardless of underlying behavior or characteristics as the major driver of reputation. This dimension suggests that negative attention could impair reputation even if the auditor was neither directly nor indirectly responsible for what caused the attention.

Using these dimensions to extend the reputation hypothesis to NAS, the “being known for something” interpretation would suggest that observable shortcomings (more precisely, perceptions of shortcomings) in any Big 4 service, including NAS, would harm the reputation of all Big 4 services, including auditing. The “general favorability” interpretation would suggest that such effects are even broader, so that any kind of attention could potentially affect the auditor’s reputation.

5.2.2 Prior Research

The empirical auditing literature has collected extensive evidence consistent with both the insurance and reputation hypothesis in analyses of auditing services only. This research relies on settings in which the effects can be analyzed separately, as they are hard to disentangle in Anglo-Saxon settings, where both effects are theoretically substantial (Brown et al., 2013; Dee et al., 2011; Skinner and Srinivasan, 2012). Recent literature has broadened this view to include all dimensions of organizational reputation. Below, we discuss these two streams of literature and discuss how our setting is particularly informative in the context of prior research.

Several studies have tried to isolate the insurance hypothesis in highly litigious settings. Menon and Williams (1994) show negative abnormal stock returns of audit clients of Laventhol & Horwath (L&H) when it filed for bankruptcy, which corresponds to a sudden cessation of the insurance. At that time, L&H argued that their bankruptcy was not due to audit quality issues, but due to excessive litigation, in line with the insurance hypothesis. Following this argument, Menon and Williams (1994) interpret their findings as evidence of the insurance hypothesis. However, the filing for bankruptcy could still be a signal to the market that there had been audit quality issues, which would support the reputation hypothesis (Baber et al., 1995; Brown et al., 2013). Using a sample of initial public offerings, Willenborg (1999) separates so-called development-stage enterprises from more mature firms. As they have little revenue, development-stage firms do not have much meaningful accounting information yet, which makes a reputable auditor relatively unimportant. However, there is considerable uncertainty about their continuing existence, exacerbating their auditor's insurance role. Consistent with the insurance hypothesis, Willenborg (1999) provides evidence that auditor size matters even for development stage enterprises. Mansi et al. (2004) find that corporate bond yields are smaller for clients of larger auditors, even when controlling for credit ratings. They also show that rating agencies incorporate auditor characteristics related to audit quality into their credit ratings. They argue that, controlling for credit ratings, the incremental effect of auditor size is consistent with the insurance hypothesis. Beyond these three studies, other analyses in Anglo-Saxon countries, which do not disentangle the insurance and reputation hypotheses, report stronger evidence for the insurance hypothesis (e.g., Khurana and Raman, 2004; Lennox, 1999).

To isolate the reputation hypothesis, researchers have often used audit failures in settings without significant auditor litigation that creates the insurance value. Such settings include, for example, Germany (Weber et al., 2008) and Japan (Skinner and Srinivasan, 2012). These studies find significant negative effects of severe audit failures on the audit firm and its clients, which they interpret as support for the reputation hypothesis. In a US setting, due to the high auditor litigation, researchers first have to exclude litigation effects to establish a relation between reputation and adverse effects on the auditor. Boone et al. (2015) use a PCAOB disciplinary order against Deloitte in 2007 and find that Deloitte loses clients and growth opportunities for their audit fees in a three-year period after the disciplinary order. As they do not find publicly observable litigation, they claim that only reputation can drive this effect.²⁹

Turning from this literature of auditing-only contexts to the literature that considers a broader conceptualization of organizational reputation, we are not aware of any research using low-litigation settings to isolate the reputation hypothesis. However, some research in the highly litigious US-setting has recently established reputation effects on auditing that come from outside the narrow auditing context. Focusing on the broadest dimension of reputation, “general favorability”, Donelson et al. (2019) analyze client failures that are out of the control of the auditor. They find that auditor bargaining power as an indicator of reputation decreases in cases of non-accounting securities violations of a client without any responsibility of the auditor, but only for other clients of the same auditor, which are in the same industry and audited by the same auditor office as the scandal client.

Another group of prior research extends the more traditional approach of using auditor failure events by way of the reputation dimension “being known for something” from audit-only contexts to broader auditor-service contexts. It uses severe misconduct in the tax advisory services of KPMG (and the other Big 4, which were not prosecuted), which led to an investigation by the Department of Justice and was finally settled with a deferred prosecution agreement in 2005. To test the insurance hypothesis, Brown et al.

²⁹ In a follow-up study, Boone et al. (2019) find that, while this first-ever PCAOB sanction had negative consequences on the commercial success of the sanctioned audit firm, the effect disappears for subsequent PCAOB sanctions. They conclude that a punitive effect of repeat sanctions through reputational damage is absent.

(2013) conduct an event study with KPMG audit clients and find significant negative market reactions when the investigation emerges and positive market reactions after the deferred prosecution agreement. Both are consistent with the insurance hypothesis. While the authors do not completely rule out the reputation hypothesis, they argue that at least the positive market reactions to the deferred prosecution agreement likely depict a clean insurance effect. The settlement decreased the expected monetary penalties for KPMG but did not contain information on KPMG's audit quality. In a similar event study, Incardona et al. (2014) find that audit clients of the other Big 4 audit firms also suffered significant negative abnormal returns at several events surrounding the tax advisory scandal. This is consistent with the perception of the events as a collective scandal spanning all Big 4 firms, not only KPMG.

Turning to the reputation hypothesis, Baugh et al. (2018) follow the approach in Boone et al. (2015) to isolate reputation effects in a litigious setting when no litigation happened after a scandal. They compare the development of KPMG's audit fees, and client gains and losses, from pre- to post-scandal periods to the other Big 4. They find no negative effects and, hence, no support for the reputation hypothesis for NAS failures. However, if the stakeholders indeed perceived the scandal as a collective event of the Big 4, such a result is not surprising in a within-Big 4 analysis. Moreover, KPMG accepted responsibility for the misconduct through the deferred prosecution agreement and committed to improving their future behavior. This could lead to an expectation of informed stakeholders, such as audit committees, that KPMG will provide higher quality services in the future, which would alleviate negative reputation effects (Corona and Randhawa, 2018).

In conclusion, recent empirical research has provided some evidence that insurance and reputation effects extend from audit-only settings to broader considerations of auditor services and the affiliation with audit clients. While there is evidence for effects of the broad "general favorability" dimension of reputation, research has failed to find evidence for effects of the more focused "being known for something" dimension of reputation. However, there is some evidence for insurance effects in the same setting. This inconclusiveness of results seems surprising. As mentioned above, one explanation could be that collective effects on all Big 4 firms made effects unobservable in a within-

Big 4 analysis. Moreover, the high-litigation setting prevents researchers from drawing conclusions from direct, short-term reactions of investors. However, the latter allow for an investigation of potential reputation effects that conceptually differs from the analytic approach in Boone et al. (2015) and Baugh et al. (2018) in two important aspects. First, it focuses on investor perceptions versus perceptions of audit committees (or clients more broadly) that decide on who to hire as an auditor and negotiate audit fees. Moreover, the long-term nature of their analysis may mute short-term reductions of reputation that may revise once investors observe that no further adverse behavior follows the negative event.

Therefore, we argue that revisiting the potential effects of the “being known for something” dimension of reputation in the NAS context by using a low-litigation setting is important to improve our understanding of reputation incentives for audit firms. Studying short-term reactions and long-term consequences to audit firms from a reputation perspective in a low-litigation setting could potentially provide a more complete picture than what prior literature could analyze. Hence, we next describe our setting and timeline of events. Then, first, we conduct a short-term event study and, second, we analyze long-term consequences to the audit firm. We describe our two analyses and their results in two separate sections.

5.2.3 German Setting

Various particularities of the German setting provide advantages and tension for the analysis of the question of whether NAS failures have negative effects on the provider’s audit practice. Litigation risks are low in Germany because auditor liability is capped and considerable third-party litigation is absent.³⁰ Therefore, if these necessary incentives are present, reputation effects could be particularly high. Moreover, Germany also provides a setting in which the importance of statutory audits and, hence, investor sensitivity to auditor reputation differ from Anglo-Saxon settings. Some arguments

³⁰ For negligent misconduct, auditor liability against the client is capped generally at € 1 m, and at € 4 m for audits of listed companies (Section 323 (2) German Commercial Code – Handelsgesetzbuch (HGB)). For third party litigation, the German Civil Code (Bürgerliches Gesetzbuch (BGB)) requires the plaintiff to prove that the auditor has acted with intent. A remote possibility to sue auditors based on negligent misconduct derives from case law, referring to a contract with a protective effect for third parties. However, it requires very specific circumstances, like direct personal contact between the auditor and a third party, and typically applies a similar liability cap (Gietzmann and Quick, 1998).

support a lower (i.e., auditor reputation is less important and reputation effects are less pronounced) and some a higher (i.e., reputation effects are more pronounced) importance of statutory audits. Weber et al. (2008) argue that the importance of auditing is comparably low and, therefore, biases against finding negative market reactions in line with the reputation hypothesis. They reason that, as banks are often both debt providers and shareholders, their representatives closely monitor companies as members of supervisory boards³¹ and are active in corporate governance. This makes Germany a setting with high monitoring activity, which may partly substitute the monitoring role of external auditing. Yet, the portion of DAX (30 largest German listed companies) shares held by domestic shareholders decreased from 49% in 2005 to 37% in 2016 (EY, 2018). This indicates a reduction of shareholders with high monitoring activity, which might lead to more demand for external monitoring. Overall, German stock ownership is concentrated in the hands of institutional investors, and strategic investors (families and companies) play a significant role, leaving only small portions of traded stock for private investors (EY, 2018; Hackethal et al., 2005; La Porta et al., 1999). As institutional investors' primary source of information comes directly from directors via analyst conferences and roadshows, whereas private investors rely solely on financial reports (Pellens and Schmidt, 2014), auditing might play a less pronounced role in such a setting.

Although the German two-tier board system theoretically provides comparably high monitoring, it may provide low monitoring quality in practice. Reasons include that board members often lack financial expertise or personal independence, serve on many boards simultaneously, and meet infrequently (Quick and Schmidt, 2018). Additionally, due to the large representation of employees on German supervisory boards, they may have interests that are very different from shareholder interests, which makes external monitoring much more important from a shareholder perspective (Elson et al., 2015).

³¹ Together with the executive board, the supervisory board makes up the German two-tier corporate governance system. Simply put, it consists of only non-executive members and is completely independent of the executive board, which is responsible for managing the company. The supervisory board is responsible for monitoring and is encouraged by the German Corporate Governance Codex to form an audit committee from its members. Over 90% of large German listed companies have an audit committee (Quick et al., 2018). For large companies, the employee (often unionists) and shareholder (typically including banks) representatives each get half of the supervisory board seats. Therefore, it is theoretically also more independent of management, which may increase monitoring quality. Refer to Quick and Schmidt (2018) for more details.

Moreover, the dual role of banks as members of the supervisory board and external stakeholders³² may also increase their demand for high audit quality. They are not only interested in monitoring from a shareholder and creditor perspective but are also obliged to exert due care as a member of the supervisory board. This includes the monitoring of the auditor and the assurance of appropriate audit quality.³³ As private investors might care little about audit quality,³⁴ the large proportion of institutional investors in the German setting may increase the demand for high audit quality. Finally, overall investor protection is low in Germany (Gul et al., 2013). Therefore, while there are arguments that the particularities of the German setting bias against finding support for the reputation hypothesis, we expect that, in practice, the demand for high audit quality and the sensitivity of investors to auditor reputation is substantial.

5.2.4 Timeline of Events

To investigate our research question, it is necessary to identify events that encompass unexpected and potentially material information. Moreover, for the short-term event study, we need to find the exact dates on which the pieces of information become public and exclude confounding events in the window of days surrounding the event date of interest (MacKinlay, 1997). First, to identify instances of potential non-audit service failure, we searched LexisNexis and the newspaper database of WISO (containing a large number of national and regional German newspapers) for a combination of the term “scandal” with either “Deloitte”, “EY”, “Ernst & Young”, “KPMG” or “PwC”. We restrict our search to the Big 4 to have sufficiently large client portfolios for our analyses. We exclude the financial crisis and search for articles beginning in 2010. While we might miss instances of NAS failure with our restriction to articles that contain the

³² This comprises their role as creditors and shareholders. Moreover, they often execute proxy votes of private investors (section 135 of the German Stock Corporation Act – Aktiengesetz (AktG)).

³³ E.g., section 117 AktG mandates the supervisory board to monitor financial reports and the audit of the financial statements and to make a statement on both at the annual shareholder meeting. Recently, the Auditing Reformation Act (Abschlussprüfungsreformgesetz (AReG)) has substantially extended the responsibility and the respective liability of the supervisory board concerning the auditor. Section D.III of the German Corporate Governance Codex further emphasizes the responsibility of the supervisory board for the quality of the financial reports and the annual audit.

³⁴ Cascino et al. (2014) report that private investors rely on filtered information, while institutional investors have a greater tendency to rely on audited financial statements as a source of verified information. The survey of Pellens and Schmidt (2014) shows that a little less than one-fifth of private investors use the auditor’s report, while they use other information sources far more extensively.

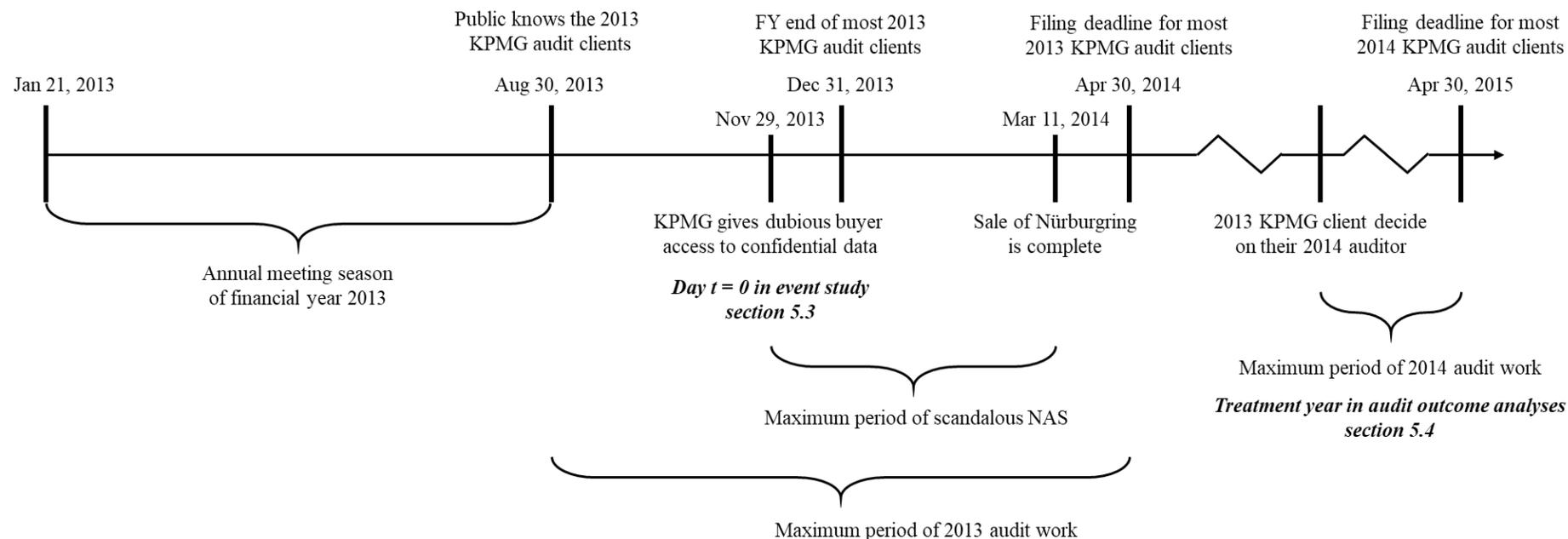
term “scandal” and explicitly mention the audit firm, this approach ensures that the press coverage is unambiguously negative and highlights the role of the auditor. Of all articles identified with this procedure, we only kept those with a clear accusation of failure or misbehavior when the NAS provision was still ongoing.³⁵

With this procedure, we identified two scandals, the *Nürburgring* and *Frankfurt-Hahn airport* sale scandal.³⁶ For each scandal, we identified the first clear event that accused KPMG (the NAS provider in both cases) of wrongdoing that led to NAS failure. We limit our analysis to these initial events because, in both cases, multiple additional accusations and additional information on existing accusations followed in close succession, making a short-term event study analysis unattainable for the later events. We made sure that the instances were relevant to the broad public. Both entailed severe monetary losses for the state Rhineland-Palatinate. We now briefly summarize the two scandals and outline their timeline in the context of the financial statements audit timeline, which is relevant for our analyses. Figure 3 summarizes the respective timelines and depicts the periods used in our analyses.

³⁵ This criterion led to the exclusion of four cases. In the first case, KPMG was investigating accusations of illegal espionage of employee information by German railway provider *Deutsche Bahn* and was accused to have helped with that espionage earlier on. Because KPMG helped to resolve this potential misbehavior at that time, we do not interpret it as negative press. In the second case, known as *LuxLeaks*, two PwC employees blew the whistle on information of tax evasion supported by, among others, the Big 4. However, the press covered PwC for the whistleblowing and not for the support of tax evasion. The third case is the *Cum-Ex* tax evasion scheme in Germany with accusations that KPMG and PwC were involved. However, they only provided legal opinions on existing schemes and never advised illegal behavior. In the fourth case, the *P+S* bankruptcy scandal, KPMG provided a positive restructuring opinion in 2009, which led the state Mecklenburg-Western Pomerania to support *P+S*. Nevertheless, *P+S* went bankrupt in 2012, turned 1750 employees unemployed, and produced multiple KPMG accusations. However, those happened long after the suspected NAS failures.

³⁶ Similarities between the two scandals are a coincidence, e.g. that KPMG was involved, Rhineland-Palatinate was involved, and the service of managing the sale of a company.

Panel A Timeline of events and analyses for Event 1



Panel B Timeline of events and analyses for Event 2

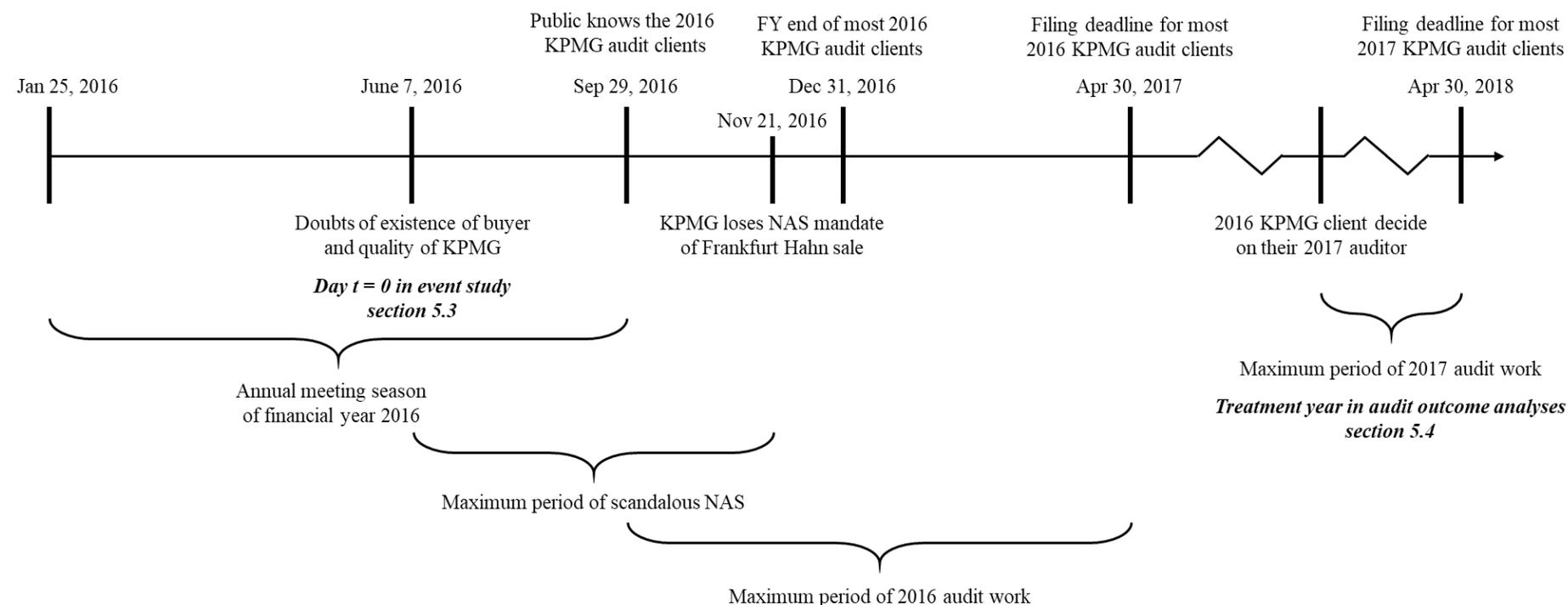


Figure 3: Timeline of events.

Before KPMG became involved in the sale of the racetrack and event location *Nürburgring*, the public had already discussed the project as a scandal of political mismanagement. The state Rhineland-Palatinate as the main owner of the racetrack had expanded it to an event location that had turned out to be a bad business decision, as the operating company had turned bankrupt in 2012, which had cost the state around € 500 million and initiated the sale process. KPMG managed the public tender. On November 29, 2013, the press reported that KPMG gave a dubious buyer access to confidential data without requesting proof of financing, which was officially necessary to gain such access. Although KPMG's advisory quality became questionable at that point, they kept managing the sale process and finally announced buyers *Capricorn* and *Getspeed* on March 11, 2014. When they could not pay the second installment and were in financial trouble themselves, it became evident that doubts of KPMG's advisory quality had been appropriate. However, just recently, the EU court in Luxemburg confirmed the conclusion of the EU commission that the KPMG-led sale was executed legally (Deutsche Presseagentur, 2019; Zerfaß, 2015, 2014).

To study short-term reputation effect of this NAS failure on KPMG's audit practice on November 29, 2013, we need to identify all KPMG audit clients. In Germany, the annual shareholder meeting elects the financial statement auditor for the current financial year. As shown in Panel A of Figure 3, all KPMG clients in our sample (described below) held their annual meeting between January 21 and August 30, 2013. Hence, the public knew they were clients of KPMG when they potentially reacted to the news on November 29, 2013. Our sample for the short-term event study comprises all 2013 KPMG audit clients. For our analysis of longer-term effects, however, we consider two dynamics. First, we estimate whether affected KPMG clients (in this case, the 2013 KPMG clients) choose to change to a different auditor (in this case, the first opportunity is the choice of the 2014 auditor). Second, changes in the audit process that may lead to changes in audit fees and audit quality. The earliest possible start of the audit process is when the client selects the auditor at the annual meeting. However, the auditor can perform some of the work only after the financial year-end, which is December 31 for most KPMG clients. The auditor has to finish no later than the filing deadline, which is April 30 for listed clients (Section 325 (4) HGB). For the scandal to influence the audit process, its effects would have to hit substantially before it is completed. In this case, the remaining 2013 audit work, if any,

was limited when the scandalous NAS period ended. Therefore, the first audit process that is likely to see any influence from the scandal is the 2014 financial statements audit.

Leading to our second scandal, a strategic realignment of airport *Frankfurt-Hahn* included the decision of majority owner Rhineland-Palatinate to sell the majority of its shares to private investors through a public tender managed by KPMG, which began in 2015. Just one day after they presented the Chinese company *SYT* as the chosen buyer, on June 7, 2016, press releases doubted the quality of KPMG's background checks and the soundness of *SYT*. Those doubts soon turned out as justified because, by the end of June, journalists from broadcaster Südwestdeutsche Rundfunk found that *SYT* was a shell company, which led to a suspension of the sale just hours later. On November 21, 2016, Rhineland-Palatinate decided to replace KPMG. The sale process was resumed with the help of Warth & Klein Grant Thornton and a buyer was found in 2017 (Deutsche Presseagentur, 2016; Petermann, 2016; Schmitt, 2016).

As Panel B of Figure 3 shows, on June 7, 2016, a considerable part of the annual meeting season for 2016 auditor selections was still outstanding. However, auditor-client relationships are sticky and investors know of the likely election of the auditor from the invitation to the annual meeting, which contains a suggestion and arrives many weeks before the annual meeting. Therefore, we use all 2016 KPMG audit clients for our event study and exclude those with annual meetings after June 7, 2016, in a robustness check. For our auditor change analyses, accordingly, we consider changes from 2016 KPMG clients to their 2017 auditor. Consistent with our design choice for the first event, we then also use the 2017 period for our analyses of audit outcomes. However, as becomes clear in Figure 3, the proportion of 2016 audit work left after the NAS engagement ended in Panel B was much greater than the proportion of 2013 audit work left in Panel A. Therefore, as a robustness check, we perform our audit outcome analyses for 2016 instead of 2017 audit outcomes.

5.3 Short-Term Event Study

As discussed above, we may extend the reputation hypothesis to a context, in which an audit firm's loss of reputation due to NAS failures affects the audit practice of the firm and its audit clients. We argue that the most direct way to observe these effects is the use of a short-term event study. It can effects that are only short-term or only manifest

in investor perceptions (but not in client or audit committee actions). The low-litigious German setting is necessary to attribute any event study results to the reputation hypothesis. Earlier research found short-term reactions to a NAS scandal in the US where it might be indicative of the insurance hypothesis (Brown et al., 2013; Incardona et al., 2014). Based on these considerations, we hypothesize:

H1: Stock returns of KPMG clients at events of KPMG NAS failures are more negative than those of a comparable market portfolio

5.3.1 Models

Our main analysis is based on the market model of Markowitz (1959) and Sharpe (1963). It assumes that the single explanatory variable for the return $R_{i,t}$ of stock i on day t is the return of the corresponding market portfolio $R_{m,t}$ on day t and presumes a linear relation between the stock return and the market portfolio return. This linear relation can be estimated with an OLS regression of the following model (MacKinlay, 1997):

$$R_{i,t} = \alpha_i + \beta_i \cdot R_{m,t} + e_{i,t} \quad (2)$$

In an empirical analysis, the market portfolio $R_{m,t}$ is not directly observable. Typically, researchers use a corresponding market index as a proxy for the market portfolio. However, in our case, readily available market indices have a significant disadvantage because a KPMG audits a large proportion of all listed firms (Audit Analytics, 2018). Therefore, if there are capital market reactions to KPMG clients, they would also strongly influence market indices and, hence, the regression would remove part of the effect and bias the estimated abnormal return. Therefore, for each calendar year, we construct a value-weighted matched portfolio of non-KPMG clients, which we use for $R_{m,t}$, as follows: For each KPMG client, we choose a non-KPMG client from the same industry which is closest to the KPMG client by total assets (with replacement). We then calculate value-weighted returns by weighting the returns of all non-KPMG clients within the resulting portfolio with their year-end market value. In robustness tests, we replace our matched portfolio with the return of the CDAX index covering all stocks listed in the regulated market of Frankfurt Stock Exchange.

We estimate the regression coefficients α_i and β_i of equation (2) in the estimation period, which starts 255 trading days (about one year) and ends 11 trading days before

the event date. We calculate abnormal returns $AR_{i,t}$ of stock i on day t as the difference between the observed return and the predicted return using estimates $\hat{\alpha}_i$ and $\hat{\beta}_i$ from the estimation of equation (2):

$$AR_{i,t} = e_{i,t} = R_{i,t} - (\hat{\alpha}_i + \hat{\beta}_i \cdot R_{m,t}) \quad (3)$$

In our main tests of abnormal returns, we use the respective event day $t = 0$. For additional analyses, we also calculate cumulative abnormal returns $CAR_{i,\tau,T}$ for firm i in the event window from $t = \tau$ to $t = T$ as the sum of the abnormal returns on each day t in this event window, as follows:

$$CAR_{i,\tau,T} = \sum_{t=\tau}^T AR_{i,t} \quad (4)$$

In our setting, event dates are clustered for all sample firms. This might bias standard t-tests of the hypothesis that abnormal returns are different from zero. Therefore, we test our hypothesis with the statistic proposed by Corrado (1989).³⁷ It is robust to event-day clustering, yet more powerful than alternative robust statistics, for short event windows (Kolari and Pynnönen, 2010). As a robustness test, we instead use a less powerful portfolio approach, the Schipper and Thompson (1983) model that earlier research on the reputation hypothesis in auditing in low-litigation settings has used (Weber et al., 2008). In this approach, we form an equally-weighted portfolio of all stocks i for each day t to arrive at

$$R_t = \frac{1}{N} \sum_{i=1}^N R_{i,t} \quad (5)$$

We then estimate the following OLS model for a period from 126 trading days before the event to 126 after the event

$$R_t = \alpha + \beta_1 \cdot R_{m,t} + \beta_2 \cdot Event_t + e_t \quad (6)$$

³⁷ The Corrado-statistic is a non-parametric approach, which does not use the abnormal return of a stock on a given day in the event window, but uses the rank of the abnormal return of a stock on a given day among all abnormal returns in the estimation and event period of that stock to construct the test statistic.

where $Event_t$ is an indicator variable that takes the value one on the event date and zero otherwise. All other variables are as defined above. If β_2 significantly differs from zero, it indicates an effect of the event on the portfolio of theoretically affected firms controlling for market returns.

To further test the robustness of our results, we use the multi-factor model introduced by Fama and French (1993) including a momentum factor (Carhart, 1997). We take the necessary daily data for this four-factor model from Brückner et al. (2015). The estimation follows the same procedure as the market model with the difference that it includes four explanatory factors instead of one and it corrects both the dependent variable and the market portfolio returns by the return of a risk-free security R_t^f :

$$R_{i,t} - R_t^f = \alpha_i + \beta_i \cdot (R_{m,t} - R_t^f) + s_i \cdot R_t^{SMB} + h_i \cdot R_t^{HML} + w_i \cdot R_t^{MOM} + e_{i,t}, \quad (7)$$

where R_t^{SMB} is the excess return of small over big companies on day t , R_t^{HML} is the excess return of companies with a high market-to-book ratio over those with a low market-to-book ratio on day t , R_t^{MOM} is the excess return of companies with good past performance over those with bad past performance on day t . We describe all other variables above.³⁸

5.3.2 Sample Selection

For each year (2013 and 2016, see discussion of the timeline of events above), we start with all German listed companies (identified by ISIN-identifiers starting with “DE”) with return data available on Thompson Reuters Datastream.³⁹ If a company has multiple listed securities, we only keep regular stock carrying voting rights. Then, we hand-collected the information on their auditor to identify all KPMG clients with available data. We drop penny stocks (less than 50 days with stock price above € 0.10) and require at least 50 days with positive trading volume in the estimation window. Next, we drop observations with no available data on total assets or year-end market value that are necessary to create our matched portfolio. For each event, we also exclude all

³⁸ Brückner et al. (2015) report details of how they calculate the single factors. Note that we use individual self-constructed value-weighted portfolios of non-KPMG clients instead of their index return data for $R_{m,t}$.

³⁹ The return for day t is the sum of the closing price of day t and the dividend paid on day t minus the closing price of day $t-1$, scaled by the closing price of day $t-1$. We require at least 50 returns in the estimation window.

companies with a confounding event.⁴⁰ Table 4 summarizes our sampling procedure and reports the final sample sizes for both events.

Table 4: Sample selection

Panel A: Sample selection for event study			
<i>Event</i>	<i>Number of firms</i>		
	Event 1	Event 2	
KPMG clients with sufficient return data on Datastream	93	98	
<i>Less</i>			
Penny stock	0	0	
Low trading volume	17	22	
Firms with confounding event	11	9	
Firms without data to match peer	7	7	
Full sample	58	60	
Firms without NAS	6	15	
NAS sample	52	45	
Firms with auditor from distant office	39	41	
Close office sample	19	19	
Panel B: Sample selection for event study			
	<i>Firm-years</i>		
	Switching sample	Fee sample	Accruals sample
Active listed German firm-years from 2010 to 2018	6,072	6,072	6,072
<i>Less</i>			
Financial firms or no industry classification	2,481	2,481	2,481
Non-positive equity	486	486	486
No auditor data available	119	119	119
Missing data	70	549	581
Firms with only one firm-year left	28		
Final sample	2,888	2,437	2,405

Notes: This table presents our sample selection for our short-term event study in Panel A and for our analysis of longer-term effects in Panel B.

Events are as follows:

Event 1: November 29, 2013; KPMG gave a dubious potential buyer access to confidential data.

Event 2: June 7, 2016; First press releases, which doubt the existence of SYT and the quality of KPMG's integrity due diligence investigation.

⁴⁰ To identify KPMG clients with confounding events, we screen all news posted on DGAP (Deutsche Gesellschaft für Ad-hoc-Publizität – German Society for Ad-hoc Publicity; <http://www.dgap.de/>; last access Nov 1, 2020), an online news service used by most companies to convey important or mandatory information to the stock market, in the (-5,5)-window. We use this larger window to ensure that we identify all confounding events that may have a significant influence lasting until or starting on the event date. We identify a confounding event if the company issues financial statements; announces large deals; reports key performance indicators; announces mergers and acquisitions; announces large capital measures; or appoints a new executive board member.

In addition to our full sample, we construct two subsamples with arguably closer proximity to the NAS scandal. Assuming that this proximity increases the strength of the reputation signal based on the “being known for something” dimension, we would expect stronger reactions in these subsamples. The first subsample includes all KPMG audit clients that also purchase NAS from KPMG. Their investors may pay closer attention to other KPMG NAS practices.⁴¹ The second subsample comprises all clients from KPMG offices that are geographically close to the company at which the NAS failure event happened. Investors of geographically closer companies may be more aware of local news. We chose a 250 km radius to achieve a sufficiently large subsample. After all selection criteria, this leaves clients from KPMG offices in Düsseldorf, Essen, Frankfurt am Main, Köln, and Mannheim.⁴²

5.3.3 Results

Table 5 reports the proportion of negative returns as well as mean returns using raw returns ($R_{i,t}$), market-adjusted returns (the simple difference $R_{i,t} - R_{m,t}$), and abnormal returns from estimating the market model in equations (2) and (3) for all three samples.

For the full sample in Panel A of Table 5, we find significant negative market reactions on the event day for both events. Specifically, for event 1, the average KPMG client experiences an abnormal return of -0.514%, and 45 of the 58 KPMG clients in our sample have a negative abnormal return. For event 2, the average abnormal return is -0.469% and 42 of the 60 KPMG clients have a negative abnormal return. Market-adjusted returns are even more negative. These results are consistent with hypothesis H1. A loss of about one-half percent of market value in one day is also economically significant. As a comparison, the mean daily raw return of the CDAX index in 2013 (2016) is 0.097% (0.033%) and the mean daily raw return of loss days only is -0.663% (-0.963%).

⁴¹ Note that, for this subsample, there might be another effect on stock prices unrelated to reputation effects on the auditing practice. As the success of NAS may determine investor perceptions of a firm's value, a signal for lower-than-expected NAS quality may be a value-decreasing signal in itself, even when perceptions of audit quality do not change. However, the effects on audit quality are arguably more important. They relate to the overall earnings quality, while NAS mostly relate to small proportions of overall earnings. As an indicator of this relatively lower importance, the average ratio of advisory NAS to audit fees for the NAS subsample is 23.3% for Event 1 and 17.4% for Event 2.

⁴² No clients from KPMG offices in Dortmund and Saarbrücken survived all selection criteria.

Table 5: Market reaction to NAS failure events for German KPMG clients

Panel A: Market reactions for the full sample				
	Percent negative	Mean	t-statistic	Corrado-statistic
Event 1 (November 29, 2013)		<i>N</i> = 58		
<i>Raw return</i>	53%	-0.092%		
<i>Market-adjusted return</i>	83%	-0.738%	-2.610***	-1.981**
<i>Market model</i>	78%	-0.514%	-1.751**	-2.187**
Event 2 (June 7, 2016)		<i>N</i> = 60		
<i>Raw return</i>	20%	0.674%		
<i>Market-adjusted return</i>	82%	-1.123%	-4.837***	-1.901**
<i>Market model</i>	70%	-0.469%	-2.215**	-1.477*
Panel B: Market reactions for the NAS sample				
	Percent negative	Mean	t-statistic	Corrado-statistic
Event 1 (November 29, 2013)		<i>N</i> = 52		
<i>Raw return</i>	58%	-0.416%		
<i>Market-adjusted return</i>	85%	-1.066%	-5.363***	-2.244**
<i>Market model</i>	79%	-0.871%	-4.455***	-2.276**
Event 2 (June 7, 2016)		<i>N</i> = 45		
<i>Raw return</i>	16%	0.902%		
<i>Market-adjusted return</i>	78%	-0.982%	-3.792***	-1.509*
<i>Market model</i>	64%	-0.284%	-1.158	-0.752
Panel C: Market reactions for the Office sample				
	Percent negative	Mean	t-statistic	Corrado-statistic
Event 1 (November 29, 2013)		<i>N</i> = 19		
<i>Raw return</i>	59%	-0.053%		
<i>Market-adjusted return</i>	89%	-1.154%	-1.585*	-1.709**
<i>Market model</i>	79%	-0.444%	-0.575	-1.898**
Event 2 (June 7, 2016)		<i>N</i> = 19		
<i>Raw return</i>	21%	0.738%		
<i>Market-adjusted return</i>	95%	-1.975%	-6.991***	-2.025**
<i>Market model</i>	68%	-0.781%	-3.232***	-1.140

Notes: This table presents sample sizes, the proportion of negative returns, mean returns, and test statistics from t-tests and Corrado-tests that the returns are different from zero for two events and three subsamples. Panel A presents the information for the full sample, Panel B for a sample of audit clients that also purchase NAS, and Panel C for a sample of firms audited by KPMG offices in Düsseldorf, Essen, Frankfurt am Main, Köln, and Mannheim.

Events are as follows:

Event 1: November 29, 2013; KPMG gave a dubious potential buyer access to confidential data.

Event 2: June 7, 2016; First press releases, which doubt the existence of SYT and the quality of KPMG's integrity due diligence investigation.

*, **, and *** denotes significance at the 10%-, 5%-, and 1%-level, respectively, of a one-tailed test that mean returns are negative.

Turning to the NAS subsample in Panel B of Table 5, results become stronger for event 1. With 41 of the remaining 52 KPMG clients experiencing negative abnormal returns, the average abnormal return is -0.871% and test-statistics are slightly stronger than in

the full sample. However, results for event 2 become weaker. The abnormal return from the market model is no longer significantly different from zero with an average of -0.284% and 29 of 45 clients having negative abnormal returns. The results of the office sample in Panel C of Table 5 also show some inconsistency. They become slightly weaker for event 1 with an average abnormal return of -0.444% that is still significant using the robust Corrado-statistic, but not using the one-sample t-test. Note, however, that the remaining sample is small with 19 clients. Of those, 15 experience negative abnormal returns. For event 2, results become substantially stronger, but the robust Corrado-test is no longer significant at conventional levels. The average abnormal return is -0.781% and 13 of the 19 remaining clients experience negative returns. Overall, despite some instances of stronger effects, our subsample results are inconclusive.

5.3.4 Robustness checks

We perform several robustness checks to corroborate our above conclusion that we find support for hypothesis H1. First, as mentioned in the discussion of the timeline of the second event, not all KPMG clients had finally elected KPMG as their 2016 auditor on June 7, 2016. Therefore, we exclude these 18 clients and repeat our event study with the remaining 42 clients. Table 6, Panel A presents the results. The average abnormal return is only -0.108% and no longer significantly different from zero.

Next, for both events, we report mean abnormal or mean cumulative abnormal returns and related t-statistics based on several different design choices in Panel B of Table 6. The first two rows for each event show the $t = 0$ abnormal return when using the Fama and French four-factor model from equation (7) and when using the market model with the CDAX index instead of our matched portfolio, respectively. The three columns show our three samples. Results remain largely robust for the Fama and French four-factor model but become weaker for the CDAX as the market portfolio. However, we expected the latter result because the idea to construct a matched portfolio instead of using a readily available index stems from the argument that using an index biases abnormal returns towards zero because KPMG clients strongly influence the index returns. The third and fourth row show cumulative abnormal returns from the market model with our matched portfolio for longer event windows, as presented in equation (4), from one day before to one day after, and ten days before to ten days after our event date. The

use of longer return windows typically addresses the possibility that information may become available earlier to investors through unobservable channels or may take time to disseminate to all investors. However, we took great care in identifying the exact event dates, the news was disseminated over multiple channels on our event days (making slow dissemination unlikely), and the news subjects just happened shortly before the news broke (making early information leakage unlikely). Recent accounting research using event studies shows that investors act very fast on new information, sometimes within seconds (e.g, Rogers et al., 2017). Consistent with these arguments, we find very similar results in the (-1,1)-window compared to our main results, except for the full sample for event 2. The latter has a small abnormal return that is statistically indistinguishable from zero. In the (-10,10)-window, results are robust for event 1 and the NAS and office sample of event 2, but not for the full sample of event 2.

Table 6: Robustness checks of market reaction to NAS failure events

Panel A: Market reactions at Event 2 restricted to clients with annual meeting before event				
	Percent negative	Mean	t-statistic	Corrado-statistic
Event 2 (June 7, 2016)		$N = 42$		
<i>Raw return</i>	17%	0.898%		
<i>Market-adjusted return</i>	81%	-0.938%	-3.963***	-1.793*
<i>Market model</i>	61%	-0.108%	-0.498	-0.949
Panel B: Alternative (cumulative) abnormal return metrics				
	Full sample	NAS sample	Office sample	
	Mean	Mean	Mean	
	(t-statistic)	(t-statistic)	(t-statistic)	
Event 1 (November 29, 2013)				
<i>Fama French three factor model</i>	-0.549%** (-1.748)	-0.864%*** (-4.405)	-0.445% (-0.582)	
<i>Market model with CDAX as market portfolio</i>	-0.279% (-0.958)	-0.623%*** (-3.139)	-0.174% (-0.229)	
<i>Market model and (-1,1) window</i>	-0.582%* (-1.605)	-0.565%* (-1.558)	0.107% (0.245)	
<i>Market model and (-10,10) window</i>	-2.312%*** (-2.858)	-2.901%*** (-3.597)	-0.845% (-0.549)	
Event 2 (June 7, 2016)				
<i>Fama French three factor model</i>	-0.318%* (-1.488)	-0.106% (-0.427)	-0.768%*** (-3.012)	
<i>Market model with CDAX as market portfolio</i>	-0.335%* -1.586	-0.160% (-0.658)	-0.184% (-0.750)	
<i>Market model and (-1,1) window</i>	-0.230% (-0.600)	-0.626%** (-1.810)	-1.123%*** (-4.009)	
<i>Market model and (-10,10) window</i>	0.460% (0.231)	-2.339%*** (-3,558)	-2.987%*** (-3.328)	

Panel C: Schipper and Thompson regression using full sample				
	Model 1	Model 2	Model 3	Model 4
	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)
<i>Intercept</i>	<0.001* (1.782)	0.001** (2.464)	<0.001* (1.808)	0.001*** (3.208)
<i>Return_{Matched_Portfolio}</i>	0.678*** (18.040)	0.641*** (24.499)	0.683*** (19.070)	0.550 (14.296)
<i>Event1</i>	-0.007* (-1.549)			
<i>Event2</i>		-0.005 (-1.009)		
<i>Pseudo-Event1</i>			0.001 (0.189)	
<i>Pseudo-Event2</i>				-0.001 (-0.157)
N	253	253	253	253
Adjusted R ²	0.563	0.705	0.590	0.446
Panel D: Schipper and Thompson regression using NAS sample				
	Model 1	Model 2	Model 3	Model 4
	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)
<i>Intercept</i>	<0.001* (1.806)	0.001** (2.089)	<0.001 (0.869)	0.001** (2.788)
<i>Return_{Matched_Portfolio}</i>	0.709*** (20.521)	0.718*** (25.544)	0.736*** (21.754)	0.613*** (14.971)
<i>Event1</i>	-0.009** (-2.191)			
<i>Event2</i>		-0.004 (-0.833)		
<i>Pseudo-Event1</i>			0.001 (0.234)	
<i>Pseudo-Event2</i>				-0.001 (-0.266)
N	253	253	253	253
Adjusted R ²	0.625	0.721	0.652	0.469
Panel E: Schipper and Thompson regression using office sample				
	Model 1	Model 2	Model 3	Model 4
	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)
<i>Intercept</i>	<0.001 (1.010)	0.001* (1.829)	0.001 (1.609)	0.001* (1.726)
<i>Return_{Matched_Portfolio}</i>	0.486*** (11.698)	0.490*** (17.742)	0.436*** (10.289)	0.372*** (9.977)
<i>Event1</i>	-0.009* (-1.398)			

<i>Event2</i>		-0.006 (-1.085)		
<i>Pseudo-Event1</i>			<0.001 (0.065)	
<i>Pseudo-Event2</i>				0.001 (0.284)
N	253	253	253	253
Adjusted R ²	0.349	0.555	0.292	0.279

Panel F: Market reactions at pseudo-events for the full sample

	Percent negative	Mean	t-statistic	Corrado-statistic
Pseudo-Event 1 (February 24, 2014)		<i>N</i> = 67		
<i>Raw return</i>	46%	0.276%		
<i>Market-adjusted return</i>	57%	0.142%	0.527	-0.056
<i>Market model</i>	57%	0.141%	0.523	-0.210
Pseudo-Event 2 (November 5, 2015)		<i>N</i> = 64		
<i>Raw return</i>	47%	0.105%	1.362	
<i>Market-adjusted return</i>	39%	0.387%	2.204**	0.703
<i>Market model</i>	47%	0.164%	0.926	0.713

Panel G: Market reactions at pseudo-events for the NAS sample

	Percent negative	Mean	t-statistic	Corrado-statistic
Pseudo-Event 1 (February 24, 2014)		<i>N</i> = 58		
<i>Raw return</i>	50%	0.249%		
<i>Market-adjusted return</i>	59%	0.114%	0.373	-0.387
<i>Market model</i>	57%	0.109%	0.355	-0.598
Pseudo-Event 2 (November 5, 2015)		<i>N</i> = 49		
<i>Raw return</i>	49%	0.045%	1.026	
<i>Market-adjusted return</i>	41%	0.452%	2.057**	0.832
<i>Market model</i>	47%	0.163%	0.734	0.426

Panel H: Market reactions at pseudo-events for the Office sample

	Percent negative	Mean	t-statistic	Corrado-statistic
Pseudo-Event 1 (February 24, 2014)		<i>N</i> = 23		
<i>Raw return</i>	39%	0.123%		
<i>Market-adjusted return</i>	39%	0.130%	0.486	0.901
<i>Market model</i>	48%	0.089%	0.328	1.039
Pseudo-Event 2 (November 5, 2015)		<i>N</i> = 21		
<i>Raw return</i>	38%	0.494%		
<i>Market-adjusted return</i>	38%	0.510%	1.473	0.561
<i>Market model</i>	43%	0.419%	1.222	0.464

Notes: Panels A, F, G, and H of this table presents sample sizes, the proportion of negative returns, mean returns, and test statistics from t-tests and Corrado-tests that the returns are different from zero. Panel A contains a subsample of clients that held their annual meeting before event 2 (Panel A). Panels F, G, and H report two pseudo-events and the three subsamples described in Table 5. Panel B contains mean returns and corresponding test statistics from a t-test (in brackets) for alternative return metrics and all three subsamples described in Table 5. Panels C, D, and E contain coefficients and t-statistics (in brackets) from OLS regressions of equation (6) and event 1, event 2, pseudo-event 1, and pseudo-event 2 in models 1, 2, 3, and 4, respectively.

Events are as follows:

Event 1: November 29, 2013; KPMG gave a dubious potential buyer access to confidential data.

Event 2: June 7, 2016; First press releases, which doubt the existence of *SYT* and the quality of KPMG's integrity due diligence investigation.

We identify Pseudo-Events with Google Trends as an event with a spike in the search term "KPMG Skandal" (German for "KPMG scandal") in a rolling window of six months, starting with the windows September 1, 2013, to March 1, 2014, and October 1, 2016 to April 1, 2017. We required that, once a spike appeared, it had to stay for two more consecutive windows and at least until the next spike appears.

*, **, and *** denotes significance at the 10%-, 5%-, and 1%-level, respectively, of a one-tailed test that the respective variable of interest is negative, and a two-tailed test otherwise.

As a final robustness check, we report the results from using the portfolio OLS regression from the Schipper and Thompson (1983) model in equation (6). Panels C, D, and E of Table 6 present the results for the full sample, NAS sample, and office sample, respectively. Model 1 contains the analysis for event 1 and model 2 the analysis for event 2 (we discuss models 3 and 4 below). In brief, effects for event 1 remain robust across all samples, and effects for event 2 become insignificant at conventional levels across all samples. To sum up, results for event 1 remain largely robust across a variety of robustness tests but results for event 2 seem sensitive to design choices. Together with the observation that results do not become stronger in subsamples, this limits the empirical support of hypothesis H1.

One possible alternative explanation that would be compatible with the ineffectiveness of using different subsamples is that overall negative attention to KPMG drives this effect rather than a focused assessment of KPMG's service quality. Theoretically, this would be consistent with the "general favorability" dimension of reputation. To rule out this alternative explanation, along with the possibility that our sample is biased for any other reason, we replace both of our events with a placebo-event and repeat our main analysis. To do so, we use Google Trends and identify the first date after the end of the annual meeting season in 2013 and 2016, respectively, on which the use of the search term "KPMG Skandal" (German for "KPMG scandal") spikes in Germany. Google Trends provides a direct and unambiguous measure of public attention (Da et al., 2011). We identify two clear spikes on February 24, 2014, and November 5, 2015.⁴³

⁴³ We use a rolling window of six months, starting with the windows September 1, 2013, to March 1, 2014, and October 1, 2016, to April 1, 2017, and moving forward one month at a time (e.g., the next step would be October 1, 2013, to April 1, 2014, and November 1, 2016, to May 1, 2017). We required that, once a spike appears, it stays for two subsequent windows and until the next spike appears. This is because Google Trends reports the relative frequency of a search term. Therefore, a spike that disappears once the next spike appears might just depict the highest search frequency in a period with an overall very low search frequency of the specific search term. With this procedure, we identified spikes on March 19, 2014, which remains when the next spike on April 28, 2014, appears, and on February 22, 2017, which remains when the next spike on March 16, 2017, appears.

We present the results of this falsification test for the full sample, NAS sample, and office sample in Panels F, G, and H of Table 6, respectively. Sample sizes differ slightly from our main analysis due to the requirements for penny stocks and trading volumes and the use of different estimation periods. Besides two significant t-tests for market-adjusted returns in the full sample and NAS sample for pseudo-event 2, there are no significant results, and abnormal returns are overall small. Moreover, the two significant results are positive and are not robust to using the Corrado-statistic. Finally, we also use the Schipper and Thompson (1983) model for both pseudo-events. The results in models 3 and 4 in Panels C, D, and E of Table 6 are all insignificant. In conclusion, the analysis of pseudo-events provides some comfort that the alternative explanation of mere negative attention does not hold. Moreover, it gives some additional indication that, despite being somewhat sensitive regarding event 2, the negative abnormal returns from our events 1 and 2 are not random.

5.4 Longer-Term Effects on KPMG's Audit Practice

As we find indications for short-term negative capital market reactions to KPMG audit clients at KPMG NAS failure events, we now investigate whether KPMG's audit practice faces persistent consequences. Theoretically, the reputation hypothesis discussed above suggests that reputational damage limits the auditor's ability to gain and retain new clients and to charge satisfactorily high audit fees (e.g., Baugh et al., 2018; DeAngelo, 1981; Klein and Leffler, 1981; Weber et al., 2008; Wilson and Grimlund, 1990). If the short-term negative reactions by investors extend to the persistent reputation perceptions of clients and audit committee members, KPMG will lose relatively more clients and gain fewer clients compared to other audit firms. Alternatively, KPMG may discount audit fees to retain clients or attract new clients to counteract potential negative reputation effects on client switching risk (Boone et al., 2015). Moreover, as discussed above, earlier research on NAS failures has not found evidence to support the hypothesis that reputation damage is strong enough to influence audit committee decisions persistently (Baugh et al., 2018). Finally, switching costs are generally high and an idiosyncratic alignment of a client and auditor can have significant value that is economically more important than the reputational damage from the NAS failure (Gerakos and Syverson, 2015). Therefore, effects on client switching risk remain an

empirical question. Nonetheless, as theory predicts an increase (but not a decrease) in switching risk, we hypothesize as follows:

H2a: KPMG clients are more likely to change auditors after events of KPMG NAS failures.

H2b: Clients of other auditors are less likely to change to KPMG after events of KPMG NAS failures.

The alternative possibility that KPMG discounts its services instead directly leads to the expectation that the NAS failure could have negative effects on audit fees. However, audit fees largely reflect audit production costs, so that KPMG may have to accept losses or lower audit quality when discounting fees (Doogar et al., 2015). Instead, KPMG may increase audit effort in an attempt to counteract reputational damage by providing higher audit quality, which would increase audit fees. Again, earlier empirical research could not establish an effect of NAS failures on audit fees (Baugh et al., 2018). Due to this ambiguous discussion, we hypothesize in the null form:

H3: Audit fees of KPMG do not change after events of KPMG NAS failures.

The above alternative that KPMG would do more work to compensate for reputational losses could be an economically optimal reaction to rebuild the prior market position (Klein and Leffler, 1981). This would lead to increases in audit quality. However, facing increasing switching risk and without fee discounts, an alternative strategy to retain clients could be to acquiesce to the client's opinion in audit negotiations. This could potentially lead to aggressive earnings management choices, generally considered as lower audit quality. In conclusion, ambiguous theory leads us to hypothesize in null form:

H4: Earnings management of KPMG clients does not change after events of KPMG NAS failures.

5.4.1 Models

We closely follow prior literature to build our client switching, audit fee, and earnings management models, which we each estimate on a panel of client-years. A description of our models follows. Table 12 in the appendix in section 5.6 contains detailed definitions of all variables. We winsorize all continuous variables at the first and 99th percentile and compute robust standard errors clustered by clients. Finally, we include *Industry* and *Year*

fixed effects in all models to control for time-invariant systematic industry differences and for systematic differences across time, respectively. We define industry based on the 16 sector indices of Frankfurt Stock Exchange and manually backfill missing industries based on SIC codes, if available, or the firm's business description.

To test hypothesis H2a, we specify the following logistic regression model:

$$\begin{aligned} Auditorchg_{i,t} = & \alpha + \beta_1 LagKPMG_{i,t} + \beta_2 LagKPMG_{i,t} \cdot 2014_i + \beta_3 LagKPMG_{i,t} \\ & \cdot 2017_i + \beta_k Controls_{i,t} + Industry + Year + e_t \end{aligned} \quad (8)$$

where $Auditorchg_{i,t}$ is an indicator variable coded as one if client i choses a new auditor for year t , and zero otherwise. $LagKPMG_{i,t}$ is an indicator variable coded as one if KPMG was the auditor of client i in the year before t . It captures whether former KPMG clients are more likely to change auditors (client losses). We interact this variable with indicators for 2014 and 2017, the first years following our NAS failure events 1 and 2, respectively, as we discussed above with the timeline of events.⁴⁴ $Controls$ is a vector of controls that we take from Weber et al. (2008). We use their parsimonious model instead of models with more control variables (such as in Boone et al., 2015) because they have established it on German data with a similar research design to ours and we would lose a significant number of observations with the addition of more controls. As auditor changes are rare, this could decrease the power of our tests.⁴⁵ Specifically, $Controls$ includes firm *Size*, changes in size (*Growth*), financial *Leverage* and changes therein ($\Delta Leverage$), a measure of profitability (*ROA*), and an indicator for firms experiencing a bottom-line *Loss*. Under hypothesis H2a, we expect positive and significant coefficients β_2 and β_3 .

⁴⁴ Because auditor-client pairings and audit fees are sticky, it may take multiple periods until a reputation effect is observable in auditor changes or changes in audit fees (Boone et al., 2015). However, in our case, we have two events that are somewhat close together in time. Specifically, once we use a two-year treatment period instead of the one-year period in our main tests, the treatment period for event 1 is 2014-2015 and for event 2 2017-2018. This considerably reduces our control periods. Given that, in our other robustness tests, we take the possibility into account that the effects of event 2 already manifest in 2016, this procedure would not leave a control period directly before event 2. Overall, this biases against results. In untabulated tests, we repeat our analyses and use 2014-2017 as an overall treatment period for both events. Inferences from our analyses concerning our hypotheses do not change with this alternative design.

⁴⁵ Nevertheless, we repeat our auditor change analysis on a smaller sample of 2,388 observations and add the control variables *DiscAcc*, *InvRec*, *LnAuditFee*, *Big4*, *NASRatio*, *IndShare* (see section 5.6) to equations (8) and (9). The results for our variables of interest are unchanged.

For our test of hypothesis H2b, we specify the following logistic regression model:

$$\begin{aligned} Auditorchg_{i,t} = & \alpha + \beta_1 KPMG_{i,t} + \beta_2 KPMG_{i,t} \cdot 2014_i + \beta_3 KPMG_{i,t} \cdot 2017_i \\ & + \beta_k Controls_{i,t} + Industry + Year + e_t \end{aligned} \quad (9)$$

where the only difference to equation (8) is that $KPMG_{i,t}$ is an indicator variable coded as one if the KPMG is the auditor of client i in year t . It captures whether current KPMG clients are more likely to have changed auditors (client gains) as opposed to client losses captured in equation (8). Under hypothesis H2b, we expect negative and significant coefficients β_2 and β_3 .

To test hypothesis H3, we specify the following OLS regression:

$$\begin{aligned} LnAuditFee_{i,t} = & \alpha + \beta_1 KPMG_{i,t} + \beta_2 KPMG_{i,t} \cdot 2014_i + \beta_3 KPMG_{i,t} \cdot 2017_i \\ & + \beta_k Controls_{i,t} + Industry + Year + e_t \end{aligned} \quad (10)$$

where $LnAuditFee$ is the natural logarithm of one plus audit fees. As in equation (9), $KPMG_{i,t}$ and the respective interactions capture the influence of currently being a KPMG client on current audit fees. From prior literature, we select the following control variables (Hay, 2013): We measure *Size* by the natural logarithm of one plus total assets. As measures of profitability, we include *ROA* and a *Loss* indicator. We use *Growth*, Altman's (1993) Z-score (*ZScore*), and the ratio of inventory and receivables to total assets (*InvRec*) as indicators of inherent risk. To account for liquidity, we include financial *Leverage* and cash flow from operations scaled by total assets (*CFO*). The market-to-book ratio (*MTB*) measures complexity. Finally, as auditor characteristics, we include *Big4*, the share of audit fees an auditor earns in the client industry in year t (*IndShare*), and the absolute value of *Assurance*, *Advisory*, and *Tax NAS* provided to the client. We will reject hypothesis H3 if coefficients β_2 or β_3 significantly differ from zero.

Finally, our tests of hypothesis H4 use the following OLS model:

$$\begin{aligned} Accruals_{i,t} = & \alpha + \beta_1 KPMG_{i,t} + \beta_2 KPMG_{i,t} \cdot 2014_i + \beta_3 KPMG_{i,t} \cdot 2017_i \\ & + \beta_k Controls_{i,t} + Industry + Year + e_t \end{aligned} \quad (11)$$

where *Accruals* represents either abnormal working capital actuals (*AWCA*), discretionary accruals (*DiscAcc*), or the absolute value of either of the two ($|AWCA|$ and $|DiscAcc|$).

We estimate *DiscAcc* with the performance-adjusted modified Jones Model (Kothari et al., 2005).⁴⁶ Our variables of interest are the same as in equation (10) to measure the influence of being a current KPMG client on earnings management. We add the following control variables, largely following prior literature (e.g., Francis and Wang, 2008; Mohrmann, 2017): We include *Size* along with *Growth* and increases in sales (*SalesGrowth*). We use *ROA* and a *Loss* indicator as profitability measures, and follow the literature to lag both measures by one year to capture prior-year profitability. We include financial *Leverage* and *MTB* to control for potential lender and shareholder pressures and for riskiness using *ZScore*. Finally, we control for auditor size (*Big4*), potential independence concerns due to NAS provision (*NASRatio*), and industry expertise (*IndShare*).

In addition to the two-stage accruals models in equation (11), we use the following one-stage OLS regression that models total accruals (*TotalAcc*) as a dependent variable and add the same variables of interest and controls as in the two-stage models (Chen et al., 2018).

$$\begin{aligned}
 TotalAcc_{i,t} = & \alpha + \beta_1 KPMG_{i,t} + \beta_2 KPMG_{i,t} \cdot 2014_i + \beta_3 KPMG_{i,t} \cdot 2017_i \\
 & + \beta_k Controls_{i,t} + \sum_{p=1}^P \gamma_{1,p} (1/Assets_{i,p,t-1}) \\
 & + \sum_{p=1}^P \gamma_{2,p} (\Delta SalesRec_{i,p,t}) + \sum_{j=1}^P \gamma_{3,p} PPE_{i,p,t} + Industry \\
 & + Year + e_t
 \end{aligned} \tag{12}$$

where all variables are as defined above and in section 5.6, and p indicates industry-years. For $p = p_0$, the indexed variables assume their actual value for observations in industry-year p_0 , and zero otherwise. We will reject hypothesis H4, if coefficients β_2 or β_3 in equations (11) or (12) significantly differ from zero.

5.4.2 Sample Selection

We start our sample selection for all three types of analyses with 6,072 firm-year observations from all active listed German firms from 2012 to 2018, where active means that the firm has a stock price on every trading day of the year. We start in 2012 to have one full financial year

⁴⁶ Specifically, we estimate the following cross-sectional OLS model by industry-year and use the residuals as our variable *DiscAcc*. For variable descriptions, see section 5.6.

$$TotalAcc_{i,t} = \beta_0 + \beta_1 (1/Assets_{i,t-1}) + \beta_2 (\Delta SalesRec_{i,t}) + \beta_3 PPE_{i,t} + \beta_4 ROA_{i,t-1} + \varepsilon_{i,t}$$

without any possible influence of our events before our first event date late in 2013. For all samples, we exclude all financial firms and firms for which we have no industry classification (2,481 observations). We drop all observations with non-positive equity (486) and no identifiable auditor (119). Finally, we exclude 70 observations with missing data and 28 firms with only one firm-year in our dataset (no switching observable) from the switching sample, 549 observations with missing data from the fee sample, and 581 observations with missing data from the accruals sample. This leaves us with 2,888 observations in the switching sample, 2,437 observations in the fee sample, and 2,405 observations in the accruals sample. Panel B of Table 4 summarizes the sample selection.

5.4.3 Results

5.4.3.1 Descriptive Statistics

Table 7 presents summary statistics for the three groups of analysis in Panels A, C, and E and correlation tables in Panels B, D, and F. KPMG audits between 15% and 16% of all firm-years, a fourth of the 60% of firm-years audited by any Big 4. Those are evenly distributed across the seven sample years so that around 2.2% of our observations assume the value of one for our interaction terms. The control variables behave similarly across all three samples. The average firm is large ($> \text{€ } 200$ million in total assets) and about 20% of the observations are loss-years.

Bivariate correlations between our variables of interest and *Auditorchg* are insignificant. Correlations with *LnAuditFee* are significantly positive but relatively small. Moreover, the variables of interest also correlate positively with *Size*, which, in turn, correlates positively and very strongly with *LnAuditFee*. Finally, the only significant correlation between our variables of interest and accruals measures is a small positive correlation between *KPMG · 2014* and $|AWCA|$. Among the independent variables, there are some very large correlations, e.g. correlations above 0.9 between *LagKPMG* and *KPMG* and the respective interactions in Panel B and negative correlations below -0.6 between *ROA* and *Loss* in all Panels. However, those are not surprising as auditor changes are rare, and *Loss*, by construction, always has the opposite sign of *ROA*. An inspection of variance inflation factors (VIFs) shows that multicollinearity is not a concern as the highest VIF (in all cases, *ROA*) is 1.98 and 1.97 when estimating equations (8) and (9), 3.09 for estimating equation (9), and 2.66 and 3.93 for estimating equations (11) and (12), respectively.

Table 7: Descriptive statistics for additional analyses

Panel A: Summary statistics of the sample used in the auditor change analysis												
	N	Mean	Standard deviation	Min	25th percentile	median	75th percentile	Max				
<i>Auditorchg</i>	2,888	0.088	0.283	0.000	0.000	0.000	0.000	1.000				
<i>LagKPMG</i>	2,888	0.154	0.361	0.000	0.000	0.000	0.000	1.000				
<i>LagKPMG · 2014</i>	2,888	0.021	0.143	0.000	0.000	0.000	0.000	1.000				
<i>LagKPMG · 2017</i>	2,888	0.022	0.146	0.000	0.000	0.000	0.000	1.000				
<i>KPMG</i>	2,888	0.149	0.356	0.000	0.000	0.000	0.000	1.000				
<i>KPMG · 2014</i>	2,888	0.021	0.143	0.000	0.000	0.000	0.000	1.000				
<i>KPMG · 2017</i>	2,888	0.021	0.144	0.000	0.000	0.000	0.000	1.000				
<i>Size</i>	2,888	12.294	2.444	6.718	10.571	12.038	13.875	18.653				
<i>Growth</i>	2,888	0.175	0.492	-0.626	-0.025	0.090	0.239	3.615				
<i>Leverage</i>	2,888	0.517	0.213	0.026	0.376	0.529	0.687	0.936				
<i>ΔLeverage</i>	2,888	0.036	0.165	-0.441	-0.026	0.019	0.074	0.924				
<i>ROA</i>	2,888	0.022	0.125	-0.665	0.005	0.036	0.071	0.304				
<i>Loss</i>	2,888	0.211	0.408	0.000	0.000	0.000	0.000	1.000				
Panel B: Correlations of the sample used in the auditor change analysis												
	(1)	(2)	(3)	(4)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1 <i>Auditorchg</i>	-											
2 <i>LagKPMG</i>	-0.030	-										
3 <i>LagKPMG · 2014</i>	-0.028	0.342	-									
4 <i>LagKPMG · 2017</i>	-0.004	0.350	-0.022	-								
5 <i>KPMG</i>	-0.075	0.934	0.334	0.323	-							
6 <i>KPMG · 2014</i>	-0.028	0.328	0.966	-0.022	0.348	-						
7 <i>KPMG · 2017</i>	-0.020	0.325	-0.021	0.934	0.351	-0.021	-					
8 <i>Size</i>	-0.114	0.269	0.087	0.104	0.276	0.087	0.108	-				
9 <i>Growth</i>	0.051	-0.029	-0.020	-0.028	-0.035	-0.014	-0.024	-0.016	-			
10 <i>Leverage</i>	-0.019	0.104	0.041	0.041	0.118	0.048	0.043	0.307	-0.081	-		
11 <i>ΔLeverage</i>	0.001	-0.022	-0.009	-0.018	-0.018	-0.011	-0.015	-0.027	0.548	-0.005	-	
12 <i>ROA</i>	-0.033	0.039	-0.001	0.030	0.042	0.003	0.032	0.214	0.011	-0.075	-0.170	-
13 <i>Loss</i>	0.035	-0.058	0.002	-0.036	-0.066	-0.010	-0.046	-0.244	-0.044	0.040	0.093	-0.651

Panel C: Summary statistics of the sample used in the audit fee and NAS fee analysis								
	N	Mean	Standard deviation	Min	25th percentile	median	75th percentile	Max
<i>LnAuditFee</i>	2,437	5.511	1.393	2.944	4.543	5.288	6.275	9.680
<i>KPMG</i>	2,437	0.160	0.366	0.000	0.000	0.000	0.000	1.000
<i>KPMG · 2014</i>	2,437	0.023	0.151	0.000	0.000	0.000	0.000	1.000
<i>KPMG · 2017</i>	2,437	0.022	0.147	0.000	0.000	0.000	0.000	1.000
<i>Size</i>	2,437	12.646	2.318	6.718	11.041	12.316	14.167	18.653
<i>ROA</i>	2,437	0.027	0.117	-0.665	0.008	0.040	0.072	0.304
<i>Growth</i>	2,437	0.176	0.466	-0.626	-0.014	0.095	0.236	3.615
<i>Leverage</i>	2,437	0.529	0.198	0.026	0.400	0.535	0.687	0.936
<i>Loss</i>	2,437	0.192	0.394	0.000	0.000	0.000	0.000	1.000
<i>CFO</i>	2,437	0.075	0.137	-0.588	0.030	0.081	0.129	0.472
<i>MTB</i>	2,437	2.687	2.539	0.299	1.159	1.877	3.139	14.429
<i>ZScore</i>	2,437	4.419	4.151	-3.965	2.386	3.449	4.941	27.607
<i>InvRec</i>	2,437	0.355	0.184	0.005	0.217	0.349	0.481	0.860
<i>Big4</i>	2,437	0.607	0.489	0.000	0.000	1.000	1.000	1.000
<i>Assurance</i>	2,437	196.245	788.781	0.000	0.000	0.000	42.000	6000.000
<i>Advisory</i>	2,437	146.511	511.852	0.000	0.000	6.000	51.000	4000.000
<i>Tax</i>	2,437	68.340	181.693	0.000	0.000	0.000	42.000	1119.560
<i>IndShare</i>	2,437	0.189	0.243	0.000	0.014	0.081	0.255	0.986

Panel D: Correlations of the sample used in the audit fee and NAS fee analysis																	
	(1)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
1 <i>LnAuditFee</i>	-																
3 <i>KPMG</i>	0.267	-															
4 <i>KPMG · 2014</i>	0.091	0.355	-														
5 <i>KPMG · 2017</i>	0.103	0.345	-0.023	-													
6 <i>Size</i>	0.900	0.248	0.080	0.099	-												
7 <i>ROA</i>	0.124	0.051	0.001	0.043	0.204	-											
8 <i>Growth</i>	-0.058	-0.037	-0.013	-0.018	-0.021	0.018	-										
9 <i>Leverage</i>	0.319	0.116	0.047	0.048	0.270	-0.095	-0.079	-									
10 <i>Loss</i>	-0.153	-0.076	-0.013	-0.052	-0.238	-0.648	-0.053	0.045	-								
11 <i>CFO</i>	0.105	0.054	0.016	0.013	0.137	0.706	-0.010	-0.011	-0.445	-							
12 <i>MTB</i>	-0.131	-0.017	-0.030	0.040	-0.197	0.000	0.068	0.111	0.014	0.079	-						

13	ZScore	-0.185	-0.079	-0.037	-0.011	-0.143	0.232	0.073	-0.532	-0.104	0.131	0.364	-					
14	InvRec	-0.064	0.013	-0.007	0.024	-0.076	0.095	-0.115	0.136	-0.084	-0.029	-0.010	0.021	-				
15	Big4	0.508	0.351	0.125	0.121	0.508	0.098	-0.059	0.090	-0.100	0.103	-0.019	-0.003	-0.083	-			
16	Assurance	0.551	0.089	0.039	0.018	0.494	0.000	-0.042	0.221	-0.016	-0.022	-0.070	-0.116	-0.054	0.187	-		
17	Advisory	0.516	0.166	0.081	0.047	0.471	0.021	-0.001	0.191	-0.055	0.006	-0.058	-0.094	-0.071	0.206	0.621	-	
18	Tax	0.491	0.199	0.103	0.051	0.439	0.044	-0.016	0.142	-0.071	0.028	-0.025	-0.051	-0.067	0.233	0.414	0.425	-
19	IndShare	0.531	0.334	0.126	0.093	0.491	0.094	-0.077	0.108	-0.128	0.114	-0.046	-0.038	-0.101	0.561	0.340	0.313	0.247

Panel E: Summary statistics of the sample used in the accruals analysis

	N	Mean	Standard deviation	Min	25th percentile	median	75th percentile	Max
TotalAcc	2,405	-0.042	0.094	-0.380	-0.078	-0.042	-0.006	0.314
1/Assets	2,405	0.000022	0.000059	<0.000001	0.000001	0.000005	0.000017	0.000471
ΔSalesRec	2,405	0.057	0.216	-0.685	-0.020	0.037	0.111	1.196
PPE	2,405	0.253	0.222	0.002	0.077	0.207	0.360	1.098
AWCA	2,405	0.010	0.126	-0.422	-0.033	0.004	0.046	0.539
AWCA	2,405	0.076	0.101	0.000	0.016	0.040	0.090	0.539
DiscAcc	2,405	-0.009	0.079	-0.387	-0.045	-0.010	0.027	0.400
DiscAcc	2,405	0.054	0.058	0.000	0.016	0.036	0.071	0.400
KPMG	2,405	0.161	0.368	0.000	0.000	0.000	0.000	1.000
KPMG · 2014	2,405	0.023	0.151	0.000	0.000	0.000	0.000	1.000
KPMG · 2017	2,405	0.022	0.148	0.000	0.000	0.000	0.000	1.000
Size	2,405	12.675	2.294	6.718	11.061	12.334	14.181	18.653
ROA	2,405	0.027	0.115	-0.584	0.009	0.040	0.070	0.277
SalesGrowth	2,405	0.065	0.250	-0.567	-0.017	0.044	0.114	1.671
Growth	2,405	0.177	0.461	-0.626	-0.012	0.095	0.237	3.615
Leverage	2,405	0.532	0.195	0.026	0.404	0.536	0.687	0.936
Loss	2,405	0.189	0.392	0.000	0.000	0.000	0.000	1.000
CFO	2,405	0.076	0.135	-0.588	0.031	0.082	0.129	0.472
MTB	2,405	2.693	2.540	0.299	1.162	1.879	3.147	14.429
ZScore	2,405	4.345	3.921	-3.965	2.387	3.445	4.904	27.607
Big4	2,405	0.609	0.488	0.000	0.000	1.000	1.000	1.000
NASRatio	2,405	0.388	0.483	0.000	0.046	0.232	0.545	2.555
IndShare	2,405	0.189	0.243	0.000	0.014	0.089	0.255	0.986

Panel F: Correlations of the sample used in the accruals analysis																			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
1 <i>TotalAcc</i>	-																		
2 <i>AWCA</i>	0.326	-																	
3 $ AWCA $	0.081	0.217	-																
4 <i>DiscAcc</i>	0.804	0.272	0.112	-															
5 $ DiscAcc $	-0.048	0.046	0.281	-0.084	-														
6 <i>KPMG</i>	-0.015	0.000	-0.045	-0.035	-0.028	-													
7 <i>KPMG · 2014</i>	-0.028	0.026	0.040	-0.031	-0.011	0.352	-												
8 <i>KPMG · 2017</i>	0.036	0.006	-0.014	0.029	-0.019	0.346	-0.023	-											
9 <i>Size</i>	0.040	0.013	-0.224	0.004	-0.211	0.252	0.084	0.098	-										
10 <i>ROA</i>	0.042	0.057	-0.122	0.026	-0.093	0.045	0.007	0.031	0.197	-									
11 <i>SalesGrowth</i>	0.014	-0.092	0.234	0.073	0.034	-0.014	-0.014	0.018	-0.048	-0.123	-								
12 <i>Growth</i>	0.045	0.085	0.348	0.106	0.131	-0.038	-0.013	-0.019	-0.023	0.019	0.447	-							
13 <i>Leverage</i>	-0.097	-0.032	-0.076	-0.061	-0.015	0.119	0.053	0.047	0.274	-0.134	-0.070	-0.080	-						
14 <i>Loss</i>	-0.018	0.003	0.150	-0.007	0.139	-0.053	0.010	-0.030	-0.219	-0.656	0.025	-0.044	0.094	-					
15 <i>CFO</i>	-0.433	-0.125	-0.116	-0.358	-0.069	0.052	0.017	0.012	0.120	0.604	0.002	0.005	-0.027	-0.396	-				
16 <i>MTB</i>	-0.063	-0.033	0.134	-0.070	0.096	-0.012	-0.029	0.040	-0.191	-0.045	0.120	0.075	0.102	0.030	0.079	-			
17 <i>ZScore</i>	0.085	-0.005	0.125	0.046	0.011	-0.078	-0.043	-0.009	-0.150	0.208	0.138	0.092	-0.514	-0.104	0.153	0.406	-		
18 <i>Big4</i>	-0.041	-0.002	-0.114	-0.046	-0.116	0.351	0.124	0.121	0.506	0.093	-0.023	-0.060	0.092	-0.083	0.094	-0.014	-0.011	-	
19 <i>NASRatio</i>	-0.041	0.018	-0.006	-0.037	0.021	0.092	0.064	0.006	0.210	-0.060	0.033	0.105	0.060	0.010	-0.060	-0.029	-0.046	0.155	-
20 <i>IndShare</i>	-0.054	0.003	-0.110	-0.022	-0.132	0.335	0.124	0.094	0.501	0.088	-0.050	-0.077	0.121	-0.105	0.106	-0.043	-0.061	0.562	0.112

Notes: Panels A, C, and E present summary statistics of our three different samples for the analysis of longer-term effects described in Panel B of Table 4. Panels B, D, and F present Pearson correlations between the variables in each sample. Correlation coefficients in bold are significant at the 5%-level. In Panel F, we do not include $1/Assets$, $\Delta SalesRec$, and PPE in the correlation analysis, because we only use them to calculate *DiscAcc* or as a factor in the untabulated interactions of the one-step total accruals model in equation (12). Variable definitions are in section 5.6.

5.4.3.2 Multiple Regression Analyses

Table 8 presents results from estimating equation (8) in Model 1 and from estimating equation (9) in Model 2. Results for Model 1 do not indicate any difference in the likelihood to change auditors for former KPMG clients compared to clients of other auditors, neither in the treatment nor in the control periods. In Model 2, the coefficient of *KPMG* is significantly negative, suggesting that current clients of KPMG are less likely in their first year with a new auditor. Put differently, KPMG is less likely to gain new clients in our sample period. However, our variables of interest, the interactions with our treatment years, are insignificant. Therefore, this overall KPMG-effect appears to be unrelated to the NAS failures we investigate. In conclusion, our results do not support hypotheses H2a and H2b.

Table 8: Logistic regression of auditor change

	Model 1		Model 2	
	Client losses		Client gains	
	Coefficient	z-statistic	Coefficient	z-statistic
<i>Intercept</i>	0.056	0.100	-0.138	-0.250
<i>LagKPMG</i>	0.097	0.410		
<i>LagKPMG · 2014</i>	-0.731	-0.940		
<i>LagKPMG · 2017</i>	-0.064	-0.120		
<i>KPMG</i>			-0.744 **	-2.350
<i>KPMG · 2014</i>			0.029	0.040
<i>KPMG · 2017</i>			0.184	0.260
<i>Size</i>	-0.212 ***	-6.010	-0.190 ***	-5.390
<i>Growth</i>	0.387 ***	3.470	0.374 ***	3.380
<i>Leverage</i>	0.264	0.720	0.297	0.820
Δ <i>Leverage</i>	-0.514	-1.380	-0.515	-1.390
<i>ROA</i>	-0.022	-0.040	-0.079	-0.130
<i>Loss</i>	-0.003	-0.010	-0.014	-0.060
Industry	Yes		Yes	
Year	Yes		Yes	
N	2,888		2,888	
Nagelkerke R ²	0.053		0.058	
Likelihood Ratio χ^2	69.01***		76.00***	

Notes: Model 1 presents logistic regression results from estimating equation (8) and Model 2 presents logistic regression results from estimation equation (9). The dependent variable is *Auditorchg*. We winsorize all continuous variables at the first and 99th percentile and compute robust standard errors clustered by client. We define industry based on the 16 sector indices of Frankfurt Stock Exchange and manually backfill missing industries based on SIC codes, if available, or a consideration of the firm's business description. *, **, and *** denotes significance at the 10%-, 5%-, and 1%-level, respectively, of a one-tailed test that the coefficients of our variables of interest are negative (Model 1) and positive (Model 2), respectively, and a two-tailed test otherwise. Variable definitions are in section 5.6.

Turning to our audit fee model, Table 9 presents the results from estimating equation (10). The main effect of KPMG as well as the interactions with our treatment years are all insignificant. Based on these results, we cannot reject hypothesis H3. There are no indications that KPMG clients, neither on average nor in reaction to the NAS failure events we study, pay more or fewer audit fees than the average client in our sample.

Table 9: OLS-regression of the natural logarithm of Audit Fees

	Model 1	
	DV = \ln_Audit_Fee	
	Coefficient	t-statistic
<i>Intercept</i>	-1.128 ***	-4.724
<i>KPMG</i>	-0.015	-0.229
<i>KPMG · 2014</i>	0.013	0.210
<i>KPMG · 2017</i>	0.036	0.728
<i>Size</i>	0.462 ***	27.949
<i>ROA</i>	-0.349	-1.380
<i>Growth</i>	-0.083 **	-2.443
<i>Leverage</i>	0.319 **	2.232
<i>Loss</i>	0.098 *	1.940
<i>CFO</i>	0.220	0.917
<i>MTB</i>	0.013	1.456
<i>ZScore</i>	-0.015 **	-1.994
<i>InvRec</i>	0.121	0.835
<i>Big4</i>	0.108 *	1.859
<i>Assurance</i>	<0.001 ***	4.760
<i>Advisory</i>	<0.001 ***	2.853
<i>Tax</i>	0.001 ***	3.959
<i>LnAuditFee</i>		
<i>IndShare</i>	0.531 ***	4.070
Industry	Yes	
Year	Yes	
N	2,437	
Adjusted R ²	0.861	
F-statistic	408.30***	

Notes: Model 1 presents OLS regression results from estimating equation (10). We winsorize all continuous variables at the first and 99th percentile and compute robust standard errors clustered by client. We define industry based on the 16 sector indices of Frankfurt Stock Exchange and manually backfill missing industries based on SIC codes, if available, or a consideration of the firm's business description. *, **, and *** denotes significance at the 10%-, 5%-, and 1%-level, respectively, of two-tailed tests. Variable definitions are in section 5.6.

Finally, Table 10 shows the results of our accruals models. Panel A presents results when estimating equation (11) with signed (Model 1) and absolute (Model 2) discretionary accruals, Panel B reports results from estimating equation (12), and Panel C depicts results from estimating equation (11) with signed (Model 1) and absolute (Model 2) working capital accruals. The main effect of *KPMG* is not significant in any of the models, suggesting that there are no differences in earnings management between KPMG clients and an average client. Turning to our variables of interest, Panel A shows that *KPMG · 2017* is positive and significant for signed discretionary accruals, but insignificant for absolute discretionary accruals, while *KPMG · 2014* is insignificant for both measures. This suggests that, for event 2 only, KPMG clients have a tendency to disclose accruals that are overall more income-increasing, but the absolute degree of earnings management does not differ from the average client. Hence, there is some indication that we can reject H4. The evidence is consistent with KPMG being less restrictive in curbing income-increasing accruals accounting by their clients after the second NAS failure. Economically, the coefficient of 0.016 corresponds to an increase of 0.203 standard deviations for KPMG 2017 clients compared to other KPMG clients.

Results in Panel B of Table 10 have the same pattern. While *KPMG · 2017* is positive and significant, *KPMG · 2014* is insignificant. The economic effect is slightly lower, with the coefficient of 0.014 corresponding to 0.149 standard deviations higher total accruals for KPMG 2017 clients compared to other KPMG clients. Finally, results in Panel C of Table 10, again, show the same pattern. *KPMG · 2017* is positive and significant for signed abnormal working capital accruals, but insignificant for absolute abnormal working capital accruals, and *KPMG · 2014* is insignificant for both measures. The economic effect is slightly stronger than the other two, as the coefficient of 0.033 suggests that abnormal working capital accruals of KPMG 2017 clients are 0.262 standard deviations higher than those of other KPMG clients. Overall, there is some evidence to reject hypothesis H4 and earnings management seems to become more aggressive after event 2, which we could interpret as an indication of lower audit quality. As we argue in our hypothesis development, this could be consistent with a strategy of KPMG to please clients in an attempt to mitigate negative consequences from the NAS failures.

Table 10: OLS-regression of accruals-based earnings management

Panel A: Discretionary Accruals from Performance Adjusted Jones Model					
	Model 1		Model 2		
	DV = <i>Disc_Acc</i>		DV = $ Disc_Acc $		
	Coefficient	t-statistic	Coefficient	t-statistic	
<i>Intercept</i>	0.004	0.252	0.092 ***	6.974	
<i>KPMG</i>	-0.002	-0.468	0.003	0.745	
<i>KPMG · 2014</i>	-0.008	-0.737	-0.004	-0.550	
<i>KPMG · 2017</i>	0.016 **	2.427	0.003	0.415	
<i>Size</i>	<0.001	0.407	-0.003 ***	-3.630	
<i>ROA</i>	0.263 ***	5.516	0.027	1.070	
<i>SalesGrowth</i>	0.028 **	2.532	-0.008	-1.141	
<i>Growth</i>	0.013 *	1.654	0.016 ***	3.173	
<i>Leverage</i>	-0.001	-0.052	0.001	0.124	
<i>Loss</i>	0.005	0.745	0.015 ***	2.963	
<i>CFO</i>	-0.356 ***	-9.097	-0.024	-1.027	
<i>MTB</i>	-0.001	-1.055	0.001	1.545	
<i>ZScore</i>	0.001 *	1.762	<0.001	-0.559	
<i>Big4</i>	-0.006	-1.097	-0.002	-0.470	
<i>NASRatio</i>	-0.010 ***	-2.855	0.004	1.440	
<i>IndShare</i>	0.009	1.114	-0.002	-0.225	
Industry	Yes		Yes		
Year	Yes		Yes		
N	2,405		2,405		
Adjusted R ²	0.247		0.121		
F-statistic	23.49***		10.45***		
Panel B: One-step Model with Total Accruals as Dependent Variable					
	Model 1				
	DV = <i>Total_Acc</i>				
	Coefficient	t-statistic			
<i>Intercept</i>	-0.026	-0.902			
<i>KPMG</i>	0.000	0.108			
<i>KPMG · 2014</i>	-0.006	-0.559			
<i>KPMG · 2017</i>	0.014 *	1.878			
<i>Size</i>	0.002	1.416			
<i>ROA</i>	0.368 ***	8.081			
<i>SalesGrowth</i>	0.013	0.773			
<i>Growth</i>	0.015	1.412			
<i>Leverage</i>	-0.017	-1.255			
<i>Loss</i>	0.008	1.234			
<i>CFO</i>	-0.561 ***	-14.422			
<i>MTB</i>	<0.001	0.198			
<i>ZScore</i>	0.002 ***	2.596			
<i>Big4</i>	-0.009 *	-1.921			
<i>NASRatio</i>	-0.009 ***	-2.811			

<i>IndShare</i>	0.012	1.398
Industry	Yes	
Year	Yes	
Interactions	Yes	
N	2,405	
Adjusted R ²	0.490	
F-statistic	7.61***	

Panel C: Abnormal working capital accruals

	Model 1		Model 2	
	DV = <i>AWCA</i>		DV = $ AWCA $	
	Coefficient	t-statistic	Coefficient	t-statistic
<i>Intercept</i>	0.026	1.304	0.125 ***	5.233
<i>KPMG</i>	-0.002	-0.327	0.001	0.169
<i>KPMG · 2014</i>	0.004	0.166	0.024	1.576
<i>KPMG · 2017</i>	0.033 *	1.693	0.012	0.917
<i>Size</i>	<0.001	0.381	-0.008 ***	-4.707
<i>ROA</i>	0.249 ***	3.053	0.034	0.683
<i>SalesGrowth</i>	-0.063 *	-1.827	0.033 *	1.895
<i>Growth</i>	0.039 **	2.009	0.068 ***	8.563
<i>Leverage</i>	-0.026	-1.338	0.017	0.920
<i>Loss</i>	0.021 **	2.113	0.026 ***	3.013
<i>CFO</i>	-0.227 ***	-3.869	-0.074 **	-2.119
<i>MTB</i>	0.001	0.730	0.002	1.348
<i>ZScore</i>	-0.001	-0.498	0.002	1.547
<i>Big4</i>	<0.001	0.013	-0.001	-0.222
<i>NASRatio</i>	<0.001	0.030	-0.002	-0.549
<i>IndShare</i>	0.008	0.894	0.010	0.874
Industry	Yes		Yes	
Year	Yes		Yes	
N	2,405		2,405	
Adjusted R ²	0.069		0.209	
F-statistic	6.06***		19.15***	

Notes: In Panel A, Model 1 presents OLS regression results from estimating equation (11) with *DiscAcc* as the dependent variable and Model 2 presents OLS regression results from estimating equation (11) with $|Disc_Acc|$ as the dependent variable. Panel B present OLS regression results from estimating equation (12). In Panel C, Model 1 presents OLS regression results from estimating equation (11) with *AWCA* as the dependent variable and Model 2 presents OLS regression results from estimating equation (11) with $|AWCA|$ as the dependent variable. We winsorize all continuous variables at the first and 99th percentile and compute robust standard errors clustered by client. We define industry based on the 16 sector indices of Frankfurt Stock Exchange and manually backfill missing industries based on SIC codes, if available, or a consideration of the firm's business description. *, **, and *** denotes significance at the 10%-, 5%-, and 1%-level, respectively, of two-tailed tests. Variable definitions are in section 5.6.

5.4.4 Additional Analyses

We conduct a robustness test and two additional analyses for our significant results in our accruals models to exclude that they are subject to omitted variables and to explore the boundaries of the effects. Table 11 summarizes the results.

Table 11: Robustness checks

Panel A: Accruals-based earnings management regressions with firm fixed effects

	Model 1		Model 2		Model 3		Model 4		Model 5	
	DV = <i>Disc_Acc</i>		DV = $ Disc_Acc $		DV = <i>Total_Acc</i>		DV = <i>AWCA</i>		DV = $ AWCA $	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
<i>Intercept</i>	-0.002	-0.014	-0.011	-0.162	-0.055	-0.344	-0.305 *	-1.789	0.266 **	2.404
<i>KPMG</i>	0.001	0.126	0.007	0.803	0.002	0.201	0.008	0.472	0.013	0.949
<i>KPMG · 2014</i>	-0.007	-0.670	-0.004	-0.435	-0.006	-0.564	0.011	0.448	0.030 *	1.923
<i>KPMG · 2017</i>	0.017 **	2.225	0.002	0.330	0.017 **	2.110	0.035	1.618	0.012	0.853
Controls	Yes		Yes		Yes		Yes		Yes	
Firm	Yes		Yes		Yes		Yes		Yes	
Year	Yes		Yes		Yes		Yes		Yes	
Interactions	-		-		Yes		-		-	
N	2,405		2,405		2,405		2,405		2,405	
Adjusted R ²	0.402		0.303		0.570		0.064		0.385	
F-statistic	4.95***		3.56***		5.40***		1.40***		4.68***	

Panel B: Accruals-based earnings management regressions with retained KPMG clients only

	Model 1		Model 2		Model 3		Model 4		Model 5	
	DV = <i>Disc_Acc</i>		DV = $ Disc_Acc $		DV = <i>Total_Acc</i>		DV = <i>AWCA</i>		DV = $ AWCA $	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
<i>Intercept</i>	0.004	0.233	0.092 ***	6.981	-0.026	-0.915	0.026	1.318	0.125 ***	5.233
<i>KPMGret</i>	-0.001	-0.262	0.003	0.670	0.002	0.336	-0.002	-0.289	0.003	0.369
<i>KPMGret · 2014</i>	-0.012	-1.096	-0.003	-0.324	-0.009	-0.783	0.014	0.620	0.019	1.328
<i>KPMGret · 2017</i>	0.012 *	1.748	0.003	0.503	0.010	1.186	0.032	1.599	0.014	1.062
Controls	Yes		Yes		Yes		Yes		Yes	
Industry	Yes		Yes		Yes		Yes		Yes	
Year	Yes		Yes		Yes		Yes		Yes	

Interactions	-	-	Yes	-	-
N	2,405	2,405	2,405	2,405	2,405
Adjusted R ²	0.247	0.121	0.490	0.069	0.209
F-statistic	23.48***	10.44***	7.61***	6.07***	19.14***

Panel C: Accruals-based earnings management regressions with 2016 as treatment period for event 2

	Model 1		Model 2		Model 3		Model 4		Model 5	
	DV = <i>Disc_Acc</i>		DV = $ Disc_Acc $		DV = <i>Total_Acc</i>		DV = <i>AWCA</i>		DV = $ AWCA $	
	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
<i>Intercept</i>	0.004	0.256	0.092 ***	6.957	-0.026	-0.901	0.025	1.248	0.125 **	5.229
<i>KPMG</i>	-0.002	-0.481	0.004	1.092	<0.001	0.073	0.005	0.718	0.002	0.358
<i>KPMG · 2014</i>	-0.007	-0.748	-0.006	-0.703	-0.006	-0.566	-0.003	-0.131	0.022	1.530
<i>KPMG · 2016</i>	0.016 **	1.985	-0.004	-0.903	0.014	1.438	-0.011	-0.737	0.004	0.443
Controls	Yes		Yes		Yes		Yes		Yes	
Industry	Yes		Yes		Yes		Yes		Yes	
Year	Yes		Yes		Yes		Yes		Yes	
Interactions	-		-		Yes		-		Yes	
N	2,405		2,405		2,405		2,405		2,405	
Adjusted R ²	0.247		0.121		0.570		0.064		0.490	
F-statistic	23.49***		10.45***		5.40***		1.40***		7.61***	

Notes: Panel A present OLS regression results from re-estimating all models from Table 10 with client firm fixed effects instead of industry fixed effects. Panel B present OLS regression results from re-estimating all models from Table 10 with different definitions of the variables of interest. $KPMGret_{i,t}$ is an indicator variable that is coded as one if KPMG is the auditor of client i in years t and $t-1$ and zero otherwise. Panel B present OLS regression results from re-estimating all models from Table 10 with 2016 instead of 2017 as the treatment year for event 2. For sake of brevity, we do not report control variables. We use the same control variables as in the respective model in Table 10. We winsorize all continuous variables at the first and 99th percentile and compute robust standard errors clustered by client. We define industry based on the 16 sector indices of Frankfurt Stock Exchange and manually backfill missing industries based on SIC codes, if available, or a consideration of the firm's business description. *, **, and *** denotes significance at the 10%-, 5%-, and 1%-level, respectively, of two-tailed tests. Variable definitions are in section 5.6.

Panel A presents the results when we re-estimate all of our models and include client firm fixed effects instead of industry fixed effects. This tight fixed effects specification creates a within-client comparison, controlling for any unobserved time-invariant omitted client characteristics. For event 2, results become slightly stronger economically, and slightly stronger (weaker) statistically for the abnormal accruals and total accruals models (abnormal working capital accruals model). In addition, *KPMG · 2014* becomes positive and significant in the absolute abnormal working capital accruals model, but this effect is not consistent across the other four models. Overall, this robustness test corroborates the results from our analysis in Table 10.

Panel B shows the results when we repeat our estimations but only include retained KPMG clients that have chosen KPMG as their auditor both in the event year and in the treatment year. The motivation for this specification is to test whether effects are merely driven by those retained clients or also by newly acquired clients who have arguably been unaffected by the short-term negative market reactions we studied above. Results for event 2 become considerably weaker economically and statistically, and the coefficient of *KPMG · 2017* only remains significant in the signed discretionary accruals model. Therefore, it seems that new clients contribute to the overall effect we find in Table 10.

Finally, Panel C of Table 11 displays the results when we use 2016 as our treatment year for event 2 instead of 2017, an alternative we introduced in our discussion of the timeline of events above. Although we argue that we expect the highest likelihood to observe effects in 2017, we cannot rule out that effects, if any, already manifest in 2016. We do not change anything in the specification for event 1. Results for event 2 (in this case, coefficients of *KPMG · 2016*) are economically and statistically weaker than in our original specification. *KPMG · 2016* is only significant for the signed discretionary accruals model. Unsurprisingly, results are unchanged for *KPMG · 2014*. Although there is some indication of an effect, this much less consistent results pattern compared to using 2017 as a treatment year gives us confidence for our original design choice.

5.5 Conclusion

We analyze stock market reactions to KPMG audit clients and consequences to KPMG's ability to retain and gain clients, their audit fees and audit quality, following two

business scandals in Germany with high public exposure, in which KPMG was involved as a provider of non-audit services. To the best of our knowledge, this is the first study in a low-litigation setting that tests whether the reputation hypothesis as a driver of audit quality also applies to the effects of NAS failures on auditor reputation. We hypothesize and find that KPMG audit clients suffer abnormal negative returns on the initial events of both scandals. For the longer-term consequences, we do not find any evidence for KPMG losing more or gaining fewer clients or for a change in their audit fees. However, we find higher signed discretionary accruals, total accruals, and signed abnormal working capital accruals for KPMG clients in the first full year after the second event. This indicates that KPMG clients in the year after the second event have a higher tendency to use accruals to report higher earnings.

Our paper makes the following contributions. As our low-litigation setting allows us to use short-term event studies to test the reputation hypothesis, we add to non-result in Baugh et al. (2018) to show that NAS failures seem to affect shareholder's short-term perceptions although they do not seem to affect audit committees and client in the longer run. This may also be relevant for regulators, as earlier research has shown that reputation may substitute for a lack of litigation in a mere audit setting, but our results indicate that it only partly substitutes potential litigation risks in a setting that considers all services of an accounting firm. Finally, we add to an emerging literature that extends the reputation hypothesis beyond the mere auditing context, as we show that the "being known for something" dimension of organizational reputation applies to accounting firms at least to some extent, in addition to earlier findings by Donelson et al. (2019) that the "general favorability" dimension matters for accounting firms.

Our paper has some limitations, which indicate future research opportunities. First, some of our robustness tests of the short-term events study indicate that results for event 2 are sensitive to design choices. Besides, we are unable to identify subsamples with particularly pronounced abnormal returns. Together, this raises some concern that we pick up a spurious relation. However, we carefully exclude confounding events and are not aware of any alternative explanation that suggests spurious results. Moreover, the use of our pseudo-events gives us additional comfort that our results are not random. Nevertheless, future research could identify a larger number of events or

additional cross-sectional variation to study subsamples with presumably larger effects to explore further the robustness of the results empirically and theoretically.

Second, as discussed above, the German setting may bias against strong reputation effects for auditors, because ownership is generally concentrated in the hands of families, banks, and companies with insider knowledge and monitoring power, which do not have to rely on auditing as much as other capital market participants (Quick et al., 2018; Weber et al., 2008). Their involvement in boards and audit committees may bias against finding longer-term economic consequences to KPMG's audit practice stemming from these non-audit service events. Moreover, although our search procedure showed that an extensive search only produced two clean events, the events are of limited scope and may be less severe and prominent than those used in earlier literature. Together with theoretical arguments for the relatively high cost of switching auditors (Gerakos and Syverson, 2015), this suggests that our events may not be severe enough to induce observable economic consequences. Therefore, we cannot make strong conclusions from our non-results. Future research could identify more severe events of NAS failure in a low-litigation environment to provide additional meaningful analyses.

Furthermore, while our findings should be relevant to comparable settings with low litigation risk such as Japan and Continental Europe (Quick et al., 2018), they may not generalize to Anglo-Saxon settings. Besides, as there are many types of NAS and, coincidentally, our events cover two very similar NAS engagements, it is unclear whether these findings also apply to other NAS that investors and audit committees may perceive very differently concerning their effects on auditing. Earlier NAS literature has shown significant differences between different types of NAS and their effects on investor perceptions of audit quality (e.g., Friedrich et al., 2020). In addition, both of our events involved KPMG and a political actor making a bad business decision. The latter is likely a reason for the high media coverage. Therefore, future research should find a procedure to identify a larger number of relevant scandals with hopefully more variation in their scandal characteristics. This might even allow for a cross-sectional analysis investigating which scandal characteristics drive potential reputation effects.

5.6 Appendix B: Variable definitions

Table 12: Variable definitions for Chapter 5

Variable	Definition
<i>Auditorchg</i>	Dummy variable, coded as 1 if a firm is audited by a new auditor for the first time, and 0 otherwise
<i>LagKPMG</i>	Dummy variable, coded as 1 if a firm has been audited by KPMG in the prior year, and 0 otherwise
<i>KPMG</i>	Dummy variable, coded as 1 if a firm is audited by KPMG, and 0 otherwise
<i>Size</i>	Natural logarithm of (1 + total assets)
<i>Growth</i>	Total assets minus lagged total assets, scaled by lagged total assets
<i>Leverage</i>	Total debt scaled by total assets
Δ <i>Leverage</i>	Leverage minus lagged leverage
<i>ROA</i>	Return on assets, measured as net income scaled by the average of total assets and lagged total assets
<i>Loss</i>	Dummy variable, coded as 1 if net income is negative, and 0 otherwise
<i>LnAuditFee</i>	Natural logarithm of (1 + audit fees)
<i>CFO</i>	Cash flow from operations scaled by total assets
<i>MTB</i>	Market value of equity scaled by book value of equity
<i>ZScore</i>	$1.2 \cdot (\text{current assets} - \text{short-term debt})/\text{total assets} + 1.4 \cdot \text{retained earnings}/\text{total assets} + 3.3 \cdot \text{EBIT}/\text{total assets} + 0.6 \cdot \text{Market value of equity}/\text{total liabilities} + 0.999 \cdot \text{sales}/\text{total assets}$
<i>InvRec</i>	Inventories + receivables, scaled by total assets
<i>Big4</i>	Dummy variable, coded as 1 if a firm is audited by a Big 4 auditor, and 0 otherwise
<i>Assurance</i>	Fees paid to the auditor for other assurance services
<i>Advisory</i>	Fees paid to the auditor other services
<i>Tax</i>	Fees paid to the auditor for tax services
<i>IndShare</i>	Sum of audit fees for all available observations audited by a given auditor in an industry-year, scaled by sum of audit fees from all available observations in an industry-year
<i>TotalAcc</i>	Total accruals, measured as cash flow from operations minus net income, scaled by lagged total assets
<i>1/Assets</i>	Inverse of lagged total assets
Δ <i>SalesRec</i>	Difference of total revenues minus lagged total revenues and total receivables minus lagged total receivables, scaled by lagged total assets
<i>PPE</i>	Property, plant, and equipment scaled by lagged total assets
<i>AWCA</i>	Abnormal working capital accruals, measured as (current assets + short-term debt – cash) – (lag of (current assets + short-term debt – cash) scaled by sales over lagged sales). The result is scaled by lagged total assets.
$ AWCA $	Absolute value of abnormal working capital accruals
<i>DiscAcc</i>	Discretionary accruals, residual from estimating a performance-adjusted Jones Model as described above
$ DiscAcc $	Absolute value of discretionary accruals

<i>SalesGrowth</i>	Sales minus lagged sales, scaled by lagged sales
<i>NASRatio</i>	Sum of Assurance, Consling, and Tax, scaled by audit fees

Chapter 6 Overall Conclusion

In the next two sections, I briefly discuss the overall contributions and limitations of my dissertation, respectively. My focus is on interdependencies between the single research papers, and I do not discuss stand-alone contributions or limitations. I present such a detailed discussion in Chapters 2, 3, 4, and 5, where I report the full papers. To reconcile my contributions with my results, consider the overview in Table 1.

6.1 Contributions

In sum, my four papers contribute empirical results from different perspectives to the accounting literature concerned with corporate fraud. Nevertheless, the results are interwoven, which provides a potentially more significant contribution than a separate consideration of the single papers. Regarding market forces and deterrence through adverse reputation effects, my first and last paper explore the boundaries of existing results in the literature that markets incentivize rule-abiding behavior through the importance of reputation. I show that these established effects spill over beyond the narrow domain of the market where a misconduct happened to other markets in which the organization is active. However, these spillovers are limited. I also show that the existing market forces could better discourage corporate misconduct if decision-makers' biased beliefs about rare fraud could be corrected.

A second contribution concerns the role detection plays in curbing misconduct. I first establish that beliefs about detection are much higher than low factual detection rates would suggest. This is important given that three of my papers study situations of detected misconduct. I then contribute that, conditional on detection, market forces may induce surprisingly desirable behaviors. Specifically, the third paper establishes a result that, if detection is certain, markets could gravitate to a transparent and high-quality investigation strategy of internal misconduct suspicions. To capitalize on such benefits, regulation could first focus on improving detection rates. One such initiative is the recent EU whistleblowing regulation (Directive (EU) 2019/1937). Results from my second paper focusing on a conceptually similar regulation indicate that such regulation may be effective in improving detection. If it is indeed, it may start multiple mutually reinforcing mechanisms. It may directly reveal more misconduct cases, improving first-hand experiences. If these are discussed more in public, and the

regulatory change raises awareness of such issues, second-hand experience may also increase. As higher detection rates may promote more transparent investigations, it may also indirectly increase the number of revelations. In sum, beliefs may update, and related mechanisms that suggest deterrence through belief-updating may activate.

Admittedly, this is a hypothetical and idealized discussion of how a single regulatory change may initiate dynamics that profit from existing market forces that are ineffective under current environmental conditions. Nevertheless, it shows how a contribution of different angles may open up a much more informed and contextualized discussion of regulatory and market effects. As such, I contribute to the literature that often focuses on single and separate analyses and has a comparably weak tradition of tying research together. However, seeing the whole picture is essential (Knechel et al., 2013). While my dissertation is only illustrative of this challenge, it is an example of how even small connections can allow for a much richer perspective.

6.2 Limitations and Directions for Future Research

Generally, the scope and topical domains of the single papers limit my dissertation. Although I discuss them in concert above, this requires the assumption that the different contexts I study and the different assumptions I make in some papers, but relax in others, will still be valid when taken together. This is by no means natural. However, in the limited scope of the dissertation, I could not address this issue. Instead, it is a natural problem in most research that has to work with limited available data from different contexts. Therefore, future research could find contexts suitable to provide more comprehensive analyses, conceptually, and in terms of data availability. This comprises the triangulation of theoretical perspectives and research methodologies.

More specifically, my dissertation is subject to several caveats regarding assumptions I make on an overall theoretical level but do not address empirically in any of the papers. First, all of my considerations of punishment (deterrence theory and reputation theory) exclude the possibility of increased regulatory punishment, either by way of experimental control or as an institutional feature of the setting (Germany as a low litigation setting in Chapter 5). I assume that regulatory and market punishments are substitutes, just as earlier audit research suggests that litigation can partly substitute for reputation, and vice versa. However, I do not empirically study this question.

Additionally, the conjecture from earlier audit research is primarily based on different settings rather than on an empirical test of whether the two mechanisms are substitutive or otherwise related. Therefore, future research could help clarify the relation between these two theories in a single setting. Ideally, it would conduct such tests in multiple settings that differ regarding their historical or natural tendency to either of the two theoretical effects. This would help understand whether these theories are universal or context-dependent.

Second, my overarching research question acknowledges the necessity that, at reasonable costs, there will always be some residual misconduct. Recent research has called for acknowledging and analyzing this phenomenon (Camfferman and Wielhouwer, 2019). While this is an important issue, only my first paper gives some indication that there is more room to curb misconduct. It shows that practice may not have reached the optimum of residual misconduct. Other than that, my research is silent to the empirical manifestation of it. Future research and future regulation could acknowledge this phenomenon and study or address it empirically.

Third, my overarching research question, and hence all of my papers, just consider whether the respective forces I study changes economic outcomes but not whether such change is optimal. For instance, in Chapter 4, I study whether unregulated forensic services provide for a valuable signal. This does not exclude that a regulated forensic service would provide an even more valuable signal. Hence, my results speak to the effectiveness of the forces I study but do not necessarily speak to these forces' relative effectiveness above alternative forces. An exception might be Chapter 3, where the generic process I develop covers a wide range of different strengths of both regulatory and market forces. Future research could go beyond investigating the effect of single forces in isolation and, instead, study the relative effects of alternative forces.

Finally, due to the limited nature of the single papers, I only study some decisions and their outcomes while assuming or prompting other decisions. In experiments, this is necessary to arrive at a clean design. Nevertheless, it limits the boundaries within which results can provide firm inferences. Of the process from the misconduct decision to its eventual outcomes, I study decisions to commit misconduct in Chapter 2. I then explore decisions to monitor in Chapter 3. I finally study outcomes in Chapter 4. However, I

never study the decisions to investigate and disclose, as I just manipulate them in Chapter 4. Although I back this choice with existing evidence, I eventually never empirically analyze whether my manipulations are viable, or even frequent, choices in practice. Future research could fill this missing link.

In a final and overall conclusion, despite its limitations, my dissertation could be of interest to accounting researchers, the profession, companies and their stakeholders that are all concerned with misconduct. I provide a simple yet comprehensive theoretical lens on the decision to misbehave. I then provide some empirical insights into how market forces, regulatory change, and the auditor as an institutional intermediary may gradually improve the economic equilibrium between costs from misconduct and costs from curbing it. My dissertation ends on the note that market forces should not be underestimated, but regulatory intervention is surely unavoidable to make markets follow the dynamic consensus of the public regarding what is desirable behavior.

References

- Agarwal, J., Osiyevskyy, O., Feldman, P.M., 2015. Corporate Reputation Measurement: Alternative Factor Structures, Nomological Validity, and Organizational Outcomes. *Journal of Business Ethics* 130 (2), 485–506.
- Akerlof, G., 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics* 84 (3), 488–500.
- Altman, E.I., 1993. *Corporate financial distress and bankruptcy: A complete guide to predicting and avoiding distress and profiting from bankruptcy*. John Wiley and Sons, New York.
- Amiram, D., Bozanic, Z., Cox, J.D., Dupont, Q., Karpoff, J.M., Sloan, R.G., 2018. Financial Reporting Fraud and Other Forms of Misconduct: A Multidisciplinary Review of the Literature. *Review of Accounting Studies* 23 (2), 732–783.
- Amiram, D., Huang, S., Rajgopal, S., 2020. Does financial reporting misconduct pay off even when discovered? *Review of Accounting Studies* 25 (3), 811–854.
- Andrade, J.A., 2015. Reconceptualising Whistleblowing in a Complex World. *Journal of Business Ethics* 128 (2), 321–335.
- Anger, H., 2018. GroKo will unternehmensinterne Ermittlungen verbessern. *Handelsblatt*.
- Arlen, J., 2012. *Corporate Criminal Liability: Theory and Evidence*, in: Harel, A., Hylton, K.N. (Eds.), *Research Handbook on the Economics of Criminal Law*. Edward Elgar, Cheltenham, UK, pp. 144–203.
- Asare, S.K., Wright, A., Zimbelman, M.F., 2015. Challenges Facing Auditors in Detecting Financial Statement Fraud: Insights from Fraud Investigations. *Journal of Forensic & Investigative Accounting* 7 (2), 63–112.
- Audit Analytics, 2018. EU Auditor Market Share: 2017 Audit Reports. <https://blog.auditanalytics.com/eu-auditor-market-share-2017-audit-reports/> (accessed November 20, 2020).
- Baber, W.R., Kumar, K.R., Verghese, T., 1995. Client Security Price Reactions to the Laventhol and Horwath Bankruptcy. *Journal of Accounting Research* 33 (2), 385–395.
- Ball, R., 2009. Market and Political/Regulatory Perspectives on the Recent Accounting Scandals. *Journal of Accounting Research* 47 (2), 277–323.

- Baron, R.M., Kenny, D.A., 1986. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51 (6), 1173–1182.
- Barton, J., Mercer, M., 2005. To blame or not to blame: Analysts' reactions to external explanations for poor financial performance. *Journal of Accounting and Economics* 39 (3), 509–533.
- Bartov, E., Marra, A., Momenté, F., 2020. Corporate Social Responsibility and the Market Reaction to Negative Events: Evidence from Inadvertent and Fraudulent Restatement Announcements. *The Accounting Review* (forthcoming).
- Baugh, M., Boone, J.P., Khurana, I.K., Raman, K.K., 2018. Did the 2005 Deferred Prosecution Agreement Adversely Impact KPMG's Audit Practice? *AUDITING: A Journal of Practice & Theory* 38 (1), 77–102.
- Becker, G.S., 1968. Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76 (2), 169–217.
- Benston, G.J., 1969. The Value of the SEC's Accounting Disclosure Requirements. *The Accounting Review* 44 (3), 515–532.
- Blackburne, T., Kepler, J.D., Quinn, P.J., Taylor, D., 2020. Undisclosed SEC Investigations. *Management Science* (forthcoming).
- Boone, J.P., Khurana, I.K., Raman, K.K., 2015. Did the 2007 PCAOB Disciplinary Order against Deloitte Impose Actual Costs on the Firm or Improve Its Audit Quality? *The Accounting Review* 90 (2), 405–441.
- Boone, J.P., Khurana, I.K., Raman, K.K., 2019. Audit Market Response to PCAOB Censures of US Big 4 Firms. *European Accounting Review* 28 (4), 621–658.
- Boritz, J.E., Kochetova-Kozloski, N., Robinson, L., 2015. Are Fraud Specialists Relatively More Effective than Auditors at Modifying Audit Programs in the Presence of Fraud Risk? *The Accounting Review* 90 (3), 881–915.
- Brody, R.G., Melendy, S.R., Perri, F.S., 2012. Commentary from the American Accounting Association's 2011 Annual Meeting Panel on Emerging Issues in Fraud Research. *Accounting Horizons* 26 (3), 513–531.
- Brown, D.L., Shu, S.Z., Soo, B.S., Trompeter, G.M., 2013. The Insurance Hypothesis: An Examination of KPMG's Audit Clients around the Investigation and Settlement of the Tax Shelter Case. *AUDITING: A Journal of Practice & Theory* 32 (4), 1–24.

- Brückner, R., Lehmann, P., Schmidt, M.H., Stehle, R., 2015. Another German Fama/French Factor Data Set. *Working Paper*.
- Button, M., Wakefield, A., Brooks, G., Lewis, C., Shepherd, D., 2015. Confronting the “fraud bottleneck”: private sanctions for fraud and their implications for justice. *Journal of Criminological Research, Policy and Practice* 1 (3), 159–174.
- Camfferman, K., Wielhouwer, J.L., 2019. 21st century scandals: towards a risk approach to financial reporting scandals. *Accounting and Business Research* 49 (5), 503–535.
- Carhart, M.M., 1997. On Persistence in Mutual Fund Performance. *The Journal of Finance* 52 (1), 57–82.
- Cascino, S., Clatworthy, M., García Osma, B., Gassen, J., Imam, S., Jeanjean, T., 2014. Who Uses Financial Reports and for What Purpose? Evidence from Capital Providers. *Accounting in Europe* 11 (2), 185–209.
- Chen, W.E., Hribar, P., Melessa, S., 2018. Incorrect Inferences When Using Residuals as Dependent Variables. *Journal of Accounting Research* 56 (3), 751–796.
- Christensen, J.A., Byington, J.R., Blalock, T.J., 2005. Sarbanes-Oxley: Will you need a forensic accountant? *Journal of Corporate Accounting & Finance* 16 (3), 69–75.
- Claeys, A.-S., 2017. Better safe than sorry: Why organizations in crisis should never hesitate to steal thunder. *Business Horizons* 60 (3), 305–311.
- Claeys, A.-S., Cauberghe, V., Pandelaere, M., 2016. Is old news no news? The impact of self-disclosure by organizations in crisis. *Journal of Business Research* 69 (10), 3963–3970.
- Coffee, J.C., 2019. Why do auditors fail? What might work? What won't? *Accounting and Business Research* 49 (5), 540–561.
- Cooper, D.J., Dacin, T., Palmer, D., 2013. Fraud in accounting, organizations and society: Extending the boundaries of research. *Accounting, Organizations and Society* 38 (6), 440–457.
- Cooper, D.J., Robson, K., 2006. Accounting, professions and regulation: Locating the sites of professionalization. *Accounting, Organizations and Society* 31 (4), 415–444.
- Cornish, D.B., Clarke, R.V., 1987. Understanding Crime Displacement: An Application of Rational Choice Theory. *Criminology* 25 (4), 933–948.
- Corona, C., Randhawa, R.S., 2018. The Value of Confession: Admitting Mistakes to Build Reputation. *The Accounting Review* 93 (3), 133–161.

- Corrado, C.J., 1989. A nonparametric test for abnormal security-price performance in event studies. *Journal of Financial Economics* 23 (2), 385–395.
- Craswell, A.T., Francis, J.R., Taylor, S.L., 1995. Auditor brand name reputations and industry specializations. *Journal of Accounting and Economics* 20 (3), 297–322.
- Da, Z.H., Engelberg, J., Gao, P., 2011. In Search of Attention. *The Journal of Finance* 66 (5), 1461–1499.
- Darby, M.R., Karni, E., 1973. Free Competition and the Optimal Amount of Fraud. *The Journal of Law and Economics* 16 (1), 67–88.
- Davis, L.R., Simon, D.T., 1992. The Impact of SEC Disciplinary Actions on Audit Fees. *AUDITING: A Journal of Practice & Theory* 11 (1), 58–68.
- DeAngelo, L.E., 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3 (3), 183–199.
- Dee, C.C., Lulseged, A., Zhang, T., 2011. Client Stock Market Reaction to PCAOB Sanctions Against a Big 4 Auditor. *Contemporary Accounting Research* 28 (1), 263–291.
- DeFond, M.L., Zhang, J., 2014. A review of archival auditing research. *Journal of Accounting and Economics* 58 (2), 275–326.
- Deutsche Presseagentur, 2016. Nur KPMG macht mit dem Hahn ein gutes Geschäft. *Frankfurter Allgemeine Zeitung*.
- Deutsche Presseagentur, 2019. EU-Gericht: Nürburgring-Verkauf war rechtens. *Zeit online*.
- DiPiazza, S.A., McDonnell, D., Parrett, W.G., Rake, M.D., Samyn, F., Turley, J.S., 2006. Global capital markets and the global economy: A vision from the CEOs of the International Audit Network: Third Global Public Policy Symposium November 2006. Paris.
- Dolnik, L., Case, T.I., Williams, K.D., 2003. Stealing thunder as a courtroom tactic revisited: Processes and boundaries. *Law and Human Behavior* 27 (3), 267–287.
- Donelson, D.C., Ege, M., Leiby, J., 2019. Audit Firm Reputational Consequences of Alleged Non-Accounting Misconduct by Clients: How Bargaining Power Temporarily Shifts Around Securities Litigation. *AUDITING: A Journal of Practice & Theory* 38 (4), 77–100.
- Doogar, R., Sivadasan, P., Solomon, I., 2015. Audit fee residuals: Costs or rents? *Review of Accounting Studies* 20 (4), 1247–1286.

- Downs, A., 1957. An Economic Theory of Political Action in a Democracy. *Journal of Political Economy* 65 (2), 135–150.
- Dyck, A., Morse, A., Zingales, L., 2017. How Pervasive is Corporate Fraud? *Working Paper*.
- Dye, R.A., 1993. Auditing Standards, Legal Liability, and Auditor Wealth. *Journal of Political Economy* 101 (5), 887–914.
- Eagly, A.H., Wood, W., Chaiken, S., 1978. Causal inferences about communicators and their effect on opinion change. *Journal of Personality and Social Psychology* 36 (4), 424–435.
- Easley, R.W., Bearden, W.O., Teel, J.E., 1995. Testing predictions derived from inoculation theory and the effectiveness of self-disclosure communications strategies. *Journal of Business Research* 34 (2), 93–105.
- Eisenhardt, K.M., 1989. Agency Theory: An Assessment and Review. *Academy of Management Review* 14 (1), 57–74.
- Elliott, W.B., Grant, S.M., Hodge, F.D., 2018. Negative News and Investor Trust: The Role of \$Firm and #CEO Twitter Use. *Journal of Accounting Research* 56 (5), 1483–1519.
- Elson, C.M., Ferrere, C.K., Goossen, N.J., 2015. The Bug At Volkswagen: Lessons in Co-Determination, Ownership, and Board Structure. *Journal of Applied Corporate Finance* 27 (4), 36–43.
- ENSSCPA, 2017. A New Growth in the Forensic Accounting Industry. <https://ensscpa.com/a-new-growth-in-the-forensic-accounting-industry/> (accessed 20 November 2020).
- EY, 2018. Wem gehört der DAX? Analyse der Aktionärsstruktur der DAX-Unternehmen 2017. http://docs.dpaq.de/13676-ey-wem-gehört-der-dax_1_.pdf (accessed 20 November 2020).
- Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1), 3–56.
- Fennis, B.M., Stroebe, W., 2014. Softening the Blow: Company Self-Disclosure of Negative Information Lessens Damaging Effects on Consumer Judgment and Decision Making. *Journal of Business Ethics* 120 (1), 109–120.
- Flesher, D.L., Flesher, T.K., 1986. Ivar Kreuger's Contribution to U.S. Financial Reporting. *The Accounting Review* 61 (3), 421–434.

- Francis, J.R., Wang, D., 2008. The Joint Effect of Investor Protection and Big 4 Audits on Earnings Quality around the World. *Contemporary Accounting Research* 25 (1), 157–191.
- Friedrich, C., Schmidt, F., Quick, R., 2020. Auditor-Provided Non-Audit Services and Perceived Audit Quality: Evidence from the Cost of Debt and Equity Capital. *Working Paper*.
- Friestad, M., Wright, P., 1994. The Persuasion Knowledge Model: How People Cope with Persuasion Attempts. *Journal of Consumer Research* 21 (1), 1–31.
- Fröndhoff, B., 2018. Die „Big Four“ der Wirtschaftsprüfer legen ein starkes Jahr hin. *Handelsblatt*.
- Gerakos, J., Syverson, C., 2015. Competition in the Audit Market: Policy Implications. *Journal of Accounting Research* 53 (4), 725–775.
- Gietzmann, M.B., Quick, R., 1998. Capping auditor liability: The German experience. *Accounting, Organizations and Society* 23 (1), 81–103.
- Glaser, B.G., Strauss, A.L., 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Aldine, New York.
- Gottschalk, P., 2017. Private police legitimacy: the case of internal investigations by fraud examiners. *Policing: An International Journal* 40 (3), 628–640.
- Grant, S.M., Hodge, F.D., Sinha, R.K., 2018. How disclosure medium affects investor reactions to CEO bragging, modesty, and humblebragging. *Accounting, Organizations and Society* 68-69, 118–134.
- Grenier, J., Pomeroy, B., Reffett, A., 2012. Speak Up or Shut Up? The Moderating Role of Credibility on Auditor Remedial Defense Tactics. *AUDITING: A Journal of Practice & Theory* 31 (4), 65–83.
- Gul, F.A., Zhou, G., Zhu, X., 2013. Investor Protection, Firm Informational Problems, Big N Auditors, and Cost of Debt around the World. *AUDITING: A Journal of Practice & Theory* 32 (3), 1–30.
- Hackethal, A., Schmidt, R.H., Tyrell, M., 2005. Banks and German Corporate Governance: on the way to a capital market-based system? *Corporate Governance: An International Review* 13 (3), 397–407.
- Hay, D.C., 2013. Further Evidence from Meta-Analysis of Audit Fee Research. *International Journal of Auditing* 17 (2), 162–176.

- Hegazy, S., Sangster, A., Kotb, A., 2017. Mapping forensic accounting in the UK. *Journal of International Accounting, Auditing and Taxation* 28 (Supplement C), 43–56.
- Heißner, S., 2014. *Erfolgsfaktor Integrität*, 2nd ed. Springer Gabler, Wiesbaden.
- Hogan, C.E., Rezaee, Z., Riley, Richard A., Velury, U.K., 2008. Financial Statement Fraud: Insights from the Academic Literature. *AUDITING: A Journal of Practice & Theory* 27 (2), 231–252.
- Huber, W.D., 2012. Is forensic accounting in the United States becoming a profession? *Journal of Forensic & Investigative Accounting* 4 (1), 255–284.
- IBISWorld, 2017. Forensic Accounting Services: Market Research Report. <https://www.ibisworld.com/united-states/market-research-reports/forensic-accounting-services-industry/> (accessed 20 November 2020).
- Incardona, J., Kannan, Y., Premuroso, R., Higgs, J.L., Huang, I., 2014. Taxing audit markets and reputation: An examination of the U.S. tax shelter controversy. *Journal of International Accounting, Auditing and Taxation* 23 (1), 18–31.
- Jamal, K., 2008. Mandatory Audit of Financial Reporting: A Failed Strategy for Dealing with Fraud. *Accounting Perspectives* 7 (2), 97–110.
- Jenkins, J.G., Negangard, E.M., Oler, M.J., 2018. Getting Comfortable on Audits: Understanding Firms' Usage of Forensic Specialists. *Contemporary Accounting Research* 35 (4), 1766–1797.
- Jensen, M.C., Meckling, W.H., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3 (4), 305–360.
- Jones, E.E., Davis, K.E., 1965. *From acts to dispositions: The attribution process in person perception*, in: Berkowitz, L. (Ed.), *Advances in experimental social psychology*, vol. 2. Academic Press, New York, pp. 219–266.
- Karpoff, J.M., Lee, D.S., Martin, G.S., 2008a. The consequences to managers for financial misrepresentation. *Journal of Financial Economics* 88 (2), 193–215.
- Karpoff, J.M., Lee, D.S., Martin, G.S., 2008b. The Cost to Firms of Cooking the Books. *Journal of Financial and Quantitative Analysis* 43 (3), 581–611.
- Karpoff, J.M., Lott, J.R., 1993. The Reputational Penalty Firms Bear from Committing Criminal Fraud. *The Journal of Law and Economics* 36 (2), 757–802.
- Kausar, A., Shroff, N., White, H., 2016. Real effects of the audit choice. *Journal of Accounting and Economics* 62 (1), 157–181.

- Khurana, I.K., Raman, K.K., 2004. Litigation Risk and the Financial Reporting Credibility of Big 4 versus Non-Big 4 Audits: Evidence from Anglo-American Countries. *The Accounting Review* 79 (2), 473–495.
- Klein, B., Leffler, K.B., 1981. The Role of Market Forces in Assuring Contractual Performance. *Journal of Political Economy* 89 (4), 615–641.
- Knechel, W.R., 2016. Audit Quality and Regulation. *International Journal of Auditing* 20 (3), 215–223.
- Knechel, W.R., Krishnan, G.V., Pevzner, M., Shefchik, L.B., Velury, U.K., 2013. Audit Quality: Insights from the Academic Literature. *AUDITING: A Journal of Practice* 32 (Supplement 1), 385–421.
- Kolari, J.W., Pynnönen, S., 2010. Event Study Testing with Cross-sectional Correlation of Abnormal Returns. *The Review of Financial Studies* 23 (11), 3996–4025.
- Koonce, L., Lipe, M.G., 2010. Earnings Trend and Performance Relative to Benchmarks: How Consistency Influences Their Joint Use. *Journal of Accounting Research* 48 (4), 859–884.
- Kothari, S.P., Leone, A.J., Wasley, C.E., 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1), 163–197.
- KPMG Law, 2019. Ein Unternehmensstrafrecht für Deutschland. <https://kpmg-law.de/mandanten-information/ein-unternehmensstrafrecht-fuer-deutschland/> (accessed 20 November 2020).
- La Porta, R., Lopez-De-Silanes, F., Shleifer, A., 1999. Corporate Ownership Around the World. *The Journal of Finance* 54 (2), 471–517.
- Lange, D., Lee, P.M., Dai, Y., 2011. Organizational Reputation: A Review. *Journal of Management* 37 (1), 153–184.
- Lee, S.Y., 2016. Weathering the crisis: Effects of stealing thunder in crisis communication. *Public Relations Review* 42 (2), 336–344.
- Lennox, C.S., 1999. Audit Quality and Auditor Size: An Evaluation of Reputation and Deep Pockets Hypotheses. *Journal of Business Finance & Accounting* 26 (7-8), 779–805.
- Lennox, C.S., Pittman, J.A., 2011. Voluntary Audits versus Mandatory Audits. *The Accounting Review* 86 (5), 1655–1678.

- Leuz, C., 2018. Evidence-based policymaking: promise, challenges and opportunities for accounting and financial markets research. *Accounting and Business Research* 48 (5), 582–608.
- Levick, R., 2017. Volkswagen Under Siege: Unprecedented Search, Unanswered Questions. *Forbes*.
- Lewis, P.V., 1985. Defining ‘business ethics’: Like nailing jello to a wall. *Journal of Business Ethics* 4 (5), 377–383.
- MacKinlay, A.C., 1997. Event Studies in Economics and Finance. *Journal of Economic Literature* 35 (1), 13–39.
- Mansi, S.A., Maxwell, W.F., Miller, D.P., 2004. Does Auditor Quality and Tenure Matter to Investors? Evidence from the Bond Market. *Journal of Accounting Research* 42 (4), 755–793.
- Markowitz, H., 1959. *Portfolio Selection, Efficient Diversification of Investments*. John Wiley & Sons, New York.
- Menon, K., Williams, D.D., 1994. The Insurance Hypothesis and Market Prices. *The Accounting Review* 69 (2), 327–342.
- Mohrmann, U., 2017. What are the Drivers of Audit Quality after an Auditor Change? Evidence from Voluntary and Mandatory Auditor Switches. *Working Paper*.
- Morales, J., Gendron, Y., Guénin-Paracini, H., 2014. The construction of the risky individual and vigilant organization: A genealogy of the fraud triangle. *Accounting, Organizations and Society* 39 (3), 170–194.
- Morris, R.D., 1987. Signalling, Agency Theory and Accounting Policy Choice. *Accounting and Business Research* 18 (69), 47–56.
- Ott, K., 2017. Wie sich VW und Audi hinter Anwaltskanzleien verstecken. *Süddeutsche Zeitung*.
- Palan, S., Schitter, C., 2018. Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance* 17, 22–27.
- Peer, E., Brandimarte, L., Samat, S., Acquisti, A., 2017. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology* 70, 153–163.
- Pellens, B., Schmidt, A., 2014. *Verhalten und Präferenzen deutscher Aktionäre*, in: Kuhn, N., Leven, F.-J. (Eds.), *Studien des deutschen Aktieninstituts*. Frankfurt.
- Petermann, A., 2016. Alarmglocken verhallten ungehört. *Deutschlandfunk*.

- Posner, R.A., 1974. Theories of Economic Regulation. *The Bell Journal of Economics and Management Science* 5 (2), 335–358.
- PWC, 2018. Wirtschaftskriminalität 2018. Mehrwert von Compliance - forensische Erfahrungen, Frankfurt am Main, Halle.
- PWC, 2020. 2020 PwC's Global Economic Crime and Fraud Survey, London.
- Quick, R., Schenk, N., Schmidt, F., Towara, T., 2018. The impact of corporate governance on auditor choice: evidence from Germany. *Journal of Management and Governance* 22 (2), 251–283.
- Quick, R., Schmidt, F., 2018. Do audit firm rotation, auditor retention, and joint audits matter? – An experimental investigation of bank directors' and institutional investors' perceptions. *Journal of Accounting Literature* 41, 1–21.
- Rapoport, M., 2018. How Did the Big Four Auditors Get \$17 Billion in Revenue Growth? Not From Auditing. *The Wall Street Journal*.
- Ratzinger-Sakel, N.V., Schönberger, M.W., 2018. The impact of auditor brand name on auditor remuneration in a large private client segment. *International Journal of Auditing* 22 (3), 536–553.
- Rezaee, Z., Crumbley, L., 2007. The Role of Forensic Auditing Techniques in Restoring Public Trust and Investor Confidence in Financial Information. *Forensic Examiner* 16 (1), 44–49.
- Rogers, J.L., Skinner, D.J., Zechman, S.L.C., 2017. Run EDGAR Run: SEC Dissemination in a High-Frequency World. *Journal of Accounting Research* 55 (2), 459–505.
- Schipper, K., Thompson, R., 1983. Evidence on the capitalized value of merger activity for acquiring firms. *Journal of Financial Economics* 11 (1), 85–119.
- Schmitt, J., 2016. KPMG verliert Mandat für Flughafen Hahn. *Finance*.
- Schmitt, J., 2019. PwC nicht mehr Marktführer im Beratungsgeschäft. *Finance*.
- Schwartz, L.H., 2020. Understanding the Statement on Standards for Forensic Services No. 1 and its implications for practice, curricula and research. *Journal of Forensic Accounting Research* (forthcoming).
- Sellhorn, T., 2020. Wirecard - der Weckruf: Skandal zeigt dringenden Bedarf nach mehr Transparenz und evidenzbasierter Regulierung auf. *Der Betrieb* (31), M4-M5.
- Selnes, F., 1993. An Examination of the Effect of Product Performance on Brand Reputation, Satisfaction and Loyalty. *European Journal of Marketing* 27 (9), 19–35.

- Serious Fraud Office, 2017. Deferred Prosecution Agreements. <https://www.sfo.gov.uk/publications/guidance-policy-and-protocols/deferred-prosecution-agreements/> (accessed 20 November 2020).
- Shapiro, C., 1983. Premiums for High Quality Products as Returns to Reputations. *The Quarterly Journal of Economics* 98 (4), 659–679.
- Sharpe, W.F., 1963. A Simplified Model for Portfolio Analysis. *Management Science* 9 (2), 277–293.
- Shibano, T., 1990. Assessing Audit Risk from Errors and Irregularities. *Journal of Accounting Research* 28 (Supplement), 110–140.
- Singleton, T.W., Singleton, A.J., 2010. *Fraud Auditing and Forensic Accounting*, 4th ed. John Wiley & Sons, Hoboken, New Jersey.
- Skinner, D.J., Srinivasan, S., 2012. Audit Quality and Auditor Reputation: Evidence from Japan. *The Accounting Review* 87 (5), 1737–1765.
- Smieliauskas, W., 2006. CAP Forum on Forensic Accounting in the Post-Enron World: Introduction and Commentary. *Canadian Accounting Perspectives* 5 (2), 239–256.
- Solomon, D.H., Soltes, E., 2019. Is “Not Guilty” the Same as “Innocent”? Evidence from SEC Financial Fraud Investigations. *Working Paper*.
- Streicher, H., Deveci, A., 2016. *Der deutsche Markt für Wirtschaftsprüfung und Steuerberatung*, in: Hossenfelder, J., Lünendonk, T. (Eds.), *Handbuch Wirtschaftsprüfung und Steuerberatung*. Haufe, Freiburg/München/Stuttgart, pp. 18–22.
- Trompeter, G.M., Carpenter, T.D., Desai, N., Jones, K.L., Riley, Richard A., 2013. A Synthesis of Fraud-Related Research. *AUDITING: A Journal of Practice & Theory* 32 (Supplement 1), 287–321.
- Watts, R.L., Zimmerman, J.L., 1978. Towards a Positive Theory of the Determination of Accounting Standards. *The Accounting Review* 53 (1), 112–134.
- Weber, J., Willenborg, M., Zhang, J., 2008. Does Auditor Reputation Matter? The Case of KPMG Germany and ComROAD AG. *Journal of Accounting Research* 46 (4), 941–972.
- Weiner, B., 1985. An attributional theory of achievement motivation and emotion. *Psychological Review* 92 (4), 548–573.
- Weiner, B., 1986. *An attributional theory of motivation and emotion*. Springer-Verlag, New York.

- Weiner, B., Graham, S., Peter, O., Zmuidinas, M., 1991. Public Confession and Forgiveness. *Journal of Personality* 59 (2), 281–312.
- Willenborg, M., 1999. Empirical Analysis of the Economic Demand for Auditing in the Initial Public Offerings Market. *Journal of Accounting Research* 37 (1), 225–238.
- Williams, J.W., 2005. Reflections on the Private Versus Public Policing of Economic Crime. *The British Journal of Criminology* 45 (3), 316–339.
- Williams, J.W., 2014. *The Private Eyes of Corporate Culture: The Forensic Accounting and Corporate Investigation Industry and the Production of Corporate Financial Security*, in: Walby, K., Lippert, R.K. (Eds.), *Corporate security in the 21st century—Theory and practice in international perspective*. Palgrave Macmillan, Hampshire, pp. 56–77.
- Williams, K.D., Bourgeois, M.J., Croyle, R.T., 1993. The effects of stealing thunder in criminal and civil trials. *Law and Human Behavior* 17 (6), 597–609.
- Wilson, T.E., Grimplund, R.A., 1990. An Examination of the Importance of an Auditor's Reputation. *AUDITING: A Journal of Practice & Theory* 9 (2), 43–59.
- Zerfaß, F., 2014. Eine Blamage für alle Beteiligten. *Wirtschaftswoche*.
- Zerfaß, F., 2015. Nächste Runde im Nürburgring-Skandal. *Wirtschaftswoche*.
- Zimbelman, M.F., 1997. The Effects of SAS No. 82 on Auditors' Attention to Fraud Risk Factors and Audit Planning Decisions. *Journal of Accounting Research* 35, 75–97.